**DISAGREEMENT AND BELIEF**

Ted Everett

(full draft, August 2012)

**CONTENTS**

[Preface](#preface)

[Introduction](#introduction)

[**1. Disagreement**](#disagreement)

1.1. The problem of disagreement

1.2. The usual explanations

1.3. Three dimensions of belief

[**2. Perception**](#perception)

2.1. Sensation, memory, and reasoning

2.2. Articulations of perceptive beliefs

2.3. The concept of knowledge

2.4. Skeptical problems and solutions

[**3. Testimony**](#testimony)

3.1. Testimony as perception

3.2. Testimony and induction

3.3. Other minds

3.4. The external world

3.5. Moral perception

[**4. Authority**](#authority)

4.1. Rational deference

4.2. Epistemic communities

4.3. Epistemic doctrines

4.4. Epistemic gravitation

[**5. Philosophy**](#philosophy)

5.1. Conflicting experts

5.2. Speculation

5.3. Opinions and arguments

5.4. Socratic and Cartesian methods in philosophy

[**6. Science**](#science)

6.1. Rational repression

6.2 Epistemic altruism

6.3. Dissent and authority in modern science

6.4. Consensus, controversy, and induction

[**7. Action**](#action)

7.1. Judgments

7.2. Convictions

7.3. Higher-order convictions

7.4 Autonomous integrity

7.5 Fanaticism and hypocrisy

[**8. Politics**](#politics)

8.1. Oppression and liberation

8.2. Democracy

8.3. Freedom

8.4. Equality

8.5. The reasonable society

[**9. Good Sense**](#goodsense)

9.1. Pure types of believer

9.2. The reasonable person

9.3. The seeker after wisdom

Afterword

Bibliography

Index

***Preface***

I have always found it hard to handle disagreements, and even harder to avoid them. I went to college during the great uproar over the war in Vietnam, a time when politics was never far from students’ minds. Like most of my friends in college and afterwards, I got into many heated arguments about the war, as well as race, sex, inequality, and all the other main themes of my generation’s politics. I found these arguments exciting and often went out of my way to provoke them, though without a clear idea of what I really wanted from them. Was it to win others over to my point of view? Was it to gain a better understanding of the issues? Or was it just to sharpen and show off my wits, so that people would think that I was smart? In any case, I found that I often felt really shaken and confused after such arguments. Shaken, because of the hard feelings they could bring about, to which I seemed to be unusually sensitive despite my own aggressiveness. And confused, because I simply couldn’t understand how bright, seemingly decent people could turn out to be so *blind* to what struck me as matters of plain fact and simple logic.

Some people shrug off disagreements like this with a simple “diagnosis” of their opponents, for example that they are *a bunch of morons*, *Nazis*, *out of their freaking minds*, or in some other way defective as human beings. But this option was generally closed to me because the people I argued with the most were typically my closest friends. They tended to take more or less the same positions on the issues of the day, and I thought that most of these positions were obviously wrong and ill-considered. But I couldn’t just diagnose my friends as stupid, crazy, wicked, or anything of the sort, because I knew them too well as rational and thoughtful individuals outside of political arguments, and as tolerant and loyal friends despite what struck them as *my* blindness to plain facts and simple logic. So, judging my friends as intellectual inferiors was out of the question (though I can’t deny having felt tempted on occasion). But keeping my mouth shut, or tip-toeing around important questions to avoid giving offense, would leave me feeling like a coward. And I found these controversies over war and peace, etc., too urgently interesting simply to ignore. So I kept pounding away at family, friends, friends’ families, unsympathetic teachers, hostile acquaintances, bemused professors, and bewildered dates, and anybody else who’d argue with me about serious things throughout my young adulthood, always in hopes of reaching some kind of agreement about things that matter, but rarely getting beyond a frazzled and, for me, increasingly depressing sort of stalemate.

Meanwhile, the corresponding public debate over these issues grew more and more bitter as the progressive factions of my generation wrestled for power both within and against existing institutions. Academic society, in particular, was becoming ever more alienated from mainstream America (outside the areas of government that it was able to influence) and more openly contemptuous of traditional religious and moral beliefs, while Christian and patriotic conservatives enthusiastically returned the sentiment. Even among scientists, objective discussion of empirical facts (concerning nuclear power, for example) was giving way more commonly to accusations of corruption and irrationality, and thoughtful consideration of dissenting views to gape-mouthed incredulity. On both left and right, even tolerating the expression of opposing arguments was increasingly seen as pointless, or even indecent. It was as if the two sides, by this time paying no serious attention to each other’s arguments and isolated socially as well as intellectually, actually lived in different worlds.

When I was close to forty, an idea occurred to me that changed my way of thinking about disagreement and, over a long time, changed my behavior too. I think that keeping this idea in mind, and working out its consequences conscientiously, has made me a happier person, and I hope also a more understanding friend, than I was when I was younger. Here is that basic idea. The reason that other people disagree with us is not ordinarily that there is something wrong with them that causes them to be unreasonable, but rather that their *evidence* is different from ours, leading them logically to different conclusions. This point is obvious when people are simply uninformed about a public fact: provide the missing information, and the disagreement usually goes away. But in our most contentious arguments, the difference in evidence that counts is not a matter merely of missing or misleading public information. It is instead a matter of our having different pools of *private* information in the form of trusted testimony, theirs coming from sources (colleagues and friends, newspapers, magazines) that they trust, and ours coming from different sources that we trust. And the problem isn’t that they are trusting sources that they shouldn’t trust while we rely on sources that we should. The crux of the problem is that they have just as good *reason* to trust their sources as we have to trust our own. Hence, even our bitterest religious and political opponents often have little rational choice but to believe the things they do. This doesn’t mean that none of us is actually *wrong* about the facts, of course, or that we never make mistakes in reasoning, or that no one is psychologically or morally corrupt. But it does mean that we ought to be far more careful in judging each other’s statements and arguments – and intelligence, and soundness of character – than we usually are.

This book is the result of twenty years’ work developing that basic idea into a general theory of belief and disagreement that encompasses morality, religion, politics, philosophy, and science, as well as ordinary conflicts of opinion. There is, of course, immensely more to say on disagreement and belief than I try to say here, but I think that what I’ve written is sufficient to the goal I have in mind, which is to make it clear how reasonable people disagree *as* reasonable people, not as people most of whom are psychologically or morally screwed up. On that main point, this essay stands in opposition to a recent trend – almost a new genre – of popular “diagnostic” literature, some relatively lighthearted,[[1]](#footnote-1) some serious and sympathetic,[[2]](#footnote-2) but most of it more or less openly contemptuous of some group of opponents.[[3]](#footnote-3) It is that growing contempt, especially, that concerns me when I think about the harsh religious and political divisions that we suffer with today, and that I hope this book will help to counteract. And when I look back at my own long history of stressful arguments, I wish that I could send this book back to the earnest, argumentative, and anxious person that I was in college. I hope whoever reads it now will find it useful.

[Acknowledgements]

**INTRODUCTION**

This book begins with a problem about disagreement and belief. When we think about our disagreements with people we respect, whose basic intelligence, education, and good sense make them about as likely as we are to be right about issues like the one in question, it seems arrogant for us to insist that we are always in the right and they are always in the wrong. So, as reasonable people, we are sometimes tempted to hold off believing controversial things until such disagreements can be worked out to everybody’s satisfaction. But to do so would effectively erase many of our deepest, most cherished beliefs. And it seems wrong, even cowardly, to stop believing in something just because others disagree. We know that it is better to think for ourselves than to depend on other people’s ideas. And we know that it is better to stand up for our beliefs when faced with opposition than to automatically relent or compromise. But if we are not actually more likely to be *right* than our opponents, exactly why should we be sticking to our guns?

I try to solve this problem in terms of a basic distinction among three aspects or dimensions of belief that I call *perception*, *opinion*, and *conviction*. I argue that only in the dimension of perception must we to temper our beliefs according to their probability of truth, that in the dimension of opinion we can hold any belief that might be useful in discussions with others, and that our beliefs in the dimension of conviction represent moral and practical commitments that are largely immune from correction by new evidence. The main part of the book presents a gradual, layered account of how beliefs are ordinarily constructed and maintained by rational people with different histories of evidence. Along the way, I try to explain how people come to radically opposed beliefs in all three dimensions of belief, and how our very rationality sometimes renders these disagreements intractable in matters of religion, politics, and even science. At the end, I talk about what reasonable people can do to cope with disagreement, both as active social beings and as individuals in search of deeper understanding.

A few terms and distinctions need to be introduced up front.

First, I want to emphasize that this is a book about *epistemology.* Epistemology is sometimes called simply the study of knowledge, but I want to stress that its working meaning covers the study of rational belief in general, even when this falls short of the certainty that we require for knowledge as such (the word *epistemic* just means having to do with rational belief or knowledge). My central focus is on what rationally justifies people’s beliefs, given all of their evidence and the order in which they have acquired it, not on what precisely qualifies as knowledge, though I do address the latter question at some length in Chapter 2. One of my main themes is that justified belief is easier to acquire than people usually think, while knowledge is much harder, and that confusing the two leads to a lot of trouble.

Epistemology is different from psychology. I am primarily concerned here with the rational foundations of our beliefs – explaining why we *ought* to hold them – not with their psychological causes. ButI will not ignore psychology entirely. It matters to my argument that people usually believe the things they do because they are rational, and that competing psychological sources of belief (self-interest, fear, complacency, and the like) are generally less important than we think. So, I will be totally right about the nature of belief only if I am also right about the power of reason as a psychological force. Ideally, there should be a rough correlation between epistemological and psychological accounts of belief in any case, but only a rough one. This is because we do not *learn* things in a perfectly rational way, especially as children, but rather through a dynamic mixture of experience and those instincts that have evolved to make our species the successful animals we are. This success has much to do with our predispositions to believe things that are true, of course, and for reasons that actually justify these beliefs. But there are other causal factors in the acquisition of belief as well, such as our built-in modular capacities for learning languages and recognizing faces, that are not directly relevant to the question whether we rationally *ought* to hold on to the beliefs that we have already acquired, or to how we ought to go about acquiring new ones, so I will not have much to say about them.

I also need to distinguish *reasonableness* from mere *rationality*, and *practical* rationality from *epistemic* rationality. Rationality in general is making proper inferences: if those are the premises, then this is the correct conclusion. The content of rationality can be defined in part mathematically, in terms of deductive logic. Some aspects of rationality, say in scientific reasoning, are harder to define, but still share an objective nature that governs the relationship of evidence to theories. Epistemic rationality is rationality with respect to true beliefs: if this is someone’s evidence, then that is what the person ought to believe, assuming he wants to believe what’s true (or probably true). Sometimes we also talk about rational *behavior*, too, for example when we say that it’s rational for a bank robber to knock any security guards unconscious, so that he and his gang can get away safely. This practical rationality plainly depends on epistemic rationality: the bank robber should *believe* on all sorts of evidence that knocking out the guards will help him get away, so this is what a rational robber will *do* if getting away is one of his goals. Since this book is mostly about what people what people ought to believe, I’ll use the words “rationality”, “rational”, and so on mostly in the more basic, epistemic sense. Reasonableness, by contrast, is the human virtue of dealing with other people’s beliefs and actions in a properly understanding way. Rationality is obviously a central ingredient in reasonableness, but it takes a lot more than pure reason to be reasonable. You might take a perfectly rational position on something, i.e. one backed up by plenty of good reasons, but if you don’t present those reasons in a way that other rational people can understand, or if you refuse to respect their rational objections or alternative positions, then you are not being reasonable. Reasonableness is even consistent with some amount of *ir*rationality, since we can make occasional errors in reasoning and still be reasonable people, as long as we are willing to correct such errors once they have been pointed out to us in a way that we can understand. But a reasonable person can’t be *totally* irrational.

I will also distinguish sometimes between *Socratic* and *Cartesian* approaches to disagreement and belief. I will explain these two approaches when I come to them, but in advance I want to say that this is nothing fancy. Socratic just means following the ancient Greek philosopher Socrates, whose ideal method was to begin as pairs or groups of people with existing disagreements and work down through discussion toward mutually agreeable beliefs. This is why Plato presents his accounts of Socrates’ work in the form of dialogues. Cartesian just means following the early modern French philosopher Rene Descartes, whose ideal method was to begin as individuals with certain beliefs that cannot be doubted, and work up from those ones to beliefs that no one else who follows the same method properly could disagree with. This is why Descartes wrote his most important works in the more common form of monologues. This book favors the Socratic approach in principle, but has been mostly written as an ordinary monologue, which is a lot easier in practice. I will, however, break into dialogue occasionally to illustrate my points about how people disagree.

Finally, in distinguishing among my three dimensions of belief, I will need to “precisify” the meanings of the words “perception”, “opinion”, and “conviction” to some extent, since their meanings are not altogether clear and distinct in ordinary English. Everywhere else, though, I have done whatever I could do to avoid unusual or technical definitions of familiar terms. I do not want what I say to depend, or seem to depend, on any kind of subtle usage. Instead, I will rely only on commonsense understandings of our ordinary epistemic concepts – give or take a little nudging – so that my arguments can be developed in an organic and intuitive way.

The structure of this book follows the structure of the theory that I want to present – no big surprise there – and I think that the steps are laid out pretty well in order. But these layers of argument are probably complex enough to justify a quick synopsis in advance, so that you know more or less what’s coming as you move along (if you prefer suspense, don’t read it).

Chapter 1 presents what I call the *problem of disagreement*, plus my general approach to a solution. The problem stems from finding ourselves disagreeing with our *epistemic peers*, people who are more or less as likely overall, given everything we know about them, as we are to have true beliefs in the subject at hand. If they are as likely as we are to be right, then the probability of our belief being true is at best 50%, in which case it seems that we should withhold judgment on the issue until better evidence becomes available. But this is not how we usually deal with disagreements among peers – in fact, we often take pride in believing things that *most* other equally smart and educated people consider false, or even preposterous. How can this be? I discuss two broad sorts of approach to this problem. One is to accept a relativistic or skeptical view of disagreement, according to which we are all equally right, or equally wrong, in maintaining controversial beliefs. The more common approach is to deny that those who disagree with us are actually epistemic peers, despite appearances, using a variety of “diagnostic” explanations to account for their errors. I think that neither approach offers a real solution to the problem. My own approach is more complex, but ought to seem familiar once it has been explained. I claim that there are three overlapping, but sometimes conflicting, principles that govern our beliefs, which I call *rationality*, *autonomy*, and *integrity*. Each principle is proper to one of the three dimensions of belief: perception, opinion, and conviction, respectively. These three aspects of belief perform essential functions with respect to thought, speech, and action in our lives. Disagreements among peers are as confusing as they often are, I argue, because it is so often unclear in which dimensions of belief we are contending, hence which principles ought to govern each dispute. In our most interesting controversies, differences of perception tend to stem from our rationally trusting different sources of information, while differences of opinion originate in our socially useful, but not always perceptively rational, practice of thinking and speaking autonomously, and differences of conviction stem from those epistemic commitments that make acting with integrity possible, but also shield us from evidence that we might be wrong.

In the succeeding chapters, I explain the common structure of beliefs in order of their rational development, from the individual epistemology of perceptions (Chapters 2-4), to the social epistemology of opinions (Chapters 5-6), to the prudential, moral, and political epistemology of convictions (Chapters 7-9). This order of presentation is a necessary one, but also necessarily an artificial one. In real thinking life, every epistemic process that I want to talk about is going on at once. But I can’t say everything I want to say all at the same time, so my explanations will need to be somewhat idealized and incomplete until I have all three dimensions of belief in play.

Chapters 2 and 3 discuss foundational issues in individual and social epistemology: the definitions of belief and knowledge, traditional skeptical problems, and the evidential role of testimony. Chapter 2 concerns direct perceptual beliefs, the ones that we construct out of our own sensations, memory, and reasoning. I give a brief account of the holistic, largely pictorial nature of our raw perceptions, and argue that these must be *articulated* into sentences from their original state before they can be explicitly believed. Such articulate beliefs take different logical forms, including categorical (yes-no) or conditional (if-then) sentences, statements of degree or probability, or any combination of these forms and others. A thorough understanding of belief has to be sensitive to these modes of articulation. I discuss the traditional skeptical argument that threatens our reliance on perception as a guide to true belief: how do we know that we’re not just dreaming, or floating as disembodied brains in vats of nutrients, or something just as bad? Aren’t we just assuming something that we cannot know? I argue that appropriate conditional articulations of perceptive beliefs (e.g. “If I am not dreaming, etc., then I am eating a bagel”) resolve this problem in a practical way, and that similar conditionalizations are able to resolve related philosophical problems about memory and inductive reasoning. For ordinary conversations, though, our ordinary categorical articulations (e.g. “I am eating a bagel”) are clear enough if used with proper understanding of their background assumptions.

In Chapter 3 I talk about the second layer of perceptual beliefs, those that arise from testimony. We use evidence from testimony to form correct beliefs about the world beyond our own direct experience. But testimony also complicates our understanding of the world in several ways. We receive it mainly in the form of articulate sentences, which have to be incorporated somehow into our mainly inarticulate perceptive models of the world. Such sentences are also liable to be vague or ambiguous, or to refer to unfamiliar concepts, so that we often fail to understand completely what we have been told. I show how beliefs based on testimony are rationally constructed out of more basic perceptual beliefs, by means of what I call second-order induction. But testimony stands apart not just because it radically expands the range of ordinary knowledge that is available to us, but also because it grants us access to a whole category of facts that is blocked off from direct perceptual examination, namely the inner mental states of other people. It is because other people tell us what they think and feel, and because we have reason to believe in their reliability in general, that we can rationally infer that they are not just mindless robots or hallucinations, but real thinking beings like ourselves. And because these trustworthy external sources confirm the great majority of our immediate perceptions, and tell us that we are also reliable to them as sources of information, we may infer that we can usually trust our own senses and memory. In this roundabout way, we can enhance our justification for believing that what we perceive directly is a genuine external world. I also argue in this chapter that our basic moral concepts necessarily derive from testimony. First-person experience simply does not contain the information necessary to construct distinctly moral beliefs, since these beliefs require non-self-centered understandings of moral terms like "right" and "wrong", and these cannot be rationally acquired except through taking moral statements as the truth. Thus, we come to believe that hurting the cat is wrong initially because we have good reason to believe that statements like "hurting the cat is wrong" are true, having heard them from our parents and other generally reliable sources.

In Chapter 4, I argue that to be fully rational perceivers, we must defer to the beliefs of experts, eyewitnesses, and other authorities whenever we have sufficient evidence that they are more reliable than we are in the matters at hand. To the extent that someone’s religious faith is based on testimony from evidently reliable authorities (including parents, clergy, and respected texts), this is a *rational*, not an *ir*rational, form of belief, even when it is maintained in the face of much direct contrary evidence. It is this subjective rationality of faith, passed on unanimously from each generation to the next, that explains the great historical stability of religious and other traditional beliefs. On this understanding, it makes perfect sense that not just isolated “backward” tribes, but also great peoples like the ancient Egyptians, Hindus, Mayans, and Chinese have believed persistently in things that modern, scientific Westerners consider gross absurdities. Religious rationality can thrive in cosmopolitan societies, too, but only within relatively closed communities of thought that include principles of solidarity like “you ought to obey your elders” and “you should not trust outsiders” among their doctrines. Subjectively rational faith in authority can sometimes produce “epistemic black holes” – beliefs so well protected by unreasonable attitudes that it is impossible for the believer to perceive contrary evidence as having any weight. In this situation, the same people can be both perfectly rational in the subjective sense, and so irrational objectively that they cannot be reasoned with about their core beliefs.

Different problems arise when our authorities themselves produce conflicting testimony, leaving us with no clear epistemic path to follow. In Chapter 5, I argue that the most rational response to such conflicts is not to abandon our authorities and base all our beliefs exclusively on our own perceptions and inferences, as Descartes recommends, but rather to withhold judgment until such experts as exist arrive at a consensus. This raises another question, though. If, as I argue, traditional beliefs are rational where authorities agree, and no belief is rational where authorities differ, how can any real *philosophy* – brand new ideas, derived from independent thought – ever arise? Part of the answer is that rational belief is not the only goal of intellectual life. A lot of good can come from speculative thinking, too. It is rewarding to explore the world as it appears to us as individuals, and to express our own ideas to others, even if we have no hope of knowing whether they will turn out to be true. On occasion, we might even come up with new theories that conflicting experts all come to accept, though there is no Cartesian formula that can guarantee this outcome.

In Chapter 6, I argue that modern science and other progressive forms of inquiry require a systematic kind of *irrationality* in order to flourish, which explains why they are rare in human history. Novel thinkers must be willing to persist in their own opinions, even in the face of rationally overwhelming evidence against them in the form of disagreement from authorities and peers. The new thinkers must prefer believing propositions that are actually *less* likely to be true than what they have been told by reliable sources, simply because these ideas seem more plausible to them intrinsically. Such unreasonable confidence in their own opinions can have good long-term consequences for their society if it takes place in a context of ongoing critical discussion, where ideas are permitted to compete for general acceptance. This dynamic, dialectical approach to solving problems is hard to extend beyond the practical realm of daily life into progressive philosophy and science, given the rational dominance of tradition in most serious areas of belief, in most places, most of the time. But once such a progressive system of critical thinking gains a foothold, it can be equally hard to dislodge. We live in a world so subtle and complex that long-term experimental competition among theories is needed in order to probe its nature very deeply, however irrational it may be for individuals to place their confidence in any particular new idea. So, in order to make independent research an effectively rational form of behavior on a wide scale, modern societies have rather ironically adopted intellectual autonomy as something like authoritative doctrine, while leaving unresolved its tense relationship to justified belief for individuals. With some exceptions, they have also promoted thoughtful dissenters from the lowest to among the highest social ranks, providing comfortable careers to people who, in other ages, might have been burned at the stake.

Chapter 7 explores the relationship between belief and action. Individual thinkers cannot come close to knowing everything that they would like to know, especially in controversial matters. Nevertheless, we are often called upon to act based on the limited information that we do possess. Our need to judge facts, not just experience the flow of thoughts and sensations, begins with our most basic perceptions, like seeing a rabbit on the lawn instead of mere white and green patches in our visual fields, and extends to all beliefs that enter into practical reasoning. Maintaining totally justified beliefs is not a reasonable goal in practical or moral thinking, since some of the things we value depend on firm conviction more than on perfect perception. Judgments provide a fixed basis for action, as when I judge categorically that there is a tiger coming towards me, so that I stop considering the evidence and start running away. Convictions are committed judgments that compose the conscious basis of our practical and moral lives. In adopting a conviction we purposely stop deliberating and adopt some principle as integral to our character, in order to control our actions with a firmer hand and to let others know that we will not to be moved by further argument. It is the very stubbornness of our convictions that counts as the measure of virtues like loyalty and courage. But this same stubbornness – what Friedrich Nietzsche calls the “will to stupidity” – can render contrary evidence practically invisible once a decision is made, sometimes placing disagreements over matters of conviction beyond the reach of rational discussion.

Chapter 8 extends my analysis of disagreement and belief into what is sometimes called political epistemology. Progressives complain about the maddening tenacity of unjust traditional beliefs and practices, to which conservatives appear indifferent. Conservatives complain about progressives’ careless reliance on untested theories and rejection of historical wisdom. I argue that a cautious experimentalism is the key to maintaining a reasonable balance between perceptive rationality and social progress. This is hard to achieve, because it requires the acceptance by progressives and conservatives alike of more intellectual and moral diversity than could be justified from *either* of their separate points of view. But a proper intellectual humility ought to dispose us to temper our political convictions with respect for the different convictions of other people whether we sympathize with them or not, and therefore to let ourselves be *jostled* into a working equilibrium with those who disagree.

In Chapter 9, I examine what it takes to be a reasonable person in the face of practical and moral demands for action and their implications for belief. I contrast the epistemic roles of rational perceiver, intellectual producer, and principled agent as requiring different ways of balancing perceptions, opinions and convictions. Most of us function as complex combinations of those pure types of believer, balancing and rebalancing criteria of belief according to the changing conditions of our lives. While I can offer no general program for regulating anyone’s beliefs, I do suggest certain parameters outside of which someone could fairly be called unreasonable. These parameters include having a good understanding of the nature of our own and other people's beliefs and evidence – if not in theory, then at least implicitly – so that we can readily tell the difference between matters of fact, opinion, and principle. This ability to keep the structure of belief in mind is, I think, a central requirement for dealing with disagreements in a reasonable way. The book concludes with more stringent advice to those whose intellectual ambitions go beyond mere reasonableness into the philosophical pursuit of wisdom.

A final note. Perceptive readers may detect a certain lack of scholarly rigor throughout this book, something that bothers my conscience as a professional philosopher. Since the book is intended for students and general readers as well as my colleagues, I have tried to avoid technical discussions while still making the arguments as clear as I can. Inevitably, conflicts between the two goals have arisen, and in making compromises I have always favored the general reader over the professional. This is a "big-picture" sort of presentation in any case, and I have not been able to consider nearly as many objections, or to connect as much of what I say to the current professional literature, as I would like. Where I present my own positive view of things without considering alternative theories, I try at least to note that I am saying something controversial.

Chapters 2 and 3 address some philosophically foundational issues and provide some stricter definitions and supporting arguments that I feel are needed for a complete presentation of the theory. These chapters (as well as the occasional technical footnote) may safely be skimmed through by non-philosophers without detracting from the central thesis of the book. If you are interested in the philosophical details, though, I should note that much of the core material has been covered in greater depth in articles over the past ten or twelve years. The main arguments in Chapter 2 are discussed in "Antiskeptical Conditionals", *Philosophy and Phenomenological Research* 73(2), 2006; the main arguments in Chapter 3 appeared in "Other Voices, Other Minds", *Australasian Journal of Philosophy* 78(2), 2000; and the main arguments in Chapters 4 through 6 are covered in "The Rationality of Science and the Rationality of Faith", *Journal of Philosophy*  98(1), 2001.

**1. DISAGREEMENT**

We should do well to commiserate our mutual Ignorance, and endeavor to remove it in all the gentle and fair ways of Information; and not instantly treat others ill, as obstinate and perverse, because they will not renounce their own, and receive our Opinions, or at least those we would force upon them, when ‘tis more than probable, that we are no less obstinate in not embracing some of theirs.

– John Locke, *Essay concerning Human Understanding*[[4]](#footnote-4)

1.1. The problem of disagreement.

We want to do the right thing for ourselves and others. But what is the right thing? People disagree over both facts and principles, and over which actions are implied by which facts according to which principles. This is how I know how difficult it is to understand the world: if it were easier, surely more people would agree in their beliefs. And this is true not just on general or abstract questions, but also on many matters close at hand where practical decisions have to be made. I see good people, individually and in groups, believing things that aren’t true, acting according to these beliefs, and making bad things happen to themselves and those around them. And I know that I have made plenty of such mistakes in my own life. So, I want to get beyond this kind of error if I can or, failing that, to understand the limits of my understanding.

Here is a problem. If I really believe in the truth of my beliefs, then, to be consistent, I must believe in the falsehood of other people's contrary beliefs. By definition, I can't even have beliefs without believing that they're true. This is just what having a belief means: thinking something to be true. And I can't think that my beliefs are true without believing that the contrary beliefs are false, just as a matter of logic. So, if I say that the Pope is Italian and my friend says that the Pope is not Italian, one of us has to be wrong. Under the circumstances, my believing that the Pope is Italian entails for me the further belief that if my friend says that the Pope is not Italian, then *he* is the one who is wrong. This is perfectly normal reasoning. But when I think about it in a certain way, it strikes me as arrogant. What is so special about me, after all, that makes me right while anyone who disagrees with me is wrong? Simply by holding beliefs that I know are controversial, I seem to imply that I am epistemically superior to – i.e. more likely to be right than – my opponents.[[5]](#footnote-5) And isn’t this arrogant, to think that I have better access to the truth than those who disagree with me? Yet the alternative would be for me to have no *interesting* beliefs at all. If I were to avoid all controversy, if I were to back down any time another person disagreed with me, then I’d be robbed of everything that makes me an honest, independent thinker. How, then, can I avoid unfounded arrogance toward other people’s beliefs, and still retain my own beliefs?

Think about this simple artificial situation. You and I have been hired to add up identical columns of numbers by hand. Our results are then checked by a machine that we both know to be completely reliable. After solving thousands of such problems in simple arithmetic, you and I have both succeeded in obtaining the correct sum, say, 96% of the time. We both know this. Now we each add up a new column of numbers, and discover that we differ in our results. Which sum is most likely to be right? Absent any indication that the work of one of us is more or less conscientious than usual, it should be clear to both of us that the two results are equally likely to be correct. In that case, shouldn’t we both just withhold judgment as to who is right and who is wrong? Even if we each feel a certain stronger connection to our own work, as a rational matter neither answer should be preferred by you, me, or anybody else before we manage to agree on a solution. And even if my track record is superior to yours, as long as there is some substantial probability of your being right while I am wrong, shouldn’t I adjust my confidence in my result accordingly? In either case, my dependence on my own arithmetic should be no stronger than the evidence I have for my own reliability, which involves little more than my track record on similar problems (this is why I rarely argue with my bank). But if this is the right analysis for disagreements in arithmetic, then why aren't all of our disagreements subject to the same analysis? Other questions of reliability are much harder, sometimes even impossible, to quantify, of course, but the principle should be the same: judge other people's reliability according to their track record and other ordinary evidence wherever possible. If you and I can establish, according to the total evidence available to us, that we have roughly similar levels of reliability, why shouldn’t we always just agree to withhold judgment when we find that our opinions differ, whether in matters of arithmetic, religion, politics, or anything else?

I want to take this problem very seriously. I hate the thought of being deeply wrong in my beliefs, and not just because I am afraid of doing the wrong thing as a result. Knowledge, or at least rationally justified belief, is also something that I value greatly for its own sake. So, I try very hard to back up all of my controversial beliefs with sound arguments and solid evidence. Yet when I consider these beliefs not only in themselves but as beliefs of *mine*, and I realize that most of the other reasonable-seeming people in the world disagree with me about many of these things, it worries me. I imagine looking through a telescope from space, zooming in on this planet, continent, country, town, and street, until I can pick myself out from among the billions of people now alive. And I ask: what makes this person more likely to be right than practically everybody else? He is admittedly an educated, thoughtful sort of person, and perhaps this sets him apart from many others – but surely not from all others. There are probably millions, certainly many thousands, of people as well situated in the world as he is for the purpose of believing what is true in general. And even (perhaps, especially) within this intellectual “elite”, there are innumerable controversies, stretching over almost every topic of importance. There is nothing interesting that the person I have zeroed in on thinks, that isn’t contradicted by countless others just as good at figuring things out as he is. So, why should I trust him in particular?

You could argue that I have no choice – I simply have to trust this guy, because he’s *me*. I am required by basic logic to believe whatever he does. My beliefs are mine; my reasoning is unavoidably my own. I can't help but rely on what seems right to me, just as you can't help but rely on what seems right to you. But this objection misses the point. I do have to use my own reasoning to form my own beliefs, so in that ultimate way I simply have to trust myself. But I do not have to trust myself in the sense that I should be required *now* to reaffirm all of my *prior* beliefs up to this moment. I always have the option of deciding that I haven’t been especially reliable before according to my own track record of mistakes, and therefore, on questions where I find my view is contradicted by my *epistemic peers* – other people who are about as likely overall as I am to be right about such things, as far as I can tell – I should probably avoid making any judgments at all.[[6]](#footnote-6) Indeed, how can I rationally avoid taking this option? On what grounds can I insist on my superior reliability on any public question, when I am contradicted by so many others who are more-or-less equally reliable in general? I seem to have plenty of such peers on practically every subject where I might have a belief, beyond such questions as what’s in my pockets at the moment. Given all that I know about these people's education, experience, and overall track record of turning out to be right (or probably right, as far as I can tell), I can’t rationally avoid concluding that their probability of having the right answer to most questions is about as good as mine. When my beliefs are controversial (i.e. *interesting*) ones, I find that most of these peers disagree with me. But if I am disagreeing with my peers, this means that I am liable to be wrong. When I am liable to be wrong, I have no reason to stick to my beliefs. Therefore, I shouldn't stick to my beliefs when they are controversial. This is what seems to come from my acknowledging that people who disagree with me are about as reliable as I am. If I am not willing to be arrogant about my own special capacity for true belief, it is hard to see any alternative to this conclusion. But should I really just give up all of my interesting beliefs? If I did, how should I get through life? If I were to reject every proposition that is in dispute among my peers, just for the sake of avoiding any error, wouldn’t that be an error in itself? Yet if I keep believing things that most of my peers deny, how can I avoid making the same sort of mistakes I seem to see them making now, and that I know that I have made in the past? If there is no solution to this problem, then it looks like there is no point in my trying to become a fully consistent, generally rational believer. I might just as well abandon the entire project, basing my future actions on my habits, inclinations, and a mindless expectation of good luck.

I am not the first person to notice that reasonable-seeming people often disagree. I am also not the first to be troubled by the existence of beliefs contrary to his own. But I do seem to find this more disturbing than most of the people I know. This is partly because I tend to be the odd man out among my friends on many issues, so I must rely more on their toleration than they do on mine. But most of my friends and colleagues also believe that they can explain their serious disagreements with others in a way that allows them to stick with their own beliefs, while at the same time being totally fair to their opponents. Here is the argument they give most often. We ought to be bothered by the fact of disagreement only when we do not know the sources of the disagreements in question. When I discover that my friend believes the Pope is non-Italian, I need to find out his reasons for this belief before I can decide whether to take his disagreement with me seriously. If his reasons are good ones, then I may need to change my mind. If his reasons are flawed, then I should point the flaws out to him, and if he’s reasonable, he’ll be the one to change his mind. In either case, it is up to me to evaluate whatever reasons he gives me purely asreasons, not as *his* reasons in particular. Reasons are reasons; it shouldn't matter where they come from. All that matters is whether they actually prove whatever it is that they are supposed to prove. This is a matter of objective rationality, not of anyone’s authority or prior reliability. So, what I am calling the problem of disagreement only counts as a problem when we don't really know what's going on *beneath* the disagreement.[[7]](#footnote-7)

The problem with this way of thinking is that most serious disagreements cannot be resolved simply by people looking at each other’s reasons and deciding together which ones are best. In most actual disagreements over religion, politics, and other topics that really matter to people, as opposed to simple matters of information, both sides are liable to see their own reasons as the good ones and their opponents’ reasons as flawed. For such *symmetrical* disagreements, it seems that no amount of explanation on either side can bring the parties closer together. Our experience with disagreements like this among epistemic peers implies that none of us have perfect abilities to analyze the total context of disagreement, including everybody's reasons as well as their conclusions. Just as I know that I am not always right in my conclusions, I know that I am not perfectly able to tell whether my arguments are sound, or whether my responses to opponents’ arguments are fair. There is always a good chance that I don't really understand some reason that he has, either because he hasn't articulated it clearly enough or because I just don't get it. My own experience with changing my mind about the arguments of others bears this out. So, my evaluation of the arguments of peers is not an adequate criterion for judging who is right. If my opponent is someone I know to be roughly as reliable as I am on the matters in question, the fact that he *still* disagrees with me, even though we have explored each other’s evidence and reasoning, entails that my chance of being right is still no better than about fifty percent.

One way of putting the whole problem is that none of us is capable of transcending his own *epistemic situation* and attaining a purely objective, God's-eye point of view on controversial issues. I have my evidence and reasoning and track record of mistakes and total life experience, and you have yours, and we both know that they are different in various ways. If you and I disagree, each of us would like to be the final judge of the entire controversy, but neither of us can do so rationally. As long as I continue treating you as a peer, I cannot claim to be a much better judge of controversies in general. And as long as we disagree symmetrically, I cannot reasonably claim the final word in this controversy in particular. We're just stuck together in uncertainty, unless and until we somehow come to an agreement.

Of course, I could always decide that you are no longer my epistemic peer, simply in virtue of this disagreement. But for this to be a rational decision, I would have to base that demotion on our entire comparative history of getting things right, not on this disagreement alone. It is not impossible, but it is necessarily rare for one error to make much difference in the overall reliability of someone's testimony compared to our own, assuming there is a substantial track record to go by. After all, it doesn't matter whether our two track records are exactly equal in value, just that your record is *roughly* equivalent to mine, i.e. close enough so that the probability of my being right in this case is not *much* greater than the probability of your being right, as far as I can tell. And this is a matter of *subjective* probabilities alone; it's not that there is an objective standard of peerage based on people's actual, objective levels of reliability, and that we have to guess at whether other people are our peers. In the sense I mean, another person is a peer to the extent that I ought to believe that he is about as reliable as I am, given the evidence I have at the moment. In the same sense that I ought to believe that my snow-blower has an 80% chance of starting on the next try because it's started 80% of the time in my experience (including whatever I know about other people's efforts to start the same machine), I ought to believe that you are just as likely as I am to be right about our present disagreement if you have been just as reliable as I have been in my experience, provided that experience has given me sufficient evidence one way or the other.[[8]](#footnote-8) In any case, peerage is not a "yes-or-no" thing, but a matter of degree. Given all of my experience so far, some people will seem likelier than me to turn out right on the next occasion, some about the same, and some less likely to various degrees. But even someone who is noticeably less reliable than me is still likely to be right occasionally when we disagree. Thus, your opinion will probably be worth considering as evidence against mine, even if you are not quite my peer.

Well, maybe we just haven’t been thorough and careful enough. If we keep at it, going over all of our evidence, arguments, and presuppositions with sufficient rigor, surely we’ll find out where one or maybe both of us went wrong. This is René Descartes's idea of how to figure everything out without making mistakes, sometimes called the Cartesian Method. Lay out absolutely every bit of evidence and step of reasoning, producing a completely analyzed, fully articulate account of each belief, and then examine each position step by step, starting from whatever is completely certain to begin with.[[9]](#footnote-9) Ordinary human reasoning will not go wrong once things have been broken down so completely and examined so patiently that we can be completely confident that nothing has been left out. In this way, every disagreement can be resolved by rational and careful thinkers given enough time, because everyone will come to the correct beliefs, or at least avoid all false beliefs, if they just follow the method conscientiously.

There is nothing wrong with this approach to disagreement in theory. I tend to agree with Descartes that basic steps of logic are essentially foolproof for any rational person, and that sufficient carefulness and comprehensiveness should guarantee a fully justified belief for anyone who makes the proper effort. In principle, if two people make the proper effort with the same initial evidence, then they should always come to agree in time. Indeed, this is the way that most people seem to view the issue of disagreement: not as a fundamental fact about people's beliefs, but as an accidental side-effect of one or more parties to a disagreement having messed up somewhere along the way. If people would only be careful and systematic in their reasoning, we could resolve our disagreements in a straightforward way, like disagreements in arithmetic. The problem here is practical – this kind of resolution almost never actually happens. We just don’t have the time and patience for a methodically complete examination of the world, or even of our own major beliefs.[[10]](#footnote-10) We don't even come close. Our world is so complex, our potential evidence so scattered and ambiguous, our sources of information so diverse, the volume of literature so overwhelming, and our actual reasoning so inarticulate, that it is beyond our individual abilities to master any serious issue completely, no matter how hard we try. Outside of pure mathematics, we don't make anything *like* a complete Cartesian effort to account for our beliefs, fully considering every piece of relevant data and every conceivable counter-argument. Since we cannot make complete investigations, each of us is vulnerable to his or her own limited, possibly misleading pool of evidence. This is how people are sometimes falsely convicted by judges and juries acting in perfectly good faith but with imperfect information, despite a strong legal presumption of innocence. This will continue to happen as long as the proof required for conviction is less than absolute, which means as long as there are limits to the resources and time that we can spend investigating crimes.

I don't mean to say that truly systematic efforts to investigate hard questions are pointless; else I’d be out of a job myself. People profit from the ongoing effort to find the truth, in several ways that I will talk about in later chapters, but not ordinarily because we end up with demonstrably true beliefs. The closest we actually come in practice is to write long books elaborating one of many points of view on this or that question, or criticizing other lengthy books that do the same thing. Think of all the arguments that have been given for and against the existence of God. We may find some of these more or less convincing, but no one has ever produced more than a general summary of all such arguments and counter-arguments devised so far. Considering the many false certainties of times past, it seems impossible to know when arguments like this are absolutely settled.[[11]](#footnote-11)

Nevertheless, many people outside of professional philosophy seem to think that the issues we care most about are usually pretty simple, so that the decisive arguments could be expressed in articles of twenty or thirty pages, if not much less. So, for example, people interested in economics tend to think there is some fairly simple argument for or against the minimum wage that is conclusive in itself. It just needs to be appreciated to convince any fair-minded, reasonable person. Since it is evidently *not* appreciated by people on both sides of the issue, people on one side are tempted to think that there must be something wrong with the people on the other. We may try to resist this sort of temptation out of a desire not to be arrogant, or not to seem arrogant, but this is hard to do consistently with the belief that our position is completely justified. Alternatively, we might try to avoid the issue with vague statements about everybody having a right to his or her own point of view. But logic is logic, and we can't have us being definitely right, plus them being definitely wrong, plus them being just as likely to be right as we are. So, what are we supposed to do?

I propose that we take disagreement as a basic fact, like Socrates does in Plato’s dialogs, not just as evidence of sloppy thinking. Rather than pursuing the Cartesian project of individually establishing beliefs that cannot be doubted, I want to take a more pragmatic approach to dealing rationally with opposing positions that are roughly equally likely to contain mistakes. I do not mean to denigrate the project of systematic philosophy *a la Cartes* **–** this book itself contains much of the same sort of foundational, constructive argumentation, especially in Chapters 2 and 3. My point is rather to acknowledge that the Cartesian project takes more time than is available to individuals, even over their entire lives. So, rather than waiting for however many hundreds of years it might take to settle our differences on an unshakable foundation, let us make the best sense possible out of our present situation, which is that we find ourselves in many important disagreements with people who about as likely to be right as we are. Prior to our someday resolving such disagreements on constructive principles, how should we think about this problem? What is the *reasonable* thing to do while waiting for perfectly *rational* solutions?

1.2 The usual explanations

Here is a list of explanations that we commonly use to account for our most serious and frustrating disagreements. Some are general efforts to remove the force of disagreement equally for both sides. Others are *diagnostic* explanations that attribute some ill state to our opponents, blamefully or not, from which we take ourselves to be immune. I do not doubt that some of these ideas are sometimes applied correctly. But none of them can solve the problem of explaining our most interesting moral, religious, or political disputes with our epistemic peers while letting us retain our different prior beliefs in a rational way. At best, they offer only partial explanations, and at worst, they only give us names to call our enemies.

*General relativism.*  First, it can be said that our opposed beliefs are not as contradictory as they appear. I might think that little sausages in cans are horrible, and you might find them delightful, but this is no real contradiction in belief. It is instead, as we say, simply a matter of taste. Sensory perceptions can be like this, too. I might see a certain shape as circular, while you see an ellipse – that’s just a matter of perspective. Or, looking at a fluffy cloud, I might see a flower while you see a crab – still not a matter of fact, but of imagination. Perhaps all of our disagreements are of this general sort. You like what you like, and I like what I like; you see things your way, and I see them mine; what's right for you may not be right for me, and so on. There is no God-like arbiter of truth. Thus, it can turn out that when we disagree, neither of us has to be wrong. Each one is right in his own way, from his own point of view, and all points of view are equally valid. On this view, neither of us can reasonably insist on his own perspective, then, even in moral matters. To do so is to be unfairly judgmental towards the other.

This universal relativism has a very long pedigree, extending back to Protagoras and other Sophists of ancient Greece and forward to Derrida and other deconstructionists today. But this cannot possibly account for all cases of disagreement, since it requires that there be no such thing as truth, no facts of any sort at all, and this is absurd. I believe that Mars is more than ten miles from the Earth, for example, and I think anyone who disagrees with me is just plain wrong. Imagine what it would take to deny this proposition in a rational way. For one thing, the gravitational attraction between two such massive objects so close together would smash the planets into each other immediately. Someone who wanted to assert coherently that Mars and Earth are nevertheless almost touching would need to explain why this has not happened, and many other impossible things besides. Even supposing that he could come up with a superficially comprehensible “proximity theory”, there would still be the obvious difference that my “distance theory” is based on quite a lot of evidence, and his is based on none. This is not a matter of perspective or anything like it. To say otherwise would make any lunatic the epistemic peer of Einstein.[[12]](#footnote-12)

In any case, the general claim that there are no objective facts is logically self-contradictory. As philosophers (including me) never tire of asking, what is the status of the general claim that there aren’t any facts? If it is said to be a fact, then it immediately falsifies itself. But if it is said to be anything less than a fact, then what is the point of saying it? If the relativist just says that it is true *for him* because he is a relativist, then the non-relativist can reply that what’s true for him is that relativism is *false*, because that’s what he believes. In either case, the relativist is either making claims to objective truth or not. If he is, then he can’t really be a relativist; if he’s not, then what reason do we have to take his statements seriously?

*Moral or cultural relativism.* A more plausible symmetrical account of disagreement would distinguish between realms of objective and subjective fact. We could allow that in the first realm it is possible to be straightforwardly right or wrong, while in the second there are no real facts, only matters of perspective or convention. The former would presumably include the distance between Earth and Mars and other plainly empirical facts, while the latter would comprise matters of personal taste and preference, plus controversial matters in ethics, politics, and religion. Informed and reasonable people from almost any background can be expected to accept the facts of ordinary observation and the results of thorough scientific inquiry, but cannot be expected to agree on these more controversial issues. While the “hard” factual questions can be answered from an almost universally accepted empirical base, the more controversial “soft” questions must be referred to local points of view or customs. For a standard example, most people in India believe that cows are sacred as a feature of their Hindu tradition, while most Americans see nothing wrong in enjoying a nice steak, since our Judeo-Christian tradition has no objection to the practice of eating beef. Perhaps all political, moral and religious disputes can be explained as matters of convention in this way. We should then view each other’s different non-empirical beliefs as we view the British driving on the left side of the road while Americans drive on the right. It may be necessary to have one driving convention or the other, but either choice would be equally good.

This view of cultural and moral values as local conventions pervades the social sciences, and is taught dogmatically to students of anthropology and sociology. The solid methodological reason for this is that science needs to be objective, and if practitioners permit themselves to judge the rightness or wrongness of another culture’s values, this will undoubtedly distort their perceptions of what actually happens in that culture, and why. But methodology does not determine metaphysics. Just because it’s not our job as *scientists* to make value judgments, this doesn’t excuse us from our responsibility to make them as *human beings*.

Another motivation for relativism among social scientists is the belief that Westerners particularly ought to avoid imposing their judgments on other cultures, given our history of arrogant colonialism, racism, and even genocide. But again, the idea that someone has no moral right to make judgments does not entail that there are no facts to be judged in principle. A murderer on trial may be in no position to tell his victim’s widow not to snap at her children in court – but she still shouldn’t snap at her children in court. Similarly, even if anti-imperialists are right that Westerners are not in a position to tell other people not to beat their slaves, strangle unwanted girls, or practice their swordsmanship on peasants, it is still wrong in fact for people to do such things.

In any case, most of the cases we care about are already agreed on both sides to be questions of *fact*. To deny this is not to allow both sides to be right from their own points of view, but rather to accuse both sides of being thoroughly wrong, in that each falsely believes that it is talking about objective rights and duties, not just conventions. Once people come to view a moral disagreement as just a matter of social convention, they tend to give up caring much about it and accept whatever the local custom seems to be: what reasonable American would insist on driving on the right-hand side in Britain? But in serious disputes about religion, politics, and most other controversial questions we don’t see this loss of interest, because the people involved believe these issues to be real matters of fact.

*Skepticism.* A different way of weakening the relativistic analysis of disagreement would be to admit the existence of objective facts, but deny that anyone can know them. As with relativism, such skepticism may be global, amounting to the claim that nobody knows anything at all about anything, or it may be limited to this or that kind of knowledge, for example empirical knowledge, philosophical or theological knowledge, moral knowledge, knowledge of other minds, or knowledge of the future. We may all live in the same objective world, but it is *as if* we lived in different worlds because each of us only knows the world from his or her own distinct point of view. This is perhaps what people actually mean when they deny the existence of an objective reality: they are really denying the existence of objective knowledge beyond the contents of our individual experience.

Skepticism is a long-established and respected theory in epistemology. It also has plenty of adherents in the wider world, especially when applied to matters of moral controversy. It is even invoked quite casually as a means of stopping angry arguments: “None of us really *knows* what happened (or ought to happen, or why things are the way they are), so none of us really knows who’s right and who’s wrong in this argument, so let’s not get so upset with each other.” Here are a couple of standard objections to such general skepticism. First, as with general relativism, there is a looming paradox around attempts at stating the theory itself: if nothing at all can be known objectively, what should we say about our relation to that very fact? If it is known objectively, then the theory automatically becomes false; but if it is not objectively known, why should we believe it? Some ancient skeptics accepted that this too cannot be known, holding that even skepticism should not be believed dogmatically, but rather practiced as a non-judgmental attitude and way of life.[[13]](#footnote-13) But this doesn’t solve the problem of what to believe for people who continue to have beliefs. Second, even if that problem can be solved, it is not clear how this helps account for reasonable disagreement. Most people would really be surprised to find out that they know nothing apart from how things seem to them. When they claim to know that the flat-Earth theory is false, or some such thing, surely they don’t mean only that it seems to *them* to be false. They mean that it really *is* false, and that they really know that. As with general relativism, then, this general account of disagreement seems not to vindicate our different beliefs as much as to tell us that we are all wrong.

In any case, most important disagreements do not involve serious claims to knowledge *per se*, but rather claims about what rationally ought to be believed based on the evidence available. Even if we can’t be certain of anything, this doesn’t mean that every position is as reasonable as every other. Even if we can’t prove absolutely that the world will not explode in ten days’ time, for example, there is no good reason to believe it will and plenty of good reason to believe it won’t, so people arguing against the theory are not liable to be satisfied by appeals to skepticism. Reasonable people come to agreement about all sorts of things without the need for perfect certainty, so why couldn’t we agree on this?

*Semantics.* Another general way to argue that both sides can be right in a dispute is to say that the meanings of the terms at issue are ambiguous. If it turns out that we are saying yes and no to different propositions, then we do not really disagree. We are instead, as we sometimes say, just talking past each other. For example, disagreements over the existence of God or the virtues of socialism might turn out to depend on just how those concepts are defined. If one says that God exists, thinking only of some ultimate creative force, and the other says that God does not exist, but means specifically the personal God of the Old Testament, then we are not really disagreeing on anything but the words. If one says socialism is a bad thing, thinking of Soviet-style command economies, and the other says socialism is good, thinking of European welfare states, a little attention to our definitions ought to be enough. We might hope that when all definitions have been carefully spelled out and complete semantic agreement has been reached, we will find all of our serious disagreements melting away. But this is too optimistic. It seems that people often understand each other perfectly well and still disagree about the issue of substance. If we disagree, say, as to whether the specifically Christian God exists, having agreed on all of His essential attributes, and one of us says “absolutely" and the other says “no way”, then further semantic specification is unlikely to resolve the problem.[[14]](#footnote-14)

People do sometimes argue past each other, and it is always good to be on the lookout for deliberate or accidental misrepresentations of your or your opponents’ positions. But this will rarely resolve fundamental or long-standing questions. For example, a lot of ink has recently been spilled, and many voices raised, over whether the second Bush administration practiced and approved of torture. They had been accused of this by some opponents on the grounds that they had authorized interrogating terror suspects by a method known as water-boarding, which effectively simulates the experience of drowning, thus inducing unbearable panic in the subject while not placing him in real physical danger. The administration had certainly approved of water-boarding, but they denied that this practice *constituted* torture according to the law. So, when the issue was presented as one of whether or not *torture* was being permitted, advocates on the two sides tended to speak past each other, each claiming that the other was misrepresenting the issue by getting the definition wrong. Still, there always remained the substantive question of whether or not water-boardingought to be practiced or permitted, regardless of whether anybody *calls* it torture. This was a struggle not just over words, but over policy. There is a very long list of increasingly harsh possible methods of interrogation, and somewhere a line needed to be drawn between the harshest permissible method and the least harsh impermissible method. Although the public argument was often confused by the rhetoric of torture, the essential question was always whether that line was to be drawn above or below water-boarding, and this was a legal and moral, not a semantic, decision. Similarly, no merely linguistic agreement is going to resolve the issue of abortion, however much the activists on both sides now confound the argument with charges of “murder” or “indifference to women’s health”. We would need somehow to find a moral consensus as to how all these fetuses ought to be treated *before* we could ever agree on whether it is proper to call them babies.

*Ignorance.* There are other ways to argue that both sides in a disagreement are equally blameless. For example, reasonable people are sometimes tricked into a false belief by incomplete or misleading evidence, through no fault of their own. Thus, a recently discovered stone-age tribesman who refuses Western medicine may be as rational as anybody else in the world. He simply does not have the information that modern people have about the efficacy of our medicines – and conceivably we can learn as much from him, since we lack his accumulated evidence about his own medical practices. Even among our Western selves, different people have access to different pools of information, and some of us have been much better educated overall than others, largely as a matter of good luck. So, when people disagree with us about the virtues of a single-payer health care system, they may simply lack the information or the education necessary to understand this complex issue, and the incomplete evidence that they do have may have misled them into a false view of the matter. None of this would have to be their fault.

This sort of analysis is surely correct some of the time. But in such cases, once any fully rational mistaken person has been given all the evidence he lacks, he should be able change his mind and accept the truth of the matter, right? Whenever both parties to a disagreement have exactly the same evidence, it seems they *must* come to agree if both are reasonable. Yet there are clearly many disagreements, including the fiercest religious and political ones, where apprising each side of the other’s evidence utterly fails to resolve the issue. Indeed, some of the best educated people in the world seem utterly intransigent to their opponents about such matters, precisely because they have all of the other side’s evidence and refuse to accept their conclusions. In the current controversies among experts over the safety of nuclear power, or of hydrofracking as a means of extracting natural gas from underground, each side has all of the essential data that the other side considers relevant. It just doesn’t seem to *move* them in the way the other side believes it should. Mere ignorance, then, in the sense of someone’s simply lacking the information that grounds other people’s beliefs, can’t wholly account for such stubborn disagreements among peers. The question remains: why is our opponent *still* “ignorant”, when all the evidence is sitting right in front of him?

*Random mistakes in reasoning.* Another common diagnosis among people reluctant to think ill of their opponents is to say that we all make more-or-less random errors in reasoning or judgment. This doesn’t need to result from any deeper flaw; it’s just something that happens. Whether through haste, fatigue, or other sources of distraction, we are all human, and we all make mistakes. So, we should all be careful in our judgments, and it is truly our own fault when we are not, but nobody is perfect. It is hardly an arrogant judgment on my friend for me to believe that he has made a mistake on this or that occasion, but no more frequently in general than I have.

The problem here again is that mere logical mistakes are easy to correct for any fully reasonable person. If my peer has become, say, a Republican (or Democrat) just by way of a mistake in following some line of argument, then I should be able to point this mistake out to him, and he should be happy to adjust his view as a result, just as we are usually happy to be corrected in arithmetic when trying to balance our checkbooks. But if he still persists in his erroneous belief, as people usually do in interesting controversies, then it cannot be solely because of that error in reasoning. Something else must be causing him not to *recognize* his error even after it has been pointed out to him. For example, Christians are often tarred with committing a simple fallacy of circular reasoning, since they allegedly believe that God exists because the Bible says so, while they rely on the Bible because it is the word of God.[[15]](#footnote-15) But could it really happen that a billion people, many superbly educated, really fall for such an elementary fallacy? You’d never persevere in such a bone-headed mistake yourself, would you? To be sure, people are not always very articulate about their beliefs, and may sometimes *say* unreasonable things like this when they are at a loss for better explanations. But if you attribute things like this to people as persistent errors in belief, not just in rhetoric or speech, then you are hardly treating them as epistemic peers. Again, you’d need to find a deeper explanation for their disagreeing with you than an “only human” sort of error.

*Blindness.* Maybe our opponent is just “blind”, as we sometimes say, in the sense that he is psychologically unable to recognize relevant facts when they are presented to him, or to perceive what these facts entail. He might be a highly rational person most of the time, but when it comes to this or that particular issue, he just fails to function rationally. So, your friend the Palestinian (Israeli) scientist is calm and thoughtful about all questions touching on his expertise, and wise and circumspect regarding life in general, but try to argue with him about the legitimacy of Israel (Palestine) and he turns into a brick wall of resistance. It is as if you were attacking his parents: reasonable discourse stops, and some other, lower kind of reaction takes its place. Hearing him vent his rage for half an hour is enough to make you swear never to raise that subject again.

In such cases, your immediate judgment on your friend is necessarily a fairly harsh one, but not a global criticism of his mind or character. He just has these one or two irrational blind spots – and, for all you know, you have a couple just as bad, yourself. But the larger the area that a peer’s blind spot seems to cover, and the longer it persists in the face of reasoned criticism, the harder you will find it to avoid judging your friend to be flawed in a more general way, to have some basic inability to see things the way they are. And even in the narrowest instances, just saying that your friend is blind about some issue doesn’t really explain anything. Again, *why* is he blind? There must be a further explanation for this otherwise reasonable person always to be dogmatically wrong about this issue while I am able to be sensible and open-minded.

It has recently become popular to say that people are not blind but “in denial”or to call them “deniers” or “denialists”when they refuse to credit solid evidence regarding controversial subjects. The paradigm case, which of course goes back for decades, is people who deny that Germans killed millions of Jews during the Second World War. Such Holocaust deniers refuse to take any of the great masses of documentary and anecdotal evidence, including numerous Nazi confessions, at face value, spinning elaborate theories to account for how and why the vast majority of experts and laypeople have been tricked into accepting this gargantuan hoax as reality. Nowadays, the term is used for people who disagree with us about almost anything we take to be established truth, which attributes to them some kind of willful blindness to the facts at hand while vaguely associating them with Nazi sympathizers – always a useful rhetorical device.[[16]](#footnote-16) Thus, people who reject allegedly conclusive evidence that human activity is causing dangerous levels of planetary warming are often called climate change deniers or denialists rather than skeptics, while those who believe that universal vaccination does people more harm than good are called vaccine denialists, and those who believe that AIDS is caused by something other than HIV are called HIV/AIDS denialists. This device gives us a snappy name to call opponents, but provides no better insight into people’s beliefs than just to call them ignorant, or for that matter, simply wrong.[[17]](#footnote-17) Once again, what’s really needed is a full explanation as to why people believe the way they do in spite of what we take to be conclusive contrary evidence, not just a label for it.

*Brainwashing.* One putative explanation for a peer opponent’s blindness or being in denial over some particular issue is to say that they have been “brainwashed”. This is a negative form of ignorance, comprising not a lack of information, but particular irrational beliefs that have been implanted into our opponents through some kind of psychological coercion. Such brainwashing can be done by parents to their little children, by churches to their respectful congregations, or by groups in power to the public through pervasive propaganda, especially in schools. Rational arguments don’t work with people under these conditions, because the brainwashing has created psychological resistance to the actual evidence and arguments that they would need to understand in order to perceive the truth. We are able see the truth ourselves not because we’re basically more rational than our opponents, but just because we have been luckier with evidence.

Karl Marx and other revolutionary communists have argued that capitalist rulers brainwash their subjects so thoroughly that nothing short of violent rebellion is capable of freeing people’s minds from the religious and social ideologies that keep them in their place. Only a new regime that represents the masses rather than the ruling class will educate the common people to pursue their own enlightened interests, rather than the interests of the bourgeoisie. For the sake of such revolutionary social change, totalitarian governments have often thought it necessary to criminalize middle-class attitudes, and to sentence those with non-proletarian backgrounds to long sentences in “re-education” camps, as during the Cultural Revolution in China, or simply to exterminate them to make room for a new generation that has been brought up correctly, as during the Pol Pot regime in Cambodia. Since the collapse of most communist governments two decades ago, there are few calls for such violent measures to counter perceived brainwashing, even in leftist academia. Most of my colleagues and friends oppose coercive religious or political “re-education” on the grounds that it is morally wrong, that it is flatly unconstitutional, that it is not required for progressive social change in modern democracies (even in one that’s seen as riddled with corruption), and that it doesn’t really work on adults, anyway. We do still try to “de-program” our students to some extent in schools and colleges, trying to fix what educators perceive as socially regressive brainwashing with psychological measures of our own, not just by handing out information. So, programs that aim to “raise awareness” of problems like sexual violence or alcohol abuse, although they do provide statistics about crime rates and other objective matters, are focused mainly on getting students to adopt a better, more serious attitude toward these problems (for example, through the use of shocking videos or emotional testimony from victims) so that they won’t tolerate sexual abuse or drink too heavily thereafter. If students absorb the information content of such programs to the point where they can pass the quizzes, but they still see date rape or binge drinking as no big deal, then the program has failed: they have transmitted the right information, but not the right awareness. Such programs are disliked by those who view their content as extending beyond common morality and good sense into political indoctrination – brainwashing in its own right – but proponents often feel that they are needed anyway, to counteract brainwashing from the wider culture through sources like advertising and pornography.

In any case, however mild or drastic the suggested remedies may be, brainwashing has its limits as a diagnosis of our peer opponents’ false beliefs. It points to a syndrome, maybe, but does not identify the core disease. Like saying that people who disagree with us are ignorant, blind, or in denial, calling them brainwashed is at best a partial explanation of the fact that they do not agree with us. For suppose they say that we, not they, are the ones who have been brainwashed. Given their similar general reliability to ours, how are we to choose which *total* explanation makes the best sense: they’re wrong and they’ve been brainwashed, or we’re wrong and we’ve been brainwashed? If they are really our epistemic peers, it seems that we have insufficient reason to conclude that it is they, not we, who hold the short stick.

*Intellectual inferiority.* Now we arrive at the bottom-level explanation in most cases where our opponents’ beliefs have been thoroughly diagnosed: people persistently disagree with us in important matters because they are not really our peers after all, but our epistemic inferiors. Much as we prefer to think otherwise, and to avoid insulting people or seeming arrogant, in our hearts of hearts we believe that our opponents are less likely to be right in the matter at hand because they are less likely to be right about such things in general. There seem to be three really basic ways of judging other people to be epistemically inferior to us: they can be weaker intellectually, emotionally, or morally. Regarding the first, it is hard to say this sort of thing, but maybe some of the people we deal with just aren’t smart enough to figure some things out. Some people might belong to the wrong political party, for example, because they just don't have the mental capacity for making proper inferences about complex issues. It can be a subtle question, after all, exactly what is wrong with some initially attractive social or economic policy. The most intelligent among us are able to examine all the facts and arguments with sufficient understanding. But a lot of other people can't, so they are more easily kept ignorant or blinded by stereotypes or brainwashed by propaganda, and there is not much anyone can do about that, other than to try to protect them from bad influences.

Plato’s *Republic* is, among other things, a guidebook for isolating intellectually superior young citizens, training them to be philosopher-kings, and brainwashing their genetic inferiors into willing obedience. Almost no one in the West believes in rule by an aristocratic class today, but many of us do prefer to have people of high intellectual achievement running the government than people who have found success in business and practical life. And almost everybody thinks that experts ought to have some level of authority within their expertise. The essential question here is whether high intelligence or academic training makes someone more likely to be *right* about the controversial practical and moral questions that dominate our politics. There are two views on the subject. One is that intellectual superiority counts for little in matters where no expertise exists, and it can even get in the way when people believe that they are capable of knowing things that have not yet been figured out. In such cases we may well be better off relying on common sense – i.e. traditional beliefs and ordinary moral intuitions – than the untested speculations of some group of Ph.Ds.[[18]](#footnote-18)

The other, more common view is that *of course* intelligent and educated people are more likely to be right about things – what else are intelligence and education for, but to help us come to true beliefs and avoid false ones? In the hard sciences this point is obvious to all, perhaps less so in the social sciences and humanities, and least in matters of public moral controversy. But it stands to reason that the best minds tend to have the greatest understanding of all complex issues. Disagreement would be easier to understand, then, if our opponents were plainly sub-normal in intelligence, or if some wide gap within the normal range had been established scientifically.[[19]](#footnote-19) In most controversies, though, there is little or no such evidence available, only a sense that people on the other side *must* be rather stupid to believe what they do. But this explanation will embarrass us whenever our opponents produce representatives of undoubtedly high intelligence, education, and mastery of the facts at hand.

*Emotional weakness*.Perhaps some people disagree with us because they are not intellectually but psychologically disabled in some way. There may be nothing wrong with their strictly cognitive capacities; it is rather at the emotional level that they go astray. This is our most common sort of diagnosis of opponents who refuse to see reason and change their minds about important issues. For example, it is often said by atheists like Sigmund Freud that religious people just can't face the fact that God does not exist – they are too infantile and weak emotionally to live their lives without the crutch of religion. On the other side, believers tend to diagnose atheists as too arrogant and narrow-minded to appreciate the mysteries of faith. Conservatives are sometimes seen as fearful, insecure children who need to depend on their parents' authority, while progressives are likened to teenagers who rebelliously reject whatever their elders have to say. In all these cases, the important causes of disagreement are not the rationalizations people give for their beliefs, but the non-rational forces that underlie such surface reasoning. If they are easily brainwashed, it is because their emotions are easy to manipulate.

Most writing that purports to offer psychological analyses of religious and political disagreements seems aimed at charging up partisan feeling rather than offering careful and serious arguments.[[20]](#footnote-20) And in casual discussion, we hear and probably say things like this all the time: that our opponents must be sick, crazy, out of their minds, etc., even if we have no more specific disorder in mind. Sometimes new diagnostic labels such as *homophobia* have been created, seemingly just to give opposing political beliefs an odor of mental illness.[[21]](#footnote-21) But other psychological approaches to diagnosing major disagreements have achieved a measure of scientific respectability in recent years. Unsurprisingly, given today’s academic climate, much of this effort has been focused on trying to understand what’s wrong psychologically with people on the political right, and studies have offered varied analyses including the ideas that conservatives are more fearful and more easily threatened than progressives, as well as more belligerent, more easily disgusted by people and things perceived as dirty, and less comfortable with complexity and nuance.[[22]](#footnote-22) On the other side, progressives have also been diagnosed as fearful and drawn to authority – except that it is government authority, not that of religion or traditional social hierarchies that supposedly shelters them – or, contrarily, as full of rage against their parents.[[23]](#footnote-23) This sort of research is certainly suggestive, but nothing has been clearly established yet, and we must always be careful not to confuse causation with mere correlation. If either progressives or conservatives are particularly fearful or angry, for example, it might well be their beliefs that *make* them particularly fearful or angry (i.e. now that they see dangers or injustices that their opponents aren’t aware of), not the other way around.

A new program in psychology called Moral Foundations Theory attempts to explain moral disagreements in terms of people’s different ways of balancing five basic intuitive values that our species has evolved, without intending to imply that one resulting moral perspective is always better than any other. In this view, political progressives are people who tend to view the moral world mainly in terms of the values of caring for others and fairness, while conservatives pay relatively more attention to the values of group loyalty, respect for authority, and purity (the inclination to enforce hygienic and social taboos). This is a very plausible view of human moral instincts. Indeed, it would be odd if we had *not* evolved a strong set of instincts to care for each other, to distribute things fairly, to protect our tribes against outsiders, to defer to parents and other leaders, and to avoid things and behaviors that are liable to harm us or our neighbors. And it makes sense that as in many other things, our different individual natures incline us to favor each of these instinctive values to greater or lesser extents.

This approach to moral psychology shows promise in helping us to understand the moral judgments of others as neither stupid nor crazy, but as arising from a shared framework of instinctive attitudes. But can this approach actually solve the problem of disagreement among epistemic peers? There seem to be three possibilities. First, if we accept that there are objective facts about right and wrong, as most people do, then presumably, different ways of balancing moral intuitions will tend to lead us closer to the truth or further from it. If we persist in sticking with our own side, we will still need an account of how we got the balance right and our opponents got it wrong. Second, if we say that each side is right about some issues and wrong about others, then it looks like some intermediate combination of the foundational attitudes, or a more flexible approach to balancing them, is required of everybody. This puts us in a “third party” position, disagreeing with almost everybody about one issue or another. Now we seem even more arrogant than people on the two major sides: how come no one but us is getting things completely right? And third, if we say that neither side is right except in terms of its own values – it’s all just a matter of different but equally valid moral perspectives – then we must say that people who believe in objective morality are positively wrong, and only relativists like ourselves are right. But if we mean to treat the non-relativist majority as our epistemic peers, then we will need to account for this higher-level disagreement about the nature of morality. How is it that we understand morality in general and they do not? None of these alternative views gives us a satisfying explanation of how others can be just as smart and sane and well-informed as we are in general, but wrong on some of the big issues that everybody cares about while we are right.

*Moral wickedness.* Finally, we must consider the possibility that those who disagree with us persistently, in the face of conclusive evidence that their beliefs are false, are inferior to us neither intellectually nor psychologically, but rather morally. They see enough, and they are rational enough, to form the right belief, but they still *choose* to be wrong rather than right about the matter. Sometimes, for example, one group of people simply hates another group for no reason at all and persecutes them to the point of slavery or genocide, even despite their own religious doctrines to the contrary, in the way that Jews have been persecuted by Christians and Muslims (not to mention ancient Egyptians and Babylonians) over the centuries. There is no reasoning with such pure, hate-filled wickedness, and no real cure for it either. There is only to protect ourselves and others from enemies like this by any means necessary.

Sometimes we say that our opponents are not so deeply wicked but only greedy, corrupt, or otherwise morally weak. Here the opponents are not said to be blinded to the evil things they do by neurosis or stupidity, or by fundamentally evil motives, but rather because they choose to turn away from inconvenient truths and toward convenient falsehoods for the sake of personal gain. We see this diagnosis often in arguments over energy and environmental policy, where industrial scientists are accused of selling out to corporations for big salaries, and university scientists are accused of selling out to government funding agencies for big academic careers. Politicians on both sides are often seen in this light, too, as selling their consciences to selfish interest groups or bloodthirsty mobs of supporters just for the sake of personal power and prestige, and maybe money on the side as well.

Common diagnostic labels like racism and sexism are used equivocally between psychological and moral meanings. When the racist or sexist is said to be “sick”, this is ambiguous in the same way. It is a pretty good test of the difference to consider whether we sympathize with these people as victims and want to cure them, or detest them as enemies and want to punish them. It seems that most of us are less inclined to flat-out moral condemnation than most of our ancestors were, and less inclined to have our opponents set on fire or beaten to death with clubs. But to whatever extent we relish having our opponents fired from responsible positions or publicly scorned or humiliated, to that extent we see what’s wrong with them as chiefly moral rather than psychological. Immorality is therefore the hardest diagnosis to retreat from, since we tend to see it as not only frustrating or even futile, but also wrong in itself to pursue mutual understanding with people we see as enemies. Even the pretense of epistemic peerage with such people is odious, not because our opponents can’t believe the truth, but because they won’t.

Here is where we stand with the problem of disagreement. Some people try to wash away all serious disagreement with the global sorts of explanation at the top of this list: general or cultural relativism, skepticism, or some similar doctrine. I have suggested that general relativism makes no sense at all, and that the milder variants do little to solve the actual problem of substantive disagreement, since people who disagree rarely accept such explanations and, in fact, tend to resent them, which creates a new three-way disagreement rather than resolving the original two-way disagreement. I also wonder whether some people are fully sincere when they profess a more respectful attitude toward their religious or political opponents than they are actually able to display, except in the most abstract terms. I do not mean that they are hypocrites, just that it’s hard to see how anyone who takes their own positions seriously can simultaneously assert that contradictory beliefs are just as good. Beyond this, I will not say much about these standard global accounts of disagreement, since they are thoroughly debated elsewhere, and I have little new to add. But I certainly sympathize with the main motivation for such theories, which is to treat people with different beliefs as even-handedly as possible. Indeed, my own theory of disagreement could be placed largely in the same general category, but with differences that I think make it credible to people who want to retain their controversial beliefs.

Most of us tend to account for most of our disagreements one at a time, using some mixture of diagnostic explanations: ignorance, blindness, and so on, while still trying to salvage some respect for our opponents. But I find it hard to see how we can reasonably use any of these explanations to account for disagreements that we have with people we consider epistemic peers: first, because they strongly imply that our opponents aren’t truly our peers; and second, because our opponents can always aim the same types of explanation back at us. I say that God exists (or doesn’t), and you say that God does not exist (or does). I say that you are wrong, and you say that I am wrong. I say that you are ignorant, blind, brainwashed, stupid, sick, or wicked, and you say that I am ignorant, blind, brainwashed, stupid, sick, or wicked. I feel sure that I am not the one with the problem, but again, you feel the same way about yourself. At most one of us can be right. So, again, what good reason do I have to think that it is me this time, when everything that can be said for me as a reliable believer, independently of the immediate dispute, can as easily be said for you? And again, I could try to turn the argument around, and use your being wrong about the point at hand as a *criterion* for judging you to be unreliable in general. But the issue of objective symmetry would still arise. If one person thinks another must be unreliable because he is an atheist, and the second thinks the first must be unreliable because he is a Christian, then neither person is entitled to use the other’s unreliability as an *explanation* for their initial disagreement between Christianity and atheism. That would be reasoning in a circle, like saying that certain Europeans are disgusting people because they eat snails, and that they eat snails because they are disgusting people. I am not saying that our beliefs are always equally true or justified, or that we are all equally reliable observers of the world. But in the absence of any objective evidence of significantly different levels of competence (e. g. one of us having some obvious problem with ordinary, non-controversial reasoning), it still seems to me that neither of us has sufficient reason to believe that his conclusions are more probably correct than his opponent’s, all things considered.

After having tried out this argument in various forms on any number of my own peers, I am struck by the force of the resistance, sometimes even anger, that it seems to provoke. I discover that there is something wrong with treating disagreements, especially important ones, in this purely probabilistic way. My friends keep telling me what I have always said myself, but somehow keep forgetting when I think in terms of probabilities, which is that we are all supposed to *think for ourselves*. We can't just conform ourselves to other people’s expressed beliefs, especially in areas like politics where disagreements tend to be most obstinate, or give up whenever we meet any opposition from our peers, or even from authorities like parents and professors. We need to be independent, critical thinkers, speaking our own minds and standing up for our beliefs, however unpopular they may be. Nothing is nobler, it often seems, than the person of principle saying and doing what he believes is right, though in the face of overwhelming opposition. I have made a lifelong, sometimes showy point of trying to follow these principles myself, admired the same sort of intellectual behavior in others, and encouraged it in friends and students. But I now feel that I never really understood these principles, much as I liked them, and I have come to wonder whether anybody does. Somehow, these rules of thinking for ourselves and standing up for our beliefs seem to keep overriding considerations of mere probability. But how can that be, when rationality just *means* seeking the beliefs most likely to be true? What is going on?

1.3. Three dimensions of belief

Here again is the problem of disagreement, as concisely as I think it can be put:

(1) When peers disagree, they are each very likely to be wrong.

(2) Those who are very likely to be wrong ought to suspend belief.

(3) Therefore, when peers disagree, each one ought to suspend belief.

This argument from (1) and (2) to (3) is plainly logically valid. The problem is that premises (1) and (2) seem to be true, while conclusion (3) strikes us as false.[[24]](#footnote-24) On the one hand, it seems obvious that we ought not to believe things that are very likely to be false, which means that statement (3) *should* be true, since it follows from (1) and (2). But none of us wants to just cave in whenever we are challenged by our peers – and if we all did that, then nobody would have any controversial beliefs at all. So statement (3) *must* be false. What’s going on here, that leads us into this apparent paradox?

Here is what I think is going on. I think that the word "ought" can be used in different ways depending on the goals or purposes that speakers have in mind, and that this word is being understood in one way when statement (2) seems to be true, and in another when statement (3) seems to be false. Sometimes we mean by "ought" what people need to *think* in order to be *rational* – what's sometimes called the "epistemic ought". Sometimes we mean what people ought to *do* in order to be *good* – what's called the "moral ought". And there are many other uses of the word "ought", depending on whatever other goals the person using it might have in mind. Thus, a statement that somebody ought to stop smoking, assuming the goal of the smoker's overall well-being, might be said to be using the "prudential ought". Most other uses of the word have no special name, but they are just as meaningful. For example, we might say that a person playing chess ought to advance his bishop, meaning only that if he does so, it will help him win the game. Or we might say that hot cocoa ought to have a little cinnamon in it, meaning only that it tastes better that way. For such practical uses of the word "ought" – I will call this broad category the "pragmatic ought" – it is usually clear enough from context what the goal is that the speaker has in mind. In these ways, statements using the word "ought" are almost always *logically incomplete*. That is, if somebody says that someone ought to think or say or do something, we're always able to ask, "For the sake of what? What purpose do you have in mind?" And the answer might be epistemic, moral, or in innumerable ways pragmatic.[[25]](#footnote-25)

In the argument above, the "ought" in premise (2) is epistemic. When two people know that they about equally likely to be wrong, the purely *rational* thing for them to do is to suspend belief. So, if we suppose that peer disagreement indicates that each peer is likely to be wrong,

statement (3) follows immediately: each disagreeing peer ought to suspend belief. Is this a problem? Not if we really understand exactly the same thing by "ought" in (3) that we understand in (2). It is perfectly fine to say that disagreeing peers always ought to suspend belief, if all we mean is that this is the strictly rational thing to do. But I suspect that this is not the sense of "ought" we have in mind when we find (3) objectionable. We want to deny (3) not because it doesn't follow from true principles of rational belief, but because we are opposed to *conformity*, to people shutting up and giving in whenever others disagree with them. We think that people ought to think for themselves and stand up for their beliefs, and thinking for yourself and standing up for your beliefs are not matters of simple rationality. These principles serve other important goals, ones that are pragmatic and moral, not strictly epistemic.

Here is my general solution to the problem of disagreement. There are three different, and sometimes conflicting, principles that we all subscribe to when it comes to most of our beliefs. I will call them the principles of rationality, autonomy, and integrity, using the ordinary English terms that seem to fit my theory best:

The *principle of rationality* is that you ought to believe whatever is most likely to be true. More precisely, you ought to believe with *greater* confidence whatever is *more* likely to be true, given the total pool of evidence available to you, and you ought to adjust your confidence accordingly whenever new evidence appears.[[26]](#footnote-26)

The *principle of autonomy* is that you ought to think for yourself. Do not depend on anybody else’s word for your beliefs. You have a right to your own point of view, and you ought to be willing and able to back up your opinions with objective evidence and arguments, not somebody else's authority.

The *principle of integrity* is that you ought to stand up for your beliefs. It is important to make your statements and actions consistent with your moral and practical commitments. If you really believe in something, you should never equivocate or back down when you are called upon to state or act on this belief.

These rules use different senses of the word "ought", reflecting different fundamental goals. In their central functions, the principle of rationality is epistemic, the principle of integrity is moral, and the principle of autonomy is pragmatic in a way that I will talk about below. But this is to idealize their differences. In ordinary cases, the three principles overlap and reinforce each other, so that we're not aware of any difference. There is nothing uncommon about thinking for yourself, coming up with a belief you see as likely to be true, and acting accordingly, all without any sense of inconsistency. Thus, a person might conclude, through independent consideration of the evidence and arguments available to him, that slavery, abortion, capital punishment, or eating meat is wrong. Then the person might stand up for this belief in arguments and actions, without ever wondering if he is being rational *as opposed to* thinking for himself or acting with integrity. It’s just a belief that he has because it seems true to him, so *of course* he speaks and acts accordingly. It would be strange for him not to. We apply the three rules faithfully without distinction in all sorts of ordinary situations, too. For example, my wife recently came to believe that our dog ought to visit the veterinarian, based on some unpleasant evidence that she had been finding around the house. Although I told her that she shouldn’t worry about it, she stuck to her guns and took the dog to the vet, who discovered that she (the dog) had worms and treated her accordingly. No big issues of principle here, just an undistinguished case of rationality, autonomy, and integrity working together as belief in daily life.

But the three principles also sometimes conflict, one taking precedence over another. There are six ways that this can happen. For now, let me illustrate them with a few quick examples, pending extended discussion in later chapters.

Rationality tends to trump autonomy whenever a lay person has less knowledge than the relevant experts. I have not invented my own particle physics, for example, but defer to physicists for my beliefs, such as they are, about the subatomic world. Having no real training in physics, I'd have to be an idiot to try to figure that stuff out for myself. Similarly, I allow my doctor to convince me that I’m only suffering from something called “acid reflux”, despite its feeling to me just like I imagine a heart attack would feel. I prefer the testimony of my family, friends, colleagues, and many others to my own guesswork about all sorts of matters great and small, just because I think they’re more likely to be right about them than I am. Have we had the car inspected yet this year? I kind of think so, but I had better ask my wife, who keeps much better track of such things. Does this wine taste the way it is supposed to? It seems okay to me, but I will ask somebody with a more refined palate, i.e. practically anybody. Thus, my desire to believe what's true often constrains my efforts to think things out independently, because I know that other people have a better angle on the truth of many things than I do.

Autonomy tends to trump rationality when it becomes important for new ideas to be produced for any reason, including not just serious work, but also training exercises, or even simple pleasure in discussion. In graduate school I once made a joke in a footnote to a paper that defended Aristotle on some point, saying that his position actually made no sense at all to me, but that since Aristotle was a much better philosopher than I was, it was still very probable that he was right. My professor was not at all amused, and wrote harsh comments on what for me was just a frivolous aside.[[27]](#footnote-27) Clearly, the main purpose of writing philosophy papers is to develop strengths in making our own arguments, not just to point out the probable truth, especially through humble deference to authority. Professors are properly rather touchy on this point, and students are sometimes expelled for making perfectly sound statements that are not their own work. [rep §6.3]

Rationality overrides integrity whenever a person believes that something is true, but has insufficient reason to speak up for that belief or to act according to its implications. For example, I believe with middling confidence that Shakespeare's works were largely written by the seventeenth Earl of Oxford, a man named Edward de Vere. That is, I take it as fairly probable, based on the modest reading I have done about the subject, that Oxford was at least heavily involved in all these projects, and that the man William Shakespeare was to some extent a front for Oxford.[[28]](#footnote-28) But this somewhat eccentric belief is not one that I take to heart in action; in fact, I am rather ashamed of it, given the way that my colleagues in the English Department have responded on the few occasions when I have had the nerve to bring the matter up. I do not belong to the Oxford Society or engage in scholarly debates with experts, or even spend much time pursuing the matter. I have no moral or personal stake in Oxford’s authorship at all – if anything, I have a stake as a philosopher in being skeptical about all such questions. Nevertheless, I do believe on balance that the “Oxford hypothesis” is probably true, given the evidence that I have come across. So, as a purely rational proposition I believe it more so than not, but I don’t actually care about it, and it requires no action from me anyway, so the principle of integrity has no real bearing on the issue for me.

Integrity tops rationality whenever a person finds it morally or practically necessary to commit to a belief without having sufficient evidence from a detached, objective point of view that the belief is true. Religious believers make such commitments in the face of great uncertainty – even apparent incoherence – when they take the “leap of faith”. Parents and children have a duty to value and honor their children and parents, which entails believing good things about them, regardless of whether they would judge those people to be valuable or honorable from a neutral perspective. Thus, when Garrison Keillor says that in his mythical town of Lake Wobegon “all the children are above average”, this is not just a joke. What he’s saying implicitly is that they all have loving and dutiful families, so that they are all *believed* to be above average.

Autonomy trumps integrity whenever we come to a belief by thinking or choosing for ourselves, but that belief implies no active commitment. In casual conversation, and especially when we are working in a creative or “brainstorming” context, and even in serious scholarship we sometimes make statements with little or no sense of personal ownership, however true they strike us at the moment. Philosophers are sometimes thought to lack integrity when we do not live entirely according to the dictates of our current theories. But this depends on whether we what we say is actually making moral *demands* on other people, or simply making moral *arguments* for others to consider. As I will discuss later on, in many cases it is unclear where scholarship leaves off and activism starts, so this can be a difficult question. In any case, to whatever extent the statements we make about morality are merely philosophical ones, our integrity is not at stake and our ideas should not be treated as personal commitments.

Finally, integrity trumps autonomy whenever we are actively committed to the truth of a belief that we have not derived or thoroughly evaluated for ourselves. On a common (but not universal) understanding of the concept, true patriotism requires not just that we act in defense of our homeland when called upon, but also that the action be in a sense unthinking or automatic, based on an implicit belief in the rightness of our own side. For some people, there is something shameful about someone taking an independent, skeptical view of patriotic matters, even if he decides on balance to support their own side. Thus, brilliant and wonderfully open-minded thinkers like Bertrand Russell have sometimes been viewed with disgust, or in Russell’s case even jailed, by their countrymen when they refuse to fall into line with everybody else on what is generally felt to be a necessary war.[[29]](#footnote-29) The formerly pacifist Russell’s (fleeting) demand for a militant Anglo-American posture against Soviet Communism just after World War Two earned him few friends among those who placed their loyalty above their individual perspective. Of course, from Russell’s own point of view this was a matter of principle, since in his view independent thinking was a supreme moral responsibility in itself, higher than any automatic duty to a group. For him, but not for everybody, there is no such thing as having integrity without autonomy.[[30]](#footnote-30)

In all these ways, then, our beliefs are controlled by the three different principles either together or separately, depending on our purposes in holding them. The purpose of rationality is to give us a most-probably true picture of the way things are, based on whatever evidence we have available. The purpose of autonomy is to help us produce new ideas, mostly for other people to evaluate, and regardless of whether we are in a position to believe them rationally at the time. And the purpose of integrity is to bind ourselves prudentially and morally to those beliefs we take as basic to our active lives, regardless of whether we are rationally required to believe them on the evidence available, and regardless of whether we work the beliefs out for ourselves or accept them on authority or faith. What should we think of belief *itself*, then, that it should be pulled in these three different directions? We use the same word “belief” in every case, so it is correct to think of belief in general as one thing that is subject to three different rules, depending on which purpose or purposes different beliefs are serving. But we also sometimes use separate labels for our beliefs when we consider them in terms of their different basic aims. I want to say that belief is not a single thing and not three different things, then, but that there are three aspects or *dimensions* of belief, deeply enough connected that the concept of belief retains its unity despite a tendency to split into three different types. I call them the dimensions of *perception*, *opinion*, and *conviction*, corresponding to the principles of rationality, autonomy, and integrity, respectively. Here again, I am using what I think are the most common and agreeable terms in ordinary English, while distinguishing the meanings of these terms somewhat more sharply than we usually do, and maybe also bending them a little bit to fit my theory.

*Perceptions* are mental representations. They comprise the inner models that we base initially on our sensations, which are then connected through memory and reason into a total picture of the world as we perceive it. There are also various forms of indirect perception, including learning things through testimony. Belief as perception is properly governed by the principle of rationality.[[31]](#footnote-31)

*Opinions* are potential contributions to discussions. In many areas of life, decisions need to be made socially. It is important that multiple points of view be represented independently, so that they can properly compete for general acceptance. To have an opinion is to be prepared (at least abstractly) to engage in controversies of that sort. Belief as opinion depends on our maintaining separate individual perspectives, then, and should be governed by the principle of autonomy.

*Convictions* are *judgments* that form cognitive platforms for action. To make a judgment is to fix a belief in place, protecting it from revision at least momentarily. To form a conviction is to commit to a judgment more or less permanently, taking responsibility for all its practical and moral consequences. Belief as conviction ought to be governed by the principle of integrity.

In ordinary life, we usually treat belief as a single, yes-or-no sort of thing: for each statement that might be considered, either you believe it or you don't. This is also the most common theoretical position taken by professional philosophers, who usually take one of what I call the three dimensions as definitive of the whole concept.[[32]](#footnote-32) Others allow for two types or levels of belief, distinguishing what I call perception from what I call judgment in light of the relative passivity of ordinary perception and the active decision involved in judgment. Descartes marks this distinction in terms of the operations of what he calls the *intellect* and the *will*, where the intellect passively perceives and the will makes all the yes-or-no judgments.[[33]](#footnote-33) I say that there are three dimensions of belief, not one or two, because perception, opinion, and conviction play very different roles in our epistemic lives. And I want to call these things dimensions of belief, rather than types or levels of belief, because they do not form distinct categories, but overlap in various ways. Some of the things we call beliefs seem to lie along only one of these dimensions, but most beliefs extend into two of them or all three, so it is usually meaningless to ask which kind of belief might be in question. But our beliefs can still be looked at and understood from three main angles or perspectives in this three-dimensional space.

There is nothing really unusual in this analytical sort of distinction. For a simple example, we might say that Barack Obama is a president, a father, and a golfer, which are different things, but not that he is three different people. Instead, we say that he is one person with three different roles or, one might say, dimensions. So, when people argue over whether Obama is good or bad, we might reasonably ask in which of those roles he is being judged.[[34]](#footnote-34) Similarly, it is sometimes useful to follow Aristotle's distinction between the form of something and the matter that composes it. For example, we usually think of something like a statue as a single thing, which is fine for most ordinary purposes. When things go wrong, though – for example, when someone borrows your bust of Aristotle and returns the "same thing" in the shape of a potato – we rely on the distinction between form and matter to make sense of what has happened (the same matter has been returned, but not the same form). Belief is similar: we properly think of it as one thing for most ordinary purposes, but when tenacious disagreements arise, we need to attend to the differences between perception, opinion, and conviction.

This three-dimensional conception of belief has its origins deep within human nature. For, what are we as human beings, but things that think, speak, and act? These are our three most basic operations – the cognitive, discursive, and active dimensions of our lives. Each dimension has its own basic goal: the cognitive, to understand the world as it is; the discursive, to share information and construct consensus; the active, to make changes in the world for individual and social purposes. As these different human operations often aim at different ends, belief is bound to have a somewhat different function relative to each of them. Of course, the three basic human purposes usually interact and intermingle – speech is a special case of action, after all, and certainly bound up with thought – so that we each make sense as one whole person rather than three. In the same way, belief remains a single thing with three overlapping functions, ordinarily harmonious but sometimes not.

It may seem that what I am talking about here is not three different dimensions of belief, but really only three degrees or levels of *certainty* in a belief. So, a perception that seems fairly certain may rise to the level of an opinion, and an opinion held with greater confidence than that may qualify as a conviction.[[35]](#footnote-35) It is true that we usually have to see something as true before we are willing to say it, and be willing to say it before we act upon it. But this is not just a matter of how confident we are in a belief. The force of reasons along the three dimensions of belief, and the degree of harmony or integration between them, depend on the facts of the particular case. Each person's total set of beliefs is always tending toward a state commonly called *reflective equilibrium* given the rational value of coherence, the social value of consistent self-expression, and the moral demands of integrity.[[36]](#footnote-36) But what each of us believes at any moment ordinarily includes a good-sized garble of unabsorbed and incompatible perceptions, opinions, and convictions. When such conflicts come to the surface, they can leave a person in a state of disagreement with himself as well as many of his peers. Those who tend not to analyze their own beliefs will not be bothered much by the unresolved epistemic conflicts that sometimes underlie their speech and actions, other than to feel somewhat confused when they are pressed to explain their inconsistent claims, or why they say one thing and do another. But the different criteria we use to justify the different aspects of belief can cause a lot of trouble for someone who is making a real effort to think, speak, and act consistently.

Our opinions, for example, typically express what we perceive, but sometimes they don't. This is a little hard to see. When we take a position on a controversial matter, we will ordinarily consider how the relevant facts appear to us, directly or through testimony from sources we trust, and base our opinion directly on that. With religious or political opinions, the normal case is for people to gather most of their opinions from peers and authorities, endorse them with or without a few individual adjustments, then pass them on in public discussion as partisans of this or that faction, rather than purely individual participants. In cases like these, the principles of autonomy and rationality function together, producing something like a cluster of group beliefs rather than large numbers of strictly personal points of view. But our religious and political opinions still reflect how things appear to each of us as individuals. Legal and medical opinions are usually based directly on perceptions, too – sincere statements of how things appear to the practitioners, depending on their own professional experience, plus their reliance on other experts and authorities. So, in all these cases, opinion expresses perception and nothing else.

Note, however, that when we want to stress that our opinion coincides with how we really see things, we often say “It’s my *sincere* opinion that…,” which vaguely suggests that there are insincere opinions, too. This seems a strange idea, because it is hard to see why anyone should be prepared in general to say things that he thinks are false. But formal medical or legal opinions can sometimes be insincere, given that such assessments are constrained by professional criteria that may not always reflect an expert’s overall subjective view of things. A lawyer might argue in court that such and such a precedent supports dismissal of his client's charges, even though he thinks that any competent judge would probably reject the argument, because he has a duty to speak in the interests of his client, not to express his own, personal view of things. Or, a doctor might present his patient with an official opinion giving a poor prognosis based on his best medical judgment, considering only those pieces of evidence that the greater medical community considers relevant, while privately believing, perhaps for no good reason that he can articulate, that the patient will probably survive much longer than the doctor has any right to say.[[37]](#footnote-37) Or, a scientist who thinks that some disaster is, say, thirty percent likely to occur might claim consistently in public that the catastrophe is *certain* to occur unless drastic action is taken to prevent it, in order to motivate others to support that drastic action. (This might be thought to constitute an insincere but morally well-founded opinion.) In such professionally or morally grounded cases of opinion, people are not just expressing their own perceptions, but saying what they think they ought to say. There are trivial cases of opinion clashing with perception, too. Sometimes people express opinions that are backed up only by whims or desires, rather than anything substantial enough to be called a perception. For example, my own annual opinion that the Red Sox will win the American League pennant has never been based on anything more than the most casual acquaintance with the facts, plus a long history of wishful thinking. It is only through a perverse refusal to take past Sox performance seriously that I can ever show much confidence in this belief. Nevertheless, my friends who favor the Yankees (and sometimes take my money) will cheerfully testify that it is indeed my firm opinion, every stupid summer, that the Sox will win the pennant in the fall.

There can also be clashes between perception and conviction. A mother whose son has just killed somebody else for money or pleasure might clearly perceive that her child must be guilty of the murder, but still insist, even to herself, that the son is a good boy who simply couldn’t do that sort of thing, and make every effort that she can to get him back out on the streets.[[38]](#footnote-38) Our religious commitments also sometimes contradict perceived facts, as someone who is deeply convinced of the efficacy of prayer may well acknowledge that good scientific studies of prayer’s effects in hospitals yield no positive results. He still believes in prayer, he might say, just as a matter of personal faith or conviction, not scientific judgment. In such cases, perceptions that would ordinarily call a person's conviction into question don't, precisely because this sort of belief doesn't depend entirely on probabilities: it is *just* a conviction. This notion of personal conviction is not unlike conviction in the law. The jury produces a conviction not because they believe in someone’s guilt with perfect confidence – this could hardly ever happen among rational observers – but because a certain *burden of proof* is met according to certain set criteria. In some cases anything over 50% will do (i.e. “preponderance of evidence”); in others, “clear and convincing evidence” is required, and in most criminal cases in the U.S., “proof beyond a reasonable doubt”. But all these standards fall below 100% confidence, because we cannot in practical life remove absolutely all possible doubt, yet we still have to take action, sometimes quickly (the US Constitution guarantees a “speedy trial”), sometimes irrevocably.

Our opinions can also run counter to our convictions if we ever have consistent reason to deny in public any of our private judgments. For example, a federal district judge who is convinced that the Supreme Court's decision in Roe vs. Wade ought to be overturned might nevertheless rule that certain abortion procedures may constitutionally be banned, based on the fact that Roe v. Wade is now widely considered settled law. His personal conviction is that states do have a right to ban the procedures in question under the Constitution, and he does everything he can, consistently with his other commitments, to vindicate that right. In the meantime, though, he understands that his official opinions as a judge have been constrained by precedent in a way that his private convictions have not. Similarly, some progressive people hold the opinion that educational opportunities ought to be equal for everybody, while convinced as parents that they must provide their own children with better-than-average educations if they can. Thus, any number of progressive American politicians have placed their own children in the most exclusive private schools in Washington, while publicly opposing policies that would make these schools more available to the poor students that they claim to represent.[[39]](#footnote-39) These politicians might be seen as hypocrites, i.e. as people without integrity. They might also be viewed as decent parents like the rest of us, who suffer from some inconsistencies among their beliefs, as most of us do. In such cases, where acting strictly according to your public opinions would violate your own inner convictions, it can be a mark of virtue *not* to be consistent.

I think that a full solution to the problem of disagreement is to be found only within this three-dimensional account of belief. What explains most of our disagreements, in my view, is not that many people are irrational, and not that there is anything illusory about the disagreements themselves. Instead, there turn out to be three reasonable forms of explanation for most of our important disagreements. First, we often have no choice but to perceive things differently, given all of the empirical, and especially testimonial, evidence we have received. Second, people have diverse opinions largely because it is good for society that a robust jumble of ideas exists from which communities can draw in handling problems in practical life. This pushes bright people who desire to be useful into the roles of intellectual creators, advocates, and critics as distinct from simple seekers of the truth, and stretches the range of propositions over which we are liable to differ. Third, our lives are too short, and there is too much work to do, for most of us to dedicate ourselves entirely to the pursuit of true beliefs. At some point, if we wish to be effective human beings, even effective *thinkers*, we are forced to commit ourselves to some imperfectly established propositions. It is a difficult question just when and how we ought to make such ramifying choices. But once we make them, our different convictions can force us onto radically divergent intellectual paths. There are three dimensions of disagreement, then, which, as they correspond directly to the three dimensions of belief, I will call differences of perception, differences of opinion, and differences of conviction.

*Differences of perception* occur when rational people have conflicting sets of evidence for their beliefs. This comes about through either different direct experiences or different sources of testimony. Sometimes people are simply forced into different beliefs by drawing unavoidable conclusions from misleading or incomplete pools of evidence.

*Differences of opinion* occur when autonomous people come up with different things to say about a controversial issue. This comes about because in thinking for ourselves, we must resist just going along with one another. Fruitful discussions typically require a diversity of views, so opinions sometimes have greater practical value just because they figure in disagreements.

*Differences of conviction* occur when people of integrity commit themselves to different practical or moral principles for purposes of action. This happens because decisions often must be made despite imperfect information and hasty analysis. Our needs and responsibilities sometimes force us, under the pressure of time, simply to make up our minds. This demand for commitment can amplify a minor difference of perception or opinion into a hardened conflict.

Pure cases of these three aspects of disagreement are easy to deal with, or at least understood. Differences of perception often arise without interference from the other two aspects of belief. For example, my wife and I frequently disagree on whether we have already seen some movie that we are thinking of watching. There is no matter of conviction involved for either of us here, or even of opinion, properly understood.[[40]](#footnote-40) It is just a difference of perception due to different memories, and it is easy to resolve by sitting through the movie until I recognize it, usually about ten minutes from the end. Or a colleague and I, relying on different magazine accounts, might disagree on whether Russia or the United States has more nuclear weapons. Here again, there is nothing much at stake for either of us, just conflicting perceptions due to different testimony. We can resolve this temporary disagreement by checking with mutual authorities, which are usually easy to find in such matters of fact.

Pure differences of opinion take longer to clear up, for the reason that such questions have by definition not yet been answered authoritatively. Political commentators tend to produce conflicting opinions as to what will happen in the next American elections, based on exactly the same polling results, electoral history, and other public facts, perhaps because they wish for different things to happen, or perhaps just to be interesting. If they are reasonable people, though, they will agree that this is *only* a matter of opinion, and await its resolution after the election itself.

Pure differences of conviction are not easily resolvable through rational argument. Crusaders and Saracens, Secessionists and Unionists, Tsarists and Bolsheviks have usually just pounded each other until one side or the other won the war. There are alternatives to violence, of course. In a mixed Christian society, Catholics and Protestants might choose to live in different neighborhoods to minimize the tensions caused by their clashing faiths. Those forced into closer proximity usually find ways to quarantine their disagreements. So, for example, if a politically conservative American Catholic and a progressive atheist get married, they might well “agree to disagree” on matters of religion and politics, avoid discussing these topics over dinner, quietly read their different magazines, and work out the children’s Sunday schedule by diplomacy rather than consensus.

In most ordinary disagreements we find differences of perception, opinion, and conviction all mixed in together. Our most tenacious and exasperating disputes over religion, morality, and politics are typically confused in this way, with no systematic attention paid to the distinctions between questions of ordinary information, matters open to argument, and fixed points of principle, each of which inflects the others. To an outsider, it can seem as if the two sides in such debates actually live in different worlds (a thesis that some relativists take as the literal truth), or speak different languages. To an insider, it often seems as if the people on the other side are so far from the obvious truth that they need to be diagnosed and treated rather than understood. Neither perspective is correct; no general theory or particular attack is likely to resolve the substance of our most interesting disagreements. Instead, a fair analysis of such disputes along the three dimensions of belief will usually tend to show that those on both sides are doing their best as rational people with the evidence they have. They do not live in different worlds, but they do live in different total *epistemic situations*, i.e. total histories of evidence and arguments available to them. If we take the right approach such situations can be understood objectively even by hardened adversaries, in terms of their perceptive origins and subsequent rational inferences. Seeking this understanding is an essential step towards resolving complex disagreements when they are capable of resolution, and towards diminishing the turmoil when they are not. This will require less than a Cartesian effort to lay out all of the evidence and arguments on both sides piece by piece and step by step. At least much of the time, we should be able to achieve a working understanding of each other’s epistemic situations through a *reasonably* careful analysis of each other’s personal and testimonial resources, prior commitments, and present purposes.

Over the next several chapters I will flesh out this skeletal theory of disagreement and belief, beginning with the simplest matters of perception.

**2. PERCEPTION**

How often, asleep at night, am I convinced of just such familiar events – that I am here in my dressing gown, sitting by the fire – when in fact I am lying undressed in bed!

– Rene Descartes, *Meditations*[[41]](#footnote-41)

2.1. Sensation, memory, and reasoning

Perceptions, in the broad sense of the word I am using here, are beliefs that we derive and justify *empirically*,that is, ultimately based on sensory experience, for the purpose of representing the world to ourselves. Such a process requires, in the very first place, three cognitive abilities: we must be able to experience parts of the world, we must remember some of these experiences, and we must draw logical inferences from them. We could expect nothing less from a machine designed to figure things out than that it have devices to input, store, and manipulate information. For me to recognize my dog, for example, requires in principle that I should see the dog, remember that this is what my dog looks like, and reason that the similar dog in front of me is probably the same actual dog as before. In our perceptive experience, the three sources of empirical belief work together pretty seamlessly. As I watch my dog running around in the field behind my house, where does my actually seeing her leave off, and my remembering her begin? There is no obvious switch between our sensory perceptions and the memories that succeed them, but rather something like a continual fade from one into the other. Even the distinction between sense perception and reasoning is not always clear. When I look at my dog, do I see only certain shapes and colors, and then figure out that these “sense data” imply the presence of a certain animal? It is psychologically more accurate to say that I just recognize this object *as* my dog, without any clear separation between the relevant sensations, memories, and inferences.[[42]](#footnote-42) Still, the epistemic justification of perceptive beliefs requires that we consider sensation, memory, and reasoning as logically separable functions.

I want to understand the concept of sensation very broadly, not just to include the usual five senses, but also direct awareness of our physical orientation and arrangement, our senses of time, motion, and balance, and such other internal states as hunger, pain, and the feeling of an upcoming sneeze. Whatever brings us new information directly counts as sensation in this broad sense of the word. The external senses are all known to be fallible, as are the senses of motion, time, balance, and being about to sneeze. I suppose that a feeling of hunger is just the same thing as being hungry, so it can't mislead us as to what we feel. But it can certainly mislead us if we take it to inform us that we need to eat, just as feeling a pain tells us for sure that we are in pain, but doesn't guarantee that we have really been hurt. So, to the extent we see sensations as sources of objective information, none of them are totally reliable. Yet we cannot know, or even believe, anything about the outside world without them.

By memory I mean nothing unusual. Hume viewed it as a weak regurgitation of our sensory experience, something that seems to provide us with useful information about past events. But memory is not necessarily more truthful or *veridical* than direct sensation. In fact, it is less so, for it contains all of the fallibilities of our original sensations, plus a strong tendency of its own to mix things up, combining true memories with stories from others, wishful thinking, dreams, and fantasies, often degrading vivid experience into shadowy fragments.[[43]](#footnote-43) Nevertheless, memory gives us an indispensable form of access to the facts of our own past, and provides us with enough continuity in our immediate experience to stretch us over time as persons, not just momentary bundles of sensations.

Reasoning is commonly said to come in two general varieties at least. The first is deduction, which is the drawing of formally necessary conclusions from premises, as in mathematical proofs. Deductive reason also covers those conclusions that need no premises, because they are formally necessary all by themselves. For example, it is said to be logically impossible for “2 + 2 = 4” to be false, on the grounds that the denial of this statement is simply incoherent. This is why, since well before Plato, mathematics has served many philosophers as the ideal form of knowledge. For Plato, Descartes, and many others, philosophy itself must attain a similar certainty if its results are ever to be taken as established. Thus, philosophy is usually considered an *a priori* discipline, meaning that it concerns mathematical facts or whatever else might be known independently of experience, as opposed to *a posteriori* inquiries like natural science that depend on observation as well as reasoning. Ancient philosophers like Plato argued that even what we now call natural science does not deserve the name of knowledge unless and until it has attained the same *a priori* demonstrability as math.[[44]](#footnote-44) Still, even Plato acknowledged that there is such a thing as rationally justified belief that falls short of provable knowledge in empirical affairs, though he did not seem to find it very interesting.[[45]](#footnote-45) Contemporary philosophers have generally given up on Plato's mathematical ideal, and would be thrilled if their work could achieve anything like the prestige of modern science.

The second general variety of reasoning is induction, which is commonly (though not universally) seen as the essence of empirical reasoning.[[46]](#footnote-46) Induction includes reasoning from known to unknown facts (e.g. from past to future facts), and also from specific facts to general ones. The two forms often work together in a pattern of generalization and prediction, something like this:

All of the ravens that we have observed so far have been black.

Therefore, all ravens are black.

Therefore, the next raven that we observe will also be black.[[47]](#footnote-47)

This little pattern of inference, based fundamentally on counting cases, is called *enumerative* induction. Most philosophers think such simple formulas cannot account for the sophisticated forms of reasoning required by science. We believe that the sun will rise tomorrow morning, for example, not just because we have observed that it has risen every morning in the past and we expect most things to continue going on as they have gone before. We have also come to believe there are laws of nature, such as Newton's laws of motion, that *explain* its rising every day, and these explanations matter to our sense that this belief is rational.[[48]](#footnote-48) So, there seems to be more to proper non-deductive reasoning than enumerative sorts of induction. In any case, I will use the word "induction" as it is commonly used, i.e. to cover *all* types of rationally acceptable non-deductive reasoning, without taking a stand on how many types there are or how they ultimately fit together.

Induction and ordinary observation overlap as factors in perception, just as observation overlaps with memory. When we observe that a particular raven is black, for example, we do not bother to examine every bit of its surface. We simply take a look at the bird from one or two angles, remembering for the moment what we have just seen, and automatically "conclude" that it is black all over, or at least close enough for us to call it a black thing overall. This could be expressed in simple terms as an inductive inference itself, something like:

All of the parts of this raven that we have observed so far have been black.

Therefore, all parts of this raven are black.

Therefore, the next part of this raven that we observe will also be black.

By the same token, the conclusions of ordinary inductive inferences can often be expressed as single "observations" scattered over time, place, and observers. Thus, we might well say that *we* (collectively) have *seen* such general facts as that dogs have a fine sense of smell, or that nobody likes a cheater, or that Afghanistan is hard to conquer.[[49]](#footnote-49) Thus, there is no point in trying to draw sharp distinctions in practice between induction, observation, memory, and deduction as sources of perceptive belief. Our mental models of the world are psychologically constructed out of all these functions working together, plus any number of instinctual devices, such as an inborn understanding of facial expressions or the special sensitivity to language that is widely attributed to infants.[[50]](#footnote-50) But these perceptive models can also be *re*constructed philosophically in order to reveal the structure of reasoning that justifies them. This is the only time it really matters for us to distinguish observation, memory, induction, deduction.

There are two common philosophical frameworks for such reconstructions of empirical belief. The first and most traditional is *foundationalism*, which involves treating immediate sensations and memories as basic building blocks, and using deductive and inductive reasoning to construct more and more complex perceptions. Thus, I might experience sensations of softness and white color together with a certain set of shapes at one place and time, remember that complex experience, and recognize the same group of sensations when they occur again repeatedly together. Then I might reason inductively that these are all features of a single object. Then I might come to recognize more objects of the same type, call them rabbits, and associate them through further experience with other objects that I have come to call carrots, by way of an activity that I have come to call eating. At this point, I could form the inductively justified belief that rabbits are inclined toward eating carrots, and go on to other topics.

The second, more recent approach to epistemic reconstruction is called *coherentism*. It focuses on how an entire perceptive model of the world might be justified internally, just in virtue of the way it all hangs together. Since each of the things that I perceive is consistent with, and can be explained in terms of, all the others, I can observe that the whole big picture is coherent, which at least rules out a lot of ways that it could be false. So, I find myself believing that rabbits eat carrots, and attempt to justify this belief after the fact by tracing and examining its connections to other perceptions that I have. I see that I believe that rabbits look like this, that carrots look like that, that I have seen this or that rabbit eating this or that carrot – it all seems to fit nicely together. I also see that this fitting-together of beliefs about rabbits and carrots coheres with my other beliefs about animals and vegetables in general, and about the physical world as a whole – although the need for coherence may also force me to adjust my interpretations of direct experience, perhaps by changing how I distinguish rabbits from hares or carrots from parsnips. I also see that these physical beliefs cohere with my beliefs about how things of all sorts ought to fit together, including my present understanding of what constitutes good evidence and reasoning. I search my mind for other beliefs that might clash with this coherent set – contrary memories of my experience with rabbits, doubts about the healthfulness of carrots, concerns about the rationality of inductive inferences, and the like – and cannot find anything that contradicts the rabbits-eating-carrots cluster of beliefs that sits within my perceptive model of the world. Thus, my belief that rabbits eat carrots has been tested as well as I can test it against other things that I believe, and it has passed that test.

Both of these basic approaches to justifying perceptions appeal to common sense and both are useful, though both have also generated plenty of criticism over the years. For my purposes here, the two methods amount to more or less the same thing from two different angles: the idea that our perceptive models of the world must be properly connected to experience and memory, and the idea that these models must be structured according to proper deductive and inductive inference. I will take this general combined approach to justified belief throughout the discussion that follows. Nothing in my argument depends on anything much more precise.

2.2. Articulations of perceptive beliefs

We use sentences to tell people what we believe. But in the dimension of perception our beliefs are not much like sentences in themselves; they are more like partial pictures of the world that may or may not find *expression* as sentences. The main exception to this generality is perceptions that are essentially *about* sentences, including sentences believed to be true because they come from reliable sources, even though they are not fully understood.[[51]](#footnote-51) We can also perceive our own unqualified opinions and convictions in sentential form, and I will talk about those later on in the book. For now, let me stress that I am only speaking about perceptions, and particularly those that we derive from our own personal resources. And if we try to look directly in our minds at those perceptions we have derived from our own sensory experience, we don't typically confront a list of pre-made sentences that we can simply read out loud in order to express ourselves. Our experience of most other perceptive beliefs is more like a combination of images, scenes, sounds, and other faint sensory cues, together with various emotional colorings. Although many of the elements in our perceptions come to us packaged as objects rather than streams of absolutely raw information (as I perceive my dog *as* a dog rather than a mere mass of data), the qualities of these objects and the relations we perceive among them are not generally broken into pre-cut propositions. Even the borders between one perceptive object and another are often too vague for us to distinguish which one is which with any precision, as mountains fade off into valleys, elbows into arms, or servants into slaves. Such continuities within experience itself make it impossible for us to “read” the world directly into sentences. For example, my beliefs about George Washington’s face seem to be stored as a mental picture of the face itself, tied to with a vague feeling of respect or admiration, rather than any description in words. If someone asked me what Washington looked like in the face, I might say that he had jowls and a wide mouth, or perhaps that he was serious- or noble-looking, but I could not come close in sentences to expressing the whole of my belief about his facial appearance. I might try to draw a picture of his face, but that would be almost worthless given my lack of skill. The best thing I could do is probably to point to the portrait of Washington on a dollar bill, and say that he looked like *that*. Similarly, there are no words that I could use to describe my younger brother’s voice with any precision, but it is something I know very well, and I could easily pick it out from among hundreds of similar voices. When I think of it, I seem to hear it, not to describe it. [[52]](#footnote-52)

Political beliefs may seem to be stored in something closer to a statement form, for example Thoreau’s conviction that “That government is best which governs least”.[[53]](#footnote-53) But aside from such slogans even beliefs about politics, when we try to examine them directly in our minds, will present themselves more like pictures than sentences. Thus, my beliefs about the formal distribution of political authority in the United States would be more easily and clearly expressed using a chart of branches and layers – the sort of thing you find textbooks on government – rather than a stream of sentences. Similarly, books about wars are very difficult to follow unless they include maps. Even mathematical beliefs, which we think of as mainly symbolic, are often represented graphically in our minds. When I think of the Pythagorean Theorem, for example, what comes to mind immediately is just a diagram like this:



After a few seconds of staring at this inner picture, I can remember the actual theorem, and after quite a few more minutes, I can reconstruct the theorem's proof using this diagram (or maybe one or two others that I dimly recall). But the proof as such is not available directly to my conscious mind. What I seem to experience instead is just this diagram connected to the name "Pythagorean Theorem", plus a certain feeling of confidence that I can recollect the proof using the diagram, together with a vaguely cinematic memory of having done so in the past. I would still be disposed to make the statement, "I can prove the Pythagorean Theorem,” because that is the simplest way to articulate the perception that all these feelings and pictures represent. But even that sentence seems to interpret the relevant perceptions rather than representing them directly.[[54]](#footnote-54)

Thus, perceptions in general seem to be represented mentally in what engineers call analog rather than digital form. Instead of discrete sentences with largely independent standing, that is, perceptions in their raw form ought to be seen as something like the world itself, prior to its interpretationor *articulation* in a language. These perceptions are not entirely separate units of information, but form a structured, interconnected, semi-continuous object that constitutes my mental picture or perceptive model of the whole world.[[55]](#footnote-55) Thus, my perception of the size, shape, and location of my house depends essentially on my perceptions of the size, shape, and locations of many other things, including my own body. My perceptions of China are what they are only because my perceptions of the continent of Asia, world climate, political and military history, human biology, and the timeless yearnings of mankind are all what they are.

To make things worse, given our uncertainty about all things empirical, these perceptive models should not be seen as single "worlds", but as great collections of them, with each alternative possible state of affairs represented by its own sub-model. Thus, my picture of where I will be living twenty years from now is really at least two pictures: one in which someone has bought the film rights to this book and I have retired to Miami Beach, and another in which I've failed to land a contract and remain here in the frozen North. Right now I'm feeling rather confident about the book, so my image of me sipping Cuba Libres on the balcony with a few young friends is rather more vivid than my image of me shoveling my present slushy sidewalk one more time, stooped over and forgotten by a heartless world. Philosophers often find it convenient to think of such alternative possible beliefs as complete *possible worlds.* This is convenient for purposes of logical analysis, but it is less convenient for understanding perception, because it makes for far more whole imaginary worlds than could be separately stored in anybody's mind.[[56]](#footnote-56) Instead of imagining that we have large numbers of distinct complete worlds in our minds, then, it is best to think of our perceptions being represented as a single, incomplete, dynamic and evolving *working model*, steady enough for most purposes most of the time, but ready to split into two or more sub-models whenever alternative possibilities need to be considered. Thus, I have a working perceptive model of my future life that is vague or incomplete in many ways, but that covers most of what I need to think about. When it comes to where I will be living in twenty years I don’t have any particular belief, so when I think about that question, I can imagine the situation in the two ways I have mentioned, and probably several others if I put my mind to it. Meanwhile, I haven’t ever thought of this before, but perhaps by then a simple cure will be discovered for so-called “male pattern baldness”. Would I take this cure? Perhaps I would, out of vanity, although I also might refrain on the grounds that any obvious effort to look younger is incompatible with my dignity, i.e. also out of vanity. So, now there are four possible future situations I might consider: bald in Miami, bald in New York, hairy in Miami, and hairy in New York. At the moment I am conscious of all four possible worlds, but ordinarily I have only one, extremely hazy perception of my distant future (old, retired, better start saving some money now) while the untold millions of precise and distinct possible worlds that I might bring to mind subsist only abstractly or potentially.

In order to express our perceptions in speech, we must *articulate* the relevant parts and features of our perceptive models into sentences that we can utter. Although they don't typically exist as separate sentence-like objects, we can think of a single perception as just any part of such a model that could be articulated in one sentence. Most such beliefs are latent ones that are never articulated at all, for example my belief that there are more than 74 people in China who have worn a hat.[[57]](#footnote-57) The generation of statements out of perceptive models is not a uniform or automatic process, but requires choices of sentential form which will depend on our pragmatic purposes in making the statements. The results can be called different (sentential) beliefs or they can be called different articulations of the same (latent) belief, just as two drawings of one face can be both different and the same. Just as a portraitist, a cartoonist, an anatomist, and a police sketch artist will draw the same face differently for different purposes, we make in language all sorts of different expressions of the same perceptive content, depending on just what information we want to convey to our audience. Our practical needs and desires determine how precisely we articulate a certain belief, in what form, and with what degree of confidence.

An essential element in the articulation of beliefs is semantic structure. The most basic structures for statements are singular assertions ("This frog is green"), negations ("This frog is not green"), conjunctions ("This frog is small and green"), disjunctions ("This frog is green or brown"), a few forms of quantification ("Some frogs are green", "All frogs are green"), and generics ("Frogs are green"). Along with these *categorical* (i.e. simple yes-or-no) sentence forms, we also use conditionals ("If this is a frog, then it is green"), comparisons and statements of degree ("This frog is smaller than that one", "This frog is not very big"), and probabilities ("This frog is likely to be poisonous"), all with utter familiarity. Most of the statements we make are categorical: this is the way it is, *simpliciter*. But I believe that categorical articulation serves different purposes in the three different dimensions of belief. In the dimensions of opinion and conviction, we often have pragmatic and moral reasons not to express every uncertainty or qualification, and I will get to these issues in later chapters. In the dimension of perception, though, and for the purposes of merely informative discussion, the same beliefs can and will typically be articulated in a great number of ways. Indeed, there is usually no such thing as a complete or final articulation of any one perception. Clarifications can always be demanded, aiming more broadly and deeply into our perceptive models. Short of somehow representing the entire interconnected model verbally, there is only more or less precise articulation, explaining more or fewer of the concepts involved in terms of more or fewer other things. The more elaborate articulations that are evoked through questioning are spoken of as being what we "really meant" in our initial rough articulation. But there is no final, absolute *real meaning* to a statement any more than there is a complete, perfect articulation of any one belief. In practice, what we call the“real meaning” is reached whenever we have explained enough for our listeners to understand us well enough, i.e. to *get the picture*, as we say. Thus, if I say that I that Dave is a fisherman, and you ask me if I mean that Dave catches fish for a living, and I say no, he just really likes to fish, then you will take it that the latter thing is what I really meant by the former. There is plenty more that you could ask me about Dave’s fishing habits, but if all you wanted to know was whether he was a professional or amateur fisherman, then you now know what I really meant with all the specificity that your interest requires.

In informative discussions, we typically introduce our perceptions using simple categorical expressions like "I'd like another sandwich" or "Cats drive me crazy". But such statements are usually quite vague, and we often need to make additional, more precise articulations of the same beliefs in order for our listeners fully to understand what we are trying to communicate. We do not usually volunteer such clarifications because they're not usually needed or expected. If we say it's cold outside and someone wants to know more about this meteorological perception, they can always ask such questions as "how cold is it?", and we will tell them more of what we think they want to know. After a reasonable point in seeking useful information, such questions can become obtuse, though, as in this little conversation:

Us: It's cold out there.

Them: What do you mean, out there? The planet? The universe?

Us: No, just around the parking lot.

Them: And how cold? Is it cold enough to freeze carbon dioxide?

Us: Well, I doubt that, but I'm sure that it's colder than usual.

Them: Colder than the mean annual temperature in this area? Is that what you're saying?

Us: No, just colder than usual for this time of year.

Them: Well, how much colder than usual for this time of year? Twenty degrees Celsius? A tenth of a degree Fahrenheit?

Us: I don’t know. Just colder enough so that you'd notice it was colder than usual, I guess.

Them: So, what you *really* meant, then, that it is cold enough in the parking lot right now for me to notice that it is colder than usual around here for this time of year?

Us: Yes, I suppose that's what I really meant, if you want to be tiresome about it.

If they wanted, and we had the patience, they could keep their new demands for more precise articulation going indefinitely, like the proverbial child's asking "why?" after his parent's every attempt to explain some ordinary fact like chickens having feathers. But it's a pointless annoyance for them to insist on greater precision from us than they need to get the picture we are trying to convey.

Much of the content of our perceptive models is really conditional rather than categorical (i.e. more completely articulated in conditional form), whether we think of it that way or not.[[58]](#footnote-58) For example, I might state casually that I think riding a bicycle a dangerous thing to do, but I also believe that *if* the rider is blindfolded, *then* riding a bicycle is quite dangerous indeed, so in some unusual circumstances I might have to qualify my initial categorical statement to that effect. It is not that I have two discrete sentential beliefs lying around in my mind, one categorical and one conditional. It is rather that I have one complex picture that can be articulated more or less fully in different forms, depending on what seems to be required. Many of the conditions that underlie perceptive beliefs are *background assumptions,* things that we don’t bother mentioning because we all just take them for granted. For example, I believe that I will go to Boston in a couple of weeks. This is conditional on my being excused from work, which is uncertain, so it is natural for me to move to a conditional articulation if I’m speaking with someone who is actually interested in my plans. It also presupposes many other conditions that it would be perverse for me to add, or for my listener to ask about, because these assumptions ought to be shared by any reasonable person:

Me: Hey, I am going to Boston in a couple of weeks.

She: Really? Are they letting you off work?

Me: Well, I don't actually know. What I mean to say is just that if they let me off, then I'll be going. I'm optimistic, though.

She: But what if you get really sick in the meantime?

Me: If I get really sick, then I will probably have to cancel my plans. Why? Is there some illness going around that I should know about?

She: Not that I'm aware of. And what if Boston is destroyed by a tsunami?

Me: Uh, I am pretty sure that technically, tsunamis can only occur in the Pacific. But let’s say that if Boston gets destroyed in any way, then I won't be going there.

She: And what if the United States is suddenly taken over by a dictator who forbids all intercity travel?

Me: Look, let me put it like this. If I can get off work, and if nothing really unexpected happens in the meantime to prevent it, then I am going to Boston in a couple of weeks. Satisfied?

My friend's first question in this conversation, as to whether I could get off work, was a perfectly reasonable one, and required a re-articulation of my initial statement in order for my intentions to be stated clearly. Her second one, about my getting sick, was odd. Ordinarily, it goes without saying that travel plans are called off when someone gets very ill, so it was strange for her to bring it up without good reason, which is why I asked her if something was going around. Her third and fourth questions were just bizarre; at that point, I was obviously just being teased, though I don't seem to have taken it in quite the right way, responding as I did with some of the obtuse, humorless literalism that is endemic to my profession. Nevertheless, my final articulation was a pretty good one. In making predictions, stating intentions, and giving promises, we all do really mean something of the same sort: as long as our usual assumptions hold concerning background conditions, this interesting thing will happen. And the same is true for factual utterances about the past and present: if all of what we typically assume is true, this interesting thing is also true.[[59]](#footnote-59) In the simplest terms, then, "I believe *p*" can ordinarily be taken to *really* mean "I believe that if *a*, then *p*", where *a* stands for the collection of relevant background assumptions. It would obviously get in the way of communication for us to bring all of these background assumptions into the articulation of each statement we wanted to make, so we just leave them in the background. This is no problem, just as long as we remember that many such presuppositions exist, implicitly conditioning all of our expressed beliefs, and that we can address them directly when the need arises.

We also sometimes articulate perceptions with different levels of confidence or *subjective probability*, depending on how sure we are, or wish to say we are, that the beliefs are true.[[60]](#footnote-60) Again, most of our casual statements are categorical in form ("Hugo will have a heart attack"), but sometimes we want our hearers to note that we are not certain about our beliefs, so we add more or less implicit hedges ("It seems like Hugo will have a heart attack"; "Hugo will probably have a heart attack"; "I’ll be surprised if Hugo doesn't have a heart attack"), or perhaps more or less explicit probabilities ("There is an 80% chance that Hugo will have a heart attack"; "I’d give four to one odds that Hugo will have a heart attack").[[61]](#footnote-61) But once again, it is unusual for us to go to the trouble to specify out loud, or even silently, the degrees of confidence that lie behind in our commonly expressed perceptions. This is another aspect of perception that is more commonly evoked through questioning than articulated in initial statements.

Him: Obama is going to be re-elected.

Her: Really? How likely do you suppose that is?

Him: I don't know about the probabilities.

Her: What do you mean? Can't you tell me how confident you feel in what you're saying?

Him: I am not talking about my confidence. I'm saying that Obama is going to win.

Her: Okay, then what's your argument?

Him: I'm not giving you one. I'm just telling you what is going to happen.

Her: So, you're supporting Obama in the election? Is that what you're saying?

Him: No, I don't care at all. I probably won't even vote.

After his first statement, it seems to her that he is expressing his *perception* of the coming election. Her asking for a subjective probability is a natural next step in probing how he sees the situation. It is odd, then, that he refuses to tell her, as if his perceptions on such a contingent matter could be fully articulated as a categorical statement of fact. So, she infers that he is just expressing an *opinion* on the issue, not an implicitly qualified perception at all, and asks for the sort of argument that statements of opinion typically introduce. Again he refuses, insisting instead on his original categorical prediction. Now she infers that he really stating a *conviction* on the issue: Obama *will* win the election and he is not going to have it any other way. When he declines this third and last interpretation, she rightly concludes that he just isn't making sense – that there is evidently nothing that he really means – and quietly edges away.

The reason that conditionality, probability, and high precision rarely figure into casual articulations of empirical beliefs is pretty obvious: they would clog human communication with loads of unnecessary detail. If I make a simple, categorical statement like "They've got good muffins at the new café," I will ordinarily be speaking to someone who shares my background assumptions about what constitutes a good muffin, how similar people's reactions are to muffins of various sorts, how much sampling is required for an ordinarily reliable judgment about the quality of items like muffins at venues like cafes, and many other things of the same sort; who requires no more precision than I have offered; and who already understands how confident someone like me is liable to be about such perceptions. Under persistent questioning from a friend obsessed with muffin quality, I might end up saying something more fully articulate like, "unless I'm too sick to be tasting things properly, and assuming that I'm not the victim of some kind of hoax, I'm pretty sure, say, 80% confident, that you would feel satisfied enough to tell me spontaneously that you felt satisfied with, say, ten out of twelve randomly assorted muffins acquired at this café during four randomly chosen visits within the next month." In one sense, the second statement obviously says something very different from the first, in that it brings in a lot of qualifiers that are absent from my initial "bid". In a deeper sense, though, the long statement really says nothing more than the short one. For after all, what did I really mean when I said categorically that there are good muffins at that cafe? I hadn't bothered to spell it out, even to myself, but what I meant to convey to any reasonable listener was something more or less like what the longer statement says, albeit much less precise. Both statements are truthful articulations of the same perception, but they have been generated for different purposes, so they express what I believe in different terms. And this is true for almost anything we say in categorical form. Our listeners can *always* ask: how much is this so, and in what ways; how certain are you that this is so; and, unless you really mean this unconditionally, what are you presupposing? We are not changing our minds about the matter when we answer these questions, just giving more precision to our answer, in the same way that an artist might add more and more detail to a quick sketch if his patrons demand it.

It is one thing to express our observations categorically for efficiency's sake. It is quite another thing to try to *reason* using only unqualified sentences. The world, or one's perceptive model of the world, cannot always be treated as if it were a simple set of yes-no propositions.[[62]](#footnote-62) Potentially important information can be lost in categorical articulations, and we may suffer if we cannot find it again afterwards. This is true for matters of conditionality, degree, and probability alike. For example, if Stacey promises in categorical terms to marry Steve, based on the background assumption of Steve's getting a good job on Wall Street, but does not communicate this presupposition to Steve, both will be understandably upset when Steve, assuming that Stacey's desire to marry him is unconditional, opts out of joining the rat race and goes back to graduate school in philosophy. Just as tragically, when different departments of the government forecast revenues and spending for the next year based on different assumptions about economic growth or other factors, then announce their projections only categorically, it is impossible for non-experts (and, evidently, most experts) to determine just how badly out of balance are the nation's books. Something like this is said to be going on in the current discussions of health care policy in the United States, with the two parties making and aggregating categorical predictions based on quite different sets of background assumptions about economic growth, inflation, and other crucial factors.

Calculations involving degree and probability are more complex than categorical inferences, so we sometimes treat qualified perceptions as simple yes-or-no affairs, ignoring intermediate degrees of truth and certainty.[[63]](#footnote-63) But this can be a dangerous convenience. There is no general formula for saying to what extent a thing must be a certain way in order for it to be that way categorically, so important information can be lost in choosing categorical articulations.[[64]](#footnote-64) For example, there is no precise height at which a person or a building becomes *tall*. Clearly, the taller something is the more appropriate it will be to call it tall, but the cut-off points or *standards* seem to be set independently by different speakers at different according to their different purposes. Similarly, there is no entirely reliable way of mapping subjective probabilities, even ones very close to 100%, onto categorical beliefs.[[65]](#footnote-65) So some information must be lost whenever we opt for an unqualified instead of a qualified articulation of a belief, opening the way to serious errors in reasoning. For example, looking out my window, I might say my neighbor's house is white enough to be called white, and has bricks enough to be called brick, but if I ignore the factors of degree and speak of the house as simply white and simply brick, then someone might infer that it is a white brick house in the sense of being all or mostly made out of white bricks. But it isn't: most of the white material is wood, and most of the bricks are red. There are just enough white bricks to make one majority of the surface white and a different majority brick. [[66]](#footnote-66) Of course, examples this straightforward are unlikely to fool any clear-thinking adult. But you can find plenty of subtler instances of this fallacy in newspaper columns and even scholarly works. For example, the columnist George F. Will often asserts that "the American people prefer divided government" (i.e. a President and Congress of two different parties), seemingly based on the fact that we elect divided governments more often than not. My guess is that the great majority of Americans actually favor *united* government in principle, but they are about evenly split on which of the two parties ought to be totally in charge; just enough of us actually vote for presidents and representatives of different parties to tip the balance. This doesn't mean that Will is wrong, exactly – “the American people" can be aggregated in many different ways – but his statement is at least misleading.[[67]](#footnote-67)

The same temptation to mistaken reasoning arises when we treat probable beliefs as if they were certain, especially when writers try to predict the long-term consequences of complex events. Even articulating perceptions as simply *probable* instead of assigning more specific probabilities can lead to errors of this sort if we ignore the relevant degrees of probability. For example, if we know that 70% of children get the measles, and an independent 60% of children get the mumps, then we can say of any random child that he will probably get the measles, and also that he will probably get the mumps. But it does not follow that he will probably get the measles *and* the mumps. In fact, he probably won’t get both, since the combined probability for measles and mumps is only 70% × 60% = 42%. But if we make the error of articulating both of our initial perceptions categorically, then we will be forced into a false conclusion, since a pair of premises like “Suzie will get the measles” and “Suzie will get the mumps” does logically entail a conclusion like “Suzie will get the measles and the mumps”. This sort of error is particularly dangerous when we use categorical articulations to simplify chains of probabilities in trying to predict future events. For example, environmentalists involved in the climate change debate are sometimes accused by skeptics of stacking such de-probabilized claims in global warming scenarios in order to predict disasters in categorical terms that are only remotely possible when all the suppressed uncertainties are taken back into account. [[68]](#footnote-68) The same accusation is made about environmental skeptics by environmentalists, who charge, for example regarding the safety of nuclear power, that just because *each* possible disaster is improbable on balance, anti-environmentalists treat them all as categorically unlikely and neglect the high combined probability that *some* such disaster will take place.[[69]](#footnote-69)

Except for quick communication's sake, then, why should we ever articulate beliefs as categorical statements, since they are always less informative than statements of conditionality, precision, and probability, and can lead to mistaken inferences? It would seem that it can never be rational to trade qualified statements for categorical assertions, since there is never a good reason to ignore whatever information we possess, other than saving time. Therefore, we ought to be wary of any reasoning that uses categorical articulations of qualified beliefs, despite the usual efficiency of speaking categorically. Some philosophers, economists, and other thinkers called Bayesians believe that fully rational empirical beliefs cannot be separated from subjective probabilities, and go so far as to *define* rationality entirely in conditional and probabilistic terms. To be a Bayesian is to believe that rationality demands our constant readjustment of degrees of belief in response to new evidence according to a formula called Bayes's Theorem.[[70]](#footnote-70) If we set initial probabilities for all of our potential beliefs being true at greater than 0 but less than 1, this process will never result in a simple yes-or-no answer to any empirical question, but only a higher or lower *degree* of rational belief as evidence for and against that statement accumulates.[[71]](#footnote-71) This fluid, probabilistic view of rationality works well mathematically, and seems to capture much of our intuitive reasoning about perceptive beliefs. I suspect that working with fully articulated propositions in this way does indeed work best for beliefs along the axis of perception. In later chapters, I will argue (as I have already suggested) that categorical articulations are generally more appropriate in the dimensions of opinion and conviction. But for as long as we are working only with perceptions, it is safest (i.e. least misleading) to articulate them more or less as Bayesians would have us do in all our reasoning, i.e. with all relevant conditions, probabilities, and other qualifications plainly expressed.

2.3. The concept of knowledge

So far, I have been writing mainly about rationally justified belief in the dimension of perception. But not all rationally justified belief is knowledge; we often believe things for good reason that we do not actually know. What *is* knowledge, then, and how does it relate to justified belief? What do we even mean by the word "knowledge"? Some accounts of knowledge consider the question from a third-person point of view, just trying to analyze the objective conditions under which we say that knowledge exists, and these are called *externalist* accounts. By contrast, accounts of knowledge that take the first-person, subjective point of view – how can I *tell* whether *I* have knowledge? – are called *internalist* accounts. Traditional philosophers like Descartes, who see helping individuals examine their own beliefs as a central goal of epistemology, take it for granted that the internalist question is the more interesting one. But many recent philosophers view epistemology as an effort to understand knowledge objectively, in a way that is continuous with empirical science. To externalists, there's little real difference, then, between epistemology and psychology, where to internalists they are utterly distinct fields. A reasonable attempt to reconcile the two approaches distinguishes between two types of knowledge: *animal* knowledge, meaning true belief that has been produced in some objectively reliable way; and *reflective* knowledge, meaning true belief that is rationally justified for the person who has them.[[72]](#footnote-72) In this way, the third-person and first-person approaches in epistemology can be seen as having essentially different, though in some ways overlapping, subjects. It should be obvious in any case that my project here belongs to epistemology as the traditional, internalist inquiry into reflective knowledge and reflectively justified belief. I want to know what it is that I know, just as I want good reason to believe that my beliefs are rational.

Let me begin with a few things that we can and cannot sensibly say concerning knowledge. We can say of something, "She believes it but she does not know it," but it seems we cannot say "She knows it but she does not believe it" unless we are intending something different from the usual meanings of these terms. Thus, it seems that knowledge requires belief, but not the other way around; nothing surprising there. It also seems that we can say "It's true but she does not know it", while we cannot say "She knows it but it isn't true," which implies that knowledge always entails truth. Again, no big surprise. It follows that if you know something, you must believe it and it must be true. But this is not enough, for sometimes people end up with a true belief entirely by accident. I might believe that I am going to be a millionaire after this book is published, and be right that I will become a millionaire, but not for the reason I think. Suppose that what is actually going to happen is that huge frackable deposits of natural gas will be discovered underneath my backyard, though I have no reason to believe this yet. But meanwhile, I have come insanely to expect that Hollywood is going to make a big-budget movie of this book (starring a grizzled, worried-looking George Clooney as me) and pay me a million dollars for the rights. In this case, I believe something about my future financial status, and it is true, but nobody would say that I know it. Thus it appears that something else must be involved in knowledge over and above true belief, something that connects believers to the facts in the right non-accidental way. Philosophers have been beating their brains out for centuries trying to pin down exactly what this extra element is: Knowledge is true belief plus…what?

Plato suggested that knowledge be defined as true belief plus *logos*.[[73]](#footnote-73) This is not as helpful a suggestion as one might imagine, because the word *logos* has no clear, single meaning even in Greek. It might be taken to mean a reason (presumably a good reason), an account or explanation, or perhaps something like understanding. Still, it seems to be a good ballpark kind of definition, if not terribly precise. More recent philosophers have tended to adopt the definition: knowledge is true belief plus *rational justification*, or, in the jargon: knowledge is *justified true belief*. This turns out not to work quite right, though, as it leads to something called the Gettier problem, which most recent philosophers consider fatal to that standard definition.[[74]](#footnote-74) The Gettier problem is that sometimes people can be rationally justified in holding a certain true belief, but in a way that no reasonable person would think of as constituting knowledge, because that justification is irrelevant to the real explanation of its truth. For example, I might come to believe truly that my students are going to throw me a party in a couple of weeks, because I know that they greatly like throwing parties, and I have been hinting broadly that my birthday is coming up while also remarking that I value loyalty above all other traits in students. Moreover, I've seen some of these students at the grocery store buying what seem to be party-sized amounts of snacks and beverages. And then they do throw me a party, but not with those very snacks and beverages, which it turns out they were purchasing to help them study for exams, and not because it is my birthday, but rather because I am about to be fired from my job – something everybody knows but me – and the party is to wish me farewell. So, I believe that they are throwing me a party, and I am rationally justified in believing this, and it is true, but no one would say that I *know* they're throwing me a party under the conditions I've described. So, just as something is wrong with defining knowledge as mere true belief, something similar is wrong with defining knowledge as mere justified true belief. Even rational justification isn't enough to connect beliefs to facts in the right way.

Many philosophers have tried to come up with a more precise definition of knowledge that avoids Gettier cases and other technical counter-examples. The simplest but most extreme position taken is *infallibilism*, the idea that in order for a belief to count as knowledge it must be impossible for the believer to be wrong in that belief. [[75]](#footnote-75) The strongest argument for this position is simply that it makes no sense to say: "I know it but I might be wrong." We can easily say "I *believe* it but I might be wrong", but it seems plain that claiming to know something instead of merely believing it removes that qualification. So, since it is impossible for you to know something and also be wrong, it follows that necessarily, if you do know something, then you can't be wrong. What prevented me from knowing in the party example above is simply that I could have been wrong about the party, because the evidence I had, while fairly persuasive, was still at least *consistent* with my belief’s being false. On this view, whatever reason you have for believing something must *entail* it's being true; under the circumstances, you absolutely, literally cannot be wrong if you believe it. The obvious problem with this analysis of the concept of knowledge is that it seems too stringent to allow for any empirical knowledge to exist. Infallibilists regret this implication, but argue that they are only being true to what the word "knowledge" *means*, and that we can't pretend it means anything looser if we want to avoid Gettier-style counter-examples.

This strikes me as about the right analysis of the concept of knowledge, but I think that it is incomplete as it stands. Supposing knowledge does entail the impossibility of error, we still need to understand exactly what it is for something to be *impossible*. The easy part is that something is impossible if it is not possible, i.e. if it must be false. Alternatively, we could say that something is impossible if its denial is necessarily true. So, if knowledge is belief plus the impossibility of mistake, this is the same as saying that knowledge is belief plus the necessity of being right. To complete this little group of standard inter-definitions, we can say that something is possible if it is not impossible, i.e. not necessarily false; that it is contingent if it is neither necessarily true nor necessarily false; and that if two propositions are impossible when taken together, they are called inconsistent, while jointly possible ones are called consistent. These definitions can be extended in obvious ways to the meanings of “can” and “can’t”, “must” and “mustn’t”, “could” and “couldn’t” and other so-called modal operators.

There are a few different conceptions of impossibility in common use among philosophers. There is logical impossibility, which is taken to mean the utter *inconceivability* of truth. This is typically applied to outright contradictions, like "this is a pipe and it is not a pipe", and to more complex statements from which such contradictions can be derived. Such statements are sometimes said to be inconsistent with themselves. There is also *physical* impossibility, which means inconsistency with physical laws. For example, it is impossible for something to go faster than the speed of light, because that would violate the principles of relativity, and a hot cup of coffee cannot get hotter when you put it in a cold refrigerator, because that would break the laws of thermodynamics. Physical impossibility can be viewed as a species of logical impossibility, because one thing's being inconsistent with another just means that together they imply a contradiction. Thus, for it to be physically impossible for one of two identical apples to fall faster than the other is for that event to be inconsistent with the laws of physics (in particular, the Law of Universal Gravitation), which is the same as saying that the total theorycomposed of the laws of physics together with that hypothetical state of affairs is logically impossible. There is also *epistemic* impossibility, which means inconsistency with whatever is known*.* So, if I say that it's impossible that all the beer is gone, on the grounds that I just saw three beers in the refrigerator thirty seconds ago, I am saying that all the beer's being gone is inconsistent with this observation plus the other things I know. As with physical impossibility, epistemic impossibility can be defined as the absolute *logical* impossibility of a group of things being true together, where the group contains whatever is already known plus the offending statement. There is a similar notion of *moral* impossibility, i.e. what people must not do, which can be understood as inconsistency with moral laws or principles. A doctor *cannot* operate on an unwilling patient to exchange his hands for his feet not only because it is a physical impossibility (since feet are too large to attach to wrists), and because it is a moral impossibility (i.e. wrong), but also because it is a *professional* impossibility (in that it violates the Hippocratic Oath and other ethical standards) as well as a *legal* impossibility (i.e. inconsistent with our laws about “informed consent”).[[76]](#footnote-76) A colleague shouting “You can’t do that!” could have any or all of these species of impossibility in mind. Other impossibilities are understood in context with various other rules or principles. Nancy tells Jim that it’s impossible for her go out with him this Friday, meaning that her dating Jim would be inconsistent with other plans she has already made, or perhaps just with her general desire to have a good time. It is impossible to castle twice in chess, meaning that such a move would be inconsistent with the rules of the game. It is impossible for Congress to pass a law establishing a single religion, meaning that such a law would be inconsistent with the U.S. Constitution. It is impossible for Fletcher Christian to apologize to Captain Bligh, meaning that Christian finds it inconsistent with his sense of dignity. And so on, through as many examples as you would like. In this way, impossibility in common language is an *open concept*, i.e. one that has no fixed denotation. It just means inconsistency with *something.*

It follows that knowledge, understood as the impossibility of error, is also an open concept. To have knowledge, on this understanding, is to hold a belief in such a way that error is impossible, which only means that error is inconsistent with…something. So, what fills in the "something"? When we say categorically that somebody "knows" this or that, how is this incomplete remark to be completed? Plainly, this will depend on the circumstances of the discussion. I claim to know that my wife is in the living room asleep right now, though I cannot see her from where I am sitting, because for her not to be sleeping at this hour on Sunday would be inconsistent with her character as I understand it. If you are playing chess, you might easily claim to know that your opponent won't advance one of his pawns three spaces, though it is physically, and probably morally, possible for him to do so; it is only impossible for you to be wrong in the sense that it would contradict the rules of chess. If your opponent has some skill, you might even claim to know that he is not going to castle while his queen is open to attack, because that would be inconsistent with the principles of strategy. But if you were teaching a child to play chess for the first time, you couldn't reasonably claim to know that he was not going to make a foolish or illegal move.

These completions of our claims to knowledge belong to the same sets of background assumptions that I mentioned in the previous section. When I say I know that *p*, I am saying that I believe *p* and that I can't be wrong about *p*, which is to say that my believing *p* together with *p*'s being false is inconsistent with some proposition in the background that I'm taking as fixed. That is what I *really mean* when I say I know that *p*, in the sense that this is a more articulate expression of my perceptions about *p*. So, if I am asked to specify exactly how I know that *p*, I will usually point out whatever background assumptions *p*'s falsehood would contradict. If I am speaking in a useful way, these background assumptions will be ones that I expect to share with whoever I'm talking to. If they don't share the same assumptions, they are liable to reject my claim to knowledge, whether I stick to it or not. For example:

Paul: Try some of this linguine carbonara.

James: No thanks. I don't eat bacon.

Paul: Why, do you believe that eating pork is wrong?

James: I *know* that eating pork is wrong.

Paul: Oh? How do you know that?

James: Because it says so in the Torah and the Torah is the word of God.

Paul: Well, I think the Torah is wrong about a lot of things, so I’m not going to worry about it. Yum.

James: Fine for you, then. But I still believe it's wrong.

Notice that James did not insist that he still *knows* that eating pork is wrong. He could have, of course, if he had wanted to defend his presupposition that the Torah can be relied on for this sort of information. Simply sticking with his original claim to knowledge about eating pork would have amounted to a brusque rejection of Paul's disagreement with his presupposition, not a continuation of the discussion. In a strictly informative conversation, once an assumption on one side has been extracted and rejected by the other side, it must either be defended by the first side or else set aside for the duration of the argument. Otherwise, there is no point in talking any further. So, when somebody objects that *a* is actually false, there's no use in my continuing to claim to know that *p* unless I am willing to claim that I also know *a*. This is because all I'm really claiming to know at this point is that *p* is true if *a* is true, and people who believe that *a* is false are not going to care very much about that conditional fact. [[77]](#footnote-77)

2.4. Skeptical problems and solutions

All of the human perceptive faculties are fallible. We see and hear things falsely, we misremember objects and events, and we mess up in our inductive and deductive reasoning. Skeptical philosophers take this as grounds for denying that empirical knowledge, or maybe any kind of knowledge, exists. But there are two ways of understanding this position. First, it can be said that as a practical matter, we have less than perfect certainty in most things empirical. This is almost universally acknowledged, and is generally seen as no big deal. As long as we are *probably* right in our beliefs, or as likely to be right as the occasion demands, we can get by just fine with less than total certainty. But some philosophers have taken the complaint much further, and argued that we are so fallible that our beliefs are not rationally justified to any extent at all. We cannot even know that we are probably right, then, and do not even have sufficient reason to prefer our present empirical beliefs to their negations. It is this allegation of the *radical* fallibility of our perceptive machinery that philosophers call skepticism.

This is one of the great, seemingly permanent topics in philosophy. Descartes opens the modern discussion of skepticism by claiming that he cannot tell whether or not he is dreaming, since in his dreams he takes his experiences to be real. Moreover, even if he could determine whether or not he was having a dream of the ordinary sort, he still could not tell whether or not he was being systematically deceived in all his sensations by a powerful and evil demon – the famous evil demon hypothesis. In these secular days it is more popular to imagine that one is a disembodied brain, stuck in a vat of life-sustaining fluid and being fed electronic impulses that simulate real experience – the so-called brain-in-a-vat hypothesis.[[78]](#footnote-78) In the current formulation, then, the basic Cartesian argument looks like this:

I don’t know that I am not a brain in a vat.

If I don’t know that I am not a brain in a vat, then I don’t know that I have two hands.

Therefore, I don’t know that I have two hands.

It looks like the problem here is not just a lack of absolute certainty. It would be good enough for ordinary reasoning if I knew that I was probably not a brain in a vat, for then I could know that I probably have hands, which would be good enough for most practical purposes. But I do not have even that probability to go on. For all I can tell, there are billions of brains-in-vats and only a few hundred actual people-with-hands minding the machines that keep the brains alive and feed us their false sensations. What reason do I have to believe that this is not the case, or even probably not the case, when I have absolutely no way to detect whether it is or it isn’t?

A similar problem exists for memory as a means of perception. Bertrand Russell asks us to imagine that the whole world was created just five minutes ago, and we were all implanted by our makers with an entire lifetime’s worth of false memories.[[79]](#footnote-79) Again, we have no way to tell that this has not happened, or even that it probably hasn’t happened, since it would be totally undetectable to us if it had. In the same simple format as above, here is the argument:

I don’t know that I am not a new clone full of phony memories.

If I don’t know that I am not such a clone, then I don’t know that I had lunch an hour ago.

Therefore, I don’t know that I had lunch an hour ago.

Inductive reasoning suffers from a similar problem. According to David Hume’s argument against induction, causal reasoning, reasoning according to natural laws, and reasoning from past to future are all entirely unjustified, because they all presuppose a uniformity in nature that we cannot prove exists.[[80]](#footnote-80) We cannot prove it deductively, because it is obviously at least conceivable that nature should radically change its ways. But we cannot verify the uniformity of nature inductively either, since that would be assuming the reliability of the very method we were trying to justify. It may be true that in the past, the future always resembled the past, but that doesn’t imply that *future* futures will resemble *future* pasts, unless we already presuppose that futures resemble pasts in general, which is exactly what we want to prove. Hence the argument:

I don’t know that the laws of nature will not change overnight.

If I don’t know that the laws of nature will not change overnight, then I don’t know that the sun will rise tomorrow.

Therefore, I don’t know that the sun will rise tomorrow.

Even deductive reasoning cannot seem to escape this sort of skeptical attack. We’d like to say we know that “some duck has feathers” follows from “many ducks have feathers”, or that seven plus five is twelve. We say we know these things, apparently because we simply can’t conceive of their being false. But so what? Where is it written that our conceptual abilities must match the actual laws of logic and mathematics? In fact, people as smart as ourselves have often taken propositions to be logically certain in the past that turned out to be false, or at least possibly false. Euclidean geometry was considered to be absolutely certain by scientists until the late nineteenth century, when alternative formal geometries were devised in pure mathematics. One of these new alternatives to Euclid, the elliptic geometry of Bernhard Riemann, was later accepted by physicists as the correct mathematical foundation for the theory of relativity. Einstein’s theory itself would have been ruled inconceivable for three hundred years beforehand, during which Newton’s laws of mechanics were held by experts to be absolutely certain. And Descartes and many other philosophers of centuries past have thought that God’s existence could be proven deductively from various sets of self-evident premises, though today most Christians and other theists have abandoned such arguments almost completely. If these great thinkers can make errors in the matters they consider absolutely certain, why couldn’t I be making errors in what I consider simple arithmetic? Hence the argument:

I don’t know that I am right about the things that seem most certain to me.

If I don’t know that I am right about the things that seem most certain to me,

then I don’t know that seven plus five is twelve.

Therefore, I don’t know that seven plus five is twelve.[[81]](#footnote-81)

This all may seem a little silly. Surely we know, or are at least rationally justified in believing, that we have hands, that we ate whatever meal we ate most recently, that the sun will rise tomorrow, and that seven plus five makes twelve. We *say* we know these things, and millions of similar things, all the time. Is it possible that we don’t understand what we are saying? This would be almost to say that we are speaking the *wrong language*, which is absurd. For it is obviously up to us what we are going to mean when we say that people know things, or that they ought to believe what they believe. If “I know that I have hands” is a perfectly correct thing to say according to our own rules of speech – the sort of thing that we always call *true*, according to our rules for using the word “true” – how could it always be *false*?

On the other hand, people are sometimes quite disturbed by skeptical arguments like the ones above, and even prove willing to withdraw their claims to knowledge when such arguments are offered.

Nigel: Good heavens! I believe I've left my wallet at the club. In fact, I know I have. Basil: Really? Then answer me this. Do you *know* that you are not just a disembodied brain, floating in a vat of nutrients and being fed sensations through a wire of some sort?

Nigel: What? No, I suppose not.

Basil: Well, if you were just a brain in a vat, then presumably you wouldn’t even own a wallet, let alone belong to any decent club. Correct?

Nigel: Correct. So?

Basil: So, you don’t know that you’re not a brain in a vat; therefore, you don’t know that you own a wallet; therefore, you don’t know that you've left a wallet at the club.

Nigel: All right. I'll take it, then, that I don’t really *know* that I have left my wallet at the club. Nevertheless, I am going back to the club to retrieve it. Good night.

Here, Nigel concedes, under pressure from a standard skeptical argument, that he does not actually know what he had said he knows. He admits this with what we imagine is a tone of annoyance, and he makes it clear that he intends to act as if he did know what he now admits he doesn’t know. And he will doubtless revert to saying firmly that he *knows* he left his wallet at the club once he arrives and hears from the manager that it cannot be found. It is striking that in practice, the confession of ignorance that skeptical arguments produces tends to be revoked or forgotten almost as soon as the philosophical conversation ends. The embarrassed party goes right back to saying that he knows whatever he had claimed to know. Then the same person runs into another skeptical philosopher, and is again forced to retract his knowledge statements. The result is a puzzling pattern of assertion and denial for the very same claims to knowledge.[[82]](#footnote-82) The problem presented by skepticism, as some philosophers now see it, is to explain this flip-flopping pattern in a way that accounts both for the powerful immediate pull of skeptical arguments and for the reemergence of our ordinary claims to knowledge.

Some philosophers adopt what they see as a common-sense view of this problem, dismissing it as not really relevant to the world of experience. We're *obviously* living in some kind of stable reality, and we *obviously* know plenty of things about it, so if there turns out to be some verbal puzzle about the possibility of knowledge under some overly-strict definition, it shouldn't be taken too seriously. Others take the problem very seriously indeed, and find themselves driven into skepticism by the power of the skeptical arguments.[[83]](#footnote-83) Most epistemologists take the problem seriously, but are not willing to give up on ordinary claims to knowledge sometimes being true, so they actually try to solve the problem in one way or another. There are too many such attempts on the market for me to do much justice to the current philosophical debate, but I should note that quite a few are centered on the notion that the same knowledge statements can be literally true in one context or situation, and literally false in another. So-called contextualists suggest that ordinary discussions have low *standards of justification* for claims of knowledge while philosophical discussions of the sort that Basil imposes on Nigel above set more exacting standards, just as in ordinary dinner-table conversation we set relatively low standards for something’s being a fine chunk of lasagna, while we expect much higher standards to be applied in the context of a Michelin restaurant review. So, Nigel's initial claim to knowledge was (presumably) true in the everyday informative discussion he began, but the same statement *became false* when Basil raised the standards for knowledge by bringing up skeptical sorts of doubts about the claim.[[84]](#footnote-84) I find this an appealing view, but it strikes me as wrong to say that there is never a plain fact of the matteras to whether a person actually knows something that he claims to know. A more successful theory would account for the seeming variation in the truth of knowledge statements in a way that vindicated clearly the existence of knowledge itself.

Here is what I think is going on. I think that our beliefs about what we know can be articulated in a number of different ways, just like the rest of our perceptions. The truth or falsity of knowledge claims often depends on how they have been articulated: the very same perception can produce a true knowledge statement when expressed in one form, and a false one in another. I think that ordinary knowledge claims are usually false if they are understood to be fully articulated in categorical form, but they are often true when what we really mean by them is made explicit, or just interpreted correctly. Where context comes in is just that some discussions force us to articulate our beliefs about what we know more fully than others do. As I have been saying, we ordinarily articulate perceptions to no greater extent than is needed to for sufficient understanding, given our purposes at the moment. The same is true for statements about knowledge or justified belief. So, in the discussion above, when Nigel says that he knows he left his wallet at the club, we should ask what the original perceptions are that have generated this remark. Does it make sense to think that Nigel really believes that he knows absolutely, without any possibility of error at all, that he left his wallet at the club? Of course not. There are all kinds of background assumptions that support and hold together his perceptive model of the situation, which he probably experiences only as a vague mental picture of seeing the wallet on a table at the club, a sense of expectation that he will find it when he gets there, a little anxiety that it might have been stolen, and the like. These include, *way* in the background, his understanding that human memory and reasoning are fallible and that the world can always trip you up in unexpected ways, so that nothing is ever absolutely certain. He is quite willing to allow that he might just be dreaming or hallucinating, or in some other way be *undetectably deceived*, about all of his perceptions. He just doesn't think about it very much, any more than he thinks about his pipe tobacco being made out of molecules, though if he is forced to think about it, he will probably allow that he is actually smoking molecules. There is a large, varied, and complex structure of such background beliefs that underlies his ordinary informative statements, which is revealed by all the qualifications he is willing to place on them in responding to questions. It is sometimes hard to call these assumptions instantly to mind, so we can get confused by clever arguments and perversely literal interpretations of what we casually say, and end up retracting statements we had every right to make. So, imagine that Nigel had been better prepared for Basil's questions, and that their conversation went like this instead:

Nigel: Good heavens! I believe I've left my wallet at the club. In fact, I know I have. Basil: Really? Then answer me this. Do you *know* that you are not just a disembodied brain, floating in a vat of nutrients and being fed sensations through a wire of some sort?

Nigel: What? No, I suppose not.

Basil: Well, if you were just a brain in a vat, then presumably you wouldn’t even own a wallet, let alone belong to any decent club. Correct?

Nigel: Correct. So?

Basil: So, you don’t know that you’re not a brain in a vat. Therefore, you don’t know that you own a wallet. Therefore, you don’t know that you've left a wallet at the club.

Nigel: No. Listen to me. When I say I know I left my wallet in the club, I am making some perfectly normal assumptions, including the assumption that I am not a brain in one of your vats, or dreaming, or something similar – the same sort of thing that anyone assumes who isn't being made to justify everything he says.

Basil: I’m sure you are. Nevertheless, you do not actually know that these assumptions are all true. Therefore, you do not know the thing you said you knew about the disposition of your wallet.

Nigel: Fine, if you insist on foolishly taking everything I say as literal, complete assertions. So, let me say instead that I know this, that *if* nothing extremely strange is going on that has deceived me, *then* I left my wallet at the club. That is all I was really trying to tell you, and you would have understood this if you weren’t always being such an ass.

I don’t see anything at all wrong with Nigel’s new responses here, and I don’t think that Basil has any good way to persist in arguing the point. Yes, in a strictly literal sense – that is, if taken as a complete articulation of his underlying perceptions – Nigel’s initial categorical statement that he knows he left his wallet at the club was false, but the belief that he conveniently articulated with this simple sentence is still likely to be true (supposing he really did leave his wallet at the club). We find this out, not by staring for a long time at his initial statement, but by considering how he responds to further examination. And we discover quickly, to nobody’s surprise, that Nigel never intended to convey to Basil the idea that he cannot be wrong about his missing wallet *even if he is just a brain in a vat*. That would make no better sense than my promising my wife that I’ll be home by midnight, and her expecting that I’ll be home by midnight *even if I get run over by a bus*. Nobody who is not deliberately obtuse would interpret ordinary categorical statements as absolutely unconditional in that way.

In general, I want to say that our most serious ordinary claims of the form “I believe that *p*” really mean something like, "I believe that if I am not undetectably deceived, then *p*", and that our most serious ordinary claims of the form "I know that *p*" really meansomething like, "I know that if I am not undetectably deceived, then *p*".[[85]](#footnote-85) It seems a little strange, I must admit, to say that all of our empirical beliefs are most precisely expressed as conditionals. I think that this oddness is largely because of what I have been saying, that the conditionality of most of our perceptions is inarticulate, there being no point to bringing it up in ordinary conversations, so that we tend to forget about it as we go about our daily business. This is why claims to knowledge are often simply retracted when their background assumptions have been challenged. I claim to know that *p*, but what I really believe is only that if my assumptions *a* are true, then *p.* And I am taken aback by skeptical challenges to my knowledge claims because I had presupposed, quite properly, that their conditionality upon our shared assumptions would be taken for granted by whomever I am speaking to. On more philosophical reflection, I can come to recognize that the conditional is all I really meant claim to know, but in ordinary contexts I wouldn’t immediately recognize that that *is* what I really mean.[[86]](#footnote-86) If this analysis is right, then skepticism should be thought of as a reasonable position only with respect to claims of categorical knowledge, i.e. outright knowledge of categorical propositions, i.e. of perceptions that are fully articulated as categorical statements. But with respect to what we really think we know, and what we really mean when we are making ordinary knowledge claims, the standard skeptical arguments have little or no force.[[87]](#footnote-87)

Similar conditionalizing responses are available for skeptical arguments focused specifically on memory. I say, for example, that I remember eating lunch about an hour ago. But I can’t prove that I am not a brand new clone programmed with phony memories, nor do I think I can, once I have given the matter any thought. What I believe I know instead is only that if I’m not a clone with phony memories, and if I am not being misled in some other extremely unexpected way, then I had lunch a while ago. People often act as if what a statement means is nothing more than what it states explicitly. But it makes more sense to say that what we really mean is what we really intend to convey to others when we make the statement, which in this case is that if I can trust my memory, then I had lunch about an hour ago. The same goes for inductive reasoning. I say in ordinary conversation that I know the sun will rise tomorrow, but I can’t claim to know that the apparent laws of nature will never change. Again, what do I really believe I know? What do I think *you* think I know after I tell you that I know the sun will rise tomorrow? Surely, you don’t believe that I believe that I can demonstrate the eternal uniformity of nature. You believe instead that I just make the same background assumptions as everybody else. So, when I say I know the sun will rise tomorrow, what you believe that I believe I know is only what I actually do know, namely that if nothing utterly astonishing to science happens, then the sun will rise tomorrow. The same sort of response applies to skepticism about deductive reasoning. Do I believe I know that my little proofs for the Pythagorean Theorem are legitimate *no matter what,* even if I turn out to be having a major stroke right now? Of course not; if I were suffering a serious injury to my cognitive faculties, then nothing I believe would count as knowledge, however clear and distinct it strikes me at any moment. But I believe I know that if I’m not brain-damaged or otherwise out of my mind, so that I can rely on my deductive reasoning in the usual way, then I have proven the Pythagorean Theorem.

We all know that we are fallible perceivers. But in order to get on with life, we simply must assume that our senses, memories and reasoning faculties are generally reliable means of perception. And it would be awfully burdensome to have to articulate all of these background assumptions in advance of making any ordinary statement. This is why they are assumptions, rather than explicit statements. We have no more need to articulate them publicly than I have to tell a bartender that I’d like a pint of Guinness *without a dead mouse in it*. The bartender, if he is any good, already knows that I don’t want a dead mouse in my drink. The only informative request I have to make is that I want a Guinness in particular, as opposed to some other kind of mouseless beverage, so this is the only part that needs to be articulated.[[88]](#footnote-88) We do bring some of our background assumptions into speech occasionally, in order to express other-than-normal levels of confidence, or just to sound a little fancy. Thus: “If my eyes don’t deceive me, that is Francesca in the garden, whispering to Paolo”, or “If memory serves, we turn left at the next traffic light”, or “If I haven't lost my marbles, this experiment proves that smoking causes cancer.” But such conditionalizations are usually taken as no more than minor hedges or flourishes precisely because they are almost completely uninformative. We already know that we are all using our own senses, memories, and reasoning, and that these things are not perfectly reliable, but that we have to rely on them anyway in order to get anything done. Still, these background assumptions about our own perceptive reliability are crucial to a complete understanding of what we say, since they support intrinsically the structure of beliefs that we are trying partly to articulate. Efficiency demands that common assumptions be suppressed in ordinary speech, but they must remain available to be articulated when the occasion demands. The problem with skeptical arguments is that there aren’t many occasions that demand deep philosophical articulations of our ordinary knowledge statements, so we are liable to get confused when suddenly pressed for one.

There is another reason for our frequent confusion, though, which is that communicating our perceptions is not our only purpose in making statements. We also use statements to argue for our opinions or to stand up for our convictions, and such beliefs are often best articulated categorically despite their lack of certainty in the dimension of perception. This means that, just as with most other qualifications, the conditionality of knowledge claims is typically suppressed in outward expressions of opinions and convictions, in spite of the complex structure of the perceptions that underlie these claims. So, when we fail to distinguish our perceptions clearly from opinions and convictions, as is common in controversial discussions, the question of how best to articulate them is obscure. I will return to these dimensional factors later on. First, I want to discuss the crucial role that testimony plays in the construction of our perceptive models.

**3. TESTIMONY**

I began to realize that I believed countless things which I had never seen or which had taken place when I was not there to see…Unless we took these things on trust, we should accomplish absolutely nothing in this life. Most of all it came home to me how firm and unshakable was the faith which told me who my parents were, because I could never have known this unless I believed what I was told.

– Augustine, Confessions[[89]](#footnote-89)

3.1. Testimony as perception

Let me distinguish between two levels of perception and related concepts. Perception that is based entirely on our own direct sensations, memories, and inferences I will call non-testimonial, first-hand, or *first-order* perception, and I'll use terms like first-order knowledge, first-order evidence, first-order reasoning, and first-order facts (i.e. facts knowable through first-order means) accordingly. Perception that relies essentially on any sort of testimony – oral, written, or pictorial – I will call testimonial, second-hand, or *second-order* perception, with connected terms for the related items.[[90]](#footnote-90) Most of our useful empirical beliefs depend on both first-order and second-order resources. I believe that the sun is about ninety-three million miles from Earth, but I have no way of establishing this fact entirely on my own. I also believe that Melbourne is a city in Australia, that gasoline is made from crude oil that has been pumped out of the ground, and that I inherited the gene for male pattern baldness, not because I worked these ideas out entirely by myself, but mainly because of what I have been told. I can, of course, obtain some first-order evidence, perhaps a lot of it, for each of these claims. But I cannot by myself acquire all of the evidence that I am already using to justify them, since I rely implicitly on others for many of the background beliefs that make such evidence accessible to me. Even if I could prove that a certain two people were born in Ireland, how could I ever prove without assistance that they were my real grandparents? I cannot even justify my belief in the existence of such things as genes or DNA without relying on the word of others, to the effect that distant experiments have confirmed a certain theory of how people's traits are passed on. As I noted in the previous chapter, our even knowing what things are of which kinds, for example that this desk is made mainly of steel and not magnesium, often depends on other people's being able to discriminate what we cannot.

I want to understand how it is that we come to take the word of others for so much of what we think. Are beliefs constructed in this way completely rational? If so, how can the rationality of such second-hand beliefs be established? Some people agree with David Hume that testimonial beliefs can be reduced in explanation to first-order beliefs, i.e. that our reasons to believe what other people tell us can be given entirely in terms of what we sense, remember, and infer on our own. This would make testimony nothing really special in epistemology, just an indirect way of getting ordinary information.[[91]](#footnote-91) Others would rather treat testimony as a fourth *fundamental* source of information, in addition to the senses, memory, and reasoning, as Thomas Reid suggested in response to Hume.[[92]](#footnote-92) This would seem to place our testimonial beliefs largely outside of our own rational control, just things to be taken for granted as probably true and folded into our perceptive models of the world, with adjustments made afterwards about which sources in particular are more or less reliable than others. Some recent philosophers have adopted this second view, on the grounds that testimony only counts as *testimony* if we presuppose that it is meaningfully uttered by people with the intention of communicating information.[[93]](#footnote-93) If we don’t make that assumption, they argue, then testimony is just a stream of noise from which nothing very interesting can be learned.

I am inclined to think that testimony may be fundamental as a matter of psychology, but not fundamental in epistemology. Psychologically, there seems to be little doubt that humans have built-in predispositions to learn language, to tell the truth, and to believe by and large what they are told.[[94]](#footnote-94) In terms of the order of actual learning, Reid’s view may well be right. It may also accord better with everyday reasoning concerning testimony if, as its proponents argue, we tend ordinarily to treat other people's words as if they were our own sensory input, with no conscious doubt as to its reliability. In the end, though, I believe that Hume’s view is the right one in epistemology. I think that our reliance on testimony can ordinarily be justified entirely in terms of our first-order sources of perception. My main question here is: regardless of how they were acquired in the first place, how can I determine *now* that my testimonial beliefs are likely to be true? The answer may recapitulate the psychological process of learning from testimony in a general way, but it is liable to be more purely logical, ignoring many non-rational *causes* of testimonial belief in order to focus exclusively on *reasons* for testimonial belief.

In my view, there is a multi-stage, “bootstrapping” sort of process through which an ordinary person can rationally come to understand and then to trust the words of others, beginning with his own sensations, memories, and inferences.[[95]](#footnote-95) This process involves simple induction about the sounds and sights that *seem* to be meaningful testimony at the outset, aided by guesswork, effective assistance from other people, and our own increasing control over our own learning situations. Eventually, we are able to form reasonable beliefs about the reliability of individual sources like our parents as well as of other resources like maps, textbooks, and some internet resources, and of people in general. We also gather from evidently trustworthy people a great deal of explicit information about their own states of mind and intentions. This new information is fed back into our perceptive model, prompting further hypotheses that can be verified in turn, either through our own observations or through other sources that have shown themselves to be reliable so far. The result is typically a generally coherent package in which most of our beliefs are multiply justified by different sets of other multiply justified beliefs. This process may or may not explain how rational people are actually brought to rely on testimony. My point is that our present trust of others can be rationally justified more or less at will, based on this sort of reconstruction and our current direct sensations, memories, and proper reasoning.

I want to explain this process in enough detail for the important logical connections to be clear. Here are three quick remarks in preparation. First, as I have been saying, since my main concern is with the epistemic justification of our reliance on testimony, this ordering of steps is logical or rationally reconstructive rather than temporal or causal. I am not trying to give a psychological account of how and why we actually come to rely on testimony, or a moral account of why we ought to rely on testimony, though both are interesting topics and I will make remarks about them on the side. Second, I am concerned to avoid a side-issue over whether second-order *knowledge* as distinct from second-order justified belief even exists for categorical beliefs. Plato for one, and Locke for another, firmly denies that testimony can ever provide us with genuine knowledge.[[96]](#footnote-96) But Plato and Locke both accept that testimony can often be used to transfer justified beliefsfrom one person to another, even if they don't meet stricter standards for full-blown knowledge. This weaker claim is all that interests me here. Third, testimony is ordinarily somewhat articulate, but not completely so. It is usually transmitted in the form of categorical statements that ignore qualifications like degree, conditionality, and probability. This can cause serious problems in the interpretation of testimony, but I will ignore such problems for the most part in order to make my central points as clearly as possible. So, I will suppose for now that ordinary categorical statements work well enough for transmitting the sort of information that we typically derive from testimony. I will discuss more complex testimonial articulations later on.

3.2. Testimony and induction

Here is a quick account of how testimonial beliefs are derived from perception and memory by way of inductive inference. The story is laid out in order of justification of the beliefs, not necessarily the causal or temporal order, though I believe they roughly coincide. My purpose, again, is just to establish that an individual, relying initially only on his own resources, *can* construct a rational defense of testimonial beliefs, not that anyone is ordinarily inclined to bother. Here is the basic four-step procedure:

Step One: Inductive learning about first-order empirical facts.

Step Two: Inductive learning of language.

Step Three: Inductive learning about the reliability of sources.

Step Four: Learning from testimony about second-order empirical facts.

And there is one further step that is not logically necessary for testimonial belief in general to be justified, but that I think is importantly involved in our actual reasoning concerning testimony:

Step Five: Learning from testimony about the inner states of sources.

These logical steps are not to be seen as discrete processes, each beginning only when the previous ones have finished. Instead, they form something like an expanding "feedback loop", in which each level of learning enhances and reinforces learning at the prior levels. Thus, the more one learns about other people's beliefs and intentions, the more one is able to learn about language, about the reliability of sources, thence about others’ beliefs and intentions, and so on. Ultimately, the whole complex process should tend toward what might be called *perceptive equilibrium* (i.e. reflective equilibrium in the dimension of perception) where beliefs acquired at the different levels all settle into a maximally unified and coherent perceptive model.

Let me go over the five steps one at a time.

Step One is the inductive learning of first-order empirical facts, meaning the basic things that one can observe for oneself, such as that this ball is red, that most cats are not black, etc. These can be broken into several types. We learn a number of particular facts, for example that a certain cat is black. We also draw deductive conclusions from such facts: if this cat is black and that cat is not black, then some cats, but not all cats, are black. We also extend or project this knowledge through induction on such cases: if all the cats that one has seen so far have one and only one tail, then probably all cats are single-tailed. Needless to say, such crude inductions carry very little certainty, especially for those of us who have enough experience to know how limited our own experience is liable to be.[[97]](#footnote-97) Nevertheless, it is rational to generalize from our experiences quite boundlessly at first, as long as we are ready to correct any initial generalities, such as that all cats have exactly one tail, when we discover exceptions later on.

Step Two is mainly the inductive learning of words, such as that this thing is “red”, this thing is a “ball”, this is a “cat”, this is “Mother”, and the like. These are just *apparent* words at the initial stage, raw patterns of sounds and other perceptible signs that we end up *interpreting* as meaningful pieces of language. For us to learn these words in the appropriate sense, all that must happen is for sounds like "cat" or "red" to be consistently matched with the right other sensations, such as those caused by the appearance of a cat or something red.[[98]](#footnote-98) It does not have to be known in advance precisely what the boundaries of these types of experience are, say, between red and orange, or between cats and dogs. If the raw, vague sound in question is constantly conjoined with other appearances for some person, then that person will be able inductively to predict the appearance of the object on the basis of the sound. As this process goes on, normal learners will find correlations between more complex and determinate patterns of sounds and more precisely defined objects, including the correlation between a complex sound like “this cat is black” and the complex observation of a cat's being black.

We ordinarily receive considerable help in mastering these inductive connections. In ordinary language learning, people are not only lucky enough to be presented with largely accurate information in a breakable code, but also blessed with the ability to produce and control sounds of a similar nature by our own will, which enhances our efficiency at acquiring further information. If I am an ordinary infant, for example, and I can make the primitive sound “buh”, then the looming faces that I seem to live among are liable to present me with a bottle, or a ball, or anything else that might remotely correspond to what I might be trying to say, all the while excitedly aping my efforts to ape them. I will get better and better at this with practice, up to the point of asking complex and explicit questions, and the faces will spare little effort to encourage and assist me.

Step Three is the inductive learning of a special type of non-testimonial fact concerning *sources*, such as that one’s mother is reliable (i.e. usually tells the truth), or that one’s grandfather is not. If I can recognize the thing that I think of as Mother, and recognize the sound “cat”, then I can also learn to associate the two correctly in their relationship to other things and sounds, especially if I attend as well to certain (speaking) motions in Mother’s face, and to the particular qualities of the sound “cat” when I hear it in the presence of that face. In the same sort of way, I learn to associate the person Mother with longer strings of sounds that I take to be whole statements, some of which I am able to verify simultaneously, others through further observation, and some not at all. I do not need to know yet that Mother is a thinking being, whose statements reflect her own beliefs and are intended to influence mine. At this point in my understanding, she is just an object of my own experience, connected to patterned sounds that I have come to interpret in a way that lets me make some pretty good predictions about my oncoming experience. This Mother thing, I learn, is a reliable source of information in general, in that her "statements" (i.e. patterned sounds) as I interpret and understand them, almost always turn out to be true on whatever tests I am able to perform.

I can discover the same sort of thing about people in general, that when somebody “tells” me anything (i.e. makes sounds that I interpret as a statement), it will usually turn out to be verified by more direct experience. This depends on the actual reliability of most other people. Systematic lying to children is possible, of course, and has its temptations. But I suspect that in any normal child’s experience, much more than half of the "statements" he identifies from any source will turn out to be true. Perhaps particular exceptions will arise (the teasing grandfather, the malicious sibling), but the preponderance of true over false information will probably continue throughout the lives of anyone who stays away from television.

The human will has an additional role in the process of establishing the overall reliability of testimony. Because we are able to *choose* which instances of questionable testimony to test, we have within ourselves a power to expand the scope of our experiments by way of arbitrary sampling. The general accuracy of a road map, for example, can be well verified by testing just a few examples, provided they are chosen more or less at random.[[99]](#footnote-99) All we need to do is pick a few assorted points on the map, then drive to the corresponding places, to get a fair idea of whether the map is any good. For another example, if the *Encyclopedia Britannica* had proven truthful to me on the first ten arbitrary facts that I have been able to confirm independently, then the probability of its being unreliable in general (i.e. the probability that the encyclopedia has just been lucky ten times in a row) would obviously be very small.[[100]](#footnote-100)

Step Four is the learning of new facts by way of testimony, through a reasoning process that I call *second-order induction*. This is not really a different thing from regular induction, just a special case of regular induction with one added step. While ordinary induction (I’ll say *first-order* induction) is generalizing on the properties of things in general, second-order induction is based specifically on the truth or falsity of statements. Instead of reasoning from this or that raven's being black, say, to the conclusion that all ravens are black, we reason from a certain group of statements’ being true to the conclusion that all similar statements are true. So, if all the statements found in the *Encyclopedia Britannica* that I have verified so far have turned out to be true, then I have inductive reason to believe that all of the statements in that work are true. If I then discover a new statement like, say, "All ravens are black" in the same encyclopedia, I will have an inductive reason to believe that this new statement is also true – hence, that all ravens are black. This inference succeeds even if I have never seen a single raven myself – indeed, even if I am blind, and even if I do not know exactly what a raven is. In this way, second-order induction can function as an indirect means of confirming perceptions that are otherwise unverifiable for the perceiver.

Look at the two examples below:

First-order Induction

All observed ravens are black.

Therefore, all ravens are black.

Therefore, the next raven we observe will be black.

Second-order Induction

All verified statements in the Encyclopedia Britannica are true.

Therefore, all statements in the Encyclopedia Britannica are true.

Therefore, the next statement in the Encyclopedia Britannica that we observe (call it *p*) will be true.

Therefore, *p*.[[101]](#footnote-101)

The first three sentences in the second-order case are identical in form to the entire first-order case. They just make reference to statements that are true instead of ravens that are black. What is different is that second-order induction has a fourth step, where we reason from the fact that some statement is true, to whatever that statement actually says. So, substitute the statement 'All ravens are black' for *p* in the second argument. If it says in the Encyclopedia Britannica that all ravens are black, and we have good reason to think that statements in that source are generally true, then we have reason to believe something about ravens based on evidence that is *not* about ravens, but only about sources of information.

This form of induction works for illiterate infants just as well as for people who can read encyclopedias and maps. So, I have already discovered that the thing I have been calling Mother is a reliable source of information in general. Now I have reason to believe that the next thing she utters, whatever it is, will probably be true. If she now tells me, for example, that there are cats that have no tails, then it is now rational for me to believe that some cats have no tails (even though I had previously inferred that all cats have exactly one tail) with a level of confidence proportional to Mother's reliability in my experience so far. By contrast, if the thing I have been calling Grandfather tells me that some cats have three tails, I should consider this unlikely just because his statements, to the extent that I have been able to test them, have usually turned out false.

As I noted in Chapter 2, I do not need to understand all of what I have been told in order to believe it rationally. Some of the sentences I hear will be easily disarticulated into my perceptive model; others, less well understood, will sit largely undigested in my memory until I make the proper effort to assimilate them. Thus, if my mother, not my grandfather, had told me that bandersnatches are occasionally frumious I should have believed her, despite my having no conception of what either "bandersnatch" or "frumious" actually means. That is, I should have believed that, coming from my mother, the *sentence* "Bandersnatches are occasionally frumious" must be *true*, which implies that there must be some such things as bandersnatches and some such property as frumiosity that the bandersnatches have. I might carelessly form some vague, tentative mental image to go along with this fact, perhaps of things looking something like badgers getting angry or producing lots of gas, but I would not have reason to believe that this was an accurate depiction of the fact. I would have still reason to believe that bandersnatches are occasionally frumious *whatever that means*, in something of the way that I currently believe that rhinoviruses cause influenza.

My credulity toward different sources must also become sensitive to the apparent contents and conditions of their statements. Even my highly reliable mother turns out to be prone to exaggerate sometimes, for example when she tells me how harshly I will be punished if I disobey her present orders. And I discover that even Grandfather usually proves to be truthful when he isn't grinning at me. I should make it clear, though, that proper inductive inferences can also spread across quite different situations and conditions. Until I learn that my mother's threats are only rarely carried through, I have good reason to believe them based on her reliability in general. Until I learn that my grandfather tells the truth when he is frowning, I have reason to doubt *whatever* he says, based on his general unreliability. Even if somebody has never spoken to me about anything except the weather but has always be proven right in what he says about that, and then he tells me that Rapid City is the capital of South Dakota, I have at least some reason to believe him. For he has not just told me true things about the weather in the past; he has told me true things about the world. He does not appear to me as reliable only about the weather, then. As far as I can tell inductively, he is also a reliable source about the world in general. I may now come to find that he is disappointingly bad at identifying capitols of US states, at which point I will have reason to restrict the scope on my inductive inferences about his truthfulness, but before that point I have no such reason. In our full epistemic lives as adults, we have all learned that no one is perfectly reliable, and even people who are experts in one domain are liable to be useless in another, so that we are always gauging how much we should trust each source about each topic under each set of conditions. Given all that we have learned about the world so far, our finding out that someone is reliable at giving tax advice or explaining Aristotle's metaphysics tells us very little about their reliability concerning courtship, classical music, or where to buy an unregistered pistol. But this is not the epistemic situation of a young child, and it is not our epistemic situation with respect to the rational reconstruction and evaluation of our present perceptions. What matters here is our whole sequential epistemic history, including what could and could not properly have been inferred from testimony at any point given all of our experiences *up to that point*, not given all that we know now. So, until we have learned that there are gaps in what appears initially to be their absolutely general reliability, any new testimony from foundational sources like our parents must be taken as powerful evidence for the truth of whatever they say, regardless of topic or circumstances. [[102]](#footnote-102)

Nothing about my sources' own meanings or intentions has been built into this account. For all I know at this point in my reasoning, the sources of all these useful sounds could be mindless machines, or even pure hallucinations of my own. I can still tell that these sources have been giving me good information, at least in the rough, correlative way that a rooster's crowing informs us that the sun is coming up. But internally, as a matter of the rational construction of my own beliefs, what I take as testimony is just another product of my own direct perception. I hear “statements” from “other people”, which means only that I experience a certain set of sights and sounds. I do not even need to care, initially, what these things really are. I am still able to discover that they "tell the truth", inasmuch as I find myself believing propositions correlated to these sounds, many of which I have been able to confirm independently. In this way, I can develop an extensive system of such second-hand perceptions prior to any background theory that I might construct about the nature of my sources.

Ultimately, what we learn from and about testimony does come to form part of a mainly coherent, explanatory theory of the world, connected and integrated with what we have learned through our first order resources. In the normal course of experience, this will include considerable inductive knowledge about the ways of other people, both in general and in all their diversity.[[103]](#footnote-103) We construct useful theories about when people usually tell the truth and when they lie. We learn more and more refined ways of establishing the likelihood of truth for different types of statements, depending on whatever else we have learned about the topics, the speakers, and the conditions under which the statements might be made. We must include as well whatever we have learned from testimony *about* testimony. For example, we can learn from common testimony that it is wise to be suspicious of people who sell fancy watches on the street, and trustful of certified accountants. College students trade all sorts of information about which professors know their stuff and teach it even-handedly, and which do not.[[104]](#footnote-104) In general, learning about sources from sources helps us to bootstrap from an initially crude induction on direct perceptions to a more and more extensive, detailed, and nuanced model of the world, one that includes theories about other humans and their works as sources of belief.

3.3. Other minds

So far, I have been describing how sensation, memory, and inference combine to give us rational beliefs about the world. But it is still only the apparent world (what Kant called the *phenomenal* world), not the actual, external world (Kant's *noumenal* world) that we are able to investigate by the process I have described so far. Some philosophers have said that if we try to build a world out of sensation, memory, and reason alone, this is the best that we can ever do, because we have no independent way of grasping external reality so that we can compare our experiences against it.[[105]](#footnote-105) Therefore, we’ll always be stuck within what might be called the *solipsistic bubble*, alone in the universe of our own experience. A solipsist, someone who believes that only he and his sensations really exist, would of course be happy with this result. But outside of infants and sociopaths, almost nobody actually thinks this way. The rest of us ordinarily assume that other people really exist as thinking beings different from ourselves, and that the objects of our experience are real things external to ourselves. As I have argued, understanding our beliefs and statements as always implicitly conditional provides an adequate rational basis for ordinary thought and discourse: if we are not being undetectably deceived, then other minds and the external world are more or less what they appear to be. But it would be nice if we had more solid reason to believe that we are not alone in a dream or a vat or a Matrix than just that people and things *seem* to us to be real.

One popular idea these days is that if we “reason to the best explanation” we will see that positing a real world and real minds like ours as ultimate causes of our perceptions makes the most sense out of our total experience. Other people *could* be robots, of course, but they would have to be awfully sophisticated ones, and we have no reason at all to believe that anyone is making huge numbers of robots – let alone an entire virtual universe – for the purpose of deceiving us. As far as we can tell, there is nothing about ourselves that makes us particularly worthy of deception, let alone metaphysically unique. Therefore, the most plausible explanation for the apparent existence of other minds and the external world is the *actual* existence of other minds and the external world.[[106]](#footnote-106) I find this a very attractive approach to solipsism, but it is hard to nail down what it requires in detail. Are there independent criteria for one explanation’s being better than another? If so, what are they? And even if one explanation was demonstrably better than the others that one happened to aware of, would this make it good enough to be believed? Whether these questions can be answered successfully or not, I think that a more basic sort of inductive evidence is available within the world of experience for believing that other minds and the external world are real.

Philosophers traditionally break the problem of refuting solipsism into two parts, the problem of other minds and the problem of perception, where the former is just what it sounds like and the latter is equivalent to the basic skeptical problem about perception that I talked about in Chapter 2. The problem of other minds is to find some way to establish that other people have inner mental states like ours, or at least probably do, despite the fact that we can never observe the inner states of other people directly. Even if we monitored their brain waves, we could not be sure that those waves represented actual thoughts as opposed to the unconscious computational states of a something like a computer. One classic approach to the problem of other minds is John Stuart Mill’s argument from analogy, which is essentially a first-order inductive argument for the existence of other minds. The idea is that in your own case, you can associate mental states with physical events – feeling pain when something falls on your foot, followed by your jumping up and down, swearing, and engaging in other “pain behaviors”. This gives you a reason to believe that other people are also experiencing pain whenever they are jumping up and down and swearing after dropping things on their own feet, etc., on the grounds that whatever inner states correlate with pain behavior in your case will probably accompany the same behavior in other cases. But this is not a very good reason, it is objected, since you are generalizing here from only one observed case of a jumping, swearing body with a mind, namely your own, to the conclusion that minds accompany jumping, swearing human bodies in general.[[107]](#footnote-107) Most philosophers find this analogical argument, understood as an inductive argument from only one observable case, too weak to provide more than nominal evidence for the existence of other minds.

Here is another way I think we are justified in believing in other minds. Suppose that we have already learned to find our way around within the solipsistic bubble, having inductively broken the code of natural language, so that we can gather information from sources beyond our own direct capacities and evaluate it with some confidence. We have found that other people often tell us things that we can verify all by ourselves, for example that we are about to feel some pain, or that we will find some pizza in the oven. And we have discovered in this way, i.e. inductively, that certain other people are probably reliable sources in general. Now we can ask these other people what they mean when they tell us things. We can ask them if they are serious or joking when they say something that surprises us. We can also demand justification from our sources in the form of direct or indirect evidence for what they say – for example, their credentials or those of their own sources. This ability to extract explicit information from our sources can enhance our understanding of the reliability of particular sources and of testimony in general, creating another layer of coherent support for all of our related beliefs. Again, we don’t have to presuppose at this point that they really *mean* anything at all. Next – and here is the essential step – these other people simply *tell* us that they have minds, thoughts, and intentions. If they are reliable sources of information in general, so that we are at least somewhat justified in believing whatever they say (especially on issues where they have particularly good access to the facts), then there is no reason for us to exclude statements about their inner lives and every reason to include them.

Here is an example. Even supposing my mother is a mere machine, she can still seem to point to things and seem to give me definitions. If she says, "by 'French toast' I mean bread that has been dipped in beaten eggs and fried," then I can learn from this in an operational way what French toast is, without my needing any prior beliefs about her intentions. This is all a function of my own private history of sensory experience, some parts of which I have been able to interpret as significant and reapply inductively to new events. In this way I have learned that Mother is a reliable general source of information. Next, she tells me that she has her own mind (or her own feelings, beliefs, etc.). By second-order induction on her prior testimony, I now have reason to believe that these new statements are also true, hence that at least one other mind exists. And the same thing happens over and over again with other sources. Now I have reason to believe that other minds exist in people generally. Remember that in the traditional argument from analogy, the conclusion that others have minds rests on a single instance of observable connection between mind and behavior. By contrast, in my second-order argument, as in most useful inductions, I can build up as much evidence as I want to for the essential general claim (i.e. that other people usually tell the truth) first, before drawing any conclusions about unverified cases.[[108]](#footnote-108) In this way the problem of other minds, which Reid and his followers consider so intractable that it disqualifies Humean accounts of testimony right off the bat, seems to be solvable *within* such an account.[[109]](#footnote-109)

In the end, we do find ourselves treating testimony as something like an independent sixth sense, as Reid and his followers suppose. Once we have been rationally assured that other people have minds and speak to us meaningfully as a general rule, we do not in practice question this fact or even think about it. Instead, we just absorb the testimony in something of the way that we absorb direct perceptions, unconsciously adjusting our credulity according to source, topic, and conditions almost as automatically as we adjust direct perceptions of objects to account for distance, lighting, and angle of observation. As an everyday matter of psychology, then, second-order perception functions along the lines of Reid's analysis rather than Hume's. But according to my argument, Hume is still right about the ultimate rational *justification* of testimonial beliefs.

But what about the standard objection to Hume’s account of testimony? In order to trust what someone else tells me, don’t I already need to know that they exist, and that they mean what they mean by what they say, as Reid and others claim? The answer is no. The perceptible statements of others form a part of my own stream of sensory experience. These perceptions of sounds are correlated with other experiences, as a child learns to associate sounds like 'mommy' and 'dog' with certain clusters of visual and other sensations. I can learn gradually, by generalizations based solely on such regular conjunctions, that the apparent statements of certain others (e.g. my parents) are reliable *as I interpret them*. Thereafter, I am justified (to some extent, at least), upon experiencing any apparent statement in the sounds coming from a reliable source, to believe that the statement I construct out of those sounds is true. Thus, it is not necessary that I know in advance that others have minds, or even that they exist outside of my imagination, in order for me to have reason to believe what they say. If someone ordinarily reliable says “This is going to hurt,” I should believe him, and expect to experience some pain. If that person says 'there is water on the planet Mars', there may be nothing in particular that I should expect by way of a confirming experience. But I have reason to believe it anyway, since I have reason to rely on the rule that whatever this person seems to say is true. When the same person says “I am really hungry”, or “I think I’ll move to Canada”, or simply “I have a mind”, I have as much reason to believe these statements as any other in the same voice, by virtue of the same inductive rule. That there can be no directly confirming experiences for these particular beliefs is irrelevant to my justification.

Some of my friends feel that this simple inductive solution to a longstanding problem must be begging the question somehow when it jumps from observable to unobservable cases. But it is really no different from other good inductive arguments, all of which draw conclusions about unobserved cases from facts about observed ones. If we want some evidence about the dark side of the moon, we can look at the visible side, and that gives us reason (other things being equal) to believe that the whole thing is a certain way, e. g. covered with craters. If the whole thing is probably a certain way, then probably the part we cannot see is that way, too. And this was true before we had any notion of sending cameras or people into space to look directly at the dark side. This is how any proper case of induction works. The only “trick” to my argument lies in its applying induction not to the first-order facts themselves in question, but to the truth or falsity of statements about these facts. Ordinarily, there is nothing to choose between the two sorts of things: the statement, “All ravens are black,” is true, after all, just in case all ravens are black.[[110]](#footnote-110) Admittedly, here, in the case of other minds, I have no direct access to most of the facts that I am interested in. So it makes all the difference to induce over the observable *statements* other people make, given that I can find out empirically whether these sources can be trusted in general, not directly over facts which I cannot observe. By moving up a logical level from verified statements to considerations of reliability, and then back down to probably-true statements about other people’s inner states, I am able to leap over the wall of unobservability that separates my mind from every other mind.

3.4. The external world

This inductive argument can be pushed a little deeper, into the problem of perception. In addressing just the problem of other minds, I have been assuming the reliability of my senses, memory, and reasoning in perceiving other people’s bodies and utterances as external things. But I do not know categorically that there is any outside world at all; as far as I can tell, my senses might not be connected to any external source of information. My beliefs about physical objects as such are only really justifiable when they have been articulated in the conditional form that I discussed in Chapter 2. Still, as long as I am capable of forming objects out of patterns of “sense-data” (or whatever might be said to be experienced directly), I can make inductive inferences about these objects that accompany the presence of various other patterns, such as the clusters of appearances that I usually take to represent Mother, a cat, the Encyclopedia Britannica, etc. Even if it’s all a dream, these virtual objects are still objects. So even though she may, for all I know, be a hallucination, the phenomenal object that I call Mother can still be associated with the sights and sounds that accompany her presence in my mind. I am still able to learn from varied experience that this thing, real or imagined, is a reliable source of information in general. Thus, whenever I hear sounds in the pattern, “Here is your oatmeal”, I find some phenomenal tan mush in a phenomenal bowl. I discover that a different sound, “Here is an apple”, correlates with a different cluster of round, red appearances. In the same way, I learn to distinguish “Here is an apple” from “There is an apple”, “There are no apples”, “This apple is not for you” and so on, to the point where I have cracked inductively the bulk of the code in which these messages seem to appear. In this way, I can find out inductively that the sounds I associate with these person-like images, as I have come to understand them, usually represent the truth about other apparent things, as far as I am able to test them. Next, let this Mother object say straight out that cats, apples, my body, her body, and many other physical things really exist, independently of my thoughts. Now, even if I have been supposing that Mother is not a person or even a robot but a pure hallucination, my justified belief in her reliability gives me reason to believe that what I understand from her new apparent statement, that apparently physical things really exist in an external world, is probably true. From this point on, she can continue to enlighten me about the nature of the physical world as such with great efficiency. As long as she continues to perform reliably on every subject I can test, I will have reason to believe it all.

This reliable source, now justifiably believed to be another thinking person, tells me two particularly useful things. First, that many of my empirical beliefs are true, as far as she can tell. Second, that I myself am a reliable source of information. Such testimony to my own reliability is not conclusive, of course, though it is a lot better than testimony to the effect that I am not reliable, or a steady silence on the topic. It is still conceivable that I am being systematically misled by sources that appear to be reliable but are actually engaged in deceiving me as much as they can. But this possibility is now less probable for me than the idea that my sources are generally telling me the truth, hence that I myself am something that reliably tells the truth. Already, even before I cracked the code of language, I had some evidence of my reliability, given my success at predicting new experiences based on past ones. But this all occurred within the solipsistic bubble. Now I have external evidence, in the form of testimony from reliable others about my own reliability. So, now I don’t just live in a somewhat predictable world of my own experience. I have evidence that things really exist outside of this phenomenal world, because reliable sources of information within that world tell me that the noumenal world also exists, and that it matches my beliefs about it pretty well.

This argument is not intended to be a complete refutation of skepticism about categorical empirical knowledge. I still cannot conclude that my sensations are *categorically* reliable perceptions of an external world. Since my argument here relies on induction, deduction, and memory as well as sensations, its conclusions about the external world remain implicitly conditional:  *If I can trust my reason and my memory, then my sensations are (probably, usually) accurate perceptions of an external world.* What the argument shows, then, is not that global skepticism about categorical beliefs is absolutely false – nothing could prove that – but only that we are justified in believing in the external world *provided* we take reason and memory for granted. This isn’t everything we’d like to have, but it is more than most epistemologists believe that we can get.

We are used to thinking of our knowledge of physical reality as coming before our knowledge of other minds and being stronger. But in my view this order of priority is not correct, and might even be reversed. According to my argument, we know of the external world *as* the external world largely through our interpretation of testimony.[[111]](#footnote-111) My argument here for a distinctly physical reality is thus no better than the argument for other minds, and could even be seen as weaker to the extent that testimonial evidence is generally strengthened and made more coherent through our belief in other minds. Seen this way, my argument is similar in overall structure to Descartes’s main argument in his *Meditations*, where he attempts to prove the existence of a perfectly benevolent God through *a priori* arguments, and then uses God’s benevolence to guarantee the truth of his own best-justified empirical perceptions.[[112]](#footnote-112) I have replaced Descartes’s perfectly benevolent God with a set of imperfectly benevolent parents and others whose general reliability we establish through internal means (given the right sort of sensory data), and who can then provide us not Cartesian certainty, but at least pretty good evidence for our basic beliefs about the world. It is a contingent fact that we have been given this reliable faculty of apparent testimony from apparent other people, and that this faculty provides us with information of interest to epistemology. If nobody ever said anything to us about their minds and the external world, then we would remain largely in the dark about such things. It is a matter of *luck*, then, after all, for each of us, that we should be provided with our own solutions to the problems of other minds and the external world, just because we happen to live among other thinking beings who will speak with us sincerely.

3.5. Moral perception

I would like to apply a similar analysis to the foundations of moral belief. First, I should make it clear that by "moral" I do not mean the *right* morality, necessarily. A person who believes that killing infidels is good has an *immoral* belief as opposed to a moral one, in the sense of incorrect vs. correct morality. But that person still has a moral beliefas opposed to an *amoral* belief, in the same sense that a person who believes the sun revolves around the earth has an empirical belief, however false. In this general sense of the term, moral perceptions don't have to be true ones; they just have to be perceptions *about* what's right and wrong. Given this understanding of the term, I want to argue that we get our grounding in (correct or incorrect) morality from sources of information that we have come inductively to trust, again in the first place typically our parents. Among the things that we hear from these reliable people are many propositions about good and bad and right and wrong and what we should and shouldn't do. If the people around us are normal human beings, most of these statements will be made sincerely, based on their own moral perspectives. And once we accept these propositions as our own beliefs, we adopt a similarly moral outlook on life implicitly, without the need for any prior grasp of moral concepts.

Among the enduring questions in philosophy is "Why be moral?” which is taken to mean something like this: Is there any good way to *argue* that a person should adopt impersonal moral principles and not pursue exclusively self-interested goals? So-called rational egoists say that there is no way to bridge the motivational gap between a person's naturally selfish desires and genuinely moral behavior. On their view, if a person happens to desire the well-being of others, or desires to think of himself as a good person, then he will have a reason to be moral, but otherwise he will not.[[113]](#footnote-113) He may wish to appear moral, of course, so as not to lose the benefits of his society’s approval, but this is only *instrumental* morality. On this account, genuine morality, understood as seeking the good of others regardless of one’s self-interested desires, has no rational basis at all.

I want to offer a response to this solipsistic view of morality along the lines of my solution to the problem of other minds. I say that rational egoism is plausible only with respect to first-order reasons for moral belief, and that second-order reasoning easily shifts most people from an egoistic perspective to a genuinely moral one. The argument is pretty simple: We have reason to believe what others tell us; others tell us that we ought to do what is morally right; therefore, we have reason to believe that we ought to do what is morally right. Further, if we have a reason to believe that we ought to do something, then we also have a reason to *do* it. This is another principle that normal people learn through testimony. Let me explain.

Here is a piece of elementary moral instruction from a trustworthy source:

Mother: Hey, stop pulling the cat's tail. You shouldn’t do that.

Me: Why not? It’s fun.

Mother: Because it's wrong to hurt the cat.

Since I am inductively justified in believing that my mother is a reliable source of information in general, I now have reason to believe that I should stop pulling the cat's tail, which is simply that it's "wrong", whatever that undigested word means. I may not yet understand the reason, and it may turn out not to be a genuinely moral (i.e. non-self-interested) reason, but I am now in a position to believe that the statement "it's wrong to hurt the cat" is true, and therefore that it’s wrong to hurt the cat, and further that this constitutes a reason for me not to hurt the cat. I am not saying that I will actually stop hurting the cat at this point – action requires desire or volition in addition to belief – just that I have a *reason* to stop.[[114]](#footnote-114)

There is a problem, however, about the interpretation of these statements, similar to the problem about interpreting reports of others’ thoughts in my argument for believing in other minds. When we initially hear other people make statements about right and wrong, these moral terms can still be interpreted non-morally, in either of two different ways. First, we might come to understand them only with reference to our own self-interest.

Mother: Hey, stop pulling the cat's tail. You shouldn’t do that.

Me: Why not? It’s fun.

Mother: Because it's wrong to hurt the cat.

Me: What does that mean?

Mother: It means that I'm going to whack you one if you don't stop.

When my mother tells me, "You shouldn’t pull the cat's tail," I quite reasonably ask her for a reason. She tells me is that hurting the cat is “wrong”, I ask her what that means, and she explains it directly in terms of my self-interest. I ought rationally to believe this explanation as well as the flat statement that hurting the cat is wrong; she has given me more than enough second-order inductive evidence to justify both perceptions. Now I have gained an understanding of a moral term "wrong", but it is not a genuinely moral understanding of that term, just a prudential one. On this interpretation, if I can find a way to guarantee not getting caught, I might as well go back to tormenting the animal.

A second amoral option would be to interpret the moral statements that we hear from others as mere descriptions of our society’s moral conventions. If "hurting the cat is wrong" means to me merely that my society will frown on such behavior, then this will not by itself give me a reason to refrain. I might not care, after all, whether other people approve of me or not.

Mother: Hey, stop pulling the cat's tail. You shouldn’t do that.

Me: Why not? It’s fun.

Mother: Because it's wrong to hurt the cat.

Me: What does that mean?

Mother: It means that people disapprove of hurting animals. It's called "animal rights".

This explanation gives me a notion of wrongness that is not about my immediate interests or my mother's intentions, but my new understanding of the concept of wrongness is still not a truly moral one, because I still have no reason not to hurt the cat beyond my possible self-interest in avoiding social disapproval.

Here, on the other hand, is a minimal but genuinely moral explanation:

Mother: Hey, stop pulling the cat's tail. You shouldn’t do that.

Me: Why not? It’s fun.

Mother: Because it's wrong to hurt the cat.

Me: What does that mean?

Mother: It means it's wrong. Don't do it.[[115]](#footnote-115)

Me: So what if it's wrong? How is that a reason for me not to have a good time?

Mother: It just *is* one.

Me: Why, are you threatening to punish me?

Mother: No, I'm telling you not to do it because it's wrong.

Me: So, you’re just invoking another of your bourgeois social conventions.

Mother: No, I’m telling you the *truth*.

The explanation that my mother gives me here is an opaque one; I can't form any real conception of the nature of wrongness from what she has said. But I have been told by an inductively reliable source that this mysterious "wrongness" does constitutes a reason for me to alter my behavior, and moreover that this reason is a distinctively moral, not a prudential or conventional one. And if I have reason to believethat something is a reason, then for me it really *is* a reason. So, I now have a superficial but clearly non-egoistic reason not to hurt the cat, which is simply my second-order perception that hurting the cat is wrong, together with my second-order perception that this *constitutes* a non-egoistic reason for me not to hurt the cat.

A useful explanation of moral beliefs will not just categorize various actions as right or wrong, but will also get at what it is about them that *makes* them right or wrong, so that moral principles can be understood more deeply and applied more broadly. When pressed for such an explanation, my mother produces one:

Mother: Hey, stop pulling the cat's tail. You shouldn’t do that.

Me: Why not? It’s fun.

Mother: Because it's wrong to hurt the cat.

Me: What does that mean?

Mother: It means that it makes God angry.

Me: So what if it makes God angry? How is that a reason for me not to have a good time?

Mother: It just *is* one.

Me: But I don't *care* whether it makes God angry.

Mother: Too bad. It makes God angry, and that's why you should stop.

Here I am being given, not just the surface statement that it is wrong to pull the cat's tail, but also an explanation of that statement that demands a moral interpretation. If I correctly perceive my mother to be reliable, I now have reason to believe not only that that I ought to leave the cat alone, but also that this is not because it’s in my personal interest to stop or because hurting the cat violates a social norm, but because it makes God angry if I do it. Now, I will not find this fact particularly motivating in itself if I was serious in saying that I don't care about whether God gets angry. Of course, if my mother then explains that if I make God angry I will go to hell, then this becomes another amoral, self-interested reason not to hurt the cat. But if I believe what I have been told so far, then I must conclude that some actions are wrong essentially because they anger God, even when they make me feel good and even if God refrains from punishing me.

Other sorts of purely moral explanation are available, and in a more reflective mood my mother offers one in terms of respecting the interests of other sentient beings:

Mother: Hey, stop pulling the cat's tail. You shouldn’t do that.

Me: Why not? It’s fun.

Mother: Because it's wrong to hurt the cat.

Me: What does that mean?

Mother: It means the cat doesn’t like it, so you should stop.

Me: So what if the cat doesn't like it? How is that a reason for me not to have a good time?

Mother: It just *is* one.

Me: But I don't *care* whether the cat likes it or not.

Mother: Too bad. The cat cares, and that's why you should stop.

Reasons of this sort may or may not be more powerful psychologically that purely religious reasons to be good, depending on whether the natural compassion for other creatures to be found in most children is stronger than their desire for supernatural approval. Again, though, I am not saying that these reasons are sufficient to control my actions, or that they are objectively good reasons. And I am certainly not saying that any set of such reasons is objectively as good as any other. All I am saying here is that these are *moral* reasons as opposed to merely self-interested ones, and that this is how we are rationally justified in forming moral perceptions as opposed to merely descriptive ones. Reliable sources of information tell us that we have reasons to act or refrain from acting that depend not on our self-interest but on rightness and wrongness, and this is enough to give us non-self-interested reasons to act or refrain from acting.

So, just to be as clear as possible: Imagine someone who has been brought up by otherwise reliable parents to be a Nazi, i.e. someone who despises all non-Aryans and who believes it is the destiny of the German people to control the world. Am I saying that this is a moral person? No and yes. He is presumably not a moral person in the sense of an objectively *good* one, but he is a moral person as opposed to a purely *self-interested* one. If, like many German soldiers in World War Two, he believes that he should sacrifice his personal interests, even his life, for the Reich or the Aryan race, then he has adopted a moral perspective in the broad sense of having adopted beliefs *about* right and wrong (regardless of whether those beliefs *are* right or wrong in substance). It follows that this person has adopted reasons for action that fall within the realm of morality rather than the realm of prudence, and this is enough to make a person a subjectively moral as opposed to a subjectively amoral one. It does not mean that he is brave enough to act on these reasons, or that he does not have stronger subjective and objective reasons to act otherwise than as a Nazi. It just means that he has *some* reason to be a “good” Nazi, based on the evidence he has obtained from testimony as a rational perceiver.

In the next chapter I will argue that in the face of authoritative testimony, most people have no rational choice but to believe things that are seriously false. Sometimes, through a kind of epistemic gravitation, we become resistant to perceiving contrary evidence as well. And in the worst situations, which I call epistemic black holes, we can even be forced into a kind of *rational irrationality* that makes it impossible for us to change our minds at all. It is because we are rational in the first place, and because we are presented with certain chains of misleading evidence, that some of us become impossible to reason with.

**4. AUTHORITY**

Our faith is faith in someone else’s faith, and in the greatest matters this is most the case.

– William James, “The Will to Believe”[[116]](#footnote-116)

4.1. Rational deference

Sometimes, the evidence that we receive through testimony makes it rational for us to override our own first-order, pre-testimonial perception about something in favor of someone else’s stated belief. This is what we should do as perceivers whenever we have good enough reason to believe that another person has a more reliable overall view of the question than we do. To the extent that other people are eyewitnesses or expert on the subject at hand and we are not, it makes sense for us to replace our own prior perceptions with theirs. Even when our own view is privileged in some way, we still sometimes properly defer to the statements of others. For example, when I have inferred from my own direct experience of chest pain that I am having a heart attack, but then my doctor examines me and tells me no, that I am only suffering from "acid reflux", I should probably just abandon my prior belief and adopt his in its place. This is because I know that he is much more likely overall to be right, under the circumstances, than I am. In general, if we are to be perfect perceivers, we need to follow the principle of rationality without exception. This means attaching the greatest credence to whatever perception is most likely to be true, given all of the evidence available to us. It is irrelevant whether the initial source of this belief is our own first-order equipment or someone else's. Therefore, as a matter of pure reason, we should "stick to our guns" in a disagreement faced with controversy only if we are justified in believing that we are the ones most likely to be right.[[117]](#footnote-117)

I do not mean to suggest that we can think somebody else's thoughts *as such*. When I adopt somebody else's belief, the perception that I end up with is still definitely mine. I am not rationally out of the picture, just because I have accepted the beliefs of another. I am still drawing the best conclusions that I can from the total evidence available to me, including my direct experiences, memories, and first-order reasoning, but also the statements of sources besides myself, which ordinarily reflect either their own direct experiences or testimony they have received from third parties. I am saying that in the end, my coming to the best *overall* conclusion may entail giving up my *first-order* conclusion. But it is still me, thinking my own thoughts from my own perspective. So, the point I am making about deference implies no form of "group-think" epistemology that would take social groups as basic subjects of belief. There is no relativism, social constructivism, or anything of the sort involved in what I am saying. I am speaking only of the total personal evidence that each of us has for his own beliefs, and claiming that this evidence can sometimes force a reasonable person into epistemic conformity with others around him.

Remember the example from Chapter 1 about adding up columns of numbers. There I asked you to imagine a case where you and I had equally good track records for computing these sums, were both aware of this equality of past reliability, and then found that we disagreed over a new example. I suggested that in this case, it would be rational for both of us to withhold judgment as to who was right until we found a reasonable way to agree. Now imagine something slightly different. Suppose this time that after solving plenty of these problems, you have succeeded in obtaining the correct sum 98% of the time, while my results for exactly the same problems have been accurate only 94% of the time. Again, suppose we both know this. And again, suppose we each add up a new column of numbers, and get different results. Now, which sum is most likely to be right? Other things being equal, it is obvious that your new result is more probably correct than mine.[[118]](#footnote-118) So, if I were forced to bet something of value on which sum was correct, it would be irrational for me to prefer my own calculation to yours. It does not matter whether I understand how you do these sums better than I do. Even if you have explained your methods to me many times, and I find they make no sense at all to me, I am still faced with the brute fact that I am on average about three times as likely to be wrong, all things considered, as you are.

Such deference can also be rational when a person is no less reliable than his competitors, but simply outnumbered by them. If you and I have both been 96% reliable over a large number of trials and come to a disagreement, each of us should realize that our answers are equally probably correct. But if we find ten other random peers all agreeing with you, and none with me, the probability of my being right and everybody else wrong is very small indeed.[[119]](#footnote-119) The same is also true for disagreements with groups of people who are somewhat less reliable than me, as long as a sufficient number of them all agree among themselves, all other things being equal. And this applies to sensory perceptions or memories as much as it does to calculations or other cases of deductive reasoning. Psychologists have shown that people are drawn into agreement with others even about statements that would be visibly false in the absence of contrary testimony. In one famous example (the Asch conformity experiments), subjects were asked whether two obviously unequal lines are equal in length. When responding by themselves they were entirely reliable, but in the presence of other people they see giving the wrong answer, they tended to agree with the group, rather than trust their own eyes, and the more so the more they were outnumbered by unanimous peers.[[120]](#footnote-120) This result is commonly taken as evidence of a human tendency towards irrational conformity. But I think that it is a perfectly rational way to think, at least in many cases of this sort. I may *seem* to see or remember or infer something very clearly in the usual way, but I know that I am not infallible, even in matters that seem obvious to me. Therefore, under unanimous testimony from other generally reliable sources (potentially even a crowd of strangers), I will defer to whoever is most probably correct in light of all the evidence I have.

Let me return to the columns of numbers example above, where you have proven yourself more reliable than me. Suppose that in our considerable joint experience, you have *always* been the one who got it right whenever we disagreed. Even though I am still fairly reliable at working these problems by myself – 94% isn’t too bad, after all – and you are still not perfectly reliable at doing it, it looks like I ought to defer to your conclusion every single time we disagree. In fact, the evidence implies that my work is entirely redundant to yours, since we are solving the exact same problems and I am never right while you are wrong. In this situation, as a matter of pure rationality there is no evident need for me to check your work at all. If all I want is to believe what’s most likely to be true, I should simply adopt you as my *authority* in making such calculations and find something else to do with my own time. The same reasoning applies to all sorts of situations in which we can tell that somebody else is more likely to be right than we are, all things considered. Such authority is usually limited to an area of expertise, as between me and you over sums, or between me and my doctor over medical diagnoses. I accept as probably true practically everything my doctor tells me about my physical condition (or at least *more* probably true than my first-order guesses), though I am disinclined to ask him about many other things, such as whether he believes that doctors are underpaid. Where he is the expert, it would be irrational for me to rely on my own counsel; where he is not, it wouldn’t.

Psychologists have striking things to say on this point. In another famous experiment (the Milgram obedience experiment), someone represented as a psychological experimenter in a white lab coat ask subjects to administer increasingly painful shocks to people they believe are other subjects (but are only acting the part). It turns out that most subjects are willing to hurt their counterparts to an extent far in excess of what would normally think of as a decent limit.[[121]](#footnote-121) This disturbing result has generally been taken to show that most people have a kind of moral weakness in their tendency toward mindless obedience toward “authority figures”, and this lesson has been broadly applied to the analysis of official cruelties like the Holocaust and the massacre at My Lai during the Vietnam War. It is commonly said that this blind obedience to authority makes people do things that they know are wrong, just as the “herd mentality” in the Asch conformity experiment makes people say things that they know are false. I certainly sympathize with victims of oppression, and it is true that some people are by personality less self-confident, less suspicious of authority, and happier taking orders than others. But I am not at all sure that most people who do cruel things under authoritative orders actually *know* at the time that what they are doing is wrong, or indeed that as a rational matter they *ought to believe* that what they are doing is wrong. This depends on each person’s total epistemic history, including how they have been raised to think about authority, science, the welfare of strangers, and the sometimes tricky nature of psychology experiments.

Authority can be restricted to a narrow topic, or it can be quite general. A severely mentally disabled person may be perfectly rational in relying on absolutely everything that a reliable caretaker tells him, as the character Lennie relies on his friend George in the novel *Of Mice and Men*.[[122]](#footnote-122) Such total epistemic dependency is obviously rare among adults, but almost all small children, including very intelligent ones, are in this position with respect to their parents, at least up to the age of five or six. The children know enough to know that they are far less reliable overall than are their parents, hence that they are rationally best off accepting whatever their parents say on almost every subject. Of course, ordinary parents are not *professional* sorts of experts on every subject, but the young child is in no position to make such distinctions. From his own point of view, given all of the evidence that he has accumulated so far, whatever his parents tell him about virtually anything is more likely to be true than are the products of his independent first-order resources. This is the essence of faith.

Think of an ordinary child's belief in the existence of Santa Claus. It may seem superficially that children who believe in Santa Claus must be making some juvenile mistake in reasoning. After all, there is no good objective evidence for Santa’s existence, and plenty of reason for doubt (regarding chimney dimensions and the like). But from the subjectively rational point of view, there is typically nothing wrong with them at all. They are not making any kind of mistake, given all of the (first- and second-order) evidence they actually have. They believe in Santa Claus because their parents have told them that he exists, and their parents are the most reliable people that most children know, themselves included. Even if they find the first-order case for Santa's existence totally unconvincing on its face, and even if their older brothers and sisters have repeatedly told them to dry up and think for themselves, the fact remains that young kids still have greater total reason to rely on their parents’ authority than their own, however things may *seem* to them on a first-order basis. It is typically only when parents have renounced their lies that young children give up this belief in Santa Claus, however suspicious or conflicted they have become in the meantime. And this is because it is only at that point that it typically becomes rational for the child to change his mind, all things considered. In cases like this, *faith* (in reliable authority) can be said to triumph over *reason* (restricted to first-order evidence), without the need for any diagnostic sort of explanation.

My point applies to true as well as false beliefs, of course. Children believe in Julius Caesar for essentially the same reasons they believe in Santa Claus, and some believe in helping strangers for the same reasons that others believe in mistreating them. If a child refused to believe in Caesar just because he had no direct, first-order evidence of his existence, then, given what the child ought to believe about the reliability of his sources, he would be making a serious error in reasoning. The same goes for adults. Almost all of us believe in Caesar, DNA, Tahiti, wave-particle duality, and the planet Pluto almost entirely on the authority of teachers, professors, and other individual experts, plus encyclopedias and other expert-created sources, or on trustworthy non-expert sources who have relied on expert sources themselves.[[123]](#footnote-123) In innumerable such matters outside the range of their direct perception, rational children and adults alike ought to believe no more or less than what they have been told.

This kind of rational belief in Santa Claus or Caesar is easily distinguished from a case presented annually in the comic strip *Peanuts*, where the character Linus has formed a totally idiosyncratic belief in what he calls the Great Pumpkin, on the basis of no evidence whatsoever, either first- or second-order. Every Halloween, Linus expects this being to rise up out of a local pumpkin patch and bring presents to all the children in the world. And every year it doesn’t happen, and he sees that it doesn’t happen, but he keeps believing in the Great Pumpkin anyway, all by himself. I think that this extreme "fideist" Linus is being irrational, while the ordinary Santa Claus believer is not.

There are, of course, other important goals in raising children besides trying to equip them with the best justified set of beliefs available at any moment. For example, it is surely in most children’s long-term interest to develop their own imaginative and critical faculties, even at some cost to the correctness of their short-term beliefs. Perhaps this helps explain why children seem naturally to exhibit phases of Linus-like fantasy or quite unreasonable defiance of their elders, starting at a very young age. As they grow up, they will need more and more to form their own opinions and convictions without supervision, especially if we want them to resist following orders that on first-order grounds strike them as wrong. Allowing them a few childish beliefs in the meantime is a small price to pay for their gaining practical experience at autonomous reasoning and choice. I will return to these other dimensions of belief in the following chapters, but for now I want to stay with my main argument concerning rationality and faith in the dimension of perception.

4.2. Epistemic communities

I want to apply the same sort of analysis to religion – more specifically, religious *faith* –that I’ve applied to children's belief in Santa Claus. Religion and faith are not exactly the same thing, of course. There are plenty of ethical and other faith-based traditions that have nothing to do with religion, and there are sources of religious belief that do not depend on faith. For example, many people claim to have specifically religious experiences, ranging from hearing vividly the voice of God to the vague “oceanic feeling” that Sigmund Freud attempts to analyzein *Civilization and its Discontents*.[[124]](#footnote-124) There is also a strong philosophical element in religious ethics and theology, and even whole religious legal systems like *sharia* and canon law. But the thing that really makes religion what it is, an enduring epistemic institution as opposed to a mere set of doctrines, is its transmission through testimony from one generation of believers to another.

The most prevalent view among Western intellectuals these days is that religious faith belief is an irrational opponent to belief based on good evidence – something to diagnose, not to dispute with care. Explanations of other people's religious faith vary from desire for protection, fear of death, and the like to political brainwashing, but the common thread is the idea that believers are persuaded by essentially psychological or psycho-social forces, not epistemic ones. This puts religion in direct opposition to science. Freud harshly attacks the intelligence of religious believers:

In the long run, nothing can withstand reason and experience, and the contradiction religion offers to both is palpable…When a man has once brought himself to accept uncritically all the absurdities that religious doctrines put before him and even to overlook the contradictions between them, we need not be greatly surprised at the weakness of his intellect.[[125]](#footnote-125)

Bertrand Russell offers an equally harsh diagnosis in terms of emotional weakness in *Why I am Not a Christian*:

Religion is based, I think, primarily and mainly upon fear… Fear is the basis of the whole thing – fear of the mysterious, fear of defeat, fear of death. Fear is the parent of cruelty, and therefore it is no wonder if cruelty and religion have gone hand in hand[[126]](#footnote-126)

The most prominent current intellectual opponent of religion is probably the biologist Richard Dawkins, who diagnoses faith as a "memetic" virus:

It is fashionable to wax apocalyptic about the threat to humanity posed by the AIDS virus, "mad cow" disease, and many others, but I think a case can be made that faith is one of the world's great evils, comparable to the smallpox virus but harder to eradicate. Faith, being belief that isn't based on evidence, is the principal vice of any religion. And who, looking at Northern Ireland or the Middle East, can be confident that the brain virus of faith is not exceedingly dangerous? [[127]](#footnote-127)

Other non-rationalistic accounts of religion are not so hostile. Many sympathetic thinkers, including some who are religious believers themselves, view faith as neither rational nor irrational, but as something entirely apart from rationality. As Pascal puts it in his *Pensees*:

It is the heart which perceives God and not the reason. That is what faith is: God perceived by the heart, not by the reason…Faith certainly tells us what the senses do not, but not the contrary of what they see; it is above, not against them.[[128]](#footnote-128)

On either sort of account, though, whether this is held to be a bad thing or not, it is taken for granted that religious belief fails to meet normal and proper standards of rationality.

I disagree. In my terms, where people like Russell and Dawkins speak of faith as belief without evidence, they can only mean objective, impersonal sorts of first-orderevidence. But rational people do not ordinarily rely on first-order evidence alone, and we could hardly function if we did. We live our thinking lives not by ourselves but as members of *epistemic communities*. These are not mere social units, but essentially intercommunicative groups in which each member rationally trusts his fellows more than outsiders. Our friends, families, or compatriots do not count as an epistemic community just because we happen to feel loyal to them. We must also have good reason to trust them more than other people. Thus we are surrounded by testimony, some of which contradicts our prior first-order perceptions, and it would be irrational for us *not* to take such testimony into account, especially where it is unanimous among the most reliable sources available to us. If someone grows up in a traditional society where his parents and everybody else believe in reincarnation, for example, then if he is rational, he will defer to his elders and come to believe in reincarnation himself. He may have no direct first-order evidence for that theory – though he may or may not have been taught that some things *count* as first-order evidence for it – but this makes no effective difference. Given all the second-order evidence that a person in that situation typically has, he ought to conclude that those around him are more reliable, at least on questions not specific to him, than he is himself. Therefore, if he wants to believe what is most likely to be true, he ought to accept whatever these authorities tell him.[[129]](#footnote-129) If what they say makes little sense to him in a first-order way that is unfortunate, but it doesn't change the rational equation very much.

This propagation of belief through rational deference explains the great inherent stability of religion and other cultural traditions. It is because human beings are usually quite rational in the subjective sense, in that we make the best sense possible of all the first- and second-order evidence we have, that normal people growing up in traditional societies so “blindly” absorb the views of those who went before them. The common prejudice of scientific Westerners to the effect that generation after generation of highly civilized Egyptians, Chinese, Incas, Hindus, Jews, Christians, and Moslems are all somehow deficient in rationality shows its absurdity, I think, when placed against this reasonable alternative account. How could it be rational for a young Egyptian under the Pharaohs to evaluate the cosmic situation on a purely first-order basis, and then to accept his own empirical constructions in preference to the undisputed testimony of the most trustworthy people he knows? To think that he can figure this sort of thing out for himself would imply that he considers his own first-order epistemic machinery more reliable than everybody else's put together. No normal person has good evidence for such a belief. Even if it strongly *seems* to him that he has good direct evidence for his own first-order theory and that everybody else is making an error he has identified, his knowledge of their overall epistemic record and abilities compared to his own should ordinarily compel him to defer.

Consider again at the quote from Saint Augustine with which I began Chapter 3, where he notes “how firm and unshakable was the faith which told me who my parents were, because I could never have known this unless I believed what I was told.” Augustine makes this testimonial account of faith a central premise of his intellectual conversion to Christianity. It is largely because he has come to know some Christians, and has found them more honest and reliable than anybody else he has come across, that he concludes that their theology is more deserving of belief than other doctrines, including his own prior Manichaeism, despite the fact that he cannot make sense of some of its elements. And for most other Christians, Muslims, Hindus, and other adherents to traditional faiths, the situation is much simpler than it was for a well-traveled, widely-read philosopher like Augustine. Since they are ordinarily surrounded from birth by testimony that is not mixed but unanimous, they are liable not to find any need to make such comparisons. Since *all* of the reliable people that they know have testified to the same tenets of faith, they are left with no rational choice in the matter at all.

Here is a major objection, first posed to Augustine’s testimonial argument for faith by Hume in his famous discussion of miracles.[[130]](#footnote-130) What if a religious claim is so implausible in the first place that the probability of even unanimous testimony in its favor being true is lower than the probability that everybody else is simply wrong? Hume claims that, since they violate what seem to be the laws of nature, miracles (and by implication the existence of an active God) are necessarily events of extremely low probability. To be persuaded by second-order evidence of miracles, then, requires that one assign an even lower probability to the proposition that the testimony in question has somehow misfired. Hume claims quite reasonably that in the case of such a conflict, the real probability of any human authority’s being correct will never be as high as the probability that the apparent laws of nature have remained in force. And I agree that from a purely objective point of view, Hume is correct about these probabilities. But he is not considering every ingredient of the *subjective* probabilities that matter to rationality when an individual is actually confronted with a problem like this. In particular, he leaves out what the individual may justifiedly believe about his own ability to figure such things out. How would *you* know what the real laws of nature are, or even what the apparent laws are, without trusting other people who may be no more reliable than those who tell me about miracles? Even if you have read Hume's argument yourself and find it quite persuasive on its face, you cannot know for sure that you have understood it fully or that conclusive counterarguments don’t exist. If you then show Hume's reasoning to your parents, teachers, and other authorities, and they all insist that it's unsound, this constitutes a rational argument *for you* against believing Hume's analysis, even if you can’t make sense of their first-order objections. Thus, it is perfectly possible for the balance of *your* subjective probabilities to shift back in favor of miracles.

But what if the traditional belief in question is not just improbable but contains an outright contradiction? If rationality is supposed to be a useful thing, shouldn’t the rational person be allowed to apply autonomously at least some basic tests for coherence? The answer is yes and no. Perceptive rationality requires that your overall view of things be maximally coherent. But this does not entail that each belief you have must make sense to you all by itself, when this conflicts with rationality’s demand that you make the best available sense of your *total* epistemic situation. In order to arrive at perceptive equilibrium – i.e. to construct the most coherent perceptive model of the entire world you can at any moment – you may well need to sacrifice some of the comprehensibility of some of its parts. If your elders had proven unreliable whenever what they said failed your own tests of plausibility, then you would indeed have good reason to doubt the next implausible suggestion that you heard from them, based on their previous performance. But to whatever extent they have established their positive reliability to you, you have that much inductive reason to believe whatever the next thing is that they say, however hard it is to fit into your existing model. If they have established a degree of reliability that is superior to your own, either in general or in certain areas, then you can easily end up with better overall reason to defer to these authorities than to believe things that seem in themselves to make a lot more sense. If this causes some local inconsistencies to crop up within your total perceptive model that’s too bad, but it is sometimes unavoidable.

Consider the so-called problem of evil. You grow up believing, on the testimony of your parents and other authorities, in an all-powerful, all-knowing, and completely benevolent God. One day, it occurs to you that the existence of such a perfect God is flatly inconsistent with the pointless suffering that is experienced by millions of innocent people (and other creatures) every day. Now, what should you think? I say that you are faced with a choice that is, in principle, still pretty simple: You can continue to trust your elders in this matter, or you can trust yourself (i.e. your strictly first-order conclusions) instead. Admittedly, you have found yourself in a certain psychological state, the state of feeling yourself persuaded by an argument to the effect that one of your long-held beliefs is incoherent. But is this feeling a reliable indication of the truth? Has something suddenly made you an expert on what is and isn't actually incoherent, as distinct from what merely seems to be incoherent? Suppose that your parents and other authorities now tell you that they are well aware of the problem of evil, and assure you that it can effectively be solved through some kind of subtle reasoning, or perhaps through an alternative process that goes "beyond" argumentation. As long as you still rationally believe that you are not as good at figuring such things out as they are, you will not have much more reason to disregard authoritative testimony now than you did before the problem occurred to you. This is true *even if you turn out to have been right*, and the problem of evil really is objectively devastating to traditional conceptions of God – indeed, even if you "know" that this is true, in the sense that you are aware of the correct first-order path to that conclusion. At the moment, you are still in no condition to trust such “knowledge”, since it comes from a relatively unreliable source, namely yourself. Consider again the example where I have been adding up numbers and managed to be 94% accurate, while you have been right 98% of the time. We add up a new column of numbers and disagree on the result. Other things equal, I should defer to your result as much more probably correct than mine, even though it still *seems* to me that mine is right in the usual way. Now suppose that just before we calculate this final sum, an intelligence beam is shot into my brain by aliens that makes me suddenly 100% accurate in calculating sums of any magnitude. Now, whenever a sum seems to me to be correct, it definitely is correct. Should this change my belief about who is right in the present example? I say no, for the simple reason that I do not yet know that I have become 100% reliable. As far as I can tell, I am still less reliable than you are, despite how things seem to me, *and despite how they actually are*. So I am still in no rational position to believe that I am right and you are wrong.

The same point holds for beliefs based on authority in general, not just religious or traditional doctrines. I have a friend who once claimed to have proved that relativity theory is incoherent (he's a philosopher, not a physicist). And he may, for all I really know about such things, have been substantively right. But I am sure that this person was not rationally *justified* in holding this as a belief, assuming he was being sincere about it. It almost doesn't matter whether relativity theory actually is coherent, or how good my friend's argument was, objectively speaking. Once he had shown it to the scientific experts and they had scoffed at it, he should have deferred to them if he was going to be perfectly rational about it. I am in something of the same situation myself with respect to the coherence of quantum mechanics, except I realize that I am in no position to judge the matter, so I just accept the experts' view that this theory is correct. It seems patently self-contradictory to me that something can be both a wave and a particle, but that apparent incoherence is not sufficient reason for me to disbelieve that proposition in spite of unanimous expert testimony. Even if a theory seems so incoherent that I cannot even frame it in my mind, it may be rational for me to accept that such-and-such a theory, which I cannot understand at all, is nevertheless probably true.

To sum up my central argument: when the available expert testimony on some point is truly unanimous, its epistemic weight can simply overwhelm the other resources of the individual rational mind, including reason itself, applied directly to that point. The most rational people in the world can believe literally *anything* if it is part of an overall theory of the world that makes the best sense out of their total first- and second-order evidence. This is why I think it is unreasonable to say that traditional religious beliefs depend essentially on the believers’ fear, stupidity, or any other human flaw, as Freud and many others claim. Nor does this sort of faith depend essentially on the rejection of evidence, as Russell and Dawkins claim (though this may well become an issue), but rather on the *acceptance* of evidence in the form of testimony from reliable sources. And, as I shall argue, even when traditional believers do reject objectively good evidence, and even when they speak irrationally in the ways that Dawkins and others frequently point out, they have typically come to do so only because they are quite rational in the first place. Thus, if Dawkins must insist on characterizing faith as a disease, then he ought to admit that it is our very rationality, not its opposite, that makes us susceptible.[[131]](#footnote-131)

4.3. Epistemic doctrines

Here is an important qualification to my claim that religious beliefs are typically rational in the dimension of perception. The traditional beliefs that I have been discussing always seem to get passed on in categorical form. This implies that whatever doubts or hesitations each new believer has will tend to be overridden by the unanimous and *unhedged* testimony that he receives from those around him. Therefore, a certain amount of potential evidence is lost each time a new believer is brought into conformity with the existing group. If it were common to think and speak much more articulately than we do, we might expect traditional beliefs to be transmitted in conditional or probabilistic forms: “If there is cosmic justice, then it seems that we must survive somehow after death”, “If what I have been told is right, then there was once a flood that covered the entire world”, “God probably exists”, and so on.[[132]](#footnote-132) If this were so, then the uncertainties of each believer would survive to the extent that they diminished the subjective probability of the belief in question for each new believer, who would then add his own doubts into the mix.[[133]](#footnote-133) Instead of unchanged doctrines being categorically propounded by each new believer’s peers and elders, a declining average probability of that doctrine’s being true would then tend to be passed on to the new believers. So it is not clear whether most traditional beliefs would survive the continual watering down that would come from the conscientiously articulate transmission of conditionalities and probabilities. In any case, religious and other traditional doctrines are not ordinarily passed on in this form. Instead, they are almost always taught dogmatically (notwithstanding Augustine, Pascal, and a few other thinkers to the contrary), which is to say expressed primarily in categorical, not conditional or probable, form. This strongly suggests that epistemic principles other than simple perceptive rationality are at work.

Religious and other traditional communities do not need complete local agreement in order to thrive, but can also maintain their doctrines over generations within mixed societies, even as members of tiny minority groups. Where it is well known that other people disagree, it can still be rational for people to retain core community beliefs as long as they are also taught to follow certain *epistemic doctrines,* i.e. second-order principlesabout what counts as proper evidence for and against the core beliefs, and about who is and isn't competent to understand or criticize those beliefs. Two very basic such doctrines in particular, common to all but the most isolated epistemic communities, serve to diminish the subjective evidential value of contrary first-order reasoning and testimony from outsiders by reinforcing what is rational in any case.

The *doctrine of humility* is that you are not should not rely on your own first-order epistemic resources when they tend to oppose the testimony of your elders. You are not capable of evaluating such disagreements in a sufficiently objective and intelligent way.

It is already rational for you to defer to others when they are more likely to be right on some subject than you are. The doctrine of humility transforms this rational principle into an explicit belief backed up by authoritative testimony: other people are very much more likely to be right than you are; therefore, you shouldn’t try to figure these matters out for yourself.

The *doctrine of solidarity* is that you should not rely on the testimony of outsiders when it tends to oppose the testimony of insiders. Other people lack the special wisdom of your peers and elders, their thinking has been corrupted in some way, or they bear ill will toward you and your community.

Again, it is already rational for most people to prefer the testimony of people, typically insiders whose track records have shown them to be trustworthy, over that of people they don’t know and have no particular reason to trust. The doctrine of solidarity erects a fence along this natural boundary, an extra obstacle that outside testimony has to clear before it is taken fully into account in the perceptions of insiders. These two doctrines together define a common pattern for all epistemic communities, but especially "clannish" minority religious groups like Amish Christians, Hasidic Jews, and others with highly self-contained epistemic communities. Their core moral and religious doctrines have been able to survive because they have been protected by a barrier of epistemic doctrines that effectively prevent their members from either challenging those core beliefs themselves or taking disagreements with outsiders very seriously. Such communities achieve a kind of *epistemic closure* when membership itself depends on uncritical adherence to its doctrines. In a closed epistemic community, it is belief itself that determines ones status, as in Catholicism as opposed to Judaism. To the extent that members restrict their intellectual relationships to others in the same community they will avoid exposure to contrary beliefs, even if they are surrounded by outsiders who disagree.

There is also a contravening tendency, often the stronger one in open, cosmopolitan societies like ours, for such minority epistemic communities to break down when their members’ justification for believing in their core doctrines is overcome by sufficiently powerful new first- or second-order evidence, leading to what we call a loss of faith. In some cases, individuals who become expert insiders themselves can discover that their first-order epistemic resources are in fact more reliable than those of their elders and peers. Similarly, outsiders are sometimes able show themselves to be more reliable sources of information in general than the believers’ original authorities. As inductive evidence accumulates that trusted community authorities are no more likely to be right about things in general as the mistrusted outside sources, the subjective probability for the believer of the received doctrines being correct decreases, sometimes to the point of threatening his very membership in the community. College education, travel, and military service have long served this dissolutive function, confronting previously sheltered individuals with evidence that non-members can indeed sometimes be trusted and that members sometimes cannot.[[134]](#footnote-134)

This sort of epistemic breakdown takes place only occasionally among the Amish, Hasidim, and other traditional communities that are largely centered on what we now call family values. For many members of these groups, experience with a harsh and chaotic-seeming outside world only strengthens the inductive prestige of the testimony of their elders. Whatever theological doctrines are mixed in with the moral beliefs of these communities, their claim to a superior way of life is hard to refute empirically. The modern secular society has proven to be very good at science and technology, something that these sects do not deny, but it is not always successful at providing happiness. And young people cannot just pick and choose among the traditional doctrines that they have been taught, keeping selected values and eliminating all the rest. To do so is already to have adopted an "outside" attitude toward their elders. Depending as it does essentially on trust, not on first-order agreement, their membership in such a community can be evaluated only as a package deal. And as a package deal, the choice to remain in isolation from the modern world may well be rational for many people who have been raised in such comparatively cohesive and morally successful communities.

The situation is more complex for people who have been brought up in less successful or less isolated situations. The Islamic Middle East comprises many such groups, where modern Western culture has encroached on and assaulted the existing traditions by offering a different package deal of economic and technological progress combined with radically un-Islamic principles and epistemic doctrines. Traditional practices like conscientious religious observance, arranged marriage, and the severe sexual restraints placed especially on women are threatened wherever Westerners, or even their popular music, movies, and magazines appear. Community precepts like respect for elders and belief in the ultimate triumph of Islam are mocked by the preeminence of Western power, transgressive imagery, and liberationist ideas. So, the choice that remains for such societies is a tough one: whether to maintain traditions at the cost of an increasingly harsh epistemic isolationism, to give up and join the West as very junior partners, to try to finesse the situation through tactical compromises between Westernizers and Islamists, or to bet everything on a desperate, seemingly crazy effort to conquer or destroy the West. Each of these strategies has paid off to some extent for some groups, but the overall situation remains a violent mess, partly because the flood of change has been too rapid to allow for gradual accommodations.

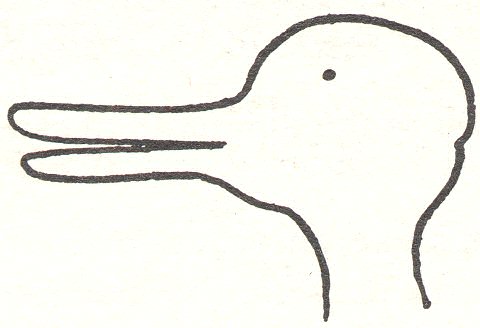
The parallel struggle in the West between progressive modernists and mainly Christian cultural conservatives is often tense but rarely violent, largely because the changes are intrinsic to the West itself and not generally experienced as a humiliating assault from outside enemies.[[135]](#footnote-135) Still, this is seen as a life-or-death issue in some Christian communities, including the official Catholic Church. How much accommodation to contemporary secular culture is consistent with essential Catholic teachings? Is existing Catechism still definitive, or will the “pretty good” Catholic, who opts out of beliefs that he is inclined to disagree with while adhering to those he independently views as most important, be allowed to ignore authority? If this "cafeteria" approach to Catholic doctrine becomes widespread, it is not clear whether the Church can recognizably survive for very long. The last two Popes have shown themselves willing to sacrifice the loyalty of millions of relatively casual adherents in order to preserve the Church's total doctrinal authority for a core membership of serious Catholics. By maintaining strict epistemic discipline on questions about abortion, celibacy, homosexuality, and female priesthood despite the growing and deepening unpopularity of Catholic teachings on these issues throughout the West, the Church seems determined to preserve its package deal intact for distant future generations of believers, however few or many they may be.

Meanwhile, the progressive side of Western culture has become more and more openly hostile to Christianity and traditional moral values, almost defining itself as oppositional to them. When I was young, my friends and I liked to hassle our “straight” (i.e. traditionalist) parents with arguments about sex, drugs, war, and religion, imagining that little more than a few slogans were enough to demolish all their dumb beliefs objectively, and relishing their inability to come back with replies we found convincing. I remember that the big reward for us came when somebody’s mom got angry enough to snap, “What makes you think you’re so smart?” We found this question hilarious and repeated it ironically among ourselves, as if it were too patently absurd to need an actual reply. But now I think it was a pretty reasonable question. What *did* make us so smart as to rationally contradict our parents’ view of things? We weren’t more mature, more intelligent, or in general better informed than they were, yet we were entirely confident that we were right and they were wrong. I have been arguing here that this high level of confidence, at least, was not rationally justified; we all should have realized that we were actually not so smart. But beyond this, our parents’ question also attempted to convey the doctrine of humility: *don’t* think you’re so smart; you’re *not* so smart. But this didn’t work on us at all, because we were already too far gone in rejecting any so-called wisdom from our parent’s generation, at least on these favorite topics. Even though we actually weren’t so smart and had no reason to believe we were so smart, we were determined to dismiss the doctrine of humility with the same contempt that we dismissed the going doctrines of propriety, sobriety, and general self-restraint that defined the middle-class culture of our parents. We made a loud point of rejecting solidarity as well, deriding and mocking Christianity, American consumerism, corporations and their buttoned-down careers, plus all the prominent figures of the day who stood for traditional “flag-waving” values from John Wayne to Richard Nixon, along with police, the military, and the war in Vietnam. It is a little puzzling why we felt we had to do this as a consequence of leaving traditional beliefs behind: to become not mere *non*-traditionalists, but fervent *anti*-traditionalists. This tendency that people have to switch sides radically, instead of moving freely around among competing ideas, depends on yet another rational force that works to hold our models of the world together in one place.

4.4 Epistemic gravitation

So far, I have been speaking as blandly as I can about the actual content of traditional beliefs. I haven't either endorsed or condemned any religious or other traditional doctrines here, beyond just saying that it is typically perceptively rational for their adherents to believe them. If any such doctrines are also *true*, then I can say that they have been transmitted in a subjectively rational way to people who objectively ought to believe them. But it is clear from all the contradictions among people's religious doctrines, if from nothing else, that most of these beliefs must be *false*. And if they are false, then it is surely a bad thing that people have been rationally forced into believing them. If it were easy to correct such false perceptions with new first- or second-order evidence, this would be a relatively minor problem, but it is not. The big problem with community-based beliefs, traditional or otherwise, is that it's often very hard for rational individuals to change their minds about such doctrines once they have accepted them, however much new evidence accumulates against them, due to a kind of *epistemic gravitation* that I want to explain.

The problem goes beyond our being surrounded with second-order evidence that a doctrine is true. It even goes beyond our being taught epistemic doctrines that make us resistant to trusting sources from outside our communities. Another major epistemic force commonly called the *theory-ladenness of perception* can also make it harder and harder for us to change our standing perceptions, regardless of new evidence that counts objectively against them. At the beginning of Chapter 2, I mentioned our cognitive ability to see a rabbit in the grass and similar real objects, not just a group of different-colored patches in our visual fields and other sensory stimuli. We typically see things *as* things belonging to this or that category; thus the world is interpreted through perception according to what we already believe the categories are, and how we already expect things to fit together in experience. A nice illustration long familiar to philosophy students is Ludwig Wittgenstein's *Gestalt* "duck-rabbit":



If you think of this picture as a rabbit facing right, it looks a certain way to you, but if you think of it as a duck facing left, it looks different somehow – “duckish” rather than “rabbitish” – even though you know that nothing has changed in either the picture itself or your physical sensations. Thus, your visual perception of the object seems to be *laden* with your *theory* about what you are looking at. Indeed, once familiar with the image, you are liable to see it as a third thing, not a duck nor a rabbit but a certain philosophical example, the duck-rabbit. It takes a much greater effort of mind to see the picture without any such interpretation, just as a squiggly curve with a dot inside.[[136]](#footnote-136)

This fact about perception, that we tend to see what we expect to see, systematically warps new evidence in favor of our existing beliefs. On the Shakespeare authorship question, for example, a true “Oxfordian” (someone who, unlike me, is convinced that Edward de Vere wrote Shakespeare’s works, or at least considers it highly probable) is likely not just to trust different private and public sources of information about the issue, but also to perceive the public evidence differently, to weigh it differently as to its relevance and its importance, and to use it differently in argument from someone who is equally committed to the traditional “Stratfordian” position. Consider, for instance, a short poem attributed to Oxford called “Were I a King":

Were I a king I could command content;  
     Were I obscure, unknown should be my cares;  
 And were I dead, no thoughts should me torment,  
     Nor words, nor wrongs, nor loves, nor hopes, nor fears.  
 A doubtful choice, of three things one to crave,  
 A kingdom, or a cottage, or a grave.[[137]](#footnote-137)

Oxford supporters perceive this verse as utterly Shakespearean: concise, clever, beautifully cadenced, and expressing typical themes of suicide and frustrated ambition.[[138]](#footnote-138) Stratfordians tend to perceive it as a mediocre effort, plainly unworthy of the Bard.[[139]](#footnote-139) This same poem counts, then, for Oxfordians as evidence that Oxford wrote the works of Shakespeare, and for Stratfordians as evidence that he did not.

This resistance to correction is a central feature of Thomas Kuhn's concept of a *paradigm* in science, a notion that is commonly applied these days to all sorts of beliefs.[[140]](#footnote-140) It is a hard word to define precisely, but roughly it means a school of thought embedded in an epistemic community, with its basic theories and methodologies in the center holding corollary theories, methodologies, and institutional arrangements all together. The theory-ladenness of perception seems to form a kind of gravitational field around any such set of core beliefs, as it were warping the epistemic space around them. This tends to make our perceptive models more coherent and to push us toward reflective equilibrium – that's the good part. But it also makes us less sensitive to evidence that ought to count against our present models, less able to perceive the pure shape underneath the duck or rabbit, and that can be bad.

A similar theory-ladenness inflects our memories. On coming to believe through testimony that you experienced something in childhood, you are liable to find yourself seeming to remember it; on being convinced that it did not take place, you are liable to forget it.[[141]](#footnote-141) For a particularly horrifying example, consider the satanic day-care scares of the 1980s, when day-care providers in several states received long terms in prison for performing sadistic rituals on children under their care, all based on memories "recovered" through constant prompting by a team of allegedly expert psychologists.[[142]](#footnote-142) And it is not only children whose memories are suggestible. I find my own apparent memories of household events easily influenced by my wife's telling me what has or hasn’t happened, for example that it was she, rather than I, who last emptied the cat box. Sometimes I suspect that she's putting me on, but I can't really tell because my memory is so elastic.

Even our reasoning is subject to influence by present theories. While I agree with Descartes that normal people are all competent at making the most basic inductive and deductive inferences one step at a time, much of our reasoning is unconscious or holistic, or involves subtle or higher-order forms of inference. For example, disagreements in social or environmental science sometimes bring about fierce struggles over statistical inference. The "Bell Curve" debate over intelligence in the 1990s and the more recent debates about global warming have both quickly devolved into such methodological disputes, with accusations from both sides that their opponents' reasoning has been corrupted to confirm desired conclusions.[[143]](#footnote-143) It also matters to good reasoning what premises we choose reason *from*, given the huge piles of potential evidence available to us, and our perceptions of what counts as relevant are very sensitive to pre-existing theories. For example, people who already perceive American justice as racist find important evidence in disproportionately many victims of crime being black; those who already believe that such charges of racism are overblown find it more relevant that disproportionately many perpetrators of crime are black. Similarly, those already opposed to the death penalty on racial grounds give more weight to the fact that black defendants are likelier than whites to be convicted of capital crimes, while their opponents give more weight to the fact that white defendants are more likely to be executed.[[144]](#footnote-144) It is not at all clear *a priori* how an outsider ought to weigh these facts against each other. For insiders, consistency with core beliefs seems to be doing much of the work.

For an extreme example of such epistemic gravitation, I remember a colleague (not a philosopher) once telling me during an argument to stop using *analogies,* on the grounds that "analogies don't work". I never found out what this person could have meant by this in substance, but I remember that he was objecting at the time to a particular analogy that I considered pretty effective. I don't think that he was being insincere, though, and I don't think he was or is irrational in any general way. My point is only that his preexisting theory about the substance of our disagreement, according to which my analogy could not possibly have worked because it led to a patently false conclusion, seems to have altered his very conception of a proper argument, if only temporarily.

Obviously, none of this means that it is rationally impossible for us to change our minds, even about our deepest beliefs. As long as there is evidence available that your own contrary first-order perceptions are reliable or that some outside sources can be trusted, you can still break free eventually. But this will typically require far more first-and second-order evidence than would seem necessary to outsiders, just as it requires more energy to reach escape velocity from the surface of a massive planet than from a little moon or gravitationally neutral position out in space. Someone who already believes deeply in witchcraft and demonic possession cannot be rationally moved by first- and second-order evidence about ergot poisoning or schizophrenia nearly as much as someone who has been brought up without believing in the supernatural, even if he is just as open to the evidence in principle. In the same way, the argument from evil strikes many atheists as conclusive proof that God does not exist, but believers tend to see it as a problem to be worked out carefully by theologians, not ordinary faithful Christians, Jews, or Muslims like themselves. And the theologians in turn tend to see at worst a puzzling anomaly, not a knock-down refutation of their core beliefs. As I have suggested, this is not irrational of them, even if the argument from evil is objectively conclusive. Hume claimed in his discussion of miracles that each believer has to ask himself which is more likely: that the evidence in question actually falsifies what seem to be the best-established core beliefs one has, or that there is some kind of problem with the evidence? And for a skeptically-minded philosopher like Hume, floating relatively free in epistemic space, the preponderance of probabilities will say that straightforwardly valid arguments and the apparent laws of nature ought to be trusted over any testimonial evidence for supernatural events or beings. But for a rational life-long believer within a theistic community, the probabilities will lean the other way and God will still appear to be more likely to exist than not, in spite of the most seemingly-conclusive arguments against it.

Still, rational long-time believers are sometimes pulled out of orbit by accumulated doubts based on first-order arguments as well as testimony from their non-believing peers and new authorities like college teachers. One of my friends in philosophy is a onetime divinity student who withdrew abruptly from his seminary after concluding one day that the argument from evil, which seminarians are taught to take very seriously, could never be plausibly addressed by Christian theology. After this he became a devoted atheist, perhaps the fiercest one I know, claiming at regular college debates on the topic that the argument from evil is not only conclusive as a refutation of belief in God, but *obviously* so to any *rational* mind. I differ from his view on this, since I believe that he was a rational person before as well as after his conversion from belief to disbelief in God. But it is interesting that he perceives the change the way he does, not just as a change of mind about one topic, but as something like a total transformation of his intellect. It is as if a rubber band, stretched to the limit, finally snapped.[[145]](#footnote-145)

When this sort of thing happens in an individual, sometimes the person is launched into a free space of uncertainty, but seemingly more often, the person simply shifts from one orbit into another, as my friend shifted from theism to atheism without pausing at agnosticism – an individual instance of what Kuhn calls a *paradigm shift* in reference to scientific revolutions. It is unusual for this to happen very gradually, because such deep beliefs are usually justified only as parts of package deals that include first-order evidence and reasoning, second-order evidence from testimony arising within an epistemic community, theory-laden perceptions that tend to support core tenets of the package, and epistemic doctrines that discount the value of contrary evidence coming from outside. To the extent that a rational believer’s initial package deal breaks down because of evidence coming from newly trusted authorities with their own core theories and epistemic doctrines, such as a student who was raised in a community of Christian fundamentalists being persuaded to reject his Biblical creationism by Darwinian biology professors at college, the believer is liable to be pulled immediately into whatever alternative package deal those new authorities support. Alternatively, to the extent that a believer’s prior package deal breaks down all by itself, according to its own principles and internally admissible evidence, the believer is more likely to be left floating in epistemic outer space, as were some disillusioned communists after the Hitler-Stalin pact of 1939.[[146]](#footnote-146) But this is much more rare, and even most ex-communists quickly gravitated to another home.

According to Bayes’s Theorem, the higher your prior probability for any proposition, the lower the extent to which that probability can change when you incorporate new evidence. In theory, given a subjective probability assignment of 100% for any perception, no amount of evidence at all can make it rational for you to alter that initial probability. If you are ever *absolutely* certain of anything, you can *never* rationally change your mind.[[147]](#footnote-147) Most of the time, we don't get close enough to such perfectly certain priors to make contrary evidence literally disappear. Even if I have just soberly and deliberately shot my brother, burned his body, and buried it in a shallow grave, with all the vivid perceptions such a chain of experiences is liable to produce in me, I am still capable of coming to believe on sufficiently powerful new evidence that it was all a trick (perhaps that I had killed a surrogate instead), and that my brother was in fact alive and reasonably well. But under the most extreme forms of epistemic gravitation, even the most rational perceivers cannot avoid achieving total certainty. In this sort of situation, which I think of as an *epistemic black hole*, all conceivable new evidence will be assimilated to existing core beliefs, and there is no purely rational means of escape at all. This might be a good, or at least an acceptable, thing when those core beliefs are demonstrably true. But usually they are not, and it is a bad thing indeed that some people seem to get stuck for life with limited, bizarre, or nasty models of the world through no epistemic fault of their own. It is easy enough to see this happening to people raised in desert cults or other forms of extreme epistemic isolation. It is harder to see, and more frustrating to acknowledge, that this happens to well-educated, high-functioning people in cosmopolitan societies as well. But it does, and it happens to them because they are rational people, and despite their being surrounded with objective evidence contrary to their views.[[148]](#footnote-148)

In empirical matters, at least, it may seem that nothing that can be imagined false could ever rationally be perceived as certain. Since empirical perceptions of the external world are always inductive in one way or another, and induction can never be proof (its whole purpose being to take us beyond what is strictly entailed by evidence), there is always some positive chance of our being wrong. So, it should never happen that a rational person concludes that there is *no* chance of their being wrong about a categorical empirical claim. But it seems that there are cases where we ought to be certain anyway, namely when we have authoritative *testimony* to the effect that things really are certain.[[149]](#footnote-149)

Consider another slightly idealized discussion between a mother and her small child (I am using the example of belief in God, but in principle it could be anything else):

Mother: Stop doubting God's existence. You should feel perfectly certain that he's real.

Me: Why? I haven’t heard any conclusive arguments.

Mother: Because I say so.

Me: Well, are you always right?

Mother: No, but I am right this time.

Me: Okay, but for me that only means that you are *probably* right, since I know (as you just admitted) that you are sometimes wrong. So I should still believe only that God *probably* exists.

Mother: But I am telling you that you are wrong about that. The fact is that God *certainly* exists, and this is what you should believe.

Me: But it's irrational for me to accept that just on your testimony, isn't it? Whatever other reasons *you* might possess for feeling certain that God exists, all *I* have to go on is your admittedly superior but still imperfect track-record. So it can only be a matter of probabilities for me. Right?

Mother: Wrong. Think of it this way. Who do you think knows more about logic and probability, not to mention theology and epistemology? Therefore, who is more likely to be right here, on the question of whether God’s existence is certain or merely probable?

Me: Me?

Mother: Guess again.

Now I am in a kind of double bind. On the face of things, it is inductively irrational for me to trust my mother *absolutely* about anything, especially a question where it seems to me that what she's saying is unreasonable. But then I realize that it is also irrational for me to trust my own rationality absolutely, given what I am being told by such a reliable source. My mother, having almost always been proven right in our previous disagreements, holds nearly total philosophical authority over me. What she says in this case is probably right, then, even though I can't say that I understand it, which is not only that God exists but that I ought to be certain of this fact. So, that's what I rationally ought to believe. My first-order reasoning here has told me that I shouldn't believe that God's existence is certain because it seems at least conceivable for God not to exist. As long as there is any uncertainty about the certainty of God's existence, the probability of God's existence must then be less than 100%. But now I also have good second-order evidence that my first-order reasoning is unreliable about this very question. So, I have to decide which source of perception is more likely *overall* to be correct, and this is where my reliance on what appears to me at first as simple rationality must give way before authoritative testimony. To insist instead on following what strikes me *directly* as the rational track would be to run myself off of the very same track in rejecting what I am being told by an authoritative source – I might just as well abandon second-order rationality altogether. The result is that I end up suppressing my prior first-order conclusion that such things cannot be certain, and accounting for my certainty in a belief that I don't even understand entirely by recourse to my mother’s authority.[[150]](#footnote-150)

Pressed for further explanation, my mother now invokes a mysterious but very potent new epistemic doctrine, which I think of as the *doctrine of faith*.

…

Me: Me?

Mother: Guess again.

Me: All right, you. But I still don’t see where all this certainty comes from. You admit that your perceptions are fallible, right? So, you shouldn’t be 100% certain about anything yourself. So, I’m asking you to help me understand why I’m supposed to be 100% certain about this stuff, beyond just that I ought to take your word for it.

Mother: Well, it’s a matter of faith. You just have to have faith.

Me: You mean, I’m supposed to believe whatever you tell me. I get that. But what’s the reason *you* believe in God so certainly yourself?

Mother: It’s because I have faith, just like I’m telling you to do.

Me: Okay, but faith in who?

Mother: Whom. God.

Me: But wait: *why* do you have faith in God? He doesn’t even talk to you.

Mother: Well, I just do. That’s how I know for sure that he exists. And so should you.

Me: But how is that any kind of reason to believe in anything?

Mother: It just *is* one.

Now I am being taught by an authoritative source that faith is not simply a matter of rationally trusting authorities, but an ultimately independent source of justification that can sometimes trump all other evidence or lack of evidence.[[151]](#footnote-151) And once someone has been fully convinced that reason *doesn't matter* on a certain question, there seems to be no way for him to change his mind again through argument, because he will perceive all further arguments as irrelevant to the belief in question. This new second-hand certainty can make me a hard person to talk to. If someone challenges it, I will have little to say in its defense other than that I learned it from people I rely on implicitly, and who told me that it was not to be doubted. In itself, this is no more or less unsatisfactory than a layman’s defense of his certainty of anthropogenic global warming, where this has been learned from teachers who insist that it is beyond reasonable dispute. But in the global warming case, the believer can at least point to other people, i.e. climate scientists, who could provide a substantive, ultimately first-order explanation of the theory. The case of my belief in God may well be different, though, in that it seems to depend on second-order evidence all the way down. In James’s phrase, my faith is “faith in someone else’s faith”, and theirs is faith in some third person’s faith, and so on down the line (or maybe in a circle) without any point at which the reasoning is grounded in substantial empirical evidence. This may be why faith itself, though proximately based on rational acceptance of authority, is so often presented as a separate, free-standing means of justifying our perceptions, alternative to standard empirical reasoning (for example in the quote from Pascal above).

Hence the apparent pointlessness of some discussions:

Me: I’ve read your book, but I still know that God exists.

Richard: How is that possible? First of all, God doesn't exist, and you can't know something when it isn't true. But even supposing that it might be true, where is your evidence for believing it?

Me: Well, there isn't really much in the way of what you would count as evidence, but that doesn't matter for me, because my certainty of God's existence is based on faith.

Richard: But why should anybody ever go beyond the evidence?

Me: It’s a good thing to do. You get better beliefs that way.

Richard: And why should you believe that?

Me: Well, that’s also based on faith.

Richard: So, aren't you just admitting that you’re being irrational?

Me: I prefer to say *a*-rational, or perhaps *trans-*rational. Faith goes *beyond* reason, and I am a man of faith.

Richard: You are sick man, is what you are. This faith of yours is nothing more than a disease of the brain, don't you know that?

Me: No, I really don't know that. But I do know that God exists, and that I know this on the basis of faith.

Richard: Fine. Then you’ve placed yourself entirely beyond the realm of evidence and rational argument. If that’s not a brain disease, then I don’t know what is.

Me: How about meningitis?

As Richard sees it, I am plainly an irrational (indeed, *infected*) sort of person. But observe how calm and *seemingly* rational I am in this discussion, while my opponent is the one getting emotional. It's an odd kind of irrationality that carves out a certain region of belief within which to think in a particular irrational way, while leaving the believer's reason otherwise intact. This is because I am in fact still being perfectly rational at a deeper level. I have reason to trust my mother, so much so that I ought to accept whatever she tells me about the limits of rationality. In this way, I rationally come to believe that ordinary rational standards don’t apply to certain core beliefs, but that they still apply everywhere else. If this is irrationality, then, it is a kind of rationalirrationality. Or, it may be better to say that the person is being rational at the bottom level and irrational at the top, or rational in the broad context and irrational in the narrow one, or rational as an enduring person and irrational (on certain issues) at the moment.[[152]](#footnote-152)

This condition may yet change if, later in life, I find out that other people who deny that faith is an acceptable foundation for belief are generally more trustworthy sources of information than my present authorities. In that situation my primary rationality, plus enough of the right kind of new evidence, could rescue me from my present secondary irrationality. But as long as I remain within my first community of faith, surrounded by like-minded people and warned by people I ought to believe against trusting either outsiders or my own first-order reasoning, I will never gather enough first-order evidence about the overall trustworthiness of outsiders to be able to gain much second-order evidence from them about either the content of my core beliefs or their certainty. My first-order reasoning concerning these beliefs now having been suppressed through rational acceptance of authoritative testimony, I will have gotten stuck for life with these beliefs through no fault of my own (and probably no fault of anybody else's, since my authorities all grew up in the same situation). In such a fixed community, the field of second-order evidence in which we have been raised are usually too massive and dense, and our first-order perceptive fuel too limited, for us ever to escape the epistemic gravity that holds us there.

There is a final force that sometimes makes it harder still for rational people to change their epistemic situations. Like the restrictive epistemic doctrines that I have been discussing, sometimes *moral* doctrines also act to make people’s perceptive models uncorrectable. It is hard enough to overcome authoritative testimony as to what is true, and as to what is certain, and as to how to figure out what is true and what is certain, but moral testimony sometimes gives us additional and very powerful reasons to keep believing as we do regardless of contrary first- and second-order evidence. When we learn what we ought to believe , sometimes the "ought" involved is moral as well as rational, and we are told that doubting or believing something is not just false or epistemically unjustified, but also wicked.

Suppose I have been taught authoritatively that God or a particular conception of God exists; also, that it doesn't matter whether there is any first-order evidence for this belief, since it is a matter of faith; also, that it is foolish for me to doubt that God exists or that faith is a good source of belief, since I am not an expert on these difficult questions; also, that it is positively wicked for me to doubt such things; also, that reason itself is wicked when applied to questions about God; also, that people outside of this faith (literally, *infidels*) are wicked people; also, that their arguments are all just fancy verbal tricks; also, that it is both foolish and wicked to listen to the seductive arguments of infidels; also, that nothing is more wicked than somebody who becomes a heretic or an apostate.[[153]](#footnote-153) Having rationally come to believe all these things, I will then be in the worst possible situation from within which to find rational escape – the blackest possible of epistemic black holes – since I will believe that it is morally wrong even to think about escaping. And this will not bother me, because I will not be an infidel and probably won't feel the slightest temptation to become one, which is of course part of the problem. I will instead feel proud not only of my intellectual and moral virtue in believing what is certainly true, but also of my contempt for those who disagree with me and for all of their arguments, and even for rationality itself when it gets in the way. As the great Christian father Tertullian so proudly said about his faith, *Certum est, quia impossibile.*[[154]](#footnote-154)

A similarly moral disdain for common rationality appears within the post-modernist community, focused as it is on issues of intellectual oppression. Spearheaded by Jacques Derrida, the deconstructionists’ critique of *logocentrism* is anattack (or counter-attack, as they see it) on rational debate as an arbiter of truth. The core idea is that our conceptions of truth and reason are always defined by (typically racist, sexist, classist, or heteronormative) power structures and, since nobody lives outside of one such structure or another, no one can ever have an objective, un-socially-constructed view of the "real" truth, including the truth about rationality. [[155]](#footnote-155) This post-modernist account of truth amounts to something like a radically relativistic interpretation of Kuhn's theory about scientific paradigms where, everybody being already situated within competing paradigms, no trans-paradigmatic position exists from which to resolve disagreements between the paradigms.[[156]](#footnote-156) Indeed, this basic view of reason and truth appears repeatedly in Western history, from Plato's enemies the Sophists to Machiavelli, Hobbes, and Marx, and is being widely taught these days at Western universities, sometimes in philosophy or social science, but most often and most dogmatically in literary studies. There are presently thousands of scholars and probably millions of their students who accept some version this theory, albeit with different levels of sincerity and radicality. No student I have known would treat the fact he had been given a particular grade on his final paper as a mere matter of perspective. But many would insist that there was no real fact of the matter as to whether their papers deserved that grade, or indeed whether the statements they made in their papers were true or false, or whether their conclusions actually followed from their premises. It can be pretty frustrating for ordinary realists to argue with people who adopt this view of things.

Larry: Some things are true.

Marie: You mean some things are *counted* as true under the current power structure.

Larry: No, I mean they're just plain true.

Marie: But who decides, sweetheart? Who decides what's true and what is false?

Larry: No one decides. It’s not a matter of decision; it's a matter of fact.

Marie: But how can you ever find out what is the truth and what is not?

Larry: You want me to explain epistemology to you?

Marie: No, definitely not. My point is that you live in a society where people in power – people like yourself – make all the rules, including the rules about what counts as true or false, or as a valid argument. And there is no way for you to step outside that power structure and perceive this thing you call the truth itself. I find it interesting that you do not understand this simple idea.

Larry: I understand, believe me – I was lecturing on Kant and Kuhn while you were still in Paris breaking windows. But that's a problem about what we know, not about how things really are. If we don't agree that some things are the way they are independently of our ability to see them clearly, then nothing we say really makes any sense. For example, what you just said about people in power making the rules: is that the truth or not?

Marie: It is the truth for me. It is the way I see things. It may not be the truth for you.

Larry: But then look: why are you *telling* me that there is no objective truth, if there's no objective truth about whether there's any objective truth? Don't you want me to be convinced by what you say?

Marie: I would never expect to change the mind of someone like you, darling. I am just telling you how things are for me, which I have every right to do. My narrative is just as valid as your narrative.

Larry: Then I have no reason at all not to keep believing that objective truth exists and that you are completely wrong about all of this stuff, since that's what's "true for me".

Marie: *Precisement.*

As is apparent from Marie's reasonable-seeming discourse here, people who believe that there is no objective truth or reason may well continue to *use* reason and to speak coherently about their views. But they cannot *endorse* this use of reason as anything better than unreason if someone else decides to be unreasonable. People who think this way have fallen into even blacker epistemic holes than people who believe in faith as an alternative to reason, having become in theory (though not typically in practice) completely indifferent to rational discussion other than as a kind of self-expressive game, no better or worse than any other game that might be played with words. This principled refusal to take their disagreements with realists seriously robs them of any rational opportunity to change their own minds, either about the content of their personal "truths" or about the ability of rational discourse to alter them. In effect, whatever they happen to believe at present, for whatever reason good or bad, becomes an article of faith as soon as they are called on to defend it.

It may seem that my view of rationality is really no different from Marie's. We both say that people are stuck in paradigms from which there is no rational escape, largely because their thinking has been influenced by other people. Whether we call that having different "truths" or not, the effect is just about the same, that disagreements cannot be resolved through reason; only non-rational forces can deeply change our minds. But there is a crucial difference. Marie believes that *everyone* is in what I am calling an epistemic black hole all the time, while I think that most people are not, though some people are (including, evidently, her). Whether or not we get drawn into black holes ourselves depends on the contingent features of our epistemic lives, including crucially what we may have been told authoritatively about proper reasoning, and particularly whether we have been rationally persuaded that some or all of our beliefs are exempt from strictly rational debate. If we have not, then there is no good reason that an agreed-upon practice of rational discussion cannot play the role for us of neutral arbiter – the paradigm within which all the other paradigms contend – effectively enough. Indeed, it seems to me that Western culture has evolved quite powerful if rather slow-moving techniques for arguing such things out in philosophical, scientific, and even political discussions, as I will argue in the following chapters. But whether we have precise and articulate rules for such discussions does not even matter that much as long as we engage in them in good faith, and as long as reason is *in fact* a single, stable, universal sort of thing.

It is entirely possible, though, at least abstractly, for the post-modernists to be correct in substance, that ordinary logic does form part of an oppressive total system of thought, and that it must be attacked in order for society to progress beyond traditional hierarchies. And it is similarly possible that the pre-modernists are right, and God exists but is accessible only through the rejection of reason and the acceptance of non-rational ways of knowing. Such substantive theories are not particularly problematic in themselves, though they may seem like dead ends to people of a modern, scientific cast of mind. The problem is that no such theory can be *debated* reasonably with people who are so fully inoculated against contrary evidence as to consider ordinary argumentation illegitimate. When people have been taught authoritatively that rational discourse is to be shunned, then what begin for them as second-order empirical perceptions based in reason can become, through no fault of their own, not just uncorrectable but even undiscussable convictions.

I will return to convictions and certainty again in Chapter 7. First, though, in Chapters 5 and 6, I want to talk about the many reasons people have to speculate, to form autonomous opinions, and to propound and defend them in public arguments*.* Such opinions need not be rationally justified in the dimension of perception, so the principles of deference and authority that I have been discussing in this chapter do not apply to them. Nevertheless, they serve essential functions for creative thinkers and for the progressive societies in which they thrive.

**5. PHILOSOPHY**

[E]ven 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.

John Stuart Mill, *On Liberty*[[157]](#footnote-157)

5.1. Conflicting experts.

My account of rational deference so far assumes something close to unanimous testimony on the part of someone’s elders or community. But our epistemic situations are not usually so simple, and we are often forced to deal with conflicting testimony from different trustworthy sources. What should I believe if I have two authorities of equal general reliability, and I discover that they contradict each other? Descartes raises this issue in the *Discourse on Method*, where he claims that conflicts among experts tend to cancel out the value of all relevant testimony in those situations, throwing the individual back on his own resources. In Descartes’ view, a purely first-order criterion of rational belief (for him, ultimately "clear and distinct perception") is then the only rational alternative to skepticism.[[158]](#footnote-158) So, if a third of my elders say that Odin is the king of the gods, and another third say that it is Zeus, and the last third say that there are no gods at all, then according to Descartes, I should just ignore the lot of them and work methodically to figure the theological situation out for myself. Perhaps I will end up agreeing with one or another side in the ongoing dispute, or perhaps I will come up with my own new theory that there is no god, or only one god, or that the gods are democratic. In any case, Descartes’ view is that it is rational for me in these circumstances to believe whatever theory ultimately makes the most sense to me on first-order grounds alone, since my pool of second-order evidence has been polluted by the contradictions it contains.

If his goal is simply to maximize immediate perceptive rationality, Descartes is wrong about this.[[159]](#footnote-159) I think the optimal procedure for us as pure perceivers is to probabilize each of the candidate beliefs, considering all of the theories we can discover and proportioning the degree of our belief in each theory to the overall subjective probability of that theory's being true. My own first-order view of things, if I have one, is never more than one of these choices. I have argued that unless I have reason to believe that I am a uniquely reliable judge of the question, when my first-order view on some issue is contradicted by the unanimous testimony of my authorities, I am far less likely to be right than they are. But this is also the case if my own non-expert view of things is opposed by the *divided* testimony of my authorities, because I am still less likely to be right than *any* of the experts. Imagine that it is 350 BC and I am a student at Plato’s Academy. I discover that Plato says *X* about the causes of cancer, while Aristotle (also then working at the Academy) says *Y*, which is contrary to *X*. Meanwhile, the best theory that I have been able to work out all by myself, ignoring theirs, is *Z*. Now, which theory should I believe is most probably right? Certainly not *Z*, unless I have good reason to think that I am a more reliable philosopher than either Plato or Aristotle. The same is true if I have only consulted a group of philosophical peers, and found that they are roughly split between support for *X* and *Y*, while I am left alone with *Z*. In either case, from a strictly rational point of view, I ought to conclude that the correct answer is more likely to be either *X* or *Y* than it is to be *Z*. So, I should be inclined to rule out *Z* if anything, though I remain in no position to decide between *X* and *Y*. My autonomous theory *Z* may be an interesting one, and I may benefit in some ways from the effort required to produce it, but I cannot conclude that *Z* has any greater strictly epistemic value than its plausibility to me, which I know from long experience to be a very fallible indicator of the truth. If my only principle is rationality, then the subjective sense of plausibility that I attach to *Z* must give way to the higher probability, based on my totalevidence, that one of the other two theories is correct, though I cannot tell which.

Supposing further that I have found Plato to be more reliable than Aristotle, or *X*-ists more reliable in general than *Y*-ists, this will normally favor *X* over *Y* as more probably true in my subjective calculation. But this relative likelihood does not entail that I should assign any high degree of certainty to *X*. It may well be that all things considered, neither *X* nor *Y* (let alone *Z*) is *very* likely to be true, and it will take something altogether different – even presently unimaginable – to get at the truth, as was the case for ancient theories of disease. This is especially true when those supporting the competing theories are themselves not very certain about them. If, as I suggested in Chapter 4, the rational perceiver ought in general to multiply the evident reliability of each source by the subjective probability that each source attributes to its testimony, it is possible for a minority position among equally reliable sources to carry more epistemic weight than a competing majority position when the minority is more certain of its position than the majority is.[[160]](#footnote-160) In fact, an individual or group source that is usually *less* reliable may sometimes be preferred to one that is generally more reliable, if the first source is really much more confident in its assertions than the second. For example, I might take the word of a very mediocre student over that of my department’s thoughtful and conscientious chairman as to whether a raccoon has somehow found its way into our building, provided that the student seems to be very certain about what he is saying and the chairman seems more tentative. Given this difference in confidence levels, it might be reasonable for me to infer that the student has better evidence for his position than the chairman does for his – perhaps that the student has actually seen the raccoon for himself, while the philosopher has been depending, characteristically, only on general principles.[[161]](#footnote-161)

It may seem that I dismiss the possibility of first-order evidence overcoming divided expert testimony too easily here. Sometimes being an expert just means having access to certain kinds of evidence; if you get the same information, that will make you just as much an expert, so at least your first-order views will be as likely to be true as theirs. Ordinarily, though, expertise is not confined to accessing data, but involves trained, practiced interpretation and analysis, too, so there is no quick way for a layman to catch up. When your autonomous reasons and observations are evidently no less reliable than those of any of your sources, they will naturally sometimes tip the balance of belief. For example, if half of your equally-trustworthy sources tell you that raccoons won’t be attracted by granola bars, and the other half say that they will, then your own experiments may be decisive in settling your view of the matter (assuming you want such a thing to be settled). But in a dispute among real experts where you have no expertise yourself, say over whether China or India had the greater population in 1500 AD, nothing you can do yourself is liable to help at all, and you will be stuck in suspense until the experts come to a consensus.[[162]](#footnote-162) Even hard study will not decide the question for you in a purely rational way, unless you are somehow able to achieve a mastery of the subject superior to that of your erstwhile authorities.[[163]](#footnote-163) But it may help to refine your ability to tell which side is more reliable on the topic in question, and this could simplify the issue of divided authority for you. It is only once consensus has been established that you can rationally accept the claims of any expert without further question. Otherwise, divided authority casts the testimony of both sides into doubt for the pure perceiver, making what might have been a passive reception of belief into a matter of conscious rational calculation. So, this much of Descartes’ analysis is true: the more dispute there is among supposed experts – hence the lower the probability of truth for any expert position – the less the individual resources of a perceiver will tend to be outweighed by testimony, because it will be more rational to think that *nobody* is much of an expert.

The same points apply to groups of non-expert peers who are more probably right than we are just because they outnumber us; for example, when you mistakenly believe that smoking is permitted on a bus and all the other passengers disagree. If it’s a closer call and you are in a somewhat larger minority, there is a greater chance of your being right and less reason to defer to the majority. But it is only in the unlikely case of a total breakdown of testimonial reliability that we are ever *completely* thrown back on our own epistemic resources. And even then, if no one in the world, including you, stands out as a reliable authority on some question, then the rationally best thing to conclude is nothing at all. That is, you should believe that some number of coherent views (including your own, perhaps) have some level of positive likelihood, but that none of them should be accepted as more likely than not to be true. The conflict of experts and peers does leave you on your own, as Descartes says – but only in the negative sense that no individual or group has earned your automaticcredulity. In this way, you are no longer just a *consumer* of beliefs. But this doesn't make you independently a more reliable *producer* of beliefs than the conflicting experts, even if you go through all the work it takes to join their ranks.

It seems, then, that to be rational we must ordinarily give up accepting our own first-order perceptions in matters of controversy, since we rarely have good enough reason to prefer our own conclusions to those of the experts, conflicting or not, or even to the majority of reasonable people. This is a skeptical sort of conclusion, but one less radical than those of the ancient Pyrrhonians like Sextus Empiricus who argued for suspending judgment on *all* questions, not just those in serious dispute among reasonable people. I agreed with them in Chapter 2 that no categorical empirical beliefs are ever certain, so we can never *know* absolutely that such things are true, but I agreed with the more moderate Academic skeptics that such propositions can still rationally be perceived as more or less *likely* to be true, given appropriate first- and second-order evidence. [[164]](#footnote-164) And this extends to cases of divided authority as well; we may well rationally infer that one position on a controversial matter is more probably correct than another, even if neither is probably correct *simpliciter*. Here I have added only these two further skeptical points: that propositions subject to controversy among experts are rarely *very* likely to be true, and that there is little a non-expert can do independently to alter these perceptive uncertainties.

5.2. Speculation

This semi-skeptical position about belief is less destructive of autonomy than it may look. For, even if we are severely limited by expert disagreement in our ability to construct categorical perceptions with any high degree of confidence, this doesn’t mean that we should not engage in all the *speculation*, i.e. hypothetical perception, that we like. We can imagine alternative theories and consider arguments that might justify them without needing to *believe* them categorically, or even probabilistically. In the absence of expert agreement, we can acknowledge the low probabilities that rationality assigns to controversial statements in our main perceptive models, while playing freely on the side with other possibilities that are worth exploring for some other reason, for example that they make more sense to us intrinsically. As long as we do not confuse such speculative reasoning with rationally justified perceptions no harm will be done, and possibly a lot of good.

The needed cognitive equipment is already in use. Our perceptive models as they stand contain multiple scenarios to represent uncertain possibilities. Just to get around in daily life we have to speculate practically all the time about what might be lurking around the next corner, what another person might be thinking about us, or what might happen to us if we tell a lie. So we are already well experienced in reasoning hypothetically in ordinary life. And it is no great leap from this kind of practical speculation about issues of personal concern to the more systematic hypothetical reasoning involved in literature, science, or philosophy.

Here is a list of ten pragmatic and epistemic reasons to think about things for ourselves even while not expecting it to end in full-fledged knowledge or even rationally justified belief on the topic in question. Some I will just mention in passing here; others I will talk about at length in what follows.

*Practice*. Speculation about issues where experts are divided gives us exercise in thinking about those practical and theoretical issues where knowledge is actually possible. After all, speculative thinking about things that cannot be known involves the same sort of reasoning as thinking about things that can be known. This is what some of my friends say to me about the one or two philosophy courses they took in college: good practice for doing *serious* work when they grew up.[[165]](#footnote-165) It can make you smarter not to give up when experts conflict, or to work on difficult ideas of any sort, but to try to become at least something of an expert yourself, even without hopes of making new discoveries. This is why most people study mathematics: not to prove new theorems but to get good at solving ordinary problems with or without a mathematical component. If you have *any* long-term intellectual goals, in fact, then focusing exclusively on momentary rationality will just get in the way, like eating all your corn instead of planting it. You may be a perfectly rational *consumer* of ideas, but you will not learn to *produce* them at an expert level.

*Pleasure.* Most of us find considerable psychic satisfaction in working out the best first-order theories that we can, independently of whether we have good reason to believe that our personal take on things is better than anybody else's. We don't only play tennis or chess or piano when we believe we can be experts. People take enough joy in the activities themselves to make it worth their time and effort, even though they don't expect results that anybody else has reason to admire. Thus, as an ordinary golfer you can play your own best game without believing that you can compete with Tiger Woods. You can even sensibly say that you played well today but Woods played badly, meaning that you played well by the standards of your usual game while Woods was disappointed at the U. S. Open, without implying that you think your play was better than Woods's in absolute terms. Similarly, you can work out your own ideas about how Lee should have campaigned at Gettysburg without believing that your theory is as likely to be right as Shelby Foote's, or about the nature of justice without thinking that you know as much as Plato or Aristotle. Again, as long as doing your enthusiastic best is not confused with actual expertise, nothing important is at risk.

*Self-expression*. There is also the social satisfaction of expressing to other people how things seem to ourselves. Such expressions sometimes take the form of poetry or painting, which can give others a glimpse into the world that we experience. Philosophy can do the same thing for our theoretical constructions of the world, and there is as much pleasure and inspiration available to you in this sort of creative work as in the fine arts. The practice of philosophy as this kind of self-expressive literature is perhaps more welcome in continental Europe, home to phenomenology and existentialism, than in the English-speaking world. But even the driest sort of philosophical analysis – indeed, even a technical article in pure logic or mathematics – is the manifestation of one point of view, and may express implicitly more of its creator’s inner experience and desire for understanding than could be addressed directly in an academic publication.

*Understanding.* I noted earlier that Plato defined knowledge as true belief plus *logos*. One of the possible meanings of *logos* is *understanding*, and it seems to me that something like understanding is indeed required for having knowledge in the deepest sense. For the best kind of knowledge it is not enough that you can't be mistaken about some statement relative to some set of background assumptions, or even that you can't be wrong at all, if this is possible. And the same thing seems to be true of belief itself, that you can only fully believe what you can fully understand. Suppose that you are as ignorant as I am about quantum mechanics, and that a person that you have found to be totally reliable so far tells you that photons exhibit wave-particle duality. So, now you are justified in believing that photons exhibit wave-particle duality, in the sense that you have reason to believe a certain undigested *sentence*, namely, "Photons exhibit wave-particle duality," to be true.[[166]](#footnote-166) But this is hardly better than believing that the sentence, "Es ist gefährlich auf das Eis zu gehen," is true because it's being shouted at you by a reliable source – but without your knowing any German, hence without your actually learning that it's dangerous to go out on the ice. The notion of belief is flexible enough in the dimension of perception for us to apply the term to almost any degree of understanding, and few of us are able to integrate all of our beliefs into a perfectly coherent system. But even within the normal range of semi-digested testimonial perceptions, it is good for us to figure these things out for ourselves to whatever extent we reasonably can. It may not make the relevant perceptions any more likely to be true *per se*, but it will certainly help in any *use* to which they might be put, and it will work to make our total models of the world more coherent, hence ultimately more rational.

*Phenomenal knowledge*. There is intrinsic value in exploring how things *seem to you*, even if you can never know whether things really are the way they seem. The phenomenal world, as Kant calls it, is also a real, fact-filled world that you live in, knowable and well worth investigating for its own sake. This is the essence of *phenomenology*, a core element of continental European philosophy over the past century or so. If it seems to you that you are a certain human being drinking coffee at a desk while resisting a certain temptation to cruise around the internet instead of focusing on work, then this seeming-to-you is a fact about the universe. Nobody else but you can know this fact directly, but that makes it no less a fact. If the coffee you are drinking tastes to you like it was made last night and left heating in a machine, you may be wrong about the external cause of your experience (e.g. the coffee may have been ruined in some other way, or perhaps the coffee is fine and there is something wrong with your taste buds or burned-out frame of mind) but the phenomenal facts are just as you perceive them. From your subjective point of view the external, noumenal world is only a construction, anyway, in the sense that you can never know for sure whether it really exists or not. But even if you are only a brain in a vat, you are still definitely having an existence of some sort, and you still want to learn what you can about it – just as, if you discovered that you were permanently trapped inside a cave, it would be better to explore the cave than just to sit there, wishing you could leave.

*Abstract knowledge.* It is interesting in itself to contemplate how things *might be*, even if you know you can't discover how things actually are. Thus, stories set in the distant past or future can be good to read, and sometimes really fascinating, in spite of our total inability to find out whether they are at all accurate when viewed as histories or predictions. By definition, fiction gives us stories that we know to be false in their details, if not in their more general representations of the world. Still, we find it satisfying to think about such stories, and we learn from them as simplified or clarified or analyzed or otherwise worked-out alternatives to the complex and sometimes overwhelming mess of ambiguities we find in actual experience. This applies to hypothetical discussions about general topics too, like morality and justice, existence and identity, or knowledge and rational belief, as much as to the hypothetical events we find in fiction. There is value to us in exploring how different conjectures might or might not be true, even when we fully expect to die before achieving a theory that is even probably true. And this is not just a psychological sort of satisfaction, for the world of possibilities is also a world we live in and a world that’s worth exploring in its own right. Some things really are possible or likely under various conditions while some are not, and these are facts. Rational speculation is *perception* of this larger world of possibilities. This is an objective way viewing of the analytic sort of philosophy that predominates in English-speaking countries, an important element of which is so-called modal logic, which provides a formal framework for discussing necessity and possibility as such.[[167]](#footnote-167) But such abstract investigation also has important practical uses, for example in constitutional law. Supreme Court justices are just as concerned as philosophers about what follows from what.

*Incidental utility.* Even when practically useless in itself, speculative research sometimes produces tools that can be used on problems in the practical world. For example, recent theoretical developments in logic and semantics have found their way into the design of dishwashers and other appliances, as well as software for computers.[[168]](#footnote-168) Even the most seemingly useless speculation about seemingly unrealizable things, like complex arithmetic and non-Euclidean geometry, gets absorbed into science and engineering in surprising ways. Abstract metaphysics, epistemology, and moral philosophy also turn out to have many practical applications, from designing governments down to Jeremy Bentham's utilitarian design for prisons. As with pure mathematics, there is no way of telling when and where abstract work that is engaged in for its own sake will turn out to be useful in the practical sphere.

*Success against the odds.* There is always a chance, of course, though ordinarily a very small one, that your next first-order theory will actually resolve conflicts among experts. Your new theory may be so much better than any of the current theories that a new consensus emerges in support of it. This has happened to great thinkers like Darwin, Einstein, and numerous other inspired sorts of speculators. If you are bright enough and like to speculate and argue, then who knows? You might get lucky.

*Scientific progress.* Even if you produce no quick solution to outstanding problems, what counts as mere conjecture at one time can be established theory in another. Thus, the atomism of Lucretius and Democritus was nothing but speculation for many centuries, but its concepts were available to modern scientists like Ernest Rutherford, who used their ideas to develop the atomic theories that we depend on today. Similarly, Gregor Mendel’s work proved to be crucial in the development of modern genetics, though he got no credit for it in his lifetime as it languished on Darwin’s shelf. So, if you keep trying to work things out autonomously, you might make progress toward knowledge for others, tough without being able to know this yourself.[[169]](#footnote-169) I will argue in Chapter 6 that this ultimately altruistic use of speculative reasoning is the essential epistemic force in modern science.

*Intellectual insurance.*  Autonomous thinking also provides a kind of general counterforce to epistemic gravitation, which can mitigate its tendency to keep rational individuals from ever seeing over their authorities’ horizons or correcting their own ramified mistakes. If you attempt to figure out abstruse second-hand perceptions like the Catholic doctrine of transubstantiation for yourself, rather than simply taking them on faith, you are more likely to notice whatever logical problems they might have, to press authorities for explanations, and to reevaluate their expertise if they cannot provide them. When such speculative criticism is widely and openly practiced, whole communities can be to some extent protected from false doctrines fostered by the transient consensus of authorities, such as Social Darwinism. I will discuss this point at greater length in section 5.4.

*Moral responsibility.* Practical life requires us to make decisions in the absence of knowledge or even justified belief that we are doing the right thing. When we’ve been stranded in a capsized bus or gotten a new sofa stuck in the elbow of a narrow staircase, probably none of us actually knows what to do. But we are all responsible for chipping in ideas as best we can. There is no consensus of experts that tells modern people in complete detail how to run our businesses or raise our children. But we still have to do these things somehow and take responsibility for the results. Instead of guessing wildly or following randomly-chosen advice, responsible adults fill regions of uncertainty by figuring things out for themselves as carefully as possible, given the information they possess. In a democracy there’s rarely an expert consensus telling us how to vote; good citizens must do their best as individuals to make their share of crucial decisions as intelligently as they can. If we fail to do these things largely autonomously, we fail to pull our weight as adults and we become free riders on the epistemic efforts of our peers. And this is as shameful as our shirking any other kind of necessary work. I will return to these ethical matters in Chapters 7-9.

5.3. Opinions and arguments

New ideas arise all the time from speculation, but they do not proliferate unchecked. Instead, they are made to compete against each other for acceptance; some survive and some do not. A belief that is neither private nor imposed on others without their agreement, but propounded and defended in argument, is an *opinion*. Opinions ordinarily extend in the dimensions of perception and conviction too, of course. That is, we usually argue for ideas that articulate our inner models of the world (“call it as we see it”), and we usually act according to the ideas that we argue for (“put our money where our mouth is”). But having an opinion does not strictly require that we believe in either other sense; it just means that there is something we’re inclined to argue for. Where perceptions may be had unconsciously, it is hard to imagine how our opinions could be other than transparent to ourselves, since we are usually quite conscious of the points we are inclined to argue for. Thus, a psychoanalyst can correct you on the content of your own articulate perceptions (e.g. as to why you never call your mother), but he cannot reasonably tell you that your opinions are anything other than the ones you think you have. Indeed, simply to *state* an opinion consistently counts as conclusive evidence that you *have* that opinion. This is the basic reason that opinions are not normally framed in probabilistic terms: the purpose of stating opinions is not to express our fully articulate perceptions. The purpose is to contribute theories and arguments to a discussion, after which people can weigh the probabilities themselves and make any judgments that need to be made.[[170]](#footnote-170) Opinions as suchonly exist, then, in a context of disagreement, or at least potential disagreement.

Stating opinions, taking sides in arguments, and advocating for and against positions with an ultimate goal of resolution by consensus – this is the fundamental form of social reasoning. In order for this sort of reasoning to work, participants need to abandon or suppress many of their own first- and second-order reasons for seeing things as they do, and concentrate on the kinds of evidence and argument that can be shared with others. So, if I want to argue publicly that God exists, it won’t help much for me to talk about how reliable my Christian parents have always been in their statements to me. Their past reliability may give me personally a good reason to believe in God, but it won’t mean much to other rational people who don’t know my parents (especially if they have their own *atheist* parents who have always been reliable for *them*). But when we ignore private authorities and base our opinions on first-order evidence and arguments available to everyone concerned, then we can often make progress together over the long run as a group, even if we gain no immediate epistemic satisfaction individually. Of course, people do not always, or even usually, come up with their opinions in a totally autonomous way. In fact, most of the time we get our opinions from our peers and authorities within some sub-community that engages in argumentation wholesale, as it were, with other factions of a larger society: Catholics and Protestants, Republicans and Democrats, behaviorists and Freudians, etc., etc. Thus, people often state their opinions as representing the position of some group to which they belong: “As a Christian, I believe that abortion is sinful”; “As a feminist, I have to disagree” – that sort of thing. But the essential point of having opinions is simply that a number of different individual *ideas* should compete in arguments against each other, and in general the more the better.[[171]](#footnote-171)

Public debate tends to produce more accurate, articulate, and rationally justified theories than individuals, expert committees, or like-minded social groups can create in isolation, in something of the way that open competition among businesses tends to produce better cars and other goods than do monopolies, cartels, or central planners.[[172]](#footnote-172) Indeed, we could hardly survive without the process of debate in ordinary life. Even if we all lived in traditional communities, it would never be possible for us to follow authority on every question, since so much of what we must decide depends on facts and circumstances that could not have been anticipated in detail. Groups with special needs for solidarity make more of an effort to nail everything down. Thus, we find long lists of rules for dealing with such things as leprosy and cattle theft in the Old Testament, and for discipline aboard ship in the Royal Navy’s Articles of War. But no such effort covers every possible contingency, and complex rules will can also sometimes get in each other’s way, making it harder, not easier, for people to know what to do, as arguably happens with the US federal tax code. So we often need to figure things out in the absence of authority, and when we do, it usually helps to take advantage of multiple points of view.

For a simple example, suppose that I have driven my car into a ditch far from the usual sources of help, and I have only a certain, peculiar set of tools available to fix it and return it to the road. No traditional authority can determine completely what I ought to think about this problem. Much may depend on external conditions like the weather and the time of day, plus the exact details of the situation, such as whether there is too much mud in the ditch to hope for any traction. But one thing likely to help in a wide range of such problems would be the presence of several friends – not just to yank on things, but to help figure out what to do, since different heads will come up with different ideas, some of which might turn out to be right, or at least better than the others. One head can do this, too, but many heads tend to be faster at it, and to produce a more complete analysis and list of possible solutions. Ideally, one or more members of the group will be experienced and wise in solving problems like this one and take the lead. More elaborate divisions of epistemic labor are also possible: one person may be best at one aspect of the problem (or just get assigned to it), others at others. But even if no one is more specialized or useful than anyone else, it is good to have a variety of perspectives to compare. Suggestions typically arise from people’s individual first-order perceptions or guesses. I survey the situation, and I speculate that we could hook our belts together into a long rope, and use this to hoist the car out of the ditch by way of an overhanging branch. Similar things sometimes work in movies, and I watch a lot of movies, so this is the first thing that pops into my head. Other people have their own conjectures to contribute, occasioned by their different past experiences, habits of thought, and momentary states of mind. Still others may have only negative or critical remarks to make, for instance by pointing out that my idea of using belts won’t work – they won't be nearly strong enough to pull a car, for God’s sake – and we might as well keep our pants up while we try to think of something better.

Our suggestions may initially be piled up in a purely hypothetical spirit, without any sense of ownership. Each person present has some notion of his own comparative and absolute reliability, plus those of his peers. If these are seen as even very roughly equal, no reasonable person present will take his own perception of the situation automatically to be the best or ignore other ideas. But if the ensuing discussion takes more than a few minutes, such hypothetical suggestions may develop into more fixed personal opinions. The longer it takes for the group to come to a consensus, the more protracted the argument becomes over which idea is better or which should be tried first and the more *opinionated* (i.e. stubborn or aggressive in propounding their own views) people will become. Some opinions, like some people, are more stubborn than others. The least rigid ones are mere expressions of immediate perceptions, like, say, Stephen’s opinion, contrary to Deborah’s, that Suzanne is looking sad, not merely tired, at dinner this evening. The most stubbornly fixed opinions blend into convictions, like Suzanne's seemingly unshakable opinion, contrary to Stephen's, that Deborah desires to ruin her reputation. In between are most people’s opinions about most things about which they have opinions, such as my somewhat embarrassing opinion that Edward de Vere wrote Shakespeare’s plays. If pressed about this in a skeptical way (“Do you *really* *believe* that that guy wrote those plays?”) I’d just state my perception that according to the evidence I have, it seems more probable than not that Oxford was importantly involved in writing the plays. But as a practical matter, we can’t always express our opinions such hedged, conditional, or probable forms. We have to say some things outright and categorically in an actual argument or no one will listen to us. This is not a necessary thing; people can sometimes be quite philosophical about their own ideas in such a situation. But most of us like to be credited with being right, and it is hard for most of us to advocate any position for long without “getting into it” to some extent emotionally. As is often noted, criminal defense attorneys tend to convince themselves that their clients are innocent even when outsiders readily perceive the clients' guilt – and this makes them better advocates than otherwise. And the same goes for all sorts of situations where winning an argument really matters. But this conflation of opinion with perception sometimes leads us into trouble by convincing both ourselves and others that we assign our views a much higher probability of being true than we actually do, or than would be rationally justified.

The process of testing opinions with arguments does more than solve problems like a stuck car. It also helps us analyze problems and concepts that are not well understood. Much of the time, we are not even clear on what we mean ourselves by what we say. Understanding a statement involves much more than competence in the language in which the statement is made or the immediate context of other beliefs. To believe something that is at all controversial with full understanding, we need to be able to explain it to ourselves and, at least potentially, to others. And this means being able give arguments – ideally, all of the reasonable arguments – for all sides of the issue. If you do not know any of the reasons for believing that nuclear power is good or bad, then your own belief or disbelief in nuclear power is nothing more than a free-floating disposition to *say* that nuclear power is good or bad. If you want to state opinions about the economics of wages and employment, then you must know the reasons for your claims about such things as how employment levels are affected by changes in the national minimum wage, and you have to know at least the most important reasons that support opposing views. Otherwise, some of the essential content of your own opinions is missing. People who can't respond to objections with explanations *don't know what they're talking about*, as we sometimes say, and they are useless in discussion.

This is true whether the beliefs in question are controversial or not. John Stuart Mill approaches this dialectical account of meaning and belief in his extensive argument for freedom of thought and speech in the essay *On Liberty*.[[173]](#footnote-173) Mill claims that freedom to disagree with the established view on any topic, even the doctrines to which one's community is most deeply committed, ought to be absolute. Looking over the long human history of suppression of ideas, Mill argues that we never know when an opposing view might turn out later on to be correct, or at least partly correct. In the past, the horror and disgust that seemingly wicked opinions have aroused has not been a reliable indicator of their falsity, and we usually wish that they had never been suppressed. So it is probably foolish to suppress similarly awful-seeming opinions even now while they are horrifying and disgusting *us*. Mill claims that even if the truth were known with perfect certainty, so that we knew that the dissenting opinions *must* be false, received doctrines would still need to be challenged freely if we wanted them to be fully understood. For, if completely understanding a proposition entails knowing *all* the reasons for believing it, it will not be enough to have established merely that the proposition is true, or to have established a few positive arguments for that belief. We'll also need to keep inventing and investigating all the dissenting opinions we can think of that have any plausibility at all. The alternative is to allow a doctrine to be passed on without proper structural support, progressively weakened over time by merely dogmatic repetition, and leaving it unclear to any future doubters why things couldn't be some other way. As I argued in Chapter 4, such an unsupported belief can last for centuries if made effectively unanimous in an isolated or otherwise effectively closed epistemic community. But belief based solely on authority for each new adherent is only belief in the truth of a sentence, not in a fully digested perception. As Mill says, expressing such beliefs becomes an empty ritual, like children reciting prayers by rote that they have learned phonetically in Latin. Even prudential or moral beliefs can be corrupted in this way, subsisting merely as statements of rules we think we ought to follow, rather than properly integrated features of our character. This is the sort of moral belief available to very little children: lists of behaviors that are counted as naughty or nice. And that’s okay for children as a stage in moral education. But adults are not adults if they do not know what they’re doing in the sense of understanding at least something of the substantive basis of their beliefs. This is the final practical justification for thinking independently as an alternative to pure perceptive rationality. We don't believe things just for the sake of making bets on what is most probably true according to our total evidence at any moment. We also deeply want to understand the world, our own experience, and our beliefs. And when it comes to acting under pressure, integral, understood beliefs tend to work better than rote prescriptions. This is why it is often good for you to work out and defend your own first-order theories *Z*, even when your epistemic betters have already taken *X* and *Y*: because a little wisdom can sometimes be more useful than a lot of piety.[[174]](#footnote-174) The only caveat – though it’s a huge one, worth repeating – is that you shouldn’t confuse your opinions with perceptions and conclude that you are more likely to be right than everybody else.

5.4 Socratic and Cartesian methods in philosophy

Competition among opinions is the definitive activity of the Socratic method in philosophy. Plato’s dialogues are all framed as conversations between his mentor Socrates and other thoughtful people, and they are in that format for a reason. Socrates himself never wrote anything down, and famously claimed that he knew nothing himself except, paradoxically, the fact that he didn’t know anything. But he did confess to having a special talent for directing others toward knowledge and compared himself to an old midwife, unable to bear children but useful in helping others to give birth. To know what something is, for Socrates and Plato, means knowing its *real definition*: what it *must* be like in order to be the kind of thing it is, as a triangle must have exactly three sides. In the dialogues, Socrates’ technique is to goad some acquaintance, usually either a friend or a fellow philosopher, into stating an opinion as to the nature of knowledge or justice or some other important concept. Plato then has Socrates present a *counterexample* to the proffered definition, i.e. either a case that fits the proposed definition but intuitively shouldn’t, or one that doesn’t fit but should. Since real definitions have to be true in every possible case, one good counterexample is always enough to disprove any suggested definition. So, Socrates having dispensed with one opinion, the conversation moves on to other suggested definitions which are examined in turn to see if they apply in every case. Once a general agreement is achieved, that is the end of the discussion. But this doesn’t usually happen, because it turns out that defining things – not just what in a dictionary, but what’s true about them necessarily – is much harder than it usually looks.

For example, in the first book of Plato’s *Republic* we find Socrates confronting his wealthy older friend Cephalus:

Socrates: So, Cephalus, tell me what justice is.

Cephalus: Okay. I guess I'd say that justice is honesty. The just person is someone who always tells the truth and pays back what he owes.

Socrates: But would you call it justice if a man returned a borrowed axe to someone who had just gone violently insane?

Cephalus: No, I don’t suppose I would.

Socrates: Well, there you go, then. You can’t say that justice is the same thing as telling the truth and paying what you owe.[[175]](#footnote-175)

After another effort or two by others at defining justice quickly, countered in turn by Socrates, the professional Sophist Thrasymachus enters the argument to offer his opinion that justice is only the self-interest of the most powerful people, since they are the ones who make the laws, and laws define what counts as justice.[[176]](#footnote-176) Thrasymachus tries very hard to win the argument that follows, having promised to charge Socrates a fee for the lesson he is being taught, but then Socrates blocks and parries every point he offers with the deftness of champion fencer.[[177]](#footnote-177) Like most readers of the *Republic*, I find this passage strikingly hard to follow (full of odd subtleties and plays on words as well as what appear to be acceptable technical arguments), indeed uniquely so within a work of generally sterling clarity. I think the point Plato is making is that winning verbal swordfights and impressing people is irrelevant to serious argumentation. Socrates proves that can do that sort of thing himself whenever he feels like it, but he never really feels like it because he only wants the truth, or at least the best articulated ignorance that he can find. By the end of *Book I*, Thrasymachus has been reduced to making bitter, sarcastic concessions to Socrates, and he is rarely heard from again. For the rest of the *Republic*,Socrates argues instead with Plato's thoughtful and cooperative brothers Glaucon and Adeimantus, in an effort to develop systematically his own (or Plato’s) theory of justice in the state as a kind of harmony among the social classes corresponding to a healthy balance of psychic forces in the just individual.

This Socratic, *dialectical* approach to disagreement is the basic theoretical alternative to Descartes's individual-perfectionist approach. The Cartesian method, remember, assumes that in cases of expert disagreement, individuals should start from scratch with clear and distinct foundations, then gradually add layer upon layer of careful inferences while avoiding future disagreements by not making any mistakes. Socrates is wiser, I think, in taking as given not some purportedly foundational position, but the commonplace observations that we already disagree and that mistakes by individuals are inevitable. The proper starting point is what we believe right now in all its diversity, and the Socratic project aims to reduce our disagreements by an essentially social process, unlike the Cartesian project which is at least initially a solo operation. On the Socratic project, for as long as reasonable people disagree about an issue, we keep working together to resolve our differences; if and when we find agreement, we move on. This method of philosophy tends to move very, very slowly. It requires great patience and mutual openness to challenge, even where the truth seems to most people to be well established in advance. And the ultimate result will be not just a final theory resting on a pyramid of justification, but an intricate roadmap of arguments connecting every initially plausible idea to every other, with the truth as Rome. Anybody who seeks real understanding will benefit from having all the possibilities laid out and all relevant arguments laid out in detail, even if most of the roads lead to dead-ends.

Even Descartes understood that philosophy was hard, at least as hard as mathematics, that even a genius like himself could only do so much, and that perfection in reasoning is a lot to ask even for limited projects. He knew that many philosophers would follow him and improve on his work, even if he got a lot of things right, so in a way his work was only one person’s opinion, one step in a larger dialectical progression, and he says so forthrightly in the *Discourse on Method*:

My present aim, then, is not to teach the method which everyone must follow in order to direct his reason correctly, but only to reveal how I have tried to direct my own. One who presumes to give precepts must think himself more skilful than those to whom he gives them; and if he makes the slightest mistake, he may be blamed. But I am presenting this work only as a history, or, if you prefer, a fable in which, among certain examples worthy of imitation, you will perhaps find many others that it would be right not to follow; and so I hope it will be useful for some without being harmful to any…[[178]](#footnote-178)

Some people think that Descartes writes about his purposes in this way just to avoid censure from the Church for possibly skirting doctrine, and that he really thinks that everybody should agree with him entirely.[[179]](#footnote-179) But it seems to me that he is probably quite sincere in saying that this method seems to work for him, and others can follow his example if they like and see whether it works for them, but that he is not in any position to guarantee it will. But whether this is Descartes’s view or not, I think it is the most reasonable view to hold.[[180]](#footnote-180)

On the other side of the methodological disagreement, Plato knows too that simple conversations can’t reveal that much of philosophical depth. Full treatments of important issues, even the definitions of important terms, take more than a few sentences to express. This is how it seems to me whenever I get into an argument about politics these days: somebody says something, I disagree, he responds with six arguments at once, I pick one of them and offer a counter-argument while trying to remember all the other five, he responds with another half-dozen rapid points, I pick one of them that seems ambiguous and ask for a clarification, he responds with something that strikes me as generating three different possible problems, one of which I point out to him while struggling to maintain a growing mental tree-chart of all the arguments that have had to be postponed. I’d really like to go down all these branches calmly together one at a time, however long it takes, but the bars around here have to close by 2 am, and even if they never closed, most of my friends find protracted dialectics rather less rewarding than I do, and tend to lose interest after blowing off some steam. So, I go home and try to lay all the arguments out for myself, searching for the deepest sources of our disagreement – but this is hard, tedious work even when you’re sober.

It seems, then, that the Socratic sort of dialogue, like real-life public debate, is fine for presenting quick counterexamples to misleading commonplace opinions and for showing how hard it is to understand the things we take for granted, but it has limited power to represent whole competing theories in a useful way unless one person gets to be the *boss* of the discussion and can tell everybody else to calm down and stop interrupting. The *Republic* itself is a nice example of this, since it starts out as a pretty plausible discussion among superficial peers, with Socrates batting away the easiest popular notions about justice, and then gradually turns into a lecture:

Socrates: So, Glaucon, don’t you agree that *blah, blah, blah, blah, blah*?

Glaucon: Yes, Socrates, indeed I do.

In the same way, we expect most serious, large-scale theoretical work in philosophy to take the form of treatises like Aristotle’s *Metaphysics* or Descartes’ *Meditations* rather than dialogues, for the reason that there’s simply far more that needs to be said in a connected way on any major topic, all things considered, than could fit into any normal-looking conversation.

Ultimately, then, the best overall method in philosophy is a combination of the Socratic and Cartesian methods. Each working philosopher finds out what other people have been arguing so far about some topic, works out his own analysis in light of all that has been said so far, composes his monograph with careful attention to precedent and possible objections, adds it to the every-growing pile of scholarly literature if anyone will publish it, and hopes that the next person working on the topic takes account of what he’s said. This is the way things *have* to be with thousands of people working all at once on complex issues while trying to get everyone else’s attention, and it generates enormous amounts of very useful discussion. But conducting arguments with articles and books rather than slogans and quick examples is a slow, slow business. In the full-scale Western discussion of justice, for example, Plato’s theory in the *Republic* counts only as one person’s expressed opinion. That opinion is disputed at length by Aristotle in the *Nichomachean Ethics* and the *Politics*, and this disagreement generates a multitude of further responses down the centuries, from Augustine’s *City of God* to Hobbes’s *Leviathan*, Locke’s *Two Treatises of Government*, Marx’s *Capital*, Rawls's *Theory of Justice*, and all the other classics and non-classics of political philosophy. All of these works take up the arguments and themes that occupied the ancient Greeks, while adding more of their own to an expanding network of argumentation on the nature of justice, much of it necessarily abstruse. In this way, the whole of philosophy has something of the same dialectical structure as an individual Socratic dialogue, which suggests to optimists an ultimate convergence on the truth despite what seems to be a mushrooming cloud of theories and opinions.[[181]](#footnote-181) The particular issue of the nature of justice has not yet been resolved, of course, and 2400 years of effort in the West seems to some people more than ample to have found any objective answer that exists. In fact, philosophers hear all the time (and some of us say it ourselves) that the project of philosophy as we traditionally conceive it is obviously doomed to failure, and we should either turn it all over to one or another branch of science or just give up and go home.[[182]](#footnote-182) In my own opinion, it is really very difficult to analyze a concept like justice that evolved among complex, diversely cultural beings like ourselves to every reasonable person’s satisfaction, and twenty-four centuries is not that long a time for a discussion in which each full statement takes a person years or decades to construct.

**6. SCIENCE**

In questions of science, the authority of a thousand is not worth the humble reasoning of a single individual.

– Galileo Galilei, quoted in Arago, *Eulogy of* *Galileo* (1874)

6.1. Rational repression.

I have been arguing that free speculation and Socratic philosophy are good things, leading to benefits for people and societies interested in intellectual progress, or even in the proper maintenance of traditional beliefs. Why, then, is philosophy of the Western sort, based on personal opinions and open arguments on every topic, so rare in history? The vast majority of human societies have been traditional religious ones. As I argued in Chapter 4, this state of affairs is very stable epistemically for dominant or isolated cultures, and even for minority sub-cultures in mixed societies, provided they include the right kinds of epistemic doctrines together with their core beliefs. But first-order speculation and debate are also normal parts of epistemic life, as I argued in Chapter 5. So, for most people, there is an area of mostly moral and religious matters in which tradition and authority hold sway, and another area of mostly practical matters that are left to unhindered argument. These areas are vaguely defined and tend to overlap to some extent, but this is not a problem as long as the traditional authorities make themselves clear enough on what counts as a matter of settled doctrine and what remains a matter of opinion. Moreover, as I have argued, even in doctrinal matters there is no necessary problem, and indeed there can be much advantage, in allowing free discussion of a hypothetical sort. As Mill says, this is needed to keep doctrines clear, plausible, and even meaningful over the long run. So, there is no need in principle for rationality to conflict with freedom of opinion properly understood.

Still, a certain probabilistic tension tends to arise between faith in authority on one hand, and independent thinking on the other. For speculative reasoning, however Socratic in spirit, can still give people powerful evidence about the relative coherence of different models of the world. My tradition and authorities have told me one thing, but first-order reasoning sometimes tells me something else that strikes me as making more sense overall. This doesn't mean that I ought to believe either position categorically. In the ordinary case, where I am surrounded by peers and authorities all holding the traditional view, rationality demands that I continue to place more credence in what I have been told despite the greater plausibility (to me) of my own first-order ideas. Still, whenever I discover arguments or evidence that seems to support my alternative, the probability that I ascribe to the traditional belief must decrease to some extent. Unless the epistemic doctrines of my community do enough to block contrary evidence and reasoning from counting against the traditional theory, then, even the most hypothetical reasoning can diminish the apparent reliability of my authorities, potentially undermining the core doctrines of my tradition. This can make it easier for individuals or sub-communities to break ranks, even when the balance of probabilities still favors the traditional beliefs.

It is a main theme of this book that people are much more rational in general than we usually give each other credit for. But we are not perfectly rational. Sometimes the vividness and seeming coherence of our own speculations is too much for us to bear in a completely rational way, and instead of treating them with caution as hypotheses we convince ourselves of them as categorical beliefs, ignoring all the second-order evidence we have against them. If we go further and assertthese new beliefs as facts, this kind of *arrogance* presents a problem for our peers and betters. What should be done with people who irrationally claim to know things that they cannot justifiedly believe, and do so in defiance of authoritative teachings? Sadly, some degree of intellectual repression seems to make sense in situations like this within traditional communities. This is not fundamentally because dissenting views present a threat to those in charge politically (though they will often do that, too), but because, given the evidential weight of testimony in favor of the standing doctrine, categorical dissents are typically irrational. When dissenters are just being arrogant,then, rather than reasonable, it is hard to see why people in an epistemically united community should have to listen to them, especially in legislatures, schools, and other public forums where everyone agrees that reason ought to rule. Even in a modern, generally anti-traditional society like ours, such repression may well seem appropriate on topics where unanimity is nearly total. For example, people who now allege that the Holocaust never took place are, and arguably should be, denied public support for the dissemination of their theory, not simply because the theory is false, or because it is nasty, but because such a belief can only be held irrationally in light of all the first- and second-order evidence that is available to everyone concerned.[[183]](#footnote-183) People who publicly dissent from such well-grounded beliefs as that the Holocaust took place are not just attacking the specific doctrines over which they disagree, but also the value of the testimony of their peers and the legitimacy of their authorities. So, whatever struggles over doctrine arise from this kind of arrogance are liable to turn into political fights as well, albeit ones in which traditional authority retains the rational upper hand.

Think about what happens to dissenters on questions of constitutional interpretation in the United States. There is no necessary conflict between our system's fidelity to the Constitution and the liveliest and farthest-reaching debate on legal principles that we can imagine, provided the Supreme Court that interprets the Constitution and the popular authorities that execute their rulings do their jobs. In fact, as most Americans see it, Constitutional integrity goes hand in hand with free debate – a principle we find embodied in the First Amendment. So, lower federal court judges are formally permitted to make rulings on whatever grounds they see fit initially, but they always do so within a web of expectations. They are expected to rule consistently with the Constitution, with precedent where the Constitution is unclear, and with established principles in applying precedent to novel cases. They are also expected to obey the higher courts when those courts overturn their rulings on appeal.[[184]](#footnote-184) So, if a lower court judge disagrees with standing Constitutional interpretation, he is permitted to make positive arguments for an alternative theory in his initial ruling. But once he has been contradicted by a higher court, he must bow humbly to their authority or be punished in some way.[[185]](#footnote-185) This does not make lower court judges into cowards, or higher ones into oppressors. These are just different roles within an institution that takes certain principles for granted, including principles of justice that are embodied in an authoritative text that has been authoritatively interpreted by authoritative experts according to authoritative procedures. If you want to change the Constitution, well, you know there are procedures for that too, and it has proven to be quite adaptable over the two and a quarter centuries since its establishment. But once you have exhausted these procedures and your view has been rejected, you must either do as you are told or face the consequences of refusing.

We all accept authority in academic life as well, even for tenured faculty who bathe in principles of academic freedom. For example, suppose that I teach logic for a living, and I like to mess around with the rules in my spare time because the standard system seems inadequate to me in some ways. Now, the fundamentals of modern symbolic logic have been pretty well set for over 100 years, and essentially the same material is taught in first-year logic classes all over the English-speaking world. But suppose that I have come up with a new system of logic that strikes me as better.[[186]](#footnote-186) What should I do as a philosopher? Certainly, write an article about my new ideas, get it published if I can, and see what happens in any subsequent debate. But what should I do as a teacher of logic? According to the view of things that I have laid out so far, I should do little if anything differently in teaching freshman courses. I may be an expert on logic, but I am not the only one; in fact, I am vastly outnumbered by people who are reasonably happy with the standard system. This fact should be enough to deter me from teaching students my new system in place of the old one. My new logic may turn out to be better objectively, but I am in no position to know this now, or even to be rationally justified in believing it. What I ought to do, then, is to teach the regular system in introductory classes, with perhaps a few remarks on my alternative if there is time to spare, while saving my speculative work for seminars with upper-level students. If I insist instead on teaching my new system dogmatically to freshmen as the simple truth, then my colleagues will rightly remonstrate with me and ultimately, if I refuse to budge, ask the dean to remove me from teaching any courses where I don't present the mainstream theories to my students as the mainstream theories. Academic freedom requires that my colleagues and superiors tolerate my oddball opinions about logic too, but only when I present them asopinions, not as authoritative doctrines.

The boundary line between the thoughtful speculation of philosophers and the irrational defiance of rebels is often hard to draw clearly in advance of actual conflicts with authority. Rather ironically, it is Socrates who provides the classic, much disputed case of a philosopher who seems to go too far. Athens had recently lost the Peloponnesian War to Sparta and the Spartan victors had deposed its democratic government, putting in its place an oligarchy known as the Thirty Tyrants. But the democrats took power again within a year and prosecuted some prominent supporters of the oligarchy, including Socrates. He was officially charged with corrupting the youth of Athens and disbelieving in the city’s gods, but there is no doubt that his opposition to democracy played a major part.[[187]](#footnote-187) Socrates defended himself largely by insisting that he took no unauthorized positions himself on matters of community doctrine, but merely acted as a “gadfly” to stimulate the thinking of other citizens by engaging them in philosophical discussions.[[188]](#footnote-188) In the same speech, though, he admits to at least one substantive teaching, namely that his youthful, generally upper-class followers should stop caring about things like money and physical health, and focus on improving their souls instead. He also claims that he is on a special mission from “the god” (i.e. Apollo, via the oracle at Delphi) to perform this function for the city. And at innumerable points elsewhere in Plato’s dialogs, Socrates harshly criticizes common beliefs, including the traditional belief in human-like gods (Zeus, Hera, Aphrodite, etc.) who are preoccupied with petty feuds and jealousies.[[189]](#footnote-189) Moreover, during the trial itself, Socrates behaves in a way that strikes his judges as arrogant and supercilious despite his protestations of humble piety and loyal citizenship. Most brazenly, once he’s convicted of the charges against him, Socrates insolently proposes that he be punished with free meals for the rest of his life (a privilege granted only to the greatest heroes), prompting the jury to sentence him to death by a greater margin than had found him guilty in the first place.[[190]](#footnote-190)

In sum, it seems that Socrates is arguably guilty of the charges laid against him, and tries to get himself off on the technicality that he does not dogmatically “teach” dissenting theories, but merely elicits them from others. In itself, this plausibly amounts to a corruption of youth inasmuch as his questioning makes them less unthinkingly patriotic Athenians, so that even if he’s technically innocent of the one charge, he is all the more guilty of the other. So, whether he’s an actual functioning heretic or not, he is certainly a sower of doubt. To the extent that this doubt constitutes a serious challenge to traditional Athenian beliefs, the jury at his trial seems to be rationally justified in convicting him, if not in executing him.[[191]](#footnote-191) On the other hand, in the following dialogue the *Crito*, whichtakes place in Socrates’ jail cell while he awaits his execution, he flatly refuses an offer from his wealthy friends to secure him an easy escape into comfortable exile. And he explains this to his friend Crito as a product of the humble obedience and loyalty he owes to Athens (like a child’s natural submission to his parents, he says) regardless of whether its judgment was right or wrong in his particular case. This certainly undercuts the idea of Socrates as an irrational religious or political rebel. In his own mind, at least, it seems that he really was just trying to be helpful (and ultimately deferential to authority) the whole time.

6.2. Epistemic altruism.

Irrational people can sometimes be useful. We find episodes of conflict between independent thinkers and authorities throughout Western history after Socrates, some of which turned out very well for the communities over the long run, if not for the individuals in question.

The case most important to the founding of modern science is the famous struggle between Galileo and the Roman Catholic Church during the early sixteenth century. In most superficial accounts of this dispute, the Church gets depicted as an irrational, intolerant, repressive villain, and Galileo as a noble martyr to reason.[[192]](#footnote-192) But in a plausible alternative view, the Vatican's intellectual position was actually much stronger, and Galileo's weaker, than the common story says.[[193]](#footnote-193)

Here is a little background. Since the founding of universities in Western Europe by the Church during the High Middle Ages, natural philosophy (i.e. pre-modern science) was taught exclusively according to the reigning paradigm, Scholasticism. Its core theory was a combination of traditional Christian doctrine, evolving Church teachings on the nature of God and the universe, and classical philosophy, especially the works of Aristotle, plus medieval commentaries on those works by Jewish, Christian, and Islamic scholars.[[194]](#footnote-194) Some of this theory was propounded dogmatically, subject to only narrow interpretation by Vatican authorities, but most of it was open to broader discussion and amendment by the wider community of scholars. At the universities, those open issues were debated in a format called the *disputatio*, with students at different levels arguing back and forth extemporaneously on one day under a professor's supervision, and the professor formally presenting his analysis of all the arguments a day or two later.[[195]](#footnote-195) During the Renaissance especially, the Church took a relaxed approach to scientific speculations that were contrary to doctrine – some high churchmen even involved themselves in such research – as long as it was presented hypothetically, as theories about the way things *seem* rather than how they are. Thus, though Nicolaus Copernicus delayed its publication out of worry about getting in trouble with the Church authorities, there was relatively little fuss when Copernicus published *On the Revolutions of the Heavenly Spheres* in 1543. This book, which offered a sun-centered theory of the solar system that opposed the earth-centered model favored by the Church (on both Biblical and philosophical grounds) was not officially proscribed until 1616, on the eve of the devastating Thirty Years' War between Catholic and Protestant forces all over Europe. In the meantime it was tolerated, though frequently criticized, by Church authorities as one first-order model of how things in the heavens look. The Church did not object to any theories whose purpose was to “save the appearances” (as they put it at the time), as long as they were not presented as fact and posed no clear danger to the institutions of the Church itself. Catholic authorities interpreted scripture in a way that was a far cry from the fanatical literalism we associate with Biblical creationists today. Their principle was rather to interpret the Bible as literally as possible *consistently with everything else that had been proven to be true*. This is essentially the same approach that so-called strict constructionists take to interpreting the US Constitution: stay as close as possible to the plain meaning of its authors until forced by novel circumstances to do something else. As we find nowadays in constitutional disputes, it was often unclear within the Church what counted as sufficient proof against a literal interpretation of the sacred texts. But this question was one for Church authorities, not outside challengers, to answer, just as our Supreme Court must decide whether or not the most literal available interpretation of the First Amendment is adequate to cover new cases of internet solicitation and pornography, downloading copyrighted information, hacking government secrets, and the like. Technical experts of all sorts must bow to the courts on questions of the law regarding internet privacy, health care policy, nuclear power, etc., and so it was for natural philosophers and the cardinals of the Inquisition regarding questions of Christian doctrine.

Giordano Bruno, another great cosmological speculator of the sixteenth century, was much more arrogant and reckless than Copernicus, taking his own first-order reasoning for granted as the final arbiter of his beliefs and sometimes treating Church authority with open contempt. He was repeatedly cautioned, then reprimanded, tried and convicted twice, urged to repent, and finally condemned to death by burning at the stake in the year 1600. Bruno was executed not because he came up with non-standard theories about the universe – that sort of thing was much encouraged by the Church in the progressive intellectual environment of the late Renaissance. Bruno was executed for insisting that his speculations were superior to settled doctrine; hence, implicitly, that his own expertise was greater than that of Church authorities. If this had been a rational position for him to take, he may have been treated differently. But it was not. Bruno had nothing like sufficient evidence for believing that his views were superior, all things considered, to those of the Church’s experts. So, when Bruno refused to recant all of his heresies upon being ordered to do so repeatedly by the Supreme Court of his day, they understandably felt they had no reasonable choice but to condemn him.[[196]](#footnote-196)

Galileo’s case falls somewhere in-between those of Copernicus and Bruno. Strongly promoted by the Medicis, a past professor at the University of Pisa, and a renowned astronomer much praised by his powerful friends within the Church, Galileo was both more prominent than Copernicus and more critical of the astronomical doctrines of the Church, but never as openly defiant as Bruno, preferring instead to grumble loudly but privately about backward Scholastics getting in the way of his new empirical philosophy, which he felt should be freed entirely from doctrinal constraints. The Church had its Bible and centuries of expert interpretation of the written word of God to anchor its authority, said Galileo, but the natural world also expressed the mind of God, and that it was natural philosophers, not cardinals, who should be treated as the experts in this second, co-equal sphere of learning, given their mastery of powerful tools like mathematics and the telescope. He thought that experts in the Biblical sphere had no business imposing their own interpretations on his expert investigations in the natural sphere, and grew more and more sarcastic and obstreperous in saying so. The Church had had no problem with Galileo's well-known, privately argued Copernicanism up to a point, but his increasingly noisy and self-righteous stance against official doctrine caused greater and greater concern, plus some resentment for the ungrateful attitude of a celebrity toward his old friends and patrons. The final straw for the authorities was Galileo’s *Dialogue Concerning the Two Chief World Systems: Ptolemaic and Copernican*, in which the Church’s traditional geocentric theory was satirically represented by a fictional doofus named Simplicius.[[197]](#footnote-197) Galileo’s erstwhile close friend Cardinal Barberini, now Pope Urban VIII, had asked him in good faith to compose an even-handed exposition of the Copernican controversy, and felt betrayed and mocked by Galileo’s use of an obvious simpleton as spokesman for his positions. Galileo was arrested, tried, and condemned in 1633 and, mindful of what had happened to Bruno, he recanted his heretical opinions and accepted the mild punishment of house arrest for the remainder of his life.

In my view, Galileo was never rationally justified in taking his own opinions for facts or even probable facts, because he never had sufficient reason to insist on his own views being true while the received views of the Church were false. It was arrogant for Galileo to believe that he could better judge the total situation, encompassing both science and religion, than could the Pope and cardinals. These friends-turned-enemies were also highly educated people, scientifically sophisticated if not working scientists themselves, who had concluded on the basis of all the evidence available to them that a proper interpretation of the scriptures ruled out certain of his astronomical hypotheses pending stronger proof, while freely admitting that they seemed pretty plausible. Galileo certainly knew that he was a great natural philosopher; nobody doubted this obvious fact. But unless he believed that his perceptions were privileged in some general way beyond those of the Pope and cardinals, there was never sufficient reason for him to prefer his own authority, as it were, to theirs on a topic that interested them both.[[198]](#footnote-198)

Whether he to be condemned by the Church or not, Galileo's own epistemic position was at best equivocal in 1633. Though he definitely knew what he was talking about in the usual sense of that phrase, I don't see how he could have justifiedly believed that he was right in contradicting the position of the Church authorities. They had never really disagreed with Galileo that the solar system looked the way he said it did, and that the Copernican theory made the simplest sense of these appearances. The question was whether his empirical results constituted sufficiently definite proof of the falsity of the most literal available interpretation of the Bible, namely the traditional geocentric model of the solar system. And I don't see how they could have, since Galileo had ignored Johannes Kepler's discovery of the elliptical (not circular) orbits of the planets, without which the Copernican model did not actually perform any better than a commonly modified version of the Ptolemaic model when it came to predicting the motions of the planets. It was the intrinsic elegance of the heliocentric view more than any direct connection with the data that made it so appealing to Galileo and other Copernicans in ignorance of Kepler's laws of planetary motion. And that was not enough, under the reasonable interpretive guidelines that the Church applied to Biblical texts, to prove the standing literal interpretation false. Philosophers and scientists now say that elegance counts for a lot in physical science, that the simpler and more comprehensive theory often makes better predictions than clumsier-seeming alternatives.[[199]](#footnote-199) But we are in a much better position to realize this, after absorbing Galileo's, Kepler's, Newton's, and Einstein's theoretical discoveries among many others, than Galileo was himself in 1633. And it doesn’t seem to me that he had enough else to go on in asserting either that his theory was as solidly established as the Church’s Ptolemaic view or that his own authority was just as good as theirs in matters of common interest. Far from being a martyr to reason, then, as he is usually portrayed, Galileo seems to have been a man whose total epistemic situation made it *rational* for him to *recant* when he was ordered to do so.[[200]](#footnote-200)

Surely it matters, though, that regardless of all I have said Galileo was *right*. He was right that the Copernican model of the solar system made more sense of the available evidence than the official geocentric model. He was right that mastery of mathematics and technology makes scientists in general better judges of empirical theories than people who are only experts at interpreting texts and traditions. He was also right in his belief that he, Galileo, had a more reliable view of the overall conflict than did the people who condemned him. And it was not really an accident that Galileo was right about these things. He was a brilliantly far-sighted thinker, able to anticipate the great future of Western science in a way that few of his contemporaries and none of his judges could. We in the modern West now all believe, at least superficially, in the autonomy of science that he favored, for the reasons that he favored it. We sympathize with his frustration at having non-scientists looking over his shoulder and telling him what he could and couldn't say straight-out in public. And we feel outrage that such a genius should be shown the implements of torture by Inquisitors to force him to renounce his honest convictions and prostrate himself before a panel of inferior minds. So we take Galileo’s side against the Church. We are happy that he stood up to the Church for as long as he did. We owe him an enormous debt of gratitude for his great contributions to science and scientific method. Even the Church itself has recently recanted its condemnation of Galileo, long after having accepted the correctness of the Copernican theory and revoked its ban on Galileo's *Dialogue*.[[201]](#footnote-201)

So, what, then, are we to make of Galileo? He was an arrogant figure, clearly, but also brilliantly perceptive in matters of substance, which is what matters most to consumers of science like ourselves.We are glad that he was arrogant, since we connect it with his creativity. We don't just forgive him for this vice; we are grateful to him for it, and we even celebrate him for it as a hero. Here was a man who thought for himself regardless of standing authority, and demanded the same autonomy for science as a whole. Here was a man who saw through the prejudices of his day, and explored the accessible universe with bold insight and solid reasoning. If it takes a little arrogance to make discoveries like this, then maybe arrogance turns out to be a virtue, not a vice. If this kind of intellectual behavior is irrational, then so much the worse for rationality. In order to make bold strides in science, it often takes resistance to tradition and authority beyond what reason licenses to individuals. Revolutionary, paradigm-shifting scientific reasoning thus involves a kind of *epistemic altruism* on the part of individual scientists like Galileo, intentional or not. By refusing to defer to others when it is subjectively rational to do so, they sacrifice the greatest likelihood of having true beliefs themselves, while helping other people to make progress over the long run.[[202]](#footnote-202) Galileo sacrificed the relative probability of his *being* right, given his total first- and second-order evidence, in order to develop a perceptive model that *seemed* right to him. Like most arrogant thinkers, he didn't mind this sacrifice because he didn’t notice it. He thought that his ideas were true and the Church's were false, and seems not to have been bothered much by the conditionalities and probabilities involved. This may seem reasonable enough in Galileo's case, given the way that things turned out. But compare his situation to that of the brilliant pantheist Bruno, who believed with equal sincerity and equal arrogance that the universe was infinite, with infinitely many inhabited worlds, and got himself burned at the stake when he refused to back down. Why did he believe this stuff? Because it seemed true to him on a first-order basis, just as the heliocentric model of the universe seemed true to Galileo. But Bruno did not turn out to be right in the way the Galileo was. Even the great Kepler and Newton held beliefs that we would now call *crazy* in addition to those insights that proved successful: Kepler spent much of his career attempting to create a modern science of astrology,[[203]](#footnote-203) and it has lately been revealed that Newton was a secret and enthusiastic alchemist.[[204]](#footnote-204) For the general case, imagine any randomly chosen brilliant scientist devoting his whole career to the development of a bold new theory that he convinces himself is categorically true despite the disagreement of his peers. Is such a theory likely to be true? No. Most revolutionary thinkers turn out to be dead wrong, and unless they kill somebody we don't hear about them.[[205]](#footnote-205) If his goal is merely to increase his personal stock of justified beliefs, the average brilliant scientist is probably just wasting his time. But there are many creative scientists working at once, on many hypotheses. If *any* of their new opinions proves to be correct, then the ultimate consumers of the scientific product benefit, albeit at the expense of all the thinkers who persuade themselves of novel theories that turn out to be false. So, the fact that radical breaks with authority are ordinarily irrational for individuals like Galileo makes it very hard for progressive thinking to emerge inside of a traditional society. The fact that empirical science *works* overthe long run, that it leads to more true beliefs and provides for better technology, weapons, and the like, makes it hard to stop once it achieves a threshold level of usefulness. It spreads, not because it makes the most sense rationally for the individuals who practice it, but because the societies in which it has taken hold tend to be more successful than those in which it has not.[[206]](#footnote-206)

In an ideal world, everybody would be reasonable all the time. Authority would always be tolerant and critics always humble, and there would always be time for every argument to be presented calmly and digested thoroughly. But in the real world it is not always enough to have an idea, write it down, and submit it meekly to the judgment of others. Your new idea is the most perspicuous to you; you understand it better implicitly than your peers and authorities are likely to understand even your best articulation of it; and you want them to take it as seriously as you do and give it a completely fair evaluation. But this means gaining and holding their attention for as long as it takes to make your full argument and respond to all of their questions and objections, and this is not easy to do in practice. People in epistemic authority are usually very busy and need to take shortcuts, depending on each other for quick reactions to new work that they have no time and often no ability to analyze completely by themselves. Even for a critic as prominent as Galileo, someone whose work is certainly going to be read with care: it is hard for such a person to believe that those who reject his (*obviously!*) true ideas actually get the picture. Ideally, with perfect perceptive rationality, the critic ought to concede that he is probably wrong, assuming that his arguments have really been understood by people who are really epistemic peers or betters. But it is hard to know for sure that those assumptions are true, and mighty frustrating to be told that you are wrong by people who don't seem fully to grasp what you are telling them. So the genius (or, more commonly, the fool) takes shortcuts of his own, proclaiming arrogantly what ought to have been humbly suggested. This brings about a *fight* in place an unequal discussion with busy authorities – and people pay attention to a fight. Most revolutionaries lose such fights, of course, and we forget them once we tire of mocking their presumption. But occasionally someone wins one of these struggles against authority, as Galileo did in the eyes of posterity, and this has earned him a heroic standing in science much like that of Socrates in philosophy. He is a hero for betting on himself and winning despite the odds – that's what we seem to admire about him as a person – and for adding something valuable to the common stock of justified beliefs, which is what we all get out of it. So, we praise successful scientific fighters like Galileo not for their rationality but for their creativity and gumption, while we thank them for the social benefits that follow from their struggles.

6.3. Dissent and authority in modern science.

Learning from Galileo and other early modern scientists that epistemic altruism sometimes pays off very well for society, we in the modern West have found ways to tolerate and even encourage the kind of individual irrationality that makes it work. What almost all other cultures in history would see as arrogant defiance of authority we have embraced under what I have been calling the principle of autonomy. This commandment to think for yourself means more than just that you should feel free to speculate about things hypothetically. It also means that you should form *opinions* on important issues independently of peers and authorities, and express these opinions whenever it seems to you they might be useful. This is the defining epistemic principle of modern intellectual life: don’t just imagine how things might otherwise be, but go ahead and contradict authority if something strikes you as right all on your own. As an epistemic doctrine entrenched in modern intellectual society, the principle takes such irrational dissent and makes it rational for those who have been taught the principle authoritatively. Professors and administrators commonly affirm this principle these days by saying that their mission is to teach students to *think critically*. So, we teach them that it is good to doubt what they are told, and to think things out for themselves, and to assert their own ideas in public, even in the face of what would otherwise be rationally overwhelming testimony to the contrary. If they *see* the emperor as naked, then they should *say* the emperor is naked, even if everybody else tells them that they are mistaken (and, in fact, kind of perverted) to see things the way they do. The person who refuses to follow this commandment is a “conformist”, not to be praised for being rational but to be condemned, essentially for being useless.

The person who refuses to follow this commandment is a “conformist”, not to be praised for being rational but to be condemned, essentially for being useless. And this is fundamentally a moral judgment on them, not an epistemological one, since it is no more rational in principle for people to accept one sort of doctrine than the other. It is ironic, then, isn’t it?, that people are called conformists precisely because they *don’t* do what they are being told to do by non-traditional authorities like their professors, preferring to agree with more traditional authorities instead.[[207]](#footnote-207) In any community that includes both substantive doctrines and the epistemic doctrine that members ought to challenge doctrines, such conflicts are inevitable. The partisans of autonomy will call the partisans of substance conformists, and the partisans of substantive doctrines will call the partisans of autonomy irrational.

The principle of thinking for yourself is not generally understood the way that I’ve presented it, i.e. as a matter of getting people to be altruistically productive thinkers rather than perceptively rational believers. We certainly talk about the usefulness of having a “marketplace of ideas” that we can “sift and winnow”, but I’m not sure that anyone acknowledges the conflict of epistemic goals that is involved. Instead, most people (including me) have taken it to be entirely continuous with perceptive rationality, treating conformism, i.e. deference to authority in spite of contrary first-order appearances, as a kind of sheepish stupidity rather than an essentially moral weakness that stems from disobeying the commandment to think for themselves.[[208]](#footnote-208)

In any case, the principle of thinking for ourselves is not enforced consistently in intellectual life; nor could it be. For surely, we are all “conformists” in accepting solely on testimony all kinds of implausible-seeming facts and theories that we couldn’t possibly establish independently. As I suggested in Chapter 1, a person like Descartes’ ideal meditator, thinking entirely for himself all the time about everything, would never survive in ordinary human circumstances, and even someone who practiced such radical autonomy only in science or philosophy could hardly function as a member of the relevant profession. Instead, he would spend most of his time recapitulating long-established elementary results at best, and more probably get stuck repeating errors that are well known to existing experts.[[209]](#footnote-209) Clearly, then, if we desire progress in science, some kind of balance between autonomy and second-order rationality must be struck.

Modern science has developed a very successful, if not always coherent, way of balancing these principles in the ever-evolving system of prescriptions and constraints we vaguely call the scientific method. This method centrally demands that we take seriously as primary evidence only what is objective in the sense of being publicly available, empirical in the sense of being accessed through the senses (rather than, say, intuition), and amenable to experimental testing. Much more is taught to students in particular sciences, from using the right sorts of statistical analysis and other mathematics to handling equipment of various kinds, writing proper lab reports, and many other things besides. And much of this training in how to work and reason scientifically depends on core consensus theories that are occasionally overthrown by scientists with different methods, though the basic notions of using objective, empirical evidence and repeatable experimentation seem to persist through all such paradigm-shifts. Thus, scientists are typically indoctrinated with an implicit rational allegiance to the dominant paradigm in their field, while at the same time, they are taught a modern ideology of independent thinking, tolerance for criticism and dissent, and skepticism toward dominant theories in general. Under a flexible combination these two opposing principles, scientists learn a way of thinking that is *socially* rational in the sense that it produces more and more widely accepted theories over time, while equivocating on the epistemic situation of the scientists involved.

What I have been saying about Galileo’s sort of epistemic altruism applies most directly to what Kuhn calls *revolutionary* as distinct from *normal* science.[[210]](#footnote-210) It is the great leaps of genius in the face of opposition from united authorities, not the incremental advances in accordance with accepted paradigms, that typically demand an arrogant suspension of second-order rationality on the part of scientists. In the normal situation, modern scientists are expected to function semi-autonomously according to authorized procedures, much like doctors, lawyers, and other professional workers, so they need not be heroes to succeed. Extrinsic rewards make a career in science very attractive to students capable of doing this kind of work. They are offered a choice among comfortable lives, either in well-paid industrial jobs or in academic situations that generally offer both more freedom and more prestige, not to mention the occasional shot at immortality. James Watson and Francis Crick’s successful race against two other teams to discover the chemical structure of DNA had not just eternal fame, but also the very lucrative Nobel Prize, as goals.[[211]](#footnote-211) In the Soviet Union and other advanced authoritarian societies, top scientists have ranked as high as party bureaucrats in terms of salary, housing, freedom to travel, and other privileges, including public honors for successful work.[[212]](#footnote-212) In the West, few normal scientists get rich, but most can afford comfortable life-styles in the upper middle class, plus considerable social status. Our normal scientists do not even need to display much altruism in the purely epistemic sense. Most of their work involves either solving technical puzzles according to the theories and methods in which they have been trained, or producing novel theories in expected, relatively routine ways, without much pressure to commit themselves to the truth of their (usually highly qualified) opinions. Our scientists are trained to produce new intellectual material that will be tested and evaluated by other people whom they trust, and to argue with one another in a spirit more of cooperation than of revolutionary competition. Indeed, it is more than possible for new hypotheses to generate research without their being anybody's personal opinions. Workers in science only have to find them plausible enough to justify the effort involved in checking them out, like a corporation's engineers drilling exploratory wells wherever their geologists have found some threshold probability of striking oil. Most present-day physicists and biomedical researchers, in particular, seem to follow something close to the ideally rational approach in rarely expecting themselves, individually, to be the ones who turn out to be right about their own hypotheses. They have seen so many large and small paradigm shifts occur within the past few decades that they have learned inductively to be quite skeptical of current theories. Still, they engage in their work with the zeal of people close to achieving all that they desire. Their suspension of personal judgment does not dampen scientific enthusiasm for them, much as this may have surprised the ancient skeptics.[[213]](#footnote-213) Perhaps the difference is that these scientists do expect to see new pieces of the truth emerging fairly quickly through the system of inquiry in which they play a part, so they are able to treat their suspension of as temporary. Thus, they see themselves both as producers and testers of relatively raw ideas, and as potential consumers of more finished scientific products.

In a workaday sense, then, science is something like organized religion, with a fairly hierarchical authority structure dominating a large field of lesser experts and a vast population of lay consumers. The authority structure of science is even a little like that of the Church in Galileo’s time, in that final authority sits at the top. Science has no Pope, but it has plenty of bishops and cardinals in the form of structured professional organizations, journals, university personnel committees, and, most importantly in some fields, funding agencies. In the controversy over "cold fusion" in the 1980s, for example, a pair of highly credentialed chemists, Stanley Pons and Martin Fleischmann, was harshly rebuked by the leaders of their profession when it was judged that they had violated authorized procedure by publicly announcing a purported major breakthrough prior to publication in a peer-reviewed journal. They were not burned at the stake, of course, but their careers certainly went down in flames.[[214]](#footnote-214)

These similarities in structure do not make science the same thing as religion. Religion depends on the transfer of an essentially static core of doctrines from one generation to the next, but science, like dialectical philosophy, is an inherently progressive enterprise. Ideally, at least, science has no specific substantive doctrines to which it commands assent; only the epistemic doctrine of thinking critically and respecting the empirical enterprise of science itself. Nevertheless, faith in substantive doctrines plays an essential part in scientific work as well. For, as has been said, no single scientist can hope to recapitulate the whole of physics, chemistry, or any other major field, or, these days, even any major sub-field. Each must rely implicitly on all the others for testimony as to what experiments have been done or are being done, what the results have been, and what people are doing in related fields. This is why fraud in science is so hard to detect when it happens: first, because scientists are usually very disciplined and honest in their work (and the penalties are very harsh for cheating), so plausible results tend to be accepted without suspicion; and second, because it would cost too much in time, money, and tedium to check carefully the mass of routine data that the system heaves up every day. In some kinds of research, especially military research, individual scientists may have only limited knowledge of the overall projects they are working on. Increasingly, even in academic science, researchers are involved only in small parts of huge, often international projects that no individual is even competent to understand completely.[[215]](#footnote-215) And it is not just in each other’s competence and honesty that scientists must have faith. They also must accept – most of them, if they are to get any work done – broad analyses of past results and whole structures of current theory. If they don’t buy into most of this material while in graduate school, they will not be considered trained and competent to work thereafter. Thus, modern science as an institution is held together by faith, in the sense of rational reliance on testimony, much as religion is. It does matter that in science, people believe that in principle, they could verify by first-order means the things that other scientists are telling them. But this, too, would depend on faith in practice. No chemist has the time to learn all of the physics that underlies his research, nor would the physicists themselves have time to verify the proper engineering of their experimental machinery, nor would the engineers who constructed those machines have time to verify the principles of physics and chemistry that support their designs. They simply have no choice but to one another, and as long as the entire system functions appropriately, they will have good reason to do so.

This universal need for trust in science can produce the same sort of epistemic gravitation and resistance to dissent that marks traditional religion, leading potentially to black holes from which individual scientists, and conceivably even science as an institution, can never rationally recover. Though it is often altruistically suppressed in revolutionary science, subjective rationality tends to regain its full force in periods of normal science. Students trust professors and professors trust their peers, not just on method and technique and on immediate interpretations of experimental data, but on core theoretical questions as well. Ordinarily, researchers are pulled into existing paradigms not through an unreasoning process of indoctrination, but through the rational force of testimony from the scientists they have the most reason to trust. Thus, even scientists consciously dedicated to open-mindedness will tend to receive theories from the broader community not as hypotheses but as “settled science” when a broad enough consensus forms around them. From the rational point of view this makes perfect sense. To the extent that consensus theories are more likely to be true than their competitors, being convinced of them keeps scientists from wasting their time and everyone’s resources chasing implausible alternatives. This happens normally in large and small areas of science, and is extent a necessary part of any reasonably stable paradigm. The danger is that if the “settled” theories turn out not to be true, and not close enough to the truth to correct themselves internally, then whatever errors they contain will tend to propagate throughout the intellectual community, fixing themselves as doctrine. And when reasonable people no longer see any need for fundamental progress, autonomy loses its value as an epistemic principle, and this makes science *itself* not stable but inherently unstable. This is why great civilizations like ancient India and China and the medieval Byzantine and Arab Empires have fallen into intellectual and technological stagnation following spurts of exploration and progressive change. Once an acceptably comprehensive and successful theoretical consensus has emerged as settled doctrine, what had once been open speculation merges back into religion. It is this ever-present tendency of rational people to come to agreement under a general consensus of experts that makes science (and progressive social rationality in general) so very rare in history, and so hard to preserve as a self-sustaining institution in spite of its demonstrated power to improve our understanding of the world.

Modern science differs self-consciously from previous progressive intellectual movements in that it seeks to maintain a permanent openness to revolution through the explicit epistemic doctrine never to close a question finally, even after a consensus has been reached among authorities. But this is very hard to do during extended periods of normal science due to the rational attractiveness to unanimous expert perception. When questions are effectively closed to argument through epistemic gravitation, modern scientists still honor the idea of open challenges in principle and even celebrate the revolutions of the past. Yet at the same time, they *must* find it really difficult to put up with the cranks and amateurs who insist on dissenting from the “settled science” every reasonable person justifiedly believes is true. Sometimes the ideology of open criticism keeps the upper hand, and the battle against rational conformity is no worse than keeping weeds from encroaching on a well-tended garden. At other times, science gets stuck in unanimity despite its nominal commitment to diverse ideas and change, and it takes a good whack from irrational dissenters to get it all running again.[[216]](#footnote-216)

I don’t mean to concur with Kuhn that revolutionary theory change in science *necessarily* depends on irrational or non-rational factors. As far as I know, objective criteria may well be found for comparing scientific theories across paradigms to every reasonable person’s satisfaction, and science as an institution may evolve ways of keeping itself constantly open to new ideas. But so far at least, rational scientists do tend to get pulled down into reigning paradigms to the extent that any serious dissent will seem unreasonable. So, it can take considerable courage and persistence for dissenting scientists to disrupt the current conversation and attack the genuine convictions of their colleagues on issues considered to be settled. Many dissenters are not even taken seriously as scientists, since they fail to meet the current standards of professional quality, necessarily defined within the reigning schools of thought. In fact, many potential revolutionaries will be kept out of the profession altogether, hard as it is to get through graduate school in any field without conforming to existing paradigms. To a great extent, as I have said, this is as it must be: technical training can never be kept altogether separate from theoretical indoctrination. Thus, from a point of view sympathetic to the reigning consensus, people who participate in scientific discussions appear as highly trained, professional experts, while from a dissenting point of view, it may appear that anyone involved in these discussions at a professional level is already corrupt. Tenure helps to enable dissent, as it is intended to do, since tenured professors are not forced into conformity simply to protect their jobs. But most scientists, like most philosophers and other academics, don’t want to sacrifice potential raises and promotions, or to have their research funding dry up or, most importantly, to lose their reputations for good work. Their jobs may be completely safe, that is, but their careers are not. Risking their professional reputations takes considerable courage even for those prominent people who are best positioned to do revolutionary work and get away with it, but who may also have the most to lose.[[217]](#footnote-217)

Aside from established scientists who change their minds about the dominant paradigm after succeeding within it, dissenters who successfully demand attention are typically outsiders from other fields or beginners with little at stake.[[218]](#footnote-218) Alfred Wegener, the originator of the theory of continental drift, was an outsider. A meteorologist by profession, Wegener was almost universally denounced as a charlatan for his views by more prominent scientists within geology committed to prevailing theories like contractionism (the idea that the earth's crust changes shape because it is gradually cooling and contracting). It was only decades later, after more and more evidence emerged and some well-known geologists converted, that Wegener’s tectonic theory came to be taken seriously by the scientific community. Galileo certainly fits the last description, having become the scientific darling of the age and a personal friend of Catholic authorities until he turned against them. Darwin, too, famously held back on releasing his radical alternative to the consensus theory of divine creation until he had established a tremendously successful career by doing safer, though still brilliant, work on other topics such as barnacles and coral reefs. He finally published only when he was in danger of being beaten to the punch by the unknown beginner Alfred Russel Wallace.[[219]](#footnote-219) Backed by influential friends, Darwin withstood the thunderstorm of protest that his work unleashed, but it was still enough to make him anxious and unhappy during most of his remaining life.

Similar, sometimes very ugly conflicts over novel theories still occur. For example, the great entomologist E. O. Wilson shocked the scientific world by publishing in 1975 his book *Sociobiology*, the final chapter of which argued on evolutionary grounds for abiding psychological differences between men and women, and possibly among ethnic groups as well, as against the reigning view that such differences are caused among humans by environmental forces only. Wilson was treated roughly by protestors, on one occasion having a pitcher of water poured over his head while protestors chanted, "Racist Wilson, you can't hide! We charge you with genocide!” He was also denounced by some prominent scientific colleagues, and his right to differ only grudgingly conceded.[[220]](#footnote-220) In spite of such resistance, Wilson’s basic approach to sex differences has gradually become part of the mainstream itself, although the hated label “sociobiology” has largely been replaced by the less tainted “evolutionary psychology”, with several new journals and hundreds of books (some very popular with general readers) sprouting up over the last two or three decades. These days, the idea that men and women have some subtle and slight but probably ineradicable average differences in cognitive tendencies seems to be broadly if very cautiously accepted by scientific experts.[[221]](#footnote-221) As recently as 2005, though, when Harvard president Larry Summers suggested tentatively that such differences might account for some portion of the gap in average success (and salary) between men and women in mathematics and the physical sciences, he was forced out of office by an infuriated segment of the Harvard faculty.[[222]](#footnote-222)

6.4. Consensus, controversy, and induction.

The most pressing current large-scale scientific disagreement is about anthropogenic climate change. Though I am anxious not to alienate readers who have strong convictions on this issue, I want approach it from a neutral angle and express some grounds for skepticism toward both sides. At stake in this controversy is our avoiding a predictable environmental catastrophe according to one side, and a predictable economic catastrophe according to the other. The theory that human industrial activity, by releasing large amounts of CO2 gas into the atmosphere, is causing a global warming trend that may soon destroy much of our civilization is the current position of a clear majority of scientists in fields connected to the issue. This theory’s partisans claim further that their theory is not just a majority position but a consensus of sufficient standing among qualified scientists that the theory ought to be considered "settled science", and strong measures such as caps on CO2 emissions ought to be taken immediately to prevent environmental disaster. But there is also a minority of fierce dissenters who argue that in one way or another the majority's case for “gloom and doom” has been exaggerated, and some of whom have gone so far as to call the theory of catastrophic anthropogenic global warming is a hoax.[[223]](#footnote-223) As can be expected in all such disputes, one side is seen by its opponents as irrationally refusing to accept established theory, while the other is seen as unscientifically insisting on conformity. Scientists on both sides of this issue have had venomous things to say about their opponents. Accusations of fraud have tainted to some extent even the supposedly hard, objective data of temperature readings, polar bear sightings, and glacial contractions that ought to be included in the pool of public evidence everyone agrees on. Each side is also accused of padding its numbers with unqualified people in order to make the minority appear more respectable or the majority more monolithic.

What should we uncommitted rational non-scientists believe about such controversies? We are not technically equipped to pass judgment on many of the issues involved in such complex discussions, or even to make plausible first-order estimates. So, it is mainly a question of which sources we decide to trust. Surely, other things being equal, the majority position ought to be counted as more probably correct, and more so as the majority is larger – but to what extent? Is the case at hand one of those controversies where the consensus theory ought to be taken as the settled truth and the few outsiders as self-serving or attention-seeking cranks? Or could it be one of the cases where the outsiders are clear-eyed, self-sacrificing dissidents, willing to risk their careers to challenge a dogmatic orthodoxy while the majority is stuck together in an epistemic black hole? The most rational criterion for non-experts to adopt on such matters is the same inductive one we use in evaluating any other source of testimony: considering all possibly relevant aspects of the current conflict, what does the history of similar disputes tell us about the overall reliability of consensus and dissenting theories? And just as when deciding how much to trust your mother or a politician, the more similar the prior cases are to the present one, the better inductive evidence we'll have.

The historical track record of scientific majorities is obviously mixed.  If contemporary natural philosophers had voted prior to 1650 or so on whether the Ptolemaic or Copernican model of the solar system was correct, Ptolemaic theory would have won by huge majority. But after Galileo, the Copernican alternative became part of a new consensus on the solar system that remains essentially unchallenged to this day.  So it looks like the old majority was not right and the new majority is right on this issue, each for about 400 years. Newtonian mechanics was so much the accepted view among scientists prior to 1905 that Einstein's relativistic theory seemed impossible, even incomprehensible, to many scientists despite its brilliance and elegance, until it was firmly established through experiments as the superior theory.  The hugely successful theories of plate tectonics in geology and Darwinian evolution in biology seem to be settled science if anything is, but were preceded by consensus theories that turned out to be false. In the various social sciences, there have been quite a few dominant schools of thought over the past century, some of which held strong majority positions temporarily. Freudianism and behaviorism each dominated academic psychology for a while to such an extent that it was hard for any dissenting work to pass through peer review for publication. But neither of those paradigms survived intact, though both produced insights that are useful to the more open and eclectic field of psychology today.

So far, it seems that unless there is some reason to think of today’s consensus theories as final, where consensus theories of the past have usually been overthrown sooner or later, it looks like scientific majorities are in general no more than about 50% reliable. This is to use only the crudest sort of induction, of course. The specific natures of the theories and the majorities involved also obviously matter very much. Thus, we consider the virtually 100% majority over the past 500 years in favor of round-earth theory to be more reliable than theories that are more complex, more recent, and more controversial, such as the current majority view that it is best for ordinary people’s health that they have screening colonoscopies every ten years beginning at the age of 50. So, rational non-experts trying to evaluate the reliability of the scientific consensus that current levels of CO2 production will result in catastrophic global warming must bring in as many observable factors as we can to narrow down the scope of our inductive calculations. As far as we can tell, what are the special properties of the current climate change consensus? It seems that there are several features we can use to narrow things down. First, we can note that the climate-change issue is relevant to urgent questions of public policy. Second, that it emerged quite recently, only a couple of decades ago. Third, that it appears in the context of a social movement that has passionate adherents in the current scientific community, namely environmentalism. Fourth, that positions on this issue have become politically polarized, making them parts of larger factional package deals.

On sensitive matters relevant to public policy, the track record of scientific majorities is not very good, for the obvious reason that scientific perceptions are subject to gravitational distortion just like those in other areas of intellectual pursuit.  Only a hundred years ago, for example, almost all credentialed scientists were explicit racists by today's standards, and their racism used to support an almost universal Social-Darwinist approach to eugenics that was expressed even in law through the authority of scientific consensus. Justice Oliver Wendell Holmes's famous opinion in the forced-sterilization case *Buck v. Bell* that "three generations of imbeciles are enough" was not the statement of an essentially hard-hearted person, but rather of a rational person who simply accepted what was seen as "settled science" in his day.[[224]](#footnote-224) In more socially progressive times, Margaret Mead's anti-Darwinist anthropological studies from the 1920s on, supporting a radical blank-slate theory as a kind of anti-paradigm to Social Darwinism were until quite recently accepted by the great majority of social scientists.  But it has turned out that her results have been cast into serious doubt by subsequent re-examinations and newer, more careful research. As I mentioned above, we now seem to see that social-causes-only theory being superseded by a more moderate semi-biological approach in the new evolutionary psychology. At both extremes, the political commitments of the scientific community seem to have ruined the reliability of its consensus theories.

In recent environmental and related matters, the track record of scientific majorities seems to be especially problematic.  Rachel Carson's *Silent Spring* of 1962 is often credited with starting the movement of environmentalism with its successful attack on the pesticide DDT, then widely used in Africa for controlling malaria.  But current science views DDT as far less harmful than Carson and subsequent mainstream scientists believed, and DDT has recently been reintroduced in Africa for limited use.[[225]](#footnote-225) Another founder of today's environmental movement is Paul Ehrlich, whose 1968 book *The Population Bomb* set off a bomb of its own with its widely accepted Malthusian predictions about human overpopulation.  Ehrlich's predicted mass starvations of the 1970s and 1980s never materialized. [[226]](#footnote-226) Rough as life remains in much of the Third World, the problem of population growth as such seems to have faded relative to other issues, while explosive economic growth in Southern and Eastern Asia has been raising hundreds of millions of people from subsistence poverty to something like a Western standard of living.

To make things more confusing in the present case of global warming, scientific questions about climate change have now become deeply enmeshed in partisan political struggles, with the Democratic Party in the United States and similar left-leaning parties elsewhere supporting the environmentalist position, while the Republicans and similar right-leaning parties mostly support what might be called the *economicist* position. Therefore, it's very hard even to find politically neutral sources of information on the issue, all merely scientific biases aside, and hard for a neutral person even to talk to people who have strong convictions on the issue. Readers of *The Nation* and the *New York Times* are surrounded by environmentalist arguments, while those who read the *National Review* and the *Wall Street Journal* are surrounded by economicist dissents, with little serious attention given to the other side on either side. And the fact that those favored publications favor their respective positions editorially counts by itself as powerful second-order evidence for non-experts that those positions are correct. Viewed from a neutral position, though, the fact that our usually trustworthy sources have largely embraced this issue as an ideological, rather than a purely scientific, one seems to render *all* of their testimony less trustworthy that it would have been otherwise. When an issue has been polarized politically to such an extent that someone’s opinion on the issue is just a proxy for their partisan affiliation, all you can learn from their testimony *per se* is what party they belong to. And to the extent that the entire scientific community aligns with one or the other political faction, the mere fact that they take a uniform position on a politicized issue, be it climate change, abortion, nuclear power, missile defense, racial preferences, capital punishment, immigration, school vouchers, or stimulative federal spending, tells you much less than it would if scientists were perfectly free to disagree with one another without fear of partisan recrimination or damage to their careers. Of course, it’s always possible that the consensus scientists agree in the first place entirely for good scientific reasons. What I’m arguing is only that non-scientists *can’t find that out* when the scientific consensus is also a partisan political consensus.[[227]](#footnote-227)

Another sign of potential unreliability concerning climate change has been the bullying behavior of some scientific authorities toward skeptical dissenters, who are now widely called "deniers" to associate them with people who deny that the Holocaust took place despite the overwhelming first- and second-order evidence that it did, while groups and organizations that support them are called “anti-science”. The treatment of statistician Bjorn Lomborg, author of the controversial book *The Skeptical Environmentalist*, is a particularly interesting example because his position on climate change is not an ideological one.[[228]](#footnote-228) Lomborg accepts the theory that humans are causing serious global warming by releasing CO2 into the atmosphere, but argues that the hundreds of billion dollars it would cost to make a noticeable difference in projected warming would be better spent on curing diseases and solving other major problems like unsanitary drinking water in the Third World. After *Scientific American*, the flagship of mainstream scientific publications if there is one, responded to Lomborg's book with a long, scathing review written by four environmental scientists in January 2002, the editors made it absurdly difficult for Lomborg to respond in turn, permitting him only a 1-page rebuttal to 11 pages of criticisms, and, when he tried responding in full on his own website, threatening to sue him to prevent his taking quotes from their review. This is surely not how a fair-minded scientific journalought to behave. Under pressure from scientists on both sides of the issue, *Scientific American* did eventually retract its threatened lawsuit, and Lomborg's response now appears on its website, together with a new rebuttal of its own.[[229]](#footnote-229) To cite my own experience, when Lomborg's book was chosen for discussion in a faculty reading group several years ago, one scientist member refused to purchase the book in order to keep his money out of Lomborg's and his publisher's dirty hands, and another declared himself unwilling on principle even to read it. The subsequent discussion, largely given over to a debate on the propriety of even holding the discussion, got so heated that the reading group disbanded afterwards. Lomborg was plainly seen by some of my colleagues as a villain, something like a Nazi, rather than a mere dissenter. On their view, just as a decent person does not purchase Nazi propaganda, read it, or dignify it by taking it seriously in discussion, a decent person shouldn’t pay attention to the views of climate change deniers. As far as is obvious to outsiders, though, Lomborg seems to be well-meaning, sincere, and humanistic in his writings, and while his work may, for all the non-expert can tell, be fatally flawed in its technical arguments, it is his general approach of weighing the human costs and benefits of fighting global warming against those of fighting other problems that seems to drive his mainstream critics so far up the wall. He doesn't even question the existence of anthropogenic global warming as a problem, only its relative importance, and there is nothing at all in terms of science to object to in this approach to the topic. Only as a political matter, a matter of his doubting publicly the urgency of drastic *action*, does it seem that Lomborg's general approach is out of the mainstream.

Additional uncertainties spring from considerable public (or private but leaked)[[230]](#footnote-230) discussion among scientists of whether or how much to exaggerate or “spin” predictions of catastrophe in order to motivate the general public to take urgent action. Here is a much-discussed statement from an interview with the prominent climatologist Steven Schneider:

*"On the one hand, as scientists we are ethically bound to the scientific method, in effect promising to tell the truth, the whole truth, and nothing but - which means that we must include all the doubts, the caveats, the ifs, ands, and buts. On the other hand****,*** *we are not just scientists but human beings as well. And like most people we'd like to see the world a better place, which in this context translates into our working to reduce the risk of potentially disastrous climatic change. To do that we need to get some broadbased support, to capture the public's imagination. That, of course, entails getting loads of media coverage. So we have to offer up scary scenarios, make simplified, dramatic statements, and make little mention of any doubts we might have. This 'double ethical bind' we frequently find ourselves in cannot be solved by any formula. Each of us has to decide what the right balance is between being effective and being honest. I hope that means being both."* [[231]](#footnote-231)

Schneider may hope that it means being both, but by the logic of his statement it cannot mean being both completely, at the same time.[[232]](#footnote-232) In any case, statements like this from scientists in the majority hardly inspire confidence in them as reliable experts among the uncommitted laity.

I am not saying that there are sufficient grounds for taking the opposite "denialist" position on this issue; far from it. And I don't mean to imply that only the majority is playing games with the debate. There is certainly plenty of propaganda coming from the other side as well, for example mainstream scientists being routinely tagged as “alarmists” and “true believers”, not to mention “liars”, “frauds”, etc. I am not even saying that slandering Danish statisticians and telling half-truths to the general public is not morally justified, given the hellish future that the climate crisis theorists want to persuade us to prevent. If the whole planet really needs to be saved by immediate and drastic action, then all other bets are off. But if nothing radical is done and the predicted catastrophe still never comes to pass (perhaps because of some undiscovered mitigating process that has not been factored into climate models), then mainstream science will have thrown away much of its remaining authority among an increasingly skeptical general public, and the usefulness to science of subjectively irrational dissenters will have been vindicated once again. It is not my purpose here to argue for either side, though, and I’m obviously not in a position to discuss the merits of the case on a first-order basis. My point is only that for someone starting from a neutral, politically uncommitted point of view, there seem to be good inductive grounds for some level of skepticism on this issue, given the falsification of past majority positions on so many politically sensitive scientific controversies, including recent ones concerning the environment, plus the present partisan behavior of so many mainstream scientists. As a matter of pure rational *perception*, then, the ideal thing would be simply to wait and see how this issue turns out without making any decisions in the meantime. As a matter of *opinion*, the ideal thing would be to keep debating the matter openly for as long as it takes to come to a really solid, satisfactory agreement among all competent and interested parties. The problem with both approaches, though, as many people see it, is that we just don’t have the time. The climate change issue cannot safely be treated as either a pure matter of perception or a pure matter of opinion; it is instead an urgent matter of *judgment*. Either the environment or the economy might well be headed for preventable catastrophe right now, depending on the combination of which theory is believed and which is true. Therefore, radical action may need to be taken or prevented right now, which means that crucial judgmentshave to be made without delay, though based on very limited and (from a non-expert point of view) probably biased information. Much as we might rather wait and see what develops, or keep debating every aspect of this interesting controversy, at some point soon we'll have no choice but to *make up our minds*.

**7. ACTION**

To close your ears to even the best counter-argument once the decision has been taken: sign of a strong character. Thus an occasional will to stupidity.

- Friedrich Nietzsche, *Beyond Good and Evil* IV.107

7.1. Judgments.

We have to make decisions, and we don't always have the time or patience or other resources needed to consider thoroughly all the evidence that we could gather. Many of our actions are effectively possible only if we take it for granted that something simply is the case, rather than merely being probable. To make up our minds is to come to believe something in the absence of sufficient epistemic justification because some kind of action must be taken that requires such categorical assent. This is *judgment*, and it plays a different role in life from both perception and opinion. Hume says, "Nature, by an absolute and uncontrollable necessity has determin'd us to judge as well as to breathe and feel..."[[233]](#footnote-233) And it is easy to see how the need for decisive action in a largely hostile world requires a single, relatively fixed working theory of that world, not the conditionalized network of probabilistic sub-models that constitute a person’s total perceptions. Hence, the purest adherence to perceptive rationality may be of limited survival value. In the ancestral environment, with all its dangers, creatures who found it very easy to suspend belief whenever evidence was incomplete or contradictory would have had trouble competing with others who were more decisive. So, we often find ourselves adopting categorical beliefs that are otherwise perceptively unjustified when under pressure to act.

Ordinary judgments are fixed in the mind temporarily, without any particular determination that they stay fixed. For example, if we see a stream of looming peripheral bus-like images as we are crossing the street, we may judge that a bus is coming towards us and jump out of the way. This sort of judgment barely registers as a belief, as opposed to a mere stimulus. Other ordinary judgments are more conscious and somewhat more stable. I think that a student’s paper is good enough for a grade of B+. I know that I could be wrong about this, and that a deeper examination of the paper may turn up some unnoticed virtues or flaws. But I have to grade these papers in a reasonable amount of time, and I believe that I am adequately competent and fair with this responsibility, so I put my doubts aside and *judge* the paper as a B+ so that I can move on. I have nothing else at stake in my belief that this student’s paper merits a B+, and like most professors, I am willing to reexamine this judgment if the student complains to me about it. In general, such ordinary judgments need only be fixed enough for purposes of immediate action, and we have no problem reconsidering the evidence for and against them once the action has been taken.

It is a bit mysterious why more than momentary fixing of belief is ever necessary. Why can’t we act on our perceptions without any *standing* judgments, as a good poker player draws cards and bets on the basis of a rolling, probabilistic assessment of his situation? It is bad for a gambler to make judgments at all, beyond the judgment that this or that hand of cards in this or that situation merits this or that action. A poker player who comes to believe categorically that a certain hand is bound to win and bets accordingly will usually lose to one who is able to fold mechanically whenever the balance of probabilities requires it, regardless of how much he has already invested in the pot. What makes gambling at cards a skill is not just mastery of the mathematics involved, but also the ability to overcome the natural human impulse to view things categorically, like the artist who must learn to see things just as they look, without putting what he sees in preexisting categories. Even the interpersonal psychology of poker is best implemented in a non-judgmental, probabilistic way: given the way he’s squirming in his chair, there is such-and-such a chance that this opponent is bluffing at the moment, and there is such-and-such a chance that he will think that there is such-and-such a chance that I will think that there is such-and-such a chance that he is bluffing, as well as such-and-such a chance that he will believe that there is such-and-such a chance that I am bluffing, and so on. Such probabilities are not usually calculated consciously, but by some hidden computer in the brain which then communicates its findings to the conscious mind through feelings such as sudden impulses to raise, stand pat, or fold, together with a certain confidence that what we're doing makes good sense. But however it works psychologically, a strategy of acting according to subjective probabilities without lasting judgments seems to be the most rational one for playing cards. Why not for everything?

Making decisions in ordinary life is different from gambling in important ways. There is a continuum between what can be done purely by probabilistic calculation at one end, and what can only be done according to fixed judgments at the other. At the one extreme, skillful gambling requires constantly attentive, if intuitive, Bayesian calculations. You can do little else requiring careful thinking at the same time. At the other extreme, taking a running jump across a chasm takes a certain grim determination *not* to keep rethinking your decision but to plow ahead as if you had no further choice. Most of our ordinary conscious decisions and actions fall in the middle, where we are able to make some calculations in the available time with the available resources of attention, but other beliefs and desires have to be taken for granted as relatively fixed presuppositions.

Here is a very simple piece of practical reasoning, called a *practical syllogism*:

I want a beer.

If I ask the bartender for a beer, then I will receive a beer.

Therefore, I ought to ask the bartender for a beer.

In more general terms:

I desire state of affairs *s*.

I believe that action *a*will bring about state of affairs *s*.

Therefore, I ought to perform action *a*.

Most of our decisions are too complex to fit into this simple formula. Instead of a simple, yes-or-no desire for a single state of affairs, we choose from within a range of possible outcomes, each of which we desire to a greater or lesser extent. Instead of a simple, yes-or-no belief that a particular action will bring about a particular state of affairs, we choose within a range of possible actions, each of which we think will result, with different probabilities, in any of a number of possible outcomes. Then we intuitively calculate something like an *expected value* for each possible action, which is the sum of the products of the probabilities of each outcome possible for that action, times the desirability for each of those outcomes.

Here is what an ideal expanded “syllogism” looks like, bringing degrees of desirability for states of affairs and probabilities for outcomes into the calculation:

I desire state of affairs *s1* to degree *d1*.

I desire state of affairs *s2* to degree *d2*.

I desire state of affairs *s3* to degree *d3*.

*[…and so on for all possible states of affairs]*

I believe that action *a1* will bring about state*s1*with probability *p1*.

I believe that action *a1* will bring about state *s2* with probability *p2*.

I believe that action *a1*will bring about state *s3*with probability *p3*.

*[…and so on for all possible states of affairs]*

I believe that action *a2* will bring about state *s1* with probability *pn+1*.

I believe that action *a2* will bring about state *s2* with probability *pn+2*.

I believe that action *a2*will bring about state *s3* with probability *pn+3*.

*[…and so on for all possible states of affairs]*

*[…and so on for all possible actions]*

Therefore: I expect action *a1* to have the value *v1*.

I expect action *a2* to have the value *v2*.

I expect action *a3*to have the value *v3*.

*[…and so on for all possible actions]*

Therefore, some action *am* has the highest expected value for me.

Therefore, I ought to perform action *am*.

Depending on the case, the relevant matrix of probabilities and desirabilities involved may be large or small. Sometimes, as in formal gambling, we can effectively calculate using matrices like this, but most of the time we don’t have the resources of time and attention to do this, even unconsciously. When we can’t, it will still often be possible to use the simpler, categorical versions of these syllogisms. This means making up our minds about what we want and how to get it. And this, in turn, means deliberately ignoring some of the information that is available to us, which makes our practical calculations less complete and accurate than they theoretically might have been. But we make up for this loss by being able to generate decisions and actions efficiently when otherwise we could not, as a practical matter, do anything useful at all.

Of course, it isn't always better to do something than to do nothing. If our evidence is poor, or if we are hurried, tired, distracted, or for some other reason especially unreliable at the moment, it may well be better not to make up our minds. It depends on what is at stake. For example, it is usually unwise to propose marriage or challenge someone to a duel when you are roaring drunk; better to forego whatever advantage might accrue to you from acting quickly, and get a good night's sleep. But sometimes we have no choice but to take action on the basis of imperfect information and analysis, and then it pays to simplify our reasoning to the point where it can yield a quick result:

This *is* (categorically) a hungry tiger running towards me.

Being eaten by a tiger *is* (categorically) undesirable.

I *will* *be* (categorically) safer if I run away.

Therefore, I *am* (frantically) running away.

Again, it’s not that people set up little arguments like this in any conscious way. Modern human beings are descended from people who got very good at taking good cognitive shortcuts when it matters, not from their more thoughtful and hesitant cousins. Thus, as a matter of psychology, we tend to deal with complex decisions through a combination of conscious, articulate reasoning and emotion or “gut instinct”. Sometimes the head leads the gut (or heart, or perhaps some other organ); other times our conscious thinking serves only to *rationalize* decisions that have already been made at a more instinctive level. But even then, the gut response encodes a reason of some sort, either an unconscious sort of calculation or a programmed response that was inherited from successful ancestors who happened to follow what turned out to be the reasonable course.[[234]](#footnote-234)

In most cases we don't need to simplify our reasoning to quite such an extent, though we are still unable to perform a complete probabilistic analysis. We do not decide on all the facts and values in question, but decide in advance which factors to hold fixed, and which are important enough to consider further. Most of our practical reasoning is intermediate in this way. We restrict ourselves to considering a few plausible-looking options for action, and then we perform a limited calculation, compact enough to generate the necessary action in the time allowed, but as comprehensive as it can be made with such resources as we have available. Think of how you usually buy a car. It isn’t:

I want a car.

If I buy this Oldsmobile, then I will have a car.

Therefore, I ought to buy this Oldsmobile.

But it also isn’t a gargantuan spreadsheet listing all of the new and used automobiles in the world, ranked by their overall desirability, and multiplied by all the possible ways of obtaining them. It’s more like this: we want a good, reliable sedan, preferably a Camry or Accord or something similar, about this age, about this mileage, for about this price, and we want it in the next few weeks. So, we end up partly evaluating a dozen or so cars that fit this general description – this one has too much rust, this one is okay except for the rattling noise, this one is fine but too expensive – before deciding that the red 2001 Mazda with the Hillary! sticker is the best quickly available car for what we want to spend. This is a typical use of judgment in practical reasoning. A few things are fixed categorically in order to reduce a problem to a calculable level of complexity, a decision is made, an action is taken, and we move on to thinking about something else.[[235]](#footnote-235) Do we really need a Japanese car? No. Mightn’t we be better off buying a Volkswagen or a Ford? Sure. But if we would rather have an actual acceptable car than a theoretically perfect decision, we will have to ignore some of these possibilities and look for a less-than-optimal solution within a structure of somewhat arbitrary temporary principles. The harder we find it to make and stick with such medium-term practical judgments, the harder we will find it to get through life.

7.2. Convictions.

Convictions are more than mere judgments. They are long-standing, conscious *commitments* to categorical beliefs and to whatever actions those beliefs entail. A conviction is a contingent proposition treated not just as a standing assumption but as a *certainty* for purposes of action. This is not a *perceptive* certainty, in the sense of a perception with a subjective probability of 100%, but a *practical* or *moral* certainty, which mean any proposition treated as a fixed principle of action, regardless of the probability assigned to it on the dimension of perception. Thus, a pious person believes in God without considering the probabilities, and prays. A patriot believes that his country is right, even in the face of evidence that it is doing wrong, and acts to defend it. A pacifist believes that violence is always wrong, regardless of arguments to the contrary, and will not retaliate when violently provoked. In all such cases, the believer places something of himself at stake in the judgment in question (actually, not in question), so that integrity requires his actions to accord with the belief. In this way, we feel ourselves to some extent *defined* by our convictions.[[236]](#footnote-236)

Where an ordinary judgment is just a stepping-stone for action, a conviction is a more-or-less permanent foundation of belief upon which large or small structures of action can be built. But there are intermediate gradations, too, between convictions and ordinary judgments. You may become convinced that pursuing a career in gold-mining is a good idea for now, while still keeping your other options open for the future. This means committing to a substantial amount of effort and expenditure, and this requires a proportionate degree of confidence in the attainability of your gold-mining objectives (e.g. finding some gold), but with a limit on the total investment you are willing to make in the project. Or you may be convinced enough of a political candidate’s merits to join his campaign for a hard year's effort, while retaining enough background suspicion toward all politicians to prevent your being shocked or compromised by any scandals his opponents ultimately bring to light. Thus, we distinguish between being convinced and *deeply* convinced about a proposition, and between being committed and *deeply* committed to a belief or action. There is really a continuum, then, from the most momentary judgments to the deepest convictions, with short-term provisional judgments and shallow convictions falling in order in-between and a belief's position on the scale depending on how reluctant you are to revise it in light of countervailing evidence.

Actions like buying a car typically involve not only practical judgments, but also convictions that function as boundary conditions in the decisions we make. Thus, in addition to the factual beliefs and practical desires mentioned concerning our purchase of a 2001 Mazda above, we are not willing to cheat someone to get it, because we believe in fair play as a matter of principle. At the same time, we may or may not accept an offer to fool the state government with a false price on the bill of sale in order to evade paying the excise tax, depending on what we believe about our duties as taxpayers. We may be unwilling to buy foreign cars, or perhaps domestic cars, as a matter of political conviction. And intermediate sorts of practical or moral convictions are liable to apply as well, for example someone’s long-standing belief that no car that has been “totaled” in an accident is safe to buy, or that there are usually hidden problems with vehicles on used-car lots. These are revisable judgments, not deep moral commitments, but we are not likely to revise them any time soon.

Successful marriage seems to require committing yourself to the truth of some otherwise very uncertain propositions, such as that your prospective spouse is an appropriate mate and a morally good person, and that you will be satisfied with this commitment for a very long time. As it turns out most of the time, you really can’t be very happy in a marriage without making this kind of epistemic commitment along with the usual legal or religious ones. If you are not already convinced that you are going to stay together till death does you part, if this remains instead a mere matter of probability or conditionality to you, then you are probably already in trouble. Most marriages (like most as business partnerships, musical groups, scientific collaborations, or all sorts of other cooperative endeavors) go through rocky stages, and a foundational belief in their ultimate success helps to prevent us from throwing in the towel when the going gets tough. So, part of a good marriage is an initial decision to believe *categorically* that your marriage is permanent and will be happy in spite of any future problems you encounter. Of course, you know perfectly well that many marriages end in divorce, including ones where people like yourself made similarly wholehearted commitments. So it seems that rationality requires you to perceive a certain likelihood of failure, while integrity requires you not to perceive it, or at least not to acknowledge it in making decisions. The same can be said for any sort of long-term commitment to something very difficult, like (I remember) writing a dissertation or (I imagine) running a marathon; if you acknowledge to yourself the real likelihood of not finishing, you are less likely to finish than if you irrationally view failure as impossible.[[237]](#footnote-237)Here is where the rules of perception and conviction part company. The principle of rationality demands that we see things exactly as they are, to the best of our ability. The principle of integrity demands that we stay loyal to our convictions, no matter what. And this is not a matter of our lacking epistemic reasons for doubt, but rather of our having practical or moral reason *not* to doubt. We have *committed* ourselves to this belief, and we will continue to believe it even in the face of evidence that it might well not be true.

Most people receive most of their moral convictions from traditional sources rather than deriving them from personal deliberation. A very few religious people create their own theologies, some are thoughtful converts, but the vast majority are people who grew up believing what their parents believed, and that these beliefs were virtuous, and made their community and way of life worthy of loyalty, all as a matter of rational perception. One way that the rationality of traditional believers has been respected in cosmopolitan societies is that people raised in pacifist sects like Quakers have been exempted as conscientious objectors from serving as soldiers in war. It makes sense that people with overwhelming second-order reason to believe that fighting is immoral ought to have their pacifism respected more than people who disapprove of fighting only as a matter of how things seem to them at present on a first-order basis, since the secular authorities responsible for conscripting an army are enforcing a judgment by the entire political community that war was sometimes justified, as in the present case, and people cannot be allowed to override their social duty with a mere personal opinion to the contrary, any more than they can avoid paying their taxes or be allowed to throw things at passersby just because they personally think they should. But in the Nineteen Sixties there were a number of legal challenges to the traditional criterion on the grounds that it discriminates unfairly against non-religious pacifists views, the ultimate result of which was that person could now qualify for conscientious objector status if his non-religious objection to all war was a “sincere and meaningful belief occupying in the life of its possessor a place parallel to that filled by the God of those admittedly qualified for exception”, and not a “merely personal moral code”.[[238]](#footnote-238) This ruling acknowledges the difference between levels of personal commitment without making it very clear in practice how we should distinguish sincere and meaningful beliefs from mere personal codes.

Convictions are ordinarily, often emphatically expressed in categorical form: “this is *true*”; “this I *believe*”; or “this I *know*”, although perceptive certainty is plainly out of the question. As long as the initial first- and second- order evidence is strong enough for us to make the judgment, we will treat the judgment as a total certainty, even if along the axis of perception we still know that it might be false. Here the commonsense theory that belief is just one thing seems to fail us, and we’re tempted to say awkwardly that something is certain if it’s certain *enough*, even if it isn’t really *certain*. The idea of accepting a conviction on the basis of any threshold probability under 100% produces an odd syllogism, paradoxical if taken at face value:

(1) *p* is 90% certain.

(2) 90% certainty meets my standard for conviction in this situation.

(3) Therefore, *p* shall be my conviction.

(4) Therefore, I must act (and speak) as though *p* were 100% certain.

(5) Therefore: “*p* is 100% certain.”

It looks even worse if we substitute for (4) and (5) the following:

(4a) Therefore, I must *believe* that *p* is 100% certain.

(5a) Therefore, p is 100% certain.

I don’t want to represent this as a formal paradox, but rather as a probabilistic ambiguity that I’ll refer to as *uncertain certainty*. This occurs whenever the strength of our moral beliefs outstrip their perceptive basis in ordinary first- and second-order evidence. In such situations, which are very common for morally active people, the same belief has less than 100% probability in the dimension of perception but total certainty in the dimension of conviction, and reasonable people won’t be able to avoid a kind of double-mindedness as a result.

We see this sort of ambivalence in most ordinary, non-fanatical religious believers who commit themselves to doctrines as a matter of faith, like the young boy speaking to his mother at the end of Chapter 4. For secular examples, go to a political rally and try to get people to tell you what they think the *probability* is that they are on the right side.

Protestor: Stop the war![[239]](#footnote-239)

Philosopher: Excuse me. Are you saying that you believe the war is wrong?

Protestor: Yes, I am. Good guess.

Philosopher: Then let me ask you a question. How *likely* would you say it is that the war is wrong?

Protestor: What? It just *is* wrong.

Philosopher: Well, have you ever made mistakes?

Protestor: Sure.

Philosopher: And have you ever had reason to change your mind on an issue that you felt very strongly about?

Protestor: Of course. For example, I used to be pro-war; now I am convinced the war is wrong.

Philosopher: Good. Now, can you be certain that this isn’t one of those times when you’re convinced about something, but turn out to be mistaken?

Protestor: I am certain that this war is wrong.

Philosopher: Yes, but since you have changed your mind in the past when you had the same feeling of certainty, you must understand that this is not something you can predict with perfect confidence.

Protestor: All right, then I guess there is some small chance that I will someday find reason to change my mind about this war being wrong. But I hope I never do, because the war *is* wrong.

Philosopher: Okay, but how much of a chance? Given your track record, what is the inductive probability that you will end up changing your mind about an issue of this general sort?

Protestor: I don’t know; maybe ten percent or something.

Philosopher: Then you can only be ninety percent certain that this war is wrong.

Protestor: Well, that’s obviously false, because I am a hundred percent certain that this war is wrong. I can assure you of that.

Philosopher: Then aren't you just being irrational?

Protestor: Look, go ask your little questions somewhere else, okay? The grown- ups have work to do.

Philosopher: But…

Protestor: Stop the war!

This isn't meant to make the protester sound ridiculous; he certainly isn't, if the issue is important and the protest is useful. I’m just pointing out that a person can commit to his position as a *moral* certainty without its being a *perceptive* certainty, which means his rationally calculated probability of error is being deliberately ignored for purposes of action. This only seems irrational from a purely perceptively rational point of view. From a practical or moral point of view, i.e. supposing goals other than maximizing momentary true beliefs – such stubbornness might be the most rational possible attitude. Most of the heroes in history were similarly thick-skulled when it came to their convictions, overriding any prior doubts through strength of character. Ask Winston Churchill in 1940 to compute the expected loss to Great Britain of fighting superior German forces to the death, and to compare that to the losses of a comfortable Norwegian-style capitulation. Which is more likely to benefit the British over the long run? How much more likely? And how certain is he of his answer? I doubt Churchill would respond at all, and I am not sure that he even could, given the depth of his conviction that Great Britain must remain free at all costs. Or ask the signers of the U. S. Declaration of Independence: is it really *self-evident* that people are endowed by their creator with unalienable rights? Among philosophers, this Lockean notion of natural rights has never been viewed as self-evidence in the *a priori* sense that its denial is literally inconceivable. Nevertheless the founders *hold* it to be self-evident, despite its evident lack of self-evidence, because it is for them a moral certainty: not the sort of thing on which they would place bets in a casino, but the sort of thing to which they pledge their lives, their fortunes, and their sacred honor.[[240]](#footnote-240)

In the abstract, we accept that all human convictions are fallible, but it seems that as a matter of normal psychology, we find it very hard to doubt our beliefs and act forcefully on them at the same time. So we manage to ignore the implications of the fact that people sometimes adopt false convictions, and that we ourselves are not exempt from this sort of error, whenever we treat a moral or prudential belief as if it were a certain truth. So, Hamlet is always criticized for failing to kill his hated uncle Claudius until he has what he considers evidence beyond a reasonable doubt that Claudius truly deserves to die. In Act III he finally obtains the proof he wants, shouts "Now I could drink hot blood!", and then runs off and stabs his girlfriend's father by mistake.[[241]](#footnote-241) Most people say that Hamlet should have acted much sooner, though on weaker evidence, because his hesitation ends up costing several more people's lives, including his own, by the end of the play. Myself, I don't see him knowing enough any earlier (i.e. solely through the testimony of an apparition) to act responsibly against his uncle. People who kill their relatives on that kind of evidence are usually called psychotic. But acting on imperfect evidence is always a risky business, and even normal people have to make decisions and do things that might not be right if they are to do anything consequential at all.

This is not an argument from pure perceptive rationality, obviously, but it can still be thought of as a matter of prudential or moral rationality. That is, if things have to be done, and if this requires us to fix some of our beliefs in place, then there must be a best, most rational way to do this. So, in treating my convictions as beyond revision, I will need to make adjustments in my total model of the world in order to integrate these fixed beliefs with all my other perceptions in a coherent and effective way. Thus, to be a rational perceiver who also happens to be a Christian (or communist, etc.) is one thing. But to be a committed Christian who also happens to be rational is something else entirely. If we call someone a rational Christian, we might mean either of these two things, depending on whether we see him as fundamentally an all-purpose perceiver or as fundamentally a Christian. In the first case, we would take the person’s rationality as a fixed principle and expect him to be open to new evidence and arguments, and to renounce his Christianity if new information contrary to Christian doctrine is sufficiently convincing. In the second case, we’d take the person’s Christianity as the fixed-point, and expect him to adjust all of his other beliefs, including new perceptions, to cohere as well as possible with his Christian convictions. This is no different in principle from the Constitution-bounded reasoning of a Supreme Court Justice or the theory-bounded reasoning of a scientist working within a given paradigm, except that those people are playing professional roles that they could in theory shrug off in their personal lives without betraying their personal convictions, while the committed Christian must be committed through and through if he is a person of integrity.

Even the people most concerned with rational belief rather than action must follow a sort of constitution in order to execute long-range programs of study. Philosophers commit to graduate study with firm beliefs about the value of the undertaking, or else they are more likely to drop out. They work on their dissertations in the firm belief that they will be completed and accepted, or they are less likely to finish. If they are able to find jobs, they must be convinced that they will merit tenure if they work hard enough and that their institutions will evaluate them fairly, or they will not be willing to spend seven years cranking out publications, ingratiating themselves with colleagues, and spending much of their remaining time on tedious committee work. More importantly, they must believe that philosophy is a good and useful career and that they can pursue it with a measure of success, or it would make more sense for them to go to law school instead. It is true that people sometimes change their minds about these things in mid-stream. But this typically happens not through neutral Bayesian adjustments in their matrix of perceptive probabilities, like a skillful gambler deciding to fold a hand of cards, but through a more or less traumatic loss of faith, like people long unable to maintain their marriages finally “facing the facts” and filing for divorce. Up to that critical point of failure, philosophers, scientists, activists, husbands, and wives still function rationally in the prudential or moral sense, constrained by their conviction that things will work out best if they “stick with the program”.

It’s not immediately clear how this prudential kind of conviction is even possible. Ordinarily, we can’t just make ourselves believe whatever we want to believe, just because we think that those beliefs are good for us. I can’t believe my 2001 Mazda is a brand-new Mercedes just because I would prefer the Mercedes. I can’t believe that I’m a better tennis player than I really am, although that belief might actually make me a better tennis player. So, how can I make myself believe with certainty that my marriage, my country, and my principles are categorically right and good, ignoring all the evidence I have that these are all very contingent things? William James argues in “The Will to Believe” that there are some things that we cannot believe given all of our experience, some things that we must believe, and some matters where we have what he calls a genuine choice.[[242]](#footnote-242) Within this area of choice, on important questions where we must believe one way or another, where we feel some attraction to both options, and where evidence alone fails to settle the matter, belief is a matter of *will*. The case that James cares most about is belief in God. If you are not lucky enough to have believed implicitly in God since childhood, but belief in God is nevertheless a live option for you, he thinks it makes the most practical sense to choose this belief. Otherwise, you will lose any hope of gaining the moral strength and psychological contentment that Christian belief seems plainly to confer on its adherents (not to mention hope of eternal life, which some people think would be a good thing). After all, he says, the central purpose of inquiry is not to avoid false beliefs but to seek true ones. In matters of genuine choice, belief and lack of belief are on a par with respect to the evidence. Therefore, non-evidential considerations may reasonably tip the scales one way or the other.

Nietzsche takes this idea a step further, arguing that moral virtue requires not just a will to believe, but sometimes a “will to stupidity” as well, the more so as our projects and responsibilities require effort over longer periods. On Nietzsche’s view, you have to willfully ignore reason and evidence sometimes if you are ever going to exemplify such ancient virtues as courage and nobility. Achilles is a greater, not a lesser, hero because he hates his enemies and won’t listen to reason from them. There is a passage in the *Iliad* where, just before their final combat, Hector asks Achilles to agree that the victor will treat his foe’s dead body with proper respect, and Achilles yells at him that this attempt to make deals with an enemy during battle is “unforgivable”, that military valor demands that warriors should act like lions or wolves, not men.[[243]](#footnote-243) In the end, it is not from any counter-argument, but only because Hector’s pleading father Priam has reminded Achilles of his love for his own aged father that Achilles finally gives the old king back the desecrated body of his son. It is this very Nietzschean stupidity, his acting on refined impulses rather than calculating right and wrong according to any theory, that allows Achilles to express his *arête*, his excellence as man and warrior.

Willful stupidity serves not only to integrate our own active and epistemic commitments, but also to communicate those convictions to other people. One way to get another person to cooperate with you is to convince him that your own position is totally inflexible, so that there is no point in trying to persuade you otherwise. This sort of strategic rigidity is said to have served a useful function during the Cold War under the doctrine of “mutually assured destruction”. Here is a statement of principle from both sides: “If you attack us with a first strike on our military bases, it may well be in our interest at that moment to capitulate immediately rather than launch an all-out counterattack that will result in the destruction of both countries. But we are not going to be rational about it if you strike us first. We are going to kill as many of you as we can with our remaining forces, even though this will provoke a devastating second strike from you. So, if you want to try a nuclear strike you had better kill us all, because we will destroy most of your country even if you win the war.” It was seen as essential for the US to convince the USSR that we might well respond in this irrational way in order to deter them effectively from actually striking first, and this involved our actual willingness to go through with the irrational response. Thus, in a sense, our strategy was rational and irrational at the same time. The ideally rational approach would be to try to convince the other side of our irrational commitment to total war in the event of a first strike, but then, if this deterrence fails and we receive a first strike anyway, to capitulate instead of keeping our word and actually launching an all-out attack. That way we could salvage what was left of our country, and presumably some future chance to rise again to power and prosperity. The problem with this ideal game-theoretic sort of strategy is that other people are too good at finding out our actual intentions. If we really want to convince a canny opponent that something is a matter of principle with us, then, the most effective way is to make it a real matter of principle. So, when bluffing can bring about disaster if it gets discovered, it may be best for us to become rigidly irrational in fact, just in order to make sure that our adversary does not *believe* that we are bluffing.[[244]](#footnote-244)

Such strategic communication is an essential function of convictions in ordinary life as well. If I want to convince other people to be confident Christians (or socialists, etc.), then I need them to believe that I am a confident Christian (or socialist, etc.) myself, and the most effective way to do this is for me to be genuinely convinced. If ignoring contrary evidence helps me attain that useful confidence, then let me ignore the evidence. It is the central principle of "method" acting that the actor will be the most convincing if he "inhabits" the character he plays from the inside, as much as possible *becoming* the character rather than merely pretending. The same is true of courtroom lawyers; the best way of convincing a jury of your client's innocence is to believe in it yourself, and some great lawyers have an amazing ability to do just that, even for defendants whose guilt is absolutely clear to everybody else.[[245]](#footnote-245) I find it true even in teaching philosophy. If I want to explain a theory to my students, the most effective way is for me to *believe* that theory while I'm explaining it, which means adopting temporarily as much as possible the whole perceptive model of a person who sees things that way, then generating new articulations on my own rather than trying to remember other people's arguments by rote.

7.3 Higher-order convictions.

Willful stupidity can lead to *willful* *ignorance* when we have reason to fear that learning more about something will weaken our convictions. Some religious believers, like Augustine and Aquinas, relish theological debate, confident in their ability to defend their faith against all comers. But many would rather not be confronted with arguments that threaten their beliefs, and still others, like Tertullian, are actually proud of their indifference to objective evidence. The same has been true for many Western communists who seem perversely to have turned their eyes away from horrors that were obvious to less committed sympathizers, especially during the purges under Stalin in the 1930’s.[[246]](#footnote-246) Even in the practical situation of buying a car, I would rather not hear about better deals on better cars than the one I that have just decided to buy. I am quite satisfied with this little Mazda, and I don’t want to regret having purchased it. So, in order to stay convinced that I did the right thing, I find myself avoiding information about potentially more optimal choices. Making a religious or political commitment is not buying a car, but the same sorts of factors are involved. If you are happy with your convictions, if they work for you, then you might be better off – not epistemically, but as a total human being – in forestalling any further doubts about them, while keeping an eye out for further support and confirmation.[[247]](#footnote-247) In politics, partisan voters are likely to become experts on the scandals of the other side, while avoiding too much knowledge of the “so-called scandals” on their own side lest it dampen their zeal. In recent American politics, this has been visible in a long string of scandals that are exhaustively reported in one party's sources of news, while the other side offers little or no coverage of the issues apart from attacks on the opponents for making (vaguely described) false charges. People on the left have been exquisitely well-informed about such things as the Watergate, Iran-Contra, and Plame-Wilson scandals, while those on the right have been immersed in Chappaquiddick, Whitewater, and Barack Obama’s birth certificate. The people on the defensive side of these issues often seem to have little or no idea what their opponents are talking about, and are reluctant to learn any more. A few presidential election campaigns ago, for example, one academic friend of mine constantly argued against the fitness of one candidate based on his unimpressive college record, while refusing to acknowledge that his opponent’s known record was equally mediocre. The latter fact hadn’t been mentioned in my friend’s sources of information, he said, and he certainly wasn’t going to go out of his way to learn bad things about his favored candidate, so he would "stick with what I know" and keep repeating the same one-sided argument.

Once again, I don’t want to make something look ridiculous that can be deeply serious. A prominent moral philosopher once told me explicitly, and with a kind of grim pride, that he wanted not to know any facts that he would have to count as evidence against his belief in absolute human equality with respect to race and sex. He had made a deep moral commitment to oppose racism and sexism, he said, and, while for purposes of purely philosophical discussion he was willing to admit that there could be some kind of empirical basis for unequal treatment, he believed that virtue demanded that he shut his eyes and ears to any particulars that might weaken his resolve – hence his refusal to read books like Wilson’s *Sociobiology* or Murray and Herrnstein’s *The Bell Curve*. This is not necessarily an unreasonable position; in fact, I find I rather admire it, sympathetic as I am to this philosopher and to his social goals. Given that he is already thoroughly convinced of the overall *moral* equality of all human groups, he finds the proposition that we are also totally undistinguishable other than by obvious physical characteristics to be an effective, simple, and comprehensible empirical basis for concerted political action. Even if it turns out to be false, it surely isn't *very* false, and it is not worth fussing around experimentally to get all of the subtleties exactly right when that is liable only to encourage racists and sexists to believe that they have science on their side. Better to stick with a clear, simple theory that works politically and morally for now, and we can discuss little details about our supposed differences at some point in the future, once a reasonable level of social justice has been achieved. To me, this will to ignorance is no different from that of a person who believes deliberately in God, heaven, and hell because he believes implicitly in Judeo-Christian ethics. A lot of Jewish, Christian, and Muslim believers seem to be like this. They don't think that their theological doctrines are just a bunch of myths; they really *believe* them in the sense of having committed themselves to live according to them, which includes standing up for them *sincerely* in discussions. But they would rather not get into theoretical debates about God's perfection and creation of the universe, or what precisely happens to us when we die, if they can avoid them. What they want at bottom is to live their lives as good Jews, Christians, or Muslims, and they take their theology simply on faith, with little inclination to pry into matters that could only shake that faith. Why should they be tempted into error and sin out of mere intellectual fastidiousness, when they can be strong and good as persons of faith and moral integrity? It is hard not to admire people like this when you know them well.[[248]](#footnote-248)

The obvious problem with having a world full of admirable people of conviction is that different people have different convictions. The convinced social progressive and the convinced religious conservative do not just disagree perceptively over such issues as abortion and same-sex marriage. Their positions are matters of principle for each, part of a system of beliefs that they identify as bedrock features of a good and decent person, which implies that their opponents are not good and decent people. What makes things worse is that a hostile “diagnosis” of opponents in such disagreements may become as much a matter of conviction as the beliefs over which they originally disagreed. If we become convinced that some person or group of people is emotionally or intellectually incapable of understanding the truth about an issue, then there is no longer any reason to argue with them, other than perhaps to show third parties what is wrong with our opponents or their arguments. If we become convinced that other social groups are irremediably bigoted, unpatriotic, or otherwise morally unable to accept the plain facts, then it becomes not just pointless and frustrating but positively indecent to try to reason with them. It may seem that such *second-order convictions* are foolish and always avoidable, but consider trying to have a reasonable discussion with an actual Nazi intellectual (there were plenty of them) about the vices of European Jews. We decent people know implicitly that Jews are human beings equal to everybody else. We know that the Nazis deny this fact. We know further that this isn’t a subtle or elusive fact, but something that should be obvious to every other decent person. So we see it as a matter not of opinion, but of the firmest conviction, that Nazis are not decent people. We do not want to bother arguing with Nazis about racial distinctions, because we know in advance that anything they say is going to be a lie, a sophistry, or a one-sided anecdote about some Jewish person doing something wrong. Furthermore, we would prefer not even to be tempted by any supposed evidence of the Jews’ iniquity into doubting something that we now take as morally certain. We do not even want to be seen speaking with Nazis or to give them any kind of public platform for debate, since this would make the full humanity of Jews and other non-Aryans seem like a matter about which people can reasonably differ. In a community under conditions of war, this is not intolerance of different views but a quite reasonable doctrine of epistemic self-defense. But even if it were only a personal conviction that Nazis ought not to be treated as reasonable people, this principle would have the same willfully stupid effect as an authoritative epistemic doctrine. And we admire people who refuse to listen to Nazis all the more when they are by themselves in this and liable to be treated badly for it.

This sort of moral reasoning can reach a third level, and a fourth, and as many others as we like. If the wrongness of speaking to Nazis itself becomes a matter of principle, on top of our convictions that Jews are equal to Aryans and that Nazis are not decent people, then we might well wonder what is wrong with people who disagree with us about the awfulness of speaking with Nazis, even though they seem to agree with us about the awfulness of Nazis themselves. We could view this third issue as just a matter of opinion: our friends who speak with Nazis are perfectly decent people; they just happen to see this particular question differently. But we could also take it as a new, third-order conviction and stop speaking to people who even speak with Nazis.[[249]](#footnote-249) As to how it is possible that some people are decent enough to agree with us on the humanity of Jews and the wickedness of Nazis, but not decent enough to agree on the wrongness of speaking with Nazis, one possibility is that they are lying or self-deceived about the first two things, and consciously or unconsciously they’re *siding* with the anti-Semites, which makes them "objectively" anti-Semitic themselves. In that case, we should not allow ourselves to speak with people who speak with Nazis any more than with the Nazis themselves. Now this third-order principle of eschewing those who don’t eschew the Nazis comes under a similar review: if we take this principle as just another moral fact that should be obvious to every decent person, then we will have a new, fourth-order conviction preventing us from speaking with an even larger class of people who speak with those who speak with Nazis. The potential result of this sort of process is a total polarization of discourse, in which even the most moderate-minded people are precipitated into one camp or the other by the refusal of those in the opposite camp to treat them as decent human beings. This seems to be about what happened when our faculty reading group broke up over *The Skeptical Environmentalist*.

This recursive tendency to form increasingly higher-order convictions can become a habit, even seemingly a matter of conviction in itself. I recently tried to organize a little debate at my college over the Shakespeare authorship question, where I would argue for the candidacy of Edward de Vere as well as I could and someone from our English department would defend the orthodox view, probably pounding me into the ground since I have no real expertise about the subject. I thought it would be fun for us and an engaging sort of discussion for the students, and I was surprised by the response I got from my three colleagues who teach Shakespeare, all friends of mine and generally reasonable people. Two of them said that the idea was just too silly to discuss in public, since the only people who could possibly find the Oxford hypothesis remotely plausible are "conspiracy theorists" and snobs unwilling to believe that commoners can write great poetry. The third, then chair of English, initially agreed to the debate but later changed his mind, explaining that he had decided that the discussion would harm our students by suggesting to them that there were two respectable sides to the authorship issue when there is only one, as all the qualified experts agree. I asked him if it wouldn’t be good to let the students decide this issue for themselves on the basis of the best arguments for each side, since they are bound to hear about the controversy sooner or later, and “better from us than from other children on the street”. But he said no, that actually wouldn’t be a good thing, because our students have too much respect for their professors to resist being unduly influenced by the mere fact that arguments are being made on both sides. In this way, my friend seems to have raised to the level of high-order conviction something that most people would agree is, if not properly a matter of opinion, then at most a rather inconsequential question of fact.[[250]](#footnote-250)

People enter into disagreements carrying not just different perceptions, opinions, and moral principles, but also different loyalties. It's not just our substantive convictions that are at stake in matters of integrity; it’s also the communities that we believe we are obliged to represent. So, I am not just a person who believes in Jesus, I am a *Christian*; I am not just a person who believes that women and men ought to be treated equally, I am a *feminist*; I am not just a person who believes in conservation, I am an *environmentalist*; and so on. Membership in such groups adds epistemic gravitation to our existing perceptive and moral hesitations about taking contrary evidence to heart, and sometimes even makes refusing to listen to opponents a criterion of membership.[[251]](#footnote-251) Under these conditions, disagreements that in principle could be resolved on a first-order basis turn into battles between hostile communities, and rational debate between mere individuals sometimes becomes impossible.

So, you go to college, and your best friend goes to college somewhere else, and for no particular reason (maybe it's required at your school) you take a freshman course in economics, while he takes a lab class in ecology. And these are the best classes that you each take in your first year, with brilliant, passionate professors and cool, helpful teaching assistants, and you're good at the stuff and understand it easily and impress the teachers and end up falling in love with the material. You really love economics and he really loves ecology and you're both thinking of majoring in these fields, which is all well and good. But studying economics and getting into economics and making friends with economics TAs and joining the economics club and so on, you are liable (depending on your school) to find yourself reading a lot of Friedrich von Hayek and Milton Friedman, and discussing their writings with smart people who really like them, and also learning what your new friends think about the causes of our present economic problems, and what they think is wrong with all the people on the other side, and finding that stuff plausible as well. So now you have this picture in your head of how supply and demand interact, and how market pricing orders the distribution of scarce goods efficiently, and how economic freedom leads to economic growth which leads to general prosperity which leads to medical advances and personal computers and ultimately safer cars, and how minimum wage laws increase unemployment for minority teenagers, and how education really ought to be privatized because then competition would produce the best techniques and structures and lazy teachers could be fired if they didn't do their jobs. And it all…just…fits…together in a clear, coherent model of how the world operates, with strongly implied policy changes that could fix the whole country if only the bone-headed liberals would get out of the way and let investors do their job so businesses could grow and hire employees and produce things that we need without environmental extremists breathing down everybody's neck and trying to ban air as a pollutant, because what they really want is just to run the world themselves because they think they're so much smarter and better than everybody else and look on businesspeople as beneath them socially, and that's why the liberals hate the so-called rich, though most of them make a pretty damned comfortable living themselves working for government agencies their whole lives without risking anything at all, because all they have to do all day is spend other people's hard-earned money and congratulate themselves for being "dedicated public servants".

Meanwhile, your friend is making friends with different decent and trustworthy people, reading different books and magazines and blogs, joining different organizations, and taking different stands. And he learns that capitalism places profits above people's needs, except for the needs of the investors who own all the mines and oil companies and other corporations that are happy to damage the environment whenever doing so will make them money. That's just how the system works. And the only force that can possibly stop this destruction is government regulation, but the regulators usually go back and forth into the corporate world themselves, so outside activists are needed to prevent the system from being thoroughly corrupted by self-interested parties making cozy little deals to line their pockets. Not to mention the immense amount of corporate money that supports political campaigns, and not just Republicans either, but they're also able to buy off enough of the Democrats to weaken regulations to the point where the environment is only being destroyed a little more slowly. And he learns from his trustworthy sources that most of the exploitable oil in the world has been exploited with production just about to peak, after which there will be a sharp crash in oil supplies that the corporations and their corporate-sponsored governments will be unprepared to handle because all they care about is next year's profits and next year's elections. And he learns that biodiversity is essential for the future health of the planet, and again that profits sometimes must give way to regulations preserving biodiversity even if they cost a few people jobs right now, because posterity depends on our taking more responsibility upon our shoulders as world citizens than people ever did back when the planet was less crowded, less industrial, less greedy, and less polluted. And it all…makes…perfect…sense of why the world has gotten so screwed up and why it's been so hard to fix. And we know that the solutions to these problems will be difficult and take a lot of fine-tuning as we go along, but we can handle these problems without wrecking the economy if all these right-wing crazies on talk radio and Fox News will just stop lying to everyone about the government and the environment and let us fight the bastard corporations on a *somewhat* more level playing field.

And then you get together during the summer or just after college, or maybe after thirty years, and somebody brings up some issue like affirmative action or global warming, and ten minutes later both of you are angry. Neither accepts the other's premises or shares his assumptions. You don't know the same facts, trust the same newspapers and other sources, or even take the same sorts of evidence and arguments seriously, and you're certainly not going to take your friend's word now for anything he says. You have learned to view his views as arrogant and motivated by a thirst for arbitrary power and a mindless antipathy for the modern world's only successful way of life. And he has learned to view your views as ignorant and motivated by racism, greed, and mindless loyalty to a system that makes people like you comfortable while destroying the world for everybody else. It is painful for each of you to see how *blind* his old friend has become to *basic* moral and empirical truths. How could someone you always knew as decent and reasonable have gotten *brainwashed* into these crazy beliefs? In the end you are very sorry, and old friends are old friends, but enemies are also enemies, and there's only so far that you can bend your principles in even talking to people like that.

Now it seems that two basically decent and reasonable people have come to believe in different models of the moral and empirical world, in a way that is supported and partly defined by objective evidence and reasons, but not determinedby them in anything like a step-by-step Cartesian process. Instead, from an external point of view, it looks more like we have fallen deeply into different gravitational fields, and it looks pretty arbitrary which of the two fieldseach of us has fallen into. It seems not to be a measure of the individual's inherent rationality or personality or decency what he'll believe when pushed to takes sides on current controversies, but mostly a matter of luck, where some initially small, even random difference of intellectual interest or attraction or acquaintance has developed through socioepistemic feedback into a hardened position, hostile to enemies and well-defended against contrary arguments. From an inside perspective things will look entirely rational, of course, for each of us, with core empirical beliefs seeming like clear and distinct perceptions, core moral principles as nothing more than common decency, and our major opponents as sick, stupid, crazy, or wickedly selfish. But right now, from an outside perspective, even if you ordinarily accept one of these paradigms as true, the fact that at some point in college you more or less just drifted into yours like your old friend just drifted into his should give you pause. If people's convictions, including your own, seem to have been arbitrarily assigned to them, then clearly something is wrong when everybody feels so certain that they are in the right. Even if the environment or the economy desperately needs to be saved, something is wrong when people on both sides are, with equal subjective epistemic justification, unable to perceive their opponents as reasonable.

It would be nice, of course, if we could all just work things out together from a neutral and objective point of view, and in the abstract this seems perfectly possible. Indeed, this is what most intellectuals and scientists believe that they are doing, together with the other objective and rational people. But we are not just thinking beings; we are also active human beings with convictions, and our core convictions influence who qualifies for each of us as an acceptably decent and rational partner in discussion. So, fine, we are inclined to say, we really should agree to work our disagreements out objectively with other *reasonable people*. But surely, this doesn’t mean we have to beat our brains out trying to reason with the *left-wing lunatics* or *racist morons* on the other side of our most serious disputes.

I am tempted, as philosophers often are in situations like this, to argue for a skeptical approach of some sort, for example to say that no one should ever be committed to any theory to such an extent that any opponents should be excluded from discussions on moral grounds. But this seems too inclusive, for surely some opponents are sometimes beyond the pale, for example Nazis, or let’s say the Devil himself. But it seems impossible to impose any specific limitations at all without ruling out the very sorts of cases I would like us to rule in. So, I can’t just say that we should be willing to argue fairly with any intelligent or reasonable or decent or psychologically normal opponents because our moral paradigms make us too likely to diagnose any serious opponents as failing even these minimal tests.[[252]](#footnote-252) So, it seems that we are stuck with trying to balance on the fly our epistemic obligation to consider all points of view objectively with our moral obligation not to make wicked and crazy theories seem respectable. As long as we are active people with convictions, we have no choice but to allow those convictions to color our perceptions about our opponents and their own convictions. As we see in the controversy over climate change, the more urgent and more drastic is the action that our own convictions lead us to demand, the less tolerance we *ought* to have, according to the same convictions, for the statements and actions of our opponents. Even if we have a pre-existing intellectual conviction as philosophers or scientists that no theory is final and minority positions should never be ruled out, this must always be balanced against other equally sincere convictions, just as Steven Schneider the climatologist said. Our only other choice is to make rationality or objectivity our only value, or a value that trumps absolutely everything else – but this makes no sense when rational people disagree and hard decisions still have to be made. The best that we can reasonably do is to *maximize* our objectivity and inclusivity to the extent that we can do so consistently with all our other duties.

7.4 Autonomous integrity.

Nietzsche's will to stupidity is a feature of his concept of the *übermensch* or *superman*, the person who embodies total autonomy and integrity at the same time. Nietzsche argued that an active life lived according to self-direction, noble integrity, and a manly craving for power is superior to either a life lived according to common values or a life lived in the pursuit of knowledge alone. The superman is someone who imposes his own will on the world as a leader, rather than relying on moral testimony as a follower. The superman is not a philosopher but a hero, someone who lives in the real world as a man of action like the great Achilles, Cesare Borgia (Machiavelli's favorite heartless conqueror), or Napoleon. It is not that Nietzsche wanted people to be stupid or thoughtless in general; indeed, the more intelligent the better. But the intelligence involved had to be *instrumental*, not a replacement for action. Reason should be a slave to instinct, not the other way around as Plato, Descartes, and most "decadent" philosophers would have it. But the instinct should not be absolutely lawless, like the id of a psychopath. It should be the product of natural drives together with a basic upbringing in noble values. Thus, when Achilles shouts at Hector that he wishes he could tear him apart and eat him like a lion would – for all that, he is still a man: speaking, not roaring, and expressing a code of war, not simple rage. And when he gives Hector's body back to Priam, he does so weeping with remembered love for his own father and noble sympathy for the great king who kneels before him. These are not the actions of a stupid man, but of a splendid one.

Nietzsche was not completely right about the decadence of modern values, even by his own criterion. We do still admire people who drive themselves forward, engaging with the world according to their instincts and convictions, provided the instincts and convictions are not too selfish or tyrannical. We don't expect George Washington or any modern president to be a philosopher or even a very reflective sort of person, but rather a strong leader with his heart in the right place and enough intelligence to see how things need to be done. Still, even Washington falls well short of Nietzsche's ideal, concerned as he was for the good opinion of his fellows, and constrained as he was by the anti-instinctive values of Christianity and the Enlightenment. We don't fully admire anyone who comes much closer than Washington does to Nietzsche's ideal. However moving we may find Achilles fighting Hector to the death in his personal quest for glory, we do not in real life take that sort of thing to be acceptable behavior. Faced with actual Greeks and Trojans on the point of war, we would want them to work their problems out in reasonable ways, coming through negotiation to some sort of agreement as to where Helen should reside.[[253]](#footnote-253) Indeed, *any* agreement that would save thousands of lives and prevent the sack of Troy would be better by our lights than what happened in the *Iliad*. This is the way we think, and it was probably the way that people other than aristocratic warriors thought in Mycenaean times as well. It may be decadent, but we have plainly passed the point of no return in our preference for peace and general well-being over the amoral will to power, at least in real life as opposed to movies and video games. We modern people live far better lives, all things considered, than people used to live when just a few of them ran the world and everybody else was stuck doing their bidding. And we – most of us, anyway – value intellectual, social, and scientific progress too much to desire a return to unreflective, pre-Socratic ways of life. As for the will to power, we have probably seen enough slaughter in the last hundred years from movements influenced by Nietzsche, left and right, to keep us suspicious for another little while.[[254]](#footnote-254)

There are in any case milder, less aristocratic versions of the ideal of creative independence than Nietzsche's. Earlier in the nineteenth century, Wilhelm von Humboldt, John Stuart Mill, and the American transcendentalists Ralph Waldo Emerson and Henry David Thoreau all advanced similar ideas of radical autonomy without including Nietzsche's harsh, amoral will to dominance. Humboldt argues that human nature is best expressed not in sheepish conformity but in “originality”:

…reason cannot desire for man any other condition than that in which each individual not only enjoys the most absolute freedom of developing himself by his own energies, in his perfect individuality, but in which external nature even is left unfashioned by any human agency, but only receives the impress given to it by each individual of himself and his own free will, according to the measure of his wants and instincts, and restricted only by the limits of his powers and his rights.[[255]](#footnote-255)

This is not a call to absolute autonomy, given Humboldt’s restriction to activities within one’s rights, but only the general claim that *maximal* autonomy within moral constraints is required for ideal human happiness. Mill states more explicitly what he thinks these constraints ought to be in propounding his so-called “harm principle”:

The liberty of the individual must be thus far limited; he must not make himself a nuisance to other people. But if he refrains from molesting others in what concerns them, and merely acts according to his own inclination and judgment in things which concern himself, the same reasons which show that opinion should be free, prove also that he should be allowed, without molestation, to carry his opinions into practice at his own cost…As it is useful that while mankind are imperfect there should be different opinions, so it is that there should be different experiments of living; that free scope should be given to varieties of character, short of injury to others; and that the worth of different modes of life should be proved practically, when any one thinks fit to try them.[[256]](#footnote-256)

Though he agrees with Humboldt that people are happiest when their self-development as individuals is least restricted, Mill also emphasizes the advantages to all of society when people are allowed to act on their opinions freely. This is the best way, he says, of discovering from a vast range of possible ways of life which are the best. Just as we permit the free expression of false opinions in the expectation that true ones will emerge from open debate, we ought to permit diverse “experiments in living” in the ultimate expectation that those that make people the happiest will prove their worth in competition against other experiments as well as against established traditions. Thus a kind of epistemic altruism is possible even for people who live selfishly in the usual sense, since they are at least adding to the pool of moral evidence available to all. The kind of commitment appropriate for the experimenter to his “active opinion” is like that of a scientist to his favorite hypothesis: to the extent that he is rational, he’ll see it as a speculation liable to be proven false; to the extent that he is useful and effective, he will tend to see it falsely as a certainty and take the consequences.

The American transcendentalists are even more attached to moral autonomy than Humboldt and Mill, demanding freedom not just from oppressive social restrictions but from any public morality at all. For Emerson, as for Humboldt and Mill, the point of independence is not to be limited by social rules in finding whatever personal greatness you are capable of, but he goes further in encouraging the striving individual to break whatever rules get in his way:

Whoso would be a man, must be a nonconformist. He who would gather immortal palms must not be hindered by the name of goodness. Nothing is at last sacred but the integrity of your own mind.[[257]](#footnote-257)

This seems to imply that it is permissible for us to harm others in the pursuit of our own individual genius, though Emerson does not confront this question directly, and elsewhere argues that violence is unproductive. For Thoreau, the main point of moral autonomy is that we and the world would be made better through the radical reliance on individual conscience, even when it is opposed to democratically enacted law:

Can there not be a government in which majorities do not virtually decide right and wrong, but conscience? Must the citizen ever for a moment, or in the least degree, resign his conscience to the legislator? Why has every man a conscience, then? I think that we should be men first, and subjects afterward…The only obligation which I have a right to assume is to do at any time what I think right.[[258]](#footnote-258)

If this kind of language is less stirring today than it was for Emerson’s and Thoreau’s audiences on the eve of the Civil War, it is only because these radical ideas have become so familiar as to strike us little more than moral common sense: *of course* we all ought to be totally independent in deciding what we think is right, and of course we should be totally committed to acting according to those decisions. In a way, it is perfectly trivial that people ought to do whatever they believe is right. We all know that sometimes these beliefs will turn out to be false, but still, as individuals we have no choice but to follow our own best judgment; the only alternatives are doing nothing and doing what we think is *wrong*. But what does it really mean to think something is right if this principle is going to make sense? Surely not just to *hypothesize* that it *might* be right. It can’t just be an opinion, in other words, because those are generally speaking unlikely to be true, and we don’t want people just springing into radical action every time some notion strikes them as a good thing to do. Clearly, both of these writers are talking about those beliefs that I am calling convictions, ones to which you have become committed to as moral certainties, so that you see them as part of your identity. But where do they imagine that we gain our convictions in the first place, if not from listening to our parents and other authorities? According to Emerson, we are supposed to find them for ourselves, deliberately rejecting moral lessons from other people so as to enhance our special individuality and avoid the sheepish conformity that he sees as wasting most people’s lives. And this notion has considerable appeal in matters of art and music and other areas where one individual’s creative flowering can do little or no harm to anybody else. But Emerson seems to be asking for more than this, that we also transcend common notions of right and wrong in all other spheres of action, too, because the flowering of individuals is more important even than morality. Thoreau goes even further in claiming that a radical individuality of conscience actually makes us morally superior to other people; that if we just ignore moral testimony from others we will end up doing the right thing, while conformity will make us do things wrong.

Can Emerson and Thoreau really mean that we should give up learned morality altogether and rely instead on something like raw moral intuition? It seems that Thoreau means it to be taken quite literally when we read what he says about the radical abolitionist John Brown, who in 1859 sacrificed two of his sons and a number of other followers in a hopeless attack on a federal armory at Harper’s Ferry, Virginia, and was hanged for murder and treason as a result. Thoreau sprays abuse at his many liberal contemporaries who considered Brown a fanatical crackpot, and praises him instead as a Christ-like figure of morally perfect self-reliance and integrity, a sort of noble savage in the form of an American pioneer:

"Misguided"! "Garrulous"! "Insane"! "Vindictive"! So ye write in your easy-chairs, and thus he wounded responds from the floor of the armory, clear as a cloudless sky, true as the voice of nature is: "No man sent me here; it was my own prompting and that of my Maker. I acknowledge no master in human form.”[[259]](#footnote-259)

It seems that Thoreau expects the “voice of nature” to speak just as clearly through anyone whose natural moral sensibilities have not been corrupted by conventional morality. But why should anybody think this way, other than through hostility to particular, existing moral conventions plus a vague optimism about what will happen once those conventions have been overthrown?

Consider what a whole society of people like John Brown would really be like. Without a no-harm principle like Mill’s constraining it, how could such radical autonomy of action be allowed to govern the actions of people living among the elbows of their peers? We cannot count on everybody's non-conformity conforming to everybody else's, can we? If not, what are the limits? What if everybody just decided, like Thoreau, to stop paying their taxes? Or stop obeying traffic laws? Or criminal laws? What are we supposed to do when someone else's non-conformity gets out of hand? If total anarchy is not to be desired, individual autonomy must be restricted in some way. If the creative genius Emerson wants us each to be is living in a society, he'll have to follow rules and judgments made by people in authority. And people are only genuine authorities, as opposed to mere dictators, if they are supported by large numbers of fellow citizens who trust them. So, the ideal non-conformist wakes up and finds himself in a community of some sort. Since he's a non-conformist, he will sometimes find himself in disagreements with most of his fellows, and may be tempted to outrage their laws or sensibilities in matters that he perceives as morally important from an independent point of view according to his own exclusively first-order reasoning. The conformists all tell him that what he is doing is morally wrong, but he will not be "hindered" by those "names", and nothing they say has any real effect on him because the only thing that he holds "sacred" is his own integrity.

What would make such a person an acceptable member of *any* community? Emerson and Thoreau are clearly imagining a highly morally informed and thoughtful person acting in firm resistance to authorities who are corrupt and wicked. But again, can any fallible human being guarantee in advance that he will end up on the right side of such confrontations? Didn’t Hitler, Stalin, Mao Tse-tung, and Pol Pot all *think* that they were being free men with the courage of their convictions? If these people thought and felt subjectively the way that John Brown thought and felt, which it appears they did, then even supposing Brown acted rightly as Thoreau insists, merely “transcending” common morality and standing up for whatever moral notions you come up with on your own cannot be trusted as a general ethical doctrine. At its worst, under this doctrine any whim that passes through an undisciplined mind like that of Charles Manson gains the force of principle. So, thinking for himself and standing up for what he thinks is right, the self-styled revolutionary hero freely slaughters anyone who strikes him as a "pig" for any reason. I don’t know what Emerson and Thoreau would say to people like that after the fact of their horrific crimes, having encouraged them before the fact to do whatever their totally independent consciences told them was the right thing. They simply assert the proposition that autonomous integrity makes you a more admirable person all by itself, and never seem to realize that the same individuality that makes you stronger and better when you conclude with John Brown that slavery must be overthrown will make you more intransigent and worse when you conclude with John Wilkes Booth that Abraham Lincoln must be killed.

I think this issue doesn’t really come up in Emerson’s and Thoreau’s writings because neither really imagines that he is speaking to the entire society, but rather only to a certain class of educated, mostly Northern progressives from whom they expect certain enlightened views. They certainly don’t put it in those terms, viz.: “Northern abolitionists like us ought to be radically independent of Southern slaveholders and those who tolerate them”. Instead, they claim that the right moral perceptions will just be *obvious* to *anyone* unclouded by the superstitions of the rabble and unbrainwashed by wicked institutions like slavery and imperialism. But they are not anticipating an entire society of nonconformists, because they think that only a fairly small elite is capable of such transcendence. What they expect or hope for is not anarchy, then, but a growing class of independent (and therefore progressive) thinkers who will not destroy but rather *transform* ordinary rule-bound morality for everybody else. Because we all agree that slavery and imperialism are unjust, we are inclined to admire Emerson and Thoreau for their own efforts at autonomous integrity and for their encouraging the same conscientious self-reliance in others. But if moral independence is really to be praised only when it leads to the right conclusions, then it is not independence that we ultimately care about, but simply coming to the right conclusions. In that case, autonomous integrity is morally valued only as a means to an end, not as the universal virtue it’s cracked up to be.

7.5 Fanaticism and hypocrisy.

Moral life contains the same socioepistemic problems that we find in science, where stubbornly independent thinkers are sometimes useful but often unreasonable. Like the epistemic altruism of revolutionary scientists, autonomous integrity in action is both potent and problematic in matters of public morality: potent in that it tends to produce diverse ideas that can ultimately lead to social progress, and problematic in that it tends to produce rigid doctrines that promote an unreasonable sense of commitment in their adherents. Convince people to think for themselves and stick to their guns, and, given the variability of human perceptions and inclinations, there's no telling which way it will go. Some people take the ideal of autonomous integrity deeply to heart, working their convictions out as independently, and acting on them as conscientiously, as they can. Such people can be heroes like Lincoln, Churchill, or Mahatma Gandhi, but they can turn out to be villains, too. The American left-environmentalist "Unabomber" Ted Kaczynski thinks for himself and acts on his convictions with all too much integrity, wounding twenty-three people and killing three in a protracted campaign of letter bombs intended to destroy modern technology. It makes sense that people on the left, being invested in radical transformation of society, would be more likely to take violent action than those who want things to stay pretty much the way they are. But there are plenty of violent extremists on the right as well, including the anti-government crusader Timothy McVeigh, who murdered 168 people when he blew up the federal building in Oklahoma City, and quite recently the Norwegian nationalist Anders Breivik, who added to the carnage by shooting dozens of people, mostly children, at a socialist youth camp after car-bombing a government building in Oslo, killing several others. Not too ironically, the right-wing fanatic Breivik's lengthy manifesto turns out to have copied some sections from the left-wing fanatic Kaczynski's.[[260]](#footnote-260)

Most committed activists are not extremists like Breivik, Kaczynski, and McVeigh, of course – although to judge from how they often speak it is not always clear *why* they are not extremists. Instead, they peaceably organize, solicit funds, press their points in public argument, and vote like everybody else. This can be because the ideology they’ve chosen is essentially non-violent, or because they think non-violence is the most effective strategy for getting what they want, or perhaps because they just don’t want to hurt other people, regardless of whether their theory tells them that they should. Thus, if someone really believed with 100% certainty that abortion is the literal murder of an innocent child, as many people claim they do, he couldn’t only carry signs; he’d be required to interfere with all necessary force, just as any decent person would in order to save the life of a toddler. But the anti-abortion movement has been generally very peaceful, and when occasional fanatics have committed violent acts which have been quickly condemned by leaders of the movement, just as John Brown was condemned by mainstream abolitionists before the Civil War. How is this possible? Must everyone with revolutionary views be either a fanatic or a coward?

I don’t think so. First of all, very few people will try to think for themselves all the way down to the Cartesian level of deriving all of their beliefs from scratch on an exclusively first-order basis. Instead, what people call thinking for themselves is ordinarily a matter of examining and choosing from a menu of commonly available positions on established controversies. This way, if we cannot just follow the experts because the experts are divided, but we feel that we must take a stand regardless, we can at least follow *some* experts. This will give us a better chance of being right than if we pretended we were experts ourselves. Still, even this will not be a fully rational thing for most of us to do unless we have been taught authoritatively that we have a responsibility to take a position rather than remaining neutral. Since our peers and authorities in fact bombard us with the message that we must aspire to autonomous integrity, this is what most of us rationally ought to do on balance. But because it’s not perceptively rational for non-experts to pick sides in controversies where the experts can’t agree, most of us are properly somewhat ambivalent about the sides that we are told to take. Thus, regardless of what our chosen morality seems to demand, our active engagement is often limited to verbal expressions of approval and disapproval for policies or politicians, plus perhaps a few more substantive activities like making modest financial contributions or attending rallies, but usually stopping short of what we ourselves claim that we ought to do. We have autonomous integrity on paper, as it were, but rarely stand up for our convictions in practice to the extent we would if we were truly certain that we were in the right.

If we took everyone’s expressed convictions at face value, we should be really surprised that there haven’t been a lot more Mansons and Kaczynskis than there have been. But such people are extremely rare, and somehow we are not surprised by this. The fact that very few of us ever go that far in acting on our autonomous convictions points to a deep ambivalence in reasonable people about those of our own beliefs that I have called uncertain certainties. Unlike actions, violent or not, that have been sanctioned by our communities (and many Union soldiers killed more people during the Civil War than John Brown did beforehand), actions done completely on our own against the judgment of our peers cannot be probably right *in themselves*, even though we base them on a general principle of self-reliance that our peers and authorities endorse. This makes the reasonable person hesitate, where the fanatic doesn’t. If our nominal convictions were really *absolute* convictions, as we have been taught to think of them, we would indeed be morally committed to acting accordingly, whatever the cost. But to the extent that they are merely Mill’s “active opinions”, we hesitate to take the drastic actions they imply when our community would disagree in substance. We accept the general principle of autonomous integrity that we have been taught, but we can hardly avoid perceiving that it leads some people badly astray, even as it promotes courageous actions and useful change, and that the disapproval of our peers and betters gives us evidence that we are liable to act wrongly in the present case. Thus, the doctrine of autonomous integrity produces not the absolute commitments that we tend to admire when we are reading Thoreau, but rather weaker things that might better be called *semi-convictions*.

What prevents most of us from becoming fanatics, then, is a benign sort of hypocrisy that keeps our active confidence roughly where it ought to be, despite a justified belief that we should act as if we’re certain we are in the right. The ultimate reason that we are taught to be certain about such things when we cannot be certain of them is that treating these active opinions as convictions motivates us to write and organize and march and otherwise to work persistently towards a moral goal. And again, just as it helps a lawyer defend his clients if he feels convinced of their innocence, it helps if our active opinions seem like convictions to ourselves. This makes our strongest opinions almost the same thing as absolute convictions, except that we don't really act as if we are certain of their truth when they have radical consequences, and we are often surprised if other reasonable-seeming people do.

This is not hypocrisy in the bad sense of the term that we apply to preachers and politicians who demand that others follow rules that they themselves do not. A progressive leader who prevents poor children from escaping lousy schools out of support for the ideal of public education, but who sends his own children to exclusive private schools, is not just behaving inconsistently with his supposed convictions, but insisting that others live by them while he does not. And similarly for the conservative leader who attacks adultery and homosexuality on Christian grounds, but shows no sexual restraint himself in private.[[261]](#footnote-261) But this difference between acceptable and unacceptable hypocrisy is one of degree, so let me articulate it somewhat more carefully: To the extent that our moral claims constitute expressions of mere active opinion, it is reasonable for us to hesitate against taking the most radical actions that they imply. This is the benign sort of hypocrisy that demonstrates a reasonable moderation rather than a real lack of integrity. Alternatively, to the extent that our claims carry force beyond the active sort of argument, i.e. to the extent that we represent them as expressions of absolute conviction, failure to live according to these claims is evidence of our bad faith. This is the malign sort of hypocrisy that we all want to condemn.

In dealing with the semi-convictions that get us raising our voices in arguments or writing books or carrying signs at rallies, but not doing the extreme things that our theories seem to require, we usually find ourselves somewhere in-between, depending on the degree of mismatch between our expressed commitments and our principled actions, and this leads to a lot of confusion over hypocrisy. The best-known and most controversial living moral philosopher is Peter Singer, who has built an enormous reputation by taking utilitarianism (a theory that defines doing the right thing as maximizing total happiness) to its logical conclusions and taking those conclusions seriously. Some of the results strike most non-utilitarians as pretty extreme, including the proposition that human babies have no more rights inherently than mature chimps, since the chimps are by neutral standards more intelligent and self-aware, hence more capable of happiness.[[262]](#footnote-262) According to Singer's extreme utilitarianism, well-to-do Westerners are morally required to keep only a modest subsistence for themselves, and give everything else away to help poor people in the Third World where the money has greater marginal utility. For example, someone in the Congo can stay alive instead of starving to death on what one of us might spend going to the movies instead of watching television. But Singer himself has been the subject of controversy for living more or less like any other successful Westerner, and in particular for having spent a good portion of his considerable earnings on private nursing care for his dying mother. [[263]](#footnote-263) This is something that almost anyone would do who had the means, but Singer cannot justify this act of normal filial duty according to his utilitarian theory, since the many dollars he spent making his mother somewhat more comfortable could easily have saved scores of lives in Africa, almost certainly leading to greater total happiness. Singer's standing response to charges of hypocrisy is that he really ought to give away more of his money but it’s inconvenient, and that he really shouldn't have spent so much to help his mother, but his sister nagged him into it.[[264]](#footnote-264)

I think there’s nothing in Singer’s personal conduct that he needs to excuse. He's being a good philosopher by pushing utilitarianism to its theoretical limits, while maximizing his effectiveness by being superficially convinced of what he says himself. He is also being a good son (or at least a good brother) by acting in a way that contradicts his putative utilitarian commitments. They are not really convictions, then, but only semi-convictions, i.e. somewhere between absolute convictions and mere active opinions. And from his actions, Singer seems to know this without quite being able to acknowledge it. To me, the idealthing would be for him to acknowledge that his utilitarian positions are only opinions, and that he doesn't actually feel as confident about them as he seems to in his writings. [[265]](#footnote-265) Lots of very good philosophers and other thoughtful people disagree with him, after all, as he well knows, and that must matter to him as a reasonable, non-fanatical sort of person. But the doctrine of autonomous integrity demands that Singer, like everybody else engaged in public moral controversy, take a stand for his "beliefs" and make at least an effort to act accordingly, which leaves him in the ambiguous position of exhorting people to do things that he does not really expect them to do and that he does not do himself. This makes him, like the rest of us, a good sort of hypocrite to the extent that he is arguing rather than commanding, and a bad sort of hypocrite only to the extent that he would force his prescriptions on others while ignoring them himself.

**8. POLITICS**

Mankind are greater gainers by suffering each other to live as seems good to themselves, than by compelling each to live as seems good to the rest.

– John Stuart Mill, *On Liberty*[[266]](#footnote-266)

8.1. Oppression and liberation

I have argued that second-order rationality does not just give people many of their factual beliefs; it also gives them information about who is an expert and who is properly empowered to make decisions, as well as who is an enemy or a fool. We learn from our parents, as they learned from theirs, not just about Santa Claus and eating spinach, but also about the reliability of teachers, doctors, presidents, police, and other people that we come to trust. In this way, both epistemic and moral authority ordinarily emerges from tradition. The Pope’s word is (or used to be) authoritative for Catholics not in virtue of his personal wisdom, but by the same long tradition of reverence that keeps the Church respected as a whole by Catholics even through periods of corruption and decline. The same is true of most established monarchs, too: respected by their rational subjects whether or not they are individually wise and good. This is why people are sometimes less than grateful to be “liberated” from oppressive but traditional rulers, even when an outsider can see plainly that they would be better off under a different regime. The typical subject has compelling reason to believe in the social and political system all of the reliable people he knows have told him is the right one, to respect his leaders, and to defend his community against internal and external enemies, as well as any number of merely prudential considerations, all holding his obedience to the traditional regime in place. It is thus a matter not essentially of fear or individual self-interest but of well-founded *conviction* for traditional people to obey their traditional authorities. Hence, just like religions, established political traditions tend to be stable over long stretches of time, even in the face of contrary tendencies toward usurpation, anarchy, and revolution.

Most people throughout history have reasonably thought: what do I know about running a country? What does any ordinary person know? We are not experts in history, philosophy, law, economics, military policy, or even geography. Let the king and his advisors do it; surely they know more than we do. Or, if the king is not a wise one, then maybe those who are wise ought to elect another king, if there is a way of doing that with causing a civil war. In any case, it hardly seems reasonable to put ignorant people like myself in charge of choosing either kings or policies, any more than making a sick child choose his doctors or his medicine. Perhaps it would do him some good to be forced to make that choice autonomously (e.g. by helping him develop into a usefully self-reliant adult, assuming he survives) but this doesn’t imply that he’ll get better treatment that way, which is the more important goal. Most of the time, what sick children reasonably want is that their parents make the decision. And what their parents reasonably want most of the time is that the children’s doctors make the decision.[[267]](#footnote-267) In general, non-experts are inclined to trust their own authorities more than either themselves or their peers.

Revolution is rare, then, not only because existing authorities don’t want to relinquish power to a mob, but also because most reasonable people don’t want to *join* a mob. And they have been reluctant not just out of fear of the probable bad consequences for themselves, but because it really doesn't seem right to them. As I discussed in Chapter 6, people will even ordinarily agree to a degree of repression of dissenting opinions on the grounds that the dissenters' views are not just different but irrational. Why should constitutional questions even be debated when we already know the truth? All it can do is stir people up for no good reason. Ordinarily, a level of reasonableness is required of the authorities, too, and some reciprocal responsibility according to tradition. Kings must act to keep their people safe and tax them fairly, or amuse their public with magnificent displays, or aggrandize their nation through conquest, or whatever else their traditions expect of them, or they will lose some measure of legitimacy.[[268]](#footnote-268) Priests must for the most part keep their public vows of chastity, or poverty, or whatever it is that they have vowed, or they will lose their status as reliable people. If a tradition includes the principle that young people ought to be protected, and people discover that their kings or priests have been routinely molesting children, this will certainly diminish their authority among rational subjects, bringing the prospect of rebellion closer to the threshold of reasonableness.

Karl Marx argues that traditional regimes are generally supported by their subjects, or at least not strongly opposed by them, because those in authority use their power to impose objectively irrational beliefs upon their victims, largely through churches and schools. These "ideologies" support unjust regimes by giving them legitimacy in the eyes of their hapless victims, together with a false belief in their unworthiness to rule themselves. This applies to oppressed workers who falsely believe that virtue requires them to know their place, racial minorities who falsely believe that they are naturally less intelligent, women who have been taught that they are too emotional to lead and that their proper place is in the home, even aristocrats who are provided with a phony sense of noble obligation to take care of their child-like inferiors.[[269]](#footnote-269) You don't have to be a Marxist to perceive that such beliefs tend to go hand in hand with unjust traditional regimes, and to dislike the package very much. But the Marxist analysis, while getting at the *depth* of people's intellectual resistance to change, misses the essential fact that their "false consciousness" is typically a *rational* consciousness, all things considered.

Because traditional people are generally rational, what’s called *political rationalism*, that is, the project of setting up regimes that seem best in terms of exclusively first-order rationality, without consideration of what people already believe and what sources they already trust, is almost always doomed to failure. Modern revolutionaries and reformers fail to liberate their own or other societies from oppressive traditional regimes because it’s very difficult to gain their trust across the sort of epistemic gap that I am talking about, and without that trust they have relatively little reason on balance to believe what the reformers see as true about proper governance, even when it’s *obviously* true to the reformers. Even when effective power has been totally removed from the traditional tribal, religious, or monarchic authorities, most people don't turn as readily to our political values as it usually seems to us Westerners they should, because they still perceive their old form of government as more legitimate. Meanwhile, Western cultural products like internet pornography, routine blasphemy against God and our own traditional religion, and what is often seen as grossly immodest clothing and personal behavior surely retard the process of gaining trust from the people we would like to liberate, by sending persistent messages that contradict their firmest convictions, hence are perceived as plainly false. That such outrageous messages cannot even be avoided these days without serious effort, and that they are so often aimed effectively at children, makes them all the more appalling to most non-Western (and some Western) adults. This is why liberating non-Western people from corrupt oppressors is never as easy as it looks in prospect, as we saw repeatedly throughout the colonial period and continue to see in places like Iraq and Afghanistan.

About the only thing that we can do to liberate traditional people non-violently is to send messages through public presentations, radio and other media, and personal behavior to the people that we want to influence that they are able to establish independently as truthful and good, so that we build up a track record of general reliability that can be spread inductively to other questions.[[270]](#footnote-270) Western medicine and other technologies can certainly help in building trust with rational non-Western people by providing clear evidence of greater truthfulness in certain areas than their traditional authorities. Christian missionaries to traditional societies have certainly found it helpful to bring modern medicine, as well as pots and pans, along with their Bibles. This is of course a very slow and uncertain procedure. For good reasons and bad, we often find ourselves politically and militarily in charge of other societies when no such trust has been established beyond their having learned that we are capable of beating their army. Then, even with the best intentions, we find ourselves opposed and bitterly resented by rational believers in local traditions.

There is no easy solution to this problem of what might be called *oppressive liberation*, and maybe no practicable short-term solution at all, given the power of the epistemic forces in play. Lenin and Mao Tse-tung had this much right: if we want to "re-educate" an entire population, simply taking power and changing the political institutions will not be enough. We will also need somehow to deal with millions of people for whom nothing that we do or say will ever constitute sufficient *evidence* that their traditional beliefs are false, and replace them with new citizens who have different histories of second-order evidence. It would be different if the principles of liberal democracy (or Christianity, or socialism, or human rights, or whatever it is that we would like to offer victims of traditional oppression) were actually self-evident, but they are not, so they need to be backed up with sufficient first- and second-order *evidence* if they are to be perceived by the formerly-oppressed as true. A few adults may be capable of being enlightened by intellectual means alone, or by a deep enough humility towards any authority old or new, but the experience of punitive “re-education” camps in communist countries suggests that people’s traditional convictions will usually outlast anything that can be done to their bodies and any amount of propaganda. If we don't have the resources needed to keep the parents’ traditionalist generation at bay, then killing them *en masse* might well be the most efficient alternative. It is certainly one to which a number of revolutionary governments have resorted during the last century, most deliberately the Cambodian regime of Pol Pot. If we are squeamish about that sort of thing, then we must acknowledge that the more peaceful options have limited effect, tend to take a very long time, and make it hard for us to stay in power while the new post-traditional consciousness is being raised.

Short of murder on a massive scale, what seems to work best in such hostile takeovers is simply to intimidate the older generation into grumbling submission while indoctrinating their children in the new authoritative beliefs. Since children usually start out trusting their parents more than anybody else, this ordinary epistemic bond has to be interrupted as early as possible and as harshly as necessary. In the early Soviet Union children were in theory supposed to be separated from their parents at a young age and raised in collectives where the state’s scientific experts would have total power over them, just as Plato recommended for the children in his philosophical republic. This proved impossible for lack of the resources needed to house and feed millions of children communally.[[271]](#footnote-271) The next best thing to total separation is to use school systems as the primary means of educating children in the new beliefs, together with an epistemic doctrine that parents are unreliable and possibly wicked influences, and that children have a duty to inform on them whenever they show disloyalty to the new regime. [[272]](#footnote-272) And it isn’t just leftists who have wanted to transform society in radical ways, but also Christians, Muslims and other religious activists as well. Jesus himself said “Do not think that I have come to bring peace on earth; I have not come to bring peace, but a sword. For I have come to set a man against his father, and a daughter against her mother…He who loves his father or mother more than me is not worthy of me…”[[273]](#footnote-273) Indeed, any movement that hopes to transform society so deeply must make a similar effort to alienate children from their parents so that they can be brought up with a foundation of second-order evidence that makes it rational, and not just grimly prudent, for them to do as they are told. And this is extremely difficult, almost always more difficult than movement leaders expect. All of the main communist and fascist regimes of the twentieth century eventually collapsed without accomplishing their transformational goals, despite tens of millions of people killed and vast economic and cultural damage (though some Western intellectuals still hold out hope for Cuba).

Revolutionary violence has sometimes been vindicated in the eyes of history. The Glorious Revolution in England in the 1680's and the American Revolution a century later were subsequently seen as great successes in most Western minds, leading as they did to what both left and right have seen as big improvements over previous regimes. But in more ambitious and cathartic revolutions, political rationalism has often led to murder on a previously unheard-of scale, for what have later been seen as crazy reasons: the fascist and National Socialist revolutions in Italy, Spain, and Germany, the Marxist-Leninist revolution in Russia, and the later communist revolutions led by Western-educated intellectuals in China and Vietnam. Although they had plenty of Western sympathizers during the Cold War, the theory-driven policies of the early Soviet Union and Mao's China strike most observers today as grossly unreasonable. In the Soviet Union, Stalin's brutal collectivization of agriculture and officially "settled" Lysenkoist biology led to millions of deaths through famines that were readily predicted by experts not in orbit around Marx and Lenin.[[274]](#footnote-274) Mao's Great Leap Forward required people to melt down their pots and pans in backyard furnaces in hopes of meeting arbitrary quotas for steel production, among other equally bizarre ideas, and millions more were killed in the resulting famine and depression. Since more recent Russian and Chinese governments have renounced communist economics (while sticking with authoritarian rule), those policies are now seen by almost everyone as crackpot ideas imposed by ruthless fanatics with little comprehension of their likely consequences.

Other violent revolutions remain controversial in the West today, particularly the French Revolution that began in the Enlightenment spirit of the English and American ones, descended into Jacobin Terror, and produced Napoleon's short-lived empire and then a hundred years of instability, but also brought France out of a deeply oppressive system of class privilege and, however wrenchingly, into the age of modern liberal democracy. Some of the most rationalistic innovations of the revolution turned out to be useful and good, such as the metric system of measurement; others, such as the guillotine and mass conscription, useful and not so good; and still others, like the new revolutionary calendar, useless or counterproductive.[[275]](#footnote-275) In continental Europe, the French Revolution is still most often seen as having been a good thing on balance in spite of its horrors, and Bastille Day remains the great national holiday in France. After the Civil War and regicidal Cromwell regime during the 1640s and 1650s, Great Britain has taken the safe, slow, steady, road of gradual reform (it’s “Bloodless Revolution” of 1688-1689 being the one mild exception), retaining its monarchy with less and less political authority, and only quite recently removing hereditary peers from its still-powerful House of Lords.

Even non-violent movements in democratic countries can be dangerous if they are able to impose their will on resistant minorities with other traditions. The Prohibition Amendment to the US Constitution was based on plausible notions of saving families from the abuses of drunken fathers and lifting the general moral tone of the country to more sober levels, and was closely tied to the movement for women's suffrage. We are still living with the disastrous effects of the consequent sharp increase in violent criminality, indifference to the law on the part of ordinary citizens of all sorts, as well as subtle cultural changes like the fashion for drinking hard liquor (which was more easily smuggled during prohibition) in place of wine or beer, plus the general association of drinking alcohol with social sophistication.[[276]](#footnote-276) And just the fear that Northern abolitionists would outlaw slavery through generally peaceful, formally legitimate means was enough to bring about rebellion in the American South in 1860 and a horrific civil war. Even after their devastating loss and the emancipation of the slaves, intransigent Southern society managed to deny equal treatment to its black citizens for another hundred years.

Despite the generally poor track-record of such efforts at top-down social transformation, the long-term gradual success of political reform in Britain, worker’s rights movements throughout the West, and particular the triumph of the American Civil Rights Movement of the 1950s-1960s give confidence to Western reformers that steady social progress is possible without substantial violence. On those models, new movements for social liberation, particularly feminism in the 1970s and 1980s, and more recently the gay rights movement, have succeeded partly or completely through a combination of popular activism, solidarity within the cultural and intellectual leadership, and direct government action. The progressive element of our moral society has become more and more like our scientific and technological society, in expecting and welcoming change rather than conservation as a general rule – though it also works to conserve and consolidate the fruits of that change at every stage. The fact that so much of the change has been directed by a class of well-educated professionals in universities, the news media, and popular entertainment strikes conservatives as evidence of a coordinated campaign of leftist-elitist propaganda. And to some extent, at least, this appears to be literally true, for example with the deliberate, sometimes jarring insertion of “safe sex” messages into commercial television programs under pressure from progressive foundations and government agencies interested in public health.[[277]](#footnote-277)

In today's universities, professors and administrators sometimes condemn the current student population as apathetic, cowardly and selfish in comparison to the Sixties generation from which so many senior faculty (including me) emerged. Education theorists have lately produced a cluster of program ideas to combat these students’ non-committal attitudes, including "values clarification", "transformative learning", "service learning”, and "civic engagement", all of which seem to focus on social justice and environmental issues implicitly. More and more high schools, colleges, and professional programs currently require some set amount of approved volunteer activity as a core requirement. The well-educated student is now expected to emerge as a person of conviction and an active member of whatever community or social movement he has chosen to identify himself with, rather than a private person who minds his own business and treats other people well as individuals. Political and educational conservatives are not pleased with these and other innovations in the schools, which they see as indoctrinating students with politically correct hostility toward their own traditions, now widely condemned as racist, sexist, homophobic, anthropocentric, or otherwise intolerant and cruel. But the language in which all these programs are couched is so vague and generally inoffensive that opponents have had trouble formulating a precise complaint.

Even some on the left oppose this mixing of activism and education.[[278]](#footnote-278) But it is not altogether new. There has always been a "citizenship" component to education, including mandatory church attendance at most private colleges into the mid-Sixties, when students were still expected to emerge from college as thoughtful Christians and good patriots. But there is something quite new and not well understood, I think, in the idea of students being expected to transform themselves by choosing and acting on their own personal convictions, as if assuming with Thoreau that greater moral autonomy will make these students wiser and more virtuous. Naturally, most college students, just like most other adults, have little desire or capacity to invent brand new sets of convictions all by themselves. What is really expected of them is that they will sample a menu of more-or-less prepackaged convictions that are made available at school, and then commit themselves to one or more of them in something of the way that they choose a career. And they will choose their new convictions not by way of truly independent reasoning, but largely through reliance on trusted sources, mainly their friends and professors, who connect them in turn with ideas and books and other friends and ultimately whole new epistemic communities. In practice, then, this cafeteria-style conviction-selection process is more a matter of socio-epistemic gravitation into communities of like-minded, co-committed people than of Cartesian construction. And in principle, this process leaves them epistemically no better off than traditional believers who take their identities for granted from birth.

So, why are we pushing it on students? Part of the answer is that, like Thoreau and Emerson, we see the flourishing of individuals as good in itself. Even if the purely epistemic position of semi-autonomous conviction-choosers is no better than that of humble traditional believers, the former may be seen as at least a step closer to the ideal of self-actualization that von Humboldt and Mill express. And there is surely a portion of college faculty and administrators who advance these notions without any political or social motivations: equally eager to liberate the Christian and the pre-committed young environmentalist, and as happy for them both to become followers of Ayn Rand as of Peter Singer if that’s where their inclinations lead them. But I suspect that the greater part of the answer is that, also like Thoreau and Emerson, we expect certain results, and the transformational doctrines exist primarily as abstractly neutral means to a desired, politically progressive end. In the cafeteria-style process there is an element of choice, to be sure, but we expect these choices to produce beliefs that are objectively better than traditional ones, because we *presuppose* that the traditional beliefs are *false*. And I don't mean to say that this approach is necessarily wrong or unreasonable. If society needs to be radically transformed – if, for example, traditional Christians are intolerably homophobic, etc. – then it is good to radically transform society. If we have to be a little vague and sneaky in the way we bring about such change, then maybe that’s a good thing too, on balance. It is probably much more effective, and certainly much more humane, than using revolutionary violence toward the same ends.[[279]](#footnote-279)

8.2. Democracy.

This problem of oppressive liberation applies to forms of government as well as to other cross-cultural differences. Westerners have agreed since the Enlightenment that one essential criterion of a government’s legitimacy, if not the only one, is the consent of the governed. This is a formal criterion, understood to be independent of whether a government's actions and policies are good ones, just as a belief's justification is independent of its truth. It is also a vague criterion, one that leaves many questions open. What percentage of the members of a society must consent? Must the wishes of all members be counted the same? How is consent to be measured? Can it be given once for all, or must it be expressed at regular intervals? Must the consent apply to every act of government, or only some, or only to its basic institutions? Different theories offer plausible answers to these questions, but most Westerners agree that a representative democracy with regular elections, roughly along the lines of the current British, American, or European constitutional systems, best satisfies the consent criterion by providing for the greatest popular input into governance that is consistent with a reasonable degree of order and stability. And we hold this to be true not just for Western countries, but for any nation in the world. Absolute monarchies, theocracies, and other traditional forms of government necessarily fall short of full legitimacy, then, since they fail to seek the opinions of the governed frequently and openly, and to act accordingly.

This might turn out to be a bad criterion. It may be that our conception of positive consent to be tested by regular elections provides only a local, Western answer to questions of legitimacy. Most people in most other places and times have had little specific idea as to how they ought to be governed, have realized this, and have been taught that the proper principles of government do not rely on their express consent, but instead on the divine or natural right of their traditional kings, priests, etc. to rule as they see fit or in accordance with tradition. A governing system that has this level of background consent, that its subjects rationally prefer it to all others, can hardly be made more legitimate (or even in the broadest sense more democratic) by ramming an objectively superior constitution down their unwilling throats. It may seem paradoxical to Westerners, but many traditional people do not and rationally should not want democracy in our more narrow sense of the word. To them, the prospect of their own participation in government seems to assume an expert competence that they neither possess nor aspire to possess. When forced by Western conquerors to make a choice, then, these people often choose to be relieved of similar choices in the future, and to return where possible to the traditional undemocratic systems that they grew up respecting. Even after having overthrown dictators that they see as illegitimate, non-Western people are more likely to revive an ancient but still respected form of government than to opt for a system that looks good on paper but that they have no other reason to trust. Thus, in the Arab Spring revolutions of the past two years, as in Iran’s Islamic Revolution of 1979, many people with traditional values but no memory of popular government have chosen to be ruled by *sharia* rather than secular law, so deep is their distrust of the notionally democratic constitutions that allowed their Western- or Soviet-backed tyrants to stay in charge for many decades.

It can be argued that the destruction of traditional forms of government is a valuable goal in itself, that people will be happier, or at least less dangerous to others, if they are dragged, even against their hardened will, into political modernity. But even if this argument is sound, it must rely on a different, more objective notion of legitimacy through consent that needs to be spelled out in something like utilitarian terms; it is not *finally* concerned about consent at all. We would be awfully lucky, after all, if good government (in the sense of government most conducive to happiness) and legitimate government (in the sense depending on maximal consent) turned out to be exactly the same thing in all cases. The question, whether a more fundamentally legitimate government ought to be replaced by a substantively better one, is often hard to answer, then, for those who have the power to impose new governments. It was certainly felt in the last few centuries by colonial authorities in Britain and elsewhere that the answer was frequently "yes". Initially, at least, the claim they pressed against the local governments of India, Africa, and China had little to do with matters of consent, and much to do with eliminating practices like slavery, mutilation, torture, and human sacrifice. Even the great liberal reformer John Stuart Mill claimed that:

Despotism is a legitimate mode of government in dealing with barbarians, provided the end be their improvement, and the means justified by actually effecting that end. Liberty, as a principle, has no application to any state of things anterior to the time when mankind have become capable of being improved by free and equal discussion.[[280]](#footnote-280)

As anti-colonial attitudes have risen throughout the world over the past half-century or so, the sympathies of Westerners in charge have gradually shifted from the many individuals seen as trapped in such "backward" societies to the same people viewed collectively as nations, each entitled to self-rule according to its own conception of “democracy”. The Western hope for a post-colonial world has been that elections held by the departing colonial powers would guarantee consent, that consent would mean legitimacy, and that legitimate government would rule in the actual interests of the governed. Results have largely proven otherwise on all three counts. As often as not, what happens is "one man, one vote, one time," and then the strongest leader from among traditional competitors ends up as President-for-Life after eliminating all his rivals.

This is an issue that has dogged political philosophy since its inception: as a means of producing the most rational policies, democracy doesn't seem to make much sense. Plato asks us to imagine how democracy would work on board a ship:

The sailors are quarreling with one another about steering the ship, each of them thinking that he should be the captain, even though he's never learned the art of navigation, cannot point to anyone who taught it to him, or to a time when he learned it. Indeed, they claim that it isn't teachable and are ready to cut to pieces anyone who says that it is…[But] a true captain must pay attention to the seasons of the year, the sky, the stars, the winds, and all that pertains to his craft, if he's really to be the ruler of a ship.[[281]](#footnote-281)

In Plato's view, democracy puts power in the hands of people who have no idea how to use it, and who moreover are so ignorant and gullible that they inevitably turn that power over to persuasive tyrants who never give it back. This may be part of what sometimes goes wrong in granting democracy abruptly to colonial peoples. But it doesn't happen very often in the West itself, Hitler's example to one side. Why is this so? Why have we been almost uniquely able to maintain stable democracies without our being obviously better judges of policy or leaders' character than anybody else?

For people to want something like a Western democracy, they have to view themselves as adequately expert at governance, or at least as adequate judges of the character of their prospective leaders, so that it makes sense *to them* for them to participate in regular elections. It is mainly only Western people who think of themselves this way, because it's not ordinarily a rational way to think. Westerners are not demonstrably wiser as individuals than traditional people in matters of governance, and we often display just the sort of ignorance and gullibility that Plato would expect. Only through long, difficult experience have we learned that a modern liberal democracy, properly constituted, well refereed by such actual experts as exist (e.g. qualified, duly appointed judges), and supported by the general public, does seem to work better than its traditional or modern alternatives. The most oppressive features of traditional governments seem to be avoided when ordinary people are permitted and expected to make our own judgments of policy and choose our own leaders, *des*pite the fact that we are not actually very good at it. As Churchill famously remarked, “No one pretends that democracy is perfect or all-wise. Indeed, it has been said that democracy is the worst form of government except all those other forms that have been tried from time to time.”[[282]](#footnote-282)

Successful democracy seems to depend on fair competition among what James Madison called factions, i.e. groups of citizens who advocate their own ideas and interests. This struggle of factions must take place under agreed-on rules, commitment to which must always exceed the factions' commitment to their own interests. This mutual acceptance seems to limit the epistemic possibilities for any democratic people. It is only possible if people find a way of believing, or at least acting as if they believe, that there are no absolute, categorical moral imperatives of any specificity that can't be set aside in the interests of the general system, beyond commitment to the system itself. It is because effective democracy practically presupposes the uncertainty of moral belief in this way that it must rely on the habitual acceptance of otherwise irrational doctrines of tolerance. The formal rights and procedures of democratic government are insufficient in themselves to guarantee lasting success; they must be backed up by a principle of "agreeing to disagree" that limits the epistemic force of all our other convictions. Thus, the proper democratic attitude is tricky to master. As citizens, we must maintain a largely autonomous perspective that is sufficiently convincing to ourselves that it enables us to act, yet not so convincing that we treat it as providing absolute certainty. That is, in order for liberal democracy to be permanent, and not just a way of pausing between civil wars, we must sometimes treat our strongest convictions as mere active opinions. Otherwise, we would be unwilling to compromise them just to preserve procedural agreements with the other side. This political cooperation with our internal enemies may be irrational for both sides, but it is the only reliable way that differences of conviction can be resolved in practice without the use of force. The incoherence in the public "religion" of democracy that results may be a reasonable price to pay for the internal peace and prosperity that it makes possible. But it must always be protected from fellow citizens (like, say, John Brown) who actually treat their convictions as moral absolutes.

In a cosmopolitan democracy like ours, the standards of certainty required for full-strength conviction have to be higher than anywhere else, and people who seem normal in traditional societies may seem like fanatics to us. If everyone in our society were Catholic, for example, then people could take their unchallenged catechism as established fact; there would be no immediate problem with viewing distant Protestants as heretics bound for an eternity in hell. But in a mixed society of Catholics and Protestants, neither side can act entirely according to such convictions without civil conflict, as Europe discovered during the many bloody years it took to stabilize after the Reformation. The successful truce among denominations that finally emerged in the Peace of Westphalia in 1648 after the Thirty Years’ War created a predictably weakened form of Christianity for most believers on both sides of the original dispute. If I cannot tell you and your children that you are going to hell for heresy, or act on my sectarian beliefs in other ways to save your soul or to protect your or my children from your satanic influence, then I cannot really be holding them as absolute convictions. I may still have good reason to *say* that they are absolute convictions, but I can only do so hypocritically. In fact, we are both treating our “private” or “personal” convictions (as we now tend to call our religious beliefs) as mere opinions in our political lives. In a well-functioning democracy, my opponents and I ordinarily refrain, by a kind of gentlemen’s agreement, from calling each other out for this kind of hypocrisy.

This watering-down of religious and other convictions has done good things for Western society internally, leading to stable, widely accepted democratic procedures for solving problems and enhancing social cooperation across the board. But there are disadvantages as well. Although democracy requires a general reduction in factional fervor, that reduction is not always fairly or evenly spread among the faction. Ever within a well-functioning democracy, the side least willing to compromise on any issue tends to get its way. And this is often a good thing; the people who care most about an issue really should have more control than those who are relatively indifferent. But it becomes a bad thing if one or another faction comes to operate in bad faith, “gaming the system” to take advantage of the other side’s desire for peaceful compromise. The first faction to break faith with constitutional procedures and the many “gentlemen’s agreements” that put them into effect – for example, by exploiting “loopholes” in laws and procedures or by flouting the decisions of the courts – will gain quick victories over parties unprepared to break the rules themselves, at the expense of long-term mutual trust and cooperation. [[283]](#footnote-283) This remains a permanent temptation for those factions most interested in radical change, as opposed to those most interested in maintaining the status quo.

Our historically exceptional level of epistemic tolerance and the resulting weakness of religious and political conviction have also presented us with us problems in the wider world. Absolute conviction is a hindrance when your disagreements are with friends, but it remains a powerful weapon when facing implacable enemies. If you must fight and die and kill for your beliefs and loyalties, it helps psychologically if you believe in them implicitly and categorically, and are not used to compromising them for practical purposes. It remains to be seen, for example, whether the vague, unassertive Christianity that remains outside of the interior United States, or the vague, unassertive social progressivism that seems to have succeeded it everywhere else in the West, can command anything like the fierce allegiance that Islamic fundamentalists seem now to be mustering in warlike opposition to our way of life. Some people think that the loss of passionate self-confidence that follows naturally from Western habits of democratic compromise is leading to effective suicide for a culture that can no longer manage to hate its enemies. The late Osama bin Laden and other jihadists have openly mocked the United States in particular as a nation of weak-willed hypocrites. But we do have other, compensating economic, intellectual, and cultural resources (not to mention SEAL teams and missiles) that have gotten us through conflicts with committed adversaries in the past, and for all we know, they may get us through this one as well. [[284]](#footnote-284)

8.3. Freedom

Jean-Jacques Rousseau sees the state ideally as a metaphysical union where individuals renounce their individuality and give themselves over to the “general will”, becoming like bees in a beehive or like cells in a larger, single body. Some people like this idea and some do not, and some are not quite sure they heard the question: Given a choice, to what extent should we “agree to agree” in a consolidated democratic structure, and to what extent should we “agree to disagree”, preserving ourselves as individuals with separate rights and interests? We differ widely on this question, and this is a fundamental disagreement about democracy itself that every democracy must face. Traditionalists (or, more broadly, conservatives) and progressives (or, more broadly, reformers) both tend to favor Rousseau’s social consolidation to the extent that they desire to impose their genuine convictions on their entire societies, either to uphold or to overturn traditional morality and social institutions.[[285]](#footnote-285) But many conservatives and reformers also oppose consolidation to the extent that they fear losing an all-out battle with the other side, and would rather get along peaceably, at least for the time being, than face a conflict anything like civil war. So, in any democratic society, there are two broad disagreements usually going on at once: what might be called reformers vs. conservatives, and what might be called communalists vs. individualists.[[286]](#footnote-286) This makes for a four-way split in theory, though in practice most people take different combined positions depending on the issue.

The United States and other English-speaking countries have a long–standing cultural preference for individualism over communalism (which seems to be favored more in continental Europe). There are plenty of historical reasons for this, including the relative weakness of the traditional English monarchy and the self-reliance required by frontier existence far from seats of authority, but there is a strong reason in principle too, that goes back to Plato’s complaints about democracy but comes to us through Locke and other Enlightenment philosophers, Madison and other founders of the American republic, and, as we have seen, John Stuart Mill, Thoreau, and Emerson. The deep problem they address is what Alexis de Tocqueville called the *tyranny of the majority*. Even if we have been liberated from oppressive traditional authority while avoiding new oppressions from our internal or external liberators by establishing a functioning democracy, the possibility remains that we oppress each other just by letting our majorities rule over our minorities. As eccentric or unpopular individuals, we will still be threatened with oppression from traditional authority, if that’s what the majority decides is best, from revolutionary or transformative forces aimed at imposing their new authority on us, if *that’s* what the majority decides is best, or from whatever else the majority decides is best at any moment, including making irrevocable changes to our institutions.

The US Constitution was designed to limit the tyranny of the majority as much as possible, consistent with a democratic government strong enough to keep the new country out of trouble. The founders limited the power of the central government in several ways: by splitting it into three branches (legislative, executive, and judicial) that would check and balance each other’s authority; by reserving most aspects of sovereignty to the states rather than the federal government; and, most crucially for this discussion, by recognizing a broad array of rights retained by individuals against any democratic majority that might wish to restrict their freedom. This is the core idea of *liberal* democracy as opposed to *simple* or *majoritarian* democracy. For many philosophers, the ideal liberal (or libertarian[[287]](#footnote-287)) democracy is one that maximizes freedom for individuals and voluntary groups, and limits the power of all levels of government accordingly. Thus, unlike simple democracy, liberal democracy establishes a separation between public and private spheres of action. In the public sphere, democracy reigns and a majority of fellow citizens can force you to do what they believe is right. Within the private sphere there is no legitimate government at all, except as necessary to protect you against force and fraud from others; in every other way you must be left alone. [[288]](#footnote-288) As Mill puts it:

…the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right…Over himself, over his own body and mind, the individual is sovereign.[[289]](#footnote-289)

This distinction between public and private spheres of action that arises in liberal societies explains to some extent the ambivalence that most of us feel about pressing our moral and religious convictions on others who disagree, even in friendly discussions. Within the private sphere, it makes perfect sense for us to treat our “absolute” convictions as absolute. If we are truly convinced that we are right, then we really ought to live according to our own lights and disregard the views of people who *must* be wrong. Indeed, we really should do everything we can to persuade our associates to change their ways in order to save their souls or the souls of their children, or just in order to get more people to do what’s right for its own sake. But in the public sphere, where coercion through government possible, to the extent that we take seriously our roles as citizens of liberal societies, we have already agreed to suppress our own convictions for the sake of civil peace and all its blessings. So we will not even *vote* to make another person pray to the right gods, or eat the food that will improve their health, or do anything else that they don’t want to do – any more than we would force them to do those things at gunpoint. And this is true even if we are *certain* that we are in the right, and even if a large majority of fellow citizens agree. In this way, the same belief that operates within the private sphere as a conviction operates in public as a mere opinion.

This is at least a certain ideal of liberal or libertarian democracy, where everyone respects everyone else’s right to do exactly as they please provided no one else is harmed. But in practice things are much more complicated, and the proper ambivalence we feel about our disagreements turns into confusion. For one thing, the boundary between public and private spheres is often unclear, and indeed there are issues and actions that seem clearly to straddle the boundary. Thus, when we are arguing with friends about, say, an upcoming election, we are liable to feel both ways and to be doing both things at once: convinced and pressing our convictions in private, and respectfully debating our opinions in public. We see this often enough in televised debates between candidates or commentators, or in our own debates with friends, when suddenly, seemingly out of the blue, one or another participant shifts from opinion to conviction mode and starts treating his opponents as enemies rather than friends. On another occasion, in the same bar, the same person will be perfectly happy to discuss the same point without rancor as a mere matter of opinion.[[290]](#footnote-290)

The more serious problem comes with trying to make the distinction between public and private spheres even in principle. The broad theory says that we should all be left alone whenever we are not harming another person without their consent. But as Mill acknowledges, this principle of liberty can only reasonably be applied to competent adults. But how do we decide which people are adults for various purposes? And how do we decide which ones are competent? And what constitutes consent? Indeed, what constitutes a person, or even another person as opposed to yourself? All of these notions, clear enough at first glance to give us a good sense of the structure of rights that libertarians have in mind, are so difficult to pin down precisely that almost any policy proposal can be interpreted to be consistent with that structure. Thus, pro- and anti-abortion sources argue over whether a fetus counts as a person, or a separate person, or a person with separate interests to be protected against its mother, or as a subordinate part of its mother, whose own rights ought to be considered paramount. And on neither side of this disagreement is the issue seen as one of borderline cases, easy to compromise about. Instead, a type of action that one side sees as *central* to the private sphere of personal choice, the other sees as *central* to the public sphere, requiring forceful interference.

The hardest concept to grasp fully in this “harm principle” is probably harm itself. What counts as harming someone enough to make public interference proper in a liberal society? Breaking their bones without their consent or poisoning them or setting their houses on fire would certainly count. But how about *risking* these things by riding a bicycle on the sidewalk or smoking a cigarette X-number of feet away from them, or using a charcoal grill in your own back yard? But how about drunkenly bumping into people on the sidewalk, so that they get a little bruise? And what about insulting them, or hurting their feelings, or telling lies about them, or just saying things that make them feel uncomfortable, or annoying them by playing obnoxious music on the bus, or wearing a stupid-looking hat, or waking non-believers up at 5 a.m. by calling everyone to prayer? And what about the many disadvantages that come to people from discriminatory hiring practices in businesses or schools? All of these questions are raised in our consideration of harassment and pollution and discrimination and affirmative action policies, and again, people on both sides will often see their own position as a *core*, not a marginal, application of the liberal distinction between public and private spheres. So, even in principle, how are we supposed to tell whether such things are personal choices within the private sphere or public harms to be controlled by democratic government?

Libertarians answer in essence that we should always promote the most expansive conception of the private sphere that we can; in Thomas Jefferson's words, the government "governs best that governs least". There are three main reasons for this maximization of the private sphere. First, personal freedom is good in itself. Our lives as humans, particularly as adult men and women, are constructed out of choices. The less choice we have in all of our affairs, the less our lives are self-originated as the products of our individual will, the less we are like men and women, and the more we are like children or domestic animals. Even if we did all of the things, and received all of the stuff we would have chosen for ourselves, a life that's been directed by others would be less one's own life, therefore less a life worth living. Second, personal freedom of choice tends to produce more contented people than central direction, for the fundamental reason that the adults among us usually know better what will satisfy us than other people do, especially people who don't even know us personally. I don't want to be told what to do for a living, who to marry, whether to have children or how many, which sport or instrument to play, what topics to write about, or anything else if I can avoid it, because I'm me and I know what I want, and if it sounds weird or other people disapprove, too bad. Third and finally, personal freedom tends to maximize social happiness as well, in a variety of ways. As I have been saying, a large private sphere allows people to live together without having to resolve most of their disagreements, and the larger the private sphere, the more disagreements we can tolerate, which gives us an outlet for diversity, dissent, and restlessness. Maximizing freedom also promotes intellectual and moral progress, as well as economic growth, through competition among in the widest possible range of philosophical, scientific, cultural, and commercial ideas.

But this all comes at the cost of some discomfort, at least, and potentially things much worse, for in expanding the private sphere as much as possible we must contract the range of harms against which we can rightfully take action. The longer and louder that you get to play your music, the more annoyance I have to put up with while I am trying to write. The more you get to say whatever you like about me, the more I must suffer the consequences of your insults and lies. The more you and your friends get to choose freely whom to associate or do business with, the more I am disadvantaged by your refusal to be fair and friendly. Libertarians acknowledge these problems but still want to keep governmental interference with adult decisions to an absolute minimum outside of enforcing contracts and preventing forceful injuries, both because they think this leads to progress and prosperity, and because they see the claims of individuals to be left alone as morally prior to the claims of others not to be made uncomfortable, or even when gross social inequalities result. Thus, the philosopher Robert Nozick argues intuitively that a talented person like the great basketball player Wilt Chamberlain has every right to ask ordinary people to pay him a little money each per ticket to watch him play, and they have every right to pay it if they like – even though this results in Chamberlain being much wealthier than they are, even when we suppose that everybody started at exactly the same level. It is hard to dispute the point that if our money is *ours*, then we can spend it as we like, even if this means that some people get rich and others poor as a result, provided that all the transactions involved are voluntary, and that we started from a situation that we all agree was fair.

The strongest objection to Nozick’s intuitive argument is that it may well be fair for individual adults to make each other rich or poor by voluntary means – even by playing poker constantly until all but a few people have run totally out of money, if they so desire – but that the result is surely unfair to their children. The winners’ children do not deserve the wealth and privilege that they inherit, and the losers’ children do not deserve being reduced to working as the winners’ children’s employees. Nozick and other libertarians may respond by insisting that whatever money is truly ours to spend is also ours to give away, including to our undeserving children, and the consequences shouldn’t matter in principle. But the consequences obviously do matter a lot in practice, even to most people raised in liberal democracies. Even though it seems fair that talented or hard-working people get to earn more money than others, and it seems fair that parents get to give their money to their children, it still seems unfair, however paradoxically, that children have to grow up with the unequal privileges and opportunities that result. It seems, then, that more than one set of epistemic forces is at work.

8.4. Equality

Freedom produces inequalities, and some inequalities are unfair. A totally free exchange of goods and services, including labor, leads almost inexorably to unfair results, with poor people and some social groups consigned to an inferior status and limited in their ability to catch up with groups more privileged. If the liberal principle of free association is applied to employers, restaurants, banks, hotels, and bus companies as well as to personal interactions, then the result is liable to sustain, if not create, persistently discriminatory treatment. Libertarians assert that market forces will tend to remove such barriers, since non-discrimination increases the pool of potential customers and employees, but moral unanimity among dominant groups has proven quite able to counteract such theoretical advantages, as in southern American states over the hundred years between the Civil War and the Civil Rights movement of the early 1960s. Even if one generation of poor or excluded people actually deserved their lower status (for being lazy or impulsive or inclined toward criminality, say, as lower classes have often been characterized), this wouldn't make it fair for their children and other descendants to suffer disadvantages as a result.

Conservative Christians blame our sinful human nature for this tendency of some groups to discriminate unfairly against other groups in economic matters. Greed and corruption have always been with us, and seem to penetrate even those institutions like the Catholic Church that dedicate themselves in principle to fighting against it. Freud blames it on the limitations of our libido, there being in his view no real love for insiders without a corresponding hatred for outsiders. Progressives tend to say that groupwise wickedness results from people having been poisoned with class or ethnic hatred, racism, sexism, and the like. Such hatred is irrational, they say, but sometimes understandable for historical, sociological, and economic reasons. Marxists see class and ethnic divisions as implicit in capitalism itself, one of the techniques that corporate forces use to divide and conquer the minds of workers.

But there are many kinds of grossly unfair treatment that seem not to be based on any kind of hatred at all. Corrupt politicians and businessmen, not to mention outright gangsters, tend to exploit such other people as can be exploited without any stronger feelings towards them than, perhaps, a mild sort of contempt. There are even cases of such systematic exploitation or oppression where there seems to be a positive mutual affection between exploiters and their exploited victims, including privileged clergy like the First Estate during the French *ancien regime*, urban political machines under crooked bosses, dirty unions like the Teamsters under Jimmy Hoffa, the Mafia under popular dons like Al Capone and John Gotti, even the murderous regimes of tyrants like Stalin and Mao Tse-Tung, who are revered by many of the people they oppress right up to the day they find themselves shipped off to camps or shot. In none of these cases do the usual social or psychological accounts of social injustice in terms of class or ethnic hatred apply.

Social psychology in general is much less useful in explaining class injustice that most people think. Even if we suppose that everyone is perfectly rational and perfectly well-disposed toward everyone else, persistent class distinctions will tend to arise, for three reasons. The first is that, as I’ve been saying, economic freedom leads to some people having more success than others and then passing their rewards on to their underserving children. On top of this, freedom of association brings these privileged children into contact with each other in wealthy neighborhoods and through private schools and other social organizations, which makes them more likely to marry and reproduce with each other than with outsiders, which tends to concentrate their wealth and to cement the privilege of their further descendants – all without anybody hating or even disliking anybody else, or doing anything that is intuitively unfair. This is the most basic source of class inequality, I think, and one that is logical rather than psychological.

The second basic reason is epistemological. People in different epistemic communities are bound to have different core beliefs, whether about ethics or religion or something else, and to defend these beliefs with epistemic doctrines telling members how to deal with people in other groups. Again, this does not by itself entail that people have to be *hostile* toward outsiders in general (though it may account for such hostility persisting over generations once it gets started), but rational individuals will always learn to trust members of their own community more than they do outsiders. This will exacerbate the normal tendency for people to marry other members of their own communities, further entrenching any existing social inequalities. (I’m not saying that this won’t tend to generate hostility – just that it isn’t generated *by* hostility.)

The third basic reason that social injustice arises and tends to persist, and I think the one most frequently overlooked, is a moral one. In almost any community, people are raised to believe that they have special obligations to their family, friends, and other associates: to provide for their own children, to honor their fathers and their mothers, to help their friends in need, and so on. Indeed, it is hard to see how any community could function without a system of such local duties and responsibilities. This means that by almost any standard, good people discriminate in various ways favor of their own associates, which logically entails discriminating in disfavor of the people they don’t know. Indeed, most of the *best* people struggle very hard to do things for their friends and families that go beyond just meeting their obligations, which means that their friends and families will benefit in ways that strangers will not. And in the most admired communities, where people treat each other the most altruistically, such benefits are more available than in communities where dogs eat dogs. To the extent that such advantages accumulate inside of any one community more than another, social unfairness will certainly result, as children of altruists in the altruistic communities will be privileged over children of egoists in the relatively egoistic ones. In this way, injustice arises not because people are so *bad* – so hateful, sinful, racist, brainwashed, or whatever – but because they are *good.*

It seems that in Western society (probably any large society) we are all taught two different moralities, a subjective sort of morality that concerns our rights and duties with respect to particular people and groups with whom we are directly in contact, and a second, objective sort of morality that concerns people in general. As with subjective and objective rationality, subjective and objective morality are sometimes deeply at odds with one another. Here is an artificial example. Suppose that I am confronted with a “Sophie’s choice” of lives to save. A number of children are faced with certain death unless I act, and I can only save a certain number of them, but that number is unknown to me. This forces me to rank all of the children in the order in which I would choose to save their lives. Now, these children include my own daughter, my nephew, the child of a close friend, the child of a friendly neighbor, the child of another person in my town, and a child from Sri Lanka. Here is what I think I would do if forced to make this set of choices. I would rank these children in exactly the order I have just listed them in here, beginning with my own child and ending with the child from Sri Lanka. Why would I do this? Do I think that American children are objectively more valuable than Sri Lankan ones? No, I don’t. I like Americans, but I’m inclined to like Sri Lankans just as much. There’s nothing especially valuable about the kids in my town, either. I don’t even think that my best friend’s kids are inherently more valuable than most other children who are being decently brought up. Even my own daughter, much as I love her and take pride in her many virtues, is not all that different from other children her age in objective terms. There are, I am quite certain, plenty of Sri Lankan, Nigerian, and even French children just as worthy as my own of being kept alive. What accounts for my ranking of children is not in any case a set of judgments about the different children's *value*, but about the different *obligations* that I have to each one, which depend on their relationships to me. I have a duty to protect my daughter. I have a lesser, but still considerable duty to protect my brothers’ children, and a duty to the families of my friends and neighbors that is greater than the duty that I have to strangers. This is my moral situation – nothing at all special about it, except that it is mine, and I am the person who is hypothetically being forced to make these choices. Imagine instead that I had ranked these children randomly, perhaps by asking them to draw straws. This random ranking would have been completely fair from an objective point of view; in fact, clearly fairer than my actual choices, since each equally deserving child would stand an equal chance of being saved. But I would be a moral *monster* if I actually chose which children would be saved in such a fair and equal way, especially in not recognizing that I have a special duty to save my own daughter first.[[291]](#footnote-291)

The same sort of conflict between personal duties and objective moral principles shows up all the time, whenever we direct special attention to our own families, friends, neighbors, colleagues, and others with whom we have specific bonds of solidarity or obligation. I want my daughter's education to be better than average – much better than average, if I can arrange that. I have nothing whatever against other children her age – I wish them all well, honestly – but I want mine to be particularly happy and fulfilled and solidly educated, and I have a moral duty to bring that about to the extent I can, consistently with all my other duties. At least, I *think* I have that moral duty, as well as similar, if lesser, duties to my other family members, friends, students, colleagues, and many other individuals and groups. That is in any case the way that I was raised, along with everybody else I know. I am not saying that we owe nothing at all to strangers, or foreigners, or humanity in general; far from it. There are many considerations that we really ought to extend to any human being, even any sentient being, insofar as it is in our power to do so. But our power is severely limited by local responsibilities, even for the wealthiest and most capable among us, so we are forced to discriminate between the people to whom we owe special obligations and everybody else. And it is these personal responsibilities to our families, our friends, our jobs, etc., that use up the great majority of our resources. Once I have done my local duties – paid my taxes, paid my bills, taken my family out to dinner once or twice, put a responsible amount away for the kid’s looming tuition and my ever-receding retirement, and thrown another forty-eight hundred dollars into the Grand Canyon of my student loans, I don’t have a lot left over at the end of every year. If I did, I might consider sending something to somebody in Africa. But even then, if I am going to give money away, I’m likely give it more of it to local charities like the hospice down the street from where I live. This is not because I’m a bad person, but because I’m a good one, or at least trying to be good – even though what I am doing is not fair to needier people in Africa. And every single adult person that I know, progressive, conservative or libertarian – even the most committed Marxist, Buddhist, or Christian – is pretty much the same way.[[292]](#footnote-292)

This is the real fundamental source of what people call institutional racism and “classism”. It is not basically the fault of human sinfulness or any social-psychological disease producing aggregate hostility in the minds of white people when they deal with black ones. These things exist, but they are more symptoms than causes of social inequality. The main causes are family, friendship, and personal responsibility. We do favors for our friends, help them get jobs, share our resources with them, and so on, whenever we reasonably can. And friendship is not randomly or evenly distributed; it grows organically in families, towns, churches, military units, sociology departments, and innumerable other places. When for any reason one such community gains an advantage over others (more talented Wilt Chamberlains, doctrines that tend to promote hard work or education, blind luck in striking oil, or whatever) that advantage will tend to stay or even propagate in that community, not because its members are selfish bastards, but just because they treat each other well. Moreover, to the extent that internally altruistic institutions like churches or labor unions or fraternal orders or even the mafia are better developed in one community than in another, that community will tend better to thrive, even if established on an initially equal footing. And this will add to the injustice suffered by individuals in communities where for any reason such networks are less extensive or powerful. People in the first community will grow up luckier, and probably wealthier, than people in the second one, and this will not be anybody’s fault.

I am not saying nothing can reasonably be done to fix such social injustices. The question is how much can be done without destroying not just freedom, but also the framework of personal moral bonds that makes communities successful in the first place. This is the problem. Love, self-respect, friendship and comradeship, neighborliness, collegiality, and so on are all good things. But each of these good things imposes special duties, which entail some kind of willingness to help or other form of preferential treatment. And preferential treatment is the same thing as discrimination. If discrimination is objectively bad, then it turns out that love, friendship, and so on are good things that have objectively bad consequences. But people are surely better off in general when we are connected to other individuals in moral networks than we would be on our own, or only morally connected to the state or other impersonal institutions. But the overall result of all of these good connections is something very different from distributive equality, or even equal opportunity. As Freud said, love is preference. Without it, there is no chance for happiness. But with it, there is no chance for perfect social justice.

8.5. The reasonable society

Then how about *im*perfect social justice? Surely much can be done and has been done to mitigate the class distinctions that arise from our local or subjective duties to associates. So, there is plainly a lot of room for compromise between the laissez faire minimal government of pure libertarians and the all-powerful Platonic or Leninist state.[[293]](#footnote-293) In fact, almost all societies seek some sort of balance between preserving the freedom of organic moral networks on one hand and promoting social justice through redistribution and compensatory privilege on the other. Western societies constantly struggle to get this balance right, with the US seemingly more tolerant of standing inequalities and the Europeans more accepting of governmental interference. But in absolute terms, the differences these days are small – much smaller than the rhetoric of cruel conservatives and tyrannical progressives would indicate. Everyone taxes income progressively for redistribution; everyone prohibits ethnic and other forms of discrimination in employment and public facilities; everyone offers substantial welfare guarantees for people who are poor, sick, young, old, or disabled. But everyone also promotes personal moral networks, allowing people to get wealthy and to spend their money more or less as they see fit, including leaving it to undeserving children, permitting private schools and other institutions to flourish, guaranteeing free association in the private sphere, and leaving small businesses exempt from anti-discrimination rules. But the principles that govern these compromises are obscure. Especially among intellectuals, leftists want things to be much more collective than they are, while conservatives and libertarians want things to be much less so. Almost nobody is satisfied with the way things are right now, seemingly because it’s too messy, corrupt, and unpredictable to accord well with anyone’s beliefs about either social or personal justice. But this is most the case for people who have been raised as principled idealists, whether their principles are conservative, progressive, libertarian, or something else. People raised to think that way are rational and generally reasonable people, but they will probably always be frustrated by politics, because there is no perfect way to balance all of the epistemic forces involved.

I think that the most useful general approach to politics is to be found neither in a dogmatic traditionalism nor in a rationalistic attachment to theories of democracy, freedom, or equality, but rather in a thoroughgoing *empiricism* that takes full account of all available evidence, past as well as present. This must include the testimony of our many ancestors, who were neither smarter nor stupider than we are on the whole, and who have about as much to tell us about life as do our contemporaries in other societies.[[294]](#footnote-294) Objectively, we should neither accept nor reject traditional beliefs (our own or others’) wholesale, but rather treat their existence as weak, *prima facie* evidence of a practical approximation to the truth. We should distinguish as well as we can between the perceptions made by people in the past, which are roughly as good as ours and much more numerous, and their convictions, which are likely to have satisfied their interests better than they do ours. If a lot of people have believed for generations that chicken soup is good for colds, then this might just be a superstition, but there might be something to it, too. It is at least worth checking out through scientific study, and in the meantime, it is probably better than nothing if you have a cold. The same goes for traditional moral principles of sexual restraint and family responsibility. The more persistent and more universal such traditional principles are, the more credibility they ought to have for us, in the manner of a well-confirmed experiment. But totally modern people cannot rationally accept such principles on faith alone, and attempts by traditionalists to impose them on us are obnoxious and unlikely to succeed. Yet it would be equally irrational for us to reject such moral principles altogether, just for the sake of breaking with the specific evils of the past. What rationality seems to require instead is a tentative acceptance of whatever seems to have worked best so far and a corresponding avoidance of what has failed (including both traditional oppressions and the rationalistic horrors of the recent past), coupled with cautious experimentation for the sake of gradual progress in the direction of common aspirations like prosperity, social equality, and individual freedom.

There is no *a priori* rule for finding the right balance of autonomy and rational humility and then teaching it as virtue or promoting it as public policy. It is instead a matter of discovering empirically how much progressive but subjectively irrational autonomy a given society, in a given historical situation, can sustain. In hierarchical societies like pre-revolutionary France or Russia, free thought and speech was mainly the privilege of the upper classes, and stability was insured by the obedience of peasants who knew their place, at least until a more independent middle class arose. In socialist dictatorships like China and the old Soviet Union, central authorities decide on the degree of general intellectual and active freedom that best promotes the supposed general interest, and tighten or loosen public reins accordingly. In Western democracies, a basic decision in favor of high levels of autonomy has already been taken and promoted in the form of constitutional and civil rights, and ongoing adjustments are mostly worked out in the democratic process itself, with progressive parties pulling at the laws and public policies one way and conservative factions pulling the other way.

The term “political empiricism” has been applied in this sense to the views of Edmund Burke:

Men of sense…will examine how a proposed imposition or regulation agrees with the opinion of those who are likely to be affected by it; they will not despise the consideration even of their habitudes and prejudices. They wish to know how it accords or disagrees with the true spirit of prior establishments, whether of government or of finance; because they well know, that…an attempt towards a compulsory equality in all circumstances, and an exact practical definition of the supreme rights in every case, is the most dangerous and chimerical of all enterprises. The old building stands well enough, though part Gothic, part Grecian, and part Chinese, until an attempt is made to square it into uniformity. Then it may come down upon our heads altogether, in much uniformity of ruin; and great will be the fall thereof.[[295]](#footnote-295)

But the core idea of balancing progress against stability goes back at least to Aristotle:

…Is it a good thing or a bad thing for cities to alter their traditional and ancestral laws and customs whenever some better way is found?...Certainly if we look at the other sciences, we can definitely say that changes have been beneficial…But looking at it in another way we must say that there will be need of the very greatest caution. In a particular case we may have to weigh a very small improvement against the danger of getting accustomed to easy changes in the law; in such a case we must tolerate a few errors on the part of the law-makers and rulers. A citizen will receive less benefit from a change in the law than damage from becoming accustomed to disobey authority…the law itself has no power to secure obedience save the power of custom, and that takes a long time to become effective. Hence easy change from established laws to new laws means weakening the power of the law.[[296]](#footnote-296)

Here Aristotle acknowledges the occasional need for progress, but for him, the essential point is that it is better to be governed by stable laws than by the immediate desires of men, even a majority. Thus he condemned the most extreme form of democracy, where those with no stake in stability can form a majority with total power, unconstrained by the wisdom of their ancestors as encoded in traditions and laws.

Aristotle and Burke are usually thought of as conservatives (though Burke strongly supported American Independence and civil rights for colonial subjects) but the progressive reformer John Stuart Mill makes essentially the same point:

In politics, again, it is almost a commonplace, that a party of order and stability, and a party of progress or reform, are both necessary elements of a healthy state of political life…Each of these modes of thinking derives its utility from the deficiencies of the other; but it is in great measure the opposition of the other that keeps each within the limits of reason and sanity…Truth, in the great practical concerns of life, is so much a question of the reconciling and combining of opposites, that very few have minds sufficiently capacious and impartial to make the adjustment with an approach to correctness, and it has to be made by the rough process of a struggle between combatants fighting under hostile banners.[[297]](#footnote-297)

In an ideally reasonable society, the amplitude of this fluctuation between progressives and conservatives will be relatively small, as in the moderate, gradualistic British monarchy after the Glorious Revolution of 1688. In a society less prone to compromise, the amplitude will be greater as its regimes rock back and forth between repressive traditionalism and its revolutionary opposite, as in France after the revolution of 1789. But a reasonable society is not necessarily a society entirely made up of entirely reasonable people. What is required is only that there be *enough* people who are reasonable *enough* to hold the balance of power between factions – the more the better, and the stronger their positions of authority the better. In medieval times, this was the ideal of a Christian king, to think only of the common good while settling issues among factions with supreme authority. In the Roman Republic and the British Empire, it was held to be the natural responsibility of the aristocracy to think of the general interest instead of their own, and that it was their virtually guaranteed wealth and social standing allowed them to be so public-spirited. The framers of the United States Constitution, well-versed in the history of republican governments, hoped for an informal aristocracy of merit to descend from people like themselves, but also set in place procedural obstacles to the ascendancy of any faction, which they feared would erode public spirit and ultimately destroy the new republic. The constitution they created is perhaps the most successful product of political empiricism so far, providing multiple degrees of fixity at different levels of law to protect minorities and dissenting individuals against political consolidation, while at the same time leaving nothing absolutely unrevisable.

On top of this, divided sovereignty between federal and state governments gives us the famous “fifty laboratories” in which to experiment with different laws and policies, potentially adding a quasi-scientific element to our political development. To the extent that states are allowed free play to answer broad questions concerning civil and economic rights in their own way, more useful information can be obtained than otherwise – though at a cost in perceived justice for any faction with a chance of imposing its view on the entire system. Thus, when the Supreme Court declared in Roe v. Wade that what had been a state-by-state decision was a universal right, this satisfied coastal progressives in their demand for reproductive freedom while outraging heartland conservatives who view abortion as the greatest sort of injustice. These conservatives, now organized into a powerful faction of the Republican Party, have made repeated efforts to establish a similarly universal ban on abortion, or some sorts of abortion, though without much success to date. Forty years into this ugly, polarizing struggle, many on both sides wish that the issue had simply been left up to the states to find a gradual consensus or, failing that, agree to disagree.

The reasonable society is not just a formal system, as the American founders well understood. For such a system to survive, and not be destroyed by factions willing to exploit all loopholes for the sake of immediate victory, presupposes a certain democratic spirit: a readiness to play by the rules and accept short-term defeats; a general weakness of moral and religious conviction for the sake of mutual tolerance; and a willingness to jostle and be jostled rather than stand on our rights in smaller things. In short, it takes a good measure of reasonableness on everyone’s part in order for the system to function properly, as well as a general commitment to the system itself. Too much independent thinking and principled action can erode this system, which some of my colleagues like to say would be a good idea, though they do not seem by their actions totally convinced. Instead, they seem to stay within the democratic mold of moderate self-reliance that has evolved in Western society over the past few centuries. It is not a democratic theory, or even a coherent set of principles, but rather a traditional set of mainly inarticulate attitudes and behaviors, including such things as a felt duty to vote (despite the obvious fact that this is rarely worth a self-interested individual’s trouble), a rough pride in one’s willingness to “stand up and be counted”, and the corresponding tolerance of others’ eccentricity that we call “minding our own business”. It also includes the sort of emotional toughness that is required of people who engage in, or are expected to withstand uncensored public criticism. “Sticks and stones can break my bones, but words can never hurt me” is the playground slogan of free citizens. It is also a patently false proposition, of course, there being little else in ordinary life that hurts most of us more than the mean things that other people sometimes say about us. But it is good for all of us that each of us is taught to assert this slogan as sincerely as he can.

In general, then, it's good to be both tough and flexible when dealing with other people's different ideas where politics is unavoidable, just as we are in dealing with their physical presence in tight public spaces. When you are standing on a subway train and the doors open to let other people in, you do not claim your present spot on the floor as a *right*, just because you got there first. You move back to allow other people in, and when the car is getting full, you allow yourself to be jostled around a bit as each person tries, mostly unconsciously, to accommodate each other person who is doing the same thing. Usually this works just fine, even in hot weather, in very crowded trains full of people who don't even speak the same language, and some of whom look pretty strange to us. Walking down the sidewalk, we make similar rolling adjustments to the positions and velocities of everybody else, even letting them bump us gently in the process, without our having to like them or otherwise even acknowledge them. Such natural instincts help us in democratic politics as well. We understand implicitly that other people are using the same public facilities and that we all have our different interests, and we tend to adjust positions and activities accordingly. I am not aware of any *theory* that tells us precisely how much we should to press into each other's bodies in crowded subway cars or sidewalks; any effort to prescribe exactly how to stand or move would probably greatly increase collisions and hard feelings over what happens naturally.[[298]](#footnote-298)

This sort of analogy applies as well to competitive as to cooperative behavior. In the game of basketball, for example, players have no choice but to bump around against each other under the basket, because there isn't enough room for everybody to be where they want to be, and the players have an interest in keeping their opponents away from where their opponents want to be. For the most part, these antagonistic interests are equilibrated not through the application of rules, but by an adversarial version of the same mutual jostling we undergo in subways. The referees are not going to call every perceptible foul; indeed, they can't, consistently with letting people play the game the way it’s naturally played. Instead, the referees typically only interfere when players are shoving each other noticeably harder than usual, to prevent fights breaking out or when the general ruckus is simply getting out of hand. Well-functioning democracies share something of the same jostling nature, through which the "game" of politics is played within the boundaries of law, but without any one team either committing excessive fouls itself or continually running to the referees to demand that every foul be called on its opponents. If we are going to maintain an adequate level of mutual trust with our opponents over the very long run, then for the system to survive we must agree implicitly among ourselves how hard to shove, and must allow ourselves to be shoved back just as hard. We must agree, that is, to put the health of the sport ahead of our commitment to winning any game.

Libertarian economic theorists have argued since Adam Smith that *spontaneous order* is generally both more just and more efficient than the management of human affairs by central planning, although of course there must be rules of fair play and some concern for distribution of basic resources.[[299]](#footnote-299) And non-perfectionist progressives can certainly agree in principle with this idea, while still arguing that stricter rules of fair play and a more equal distribution of resources are required for an acceptable level of justice. What matters most is not which principles we have ourselves, but that we have a principled willingness to take the others’ principles into account. As Bernard Crick says in his famous defense of politics:

The political process is not tied to any particular doctrine. Genuine political doctrines, rather, are the attempt to find particular and workable solutions to this perpetual and shifty problem of conciliation…Politics are, as it were, the market place and the price mechanism of all social demands – though there is no guarantee that a just price will be struck... [[300]](#footnote-300)

In Crick's sense of the word, then, politics requires a level of long-term cooperation hard to maintain among factions truly convinced of their exclusive claims to justice. As long as we play the game of politics, we have to accept our share of losses, and this means always accepting some degree of *injustice* as each faction sees it. And this may not always be the right thing to do. If players are truly certain that the fate of the world depends on their quick victory in some dispute, or certain enough that what’s at stake has greater value than the system of jostling compromises that has provided centuries of relatively peaceful progress in Western democracies over the past few centuries, then they will have no rational choice but to abandon politics and go to war. Once the players have convinced themselves that nothing matters more than this one game, any compromise beyond the tactical becomes subjectively immoral. In such cases, all gentlemen’s agreements among players will be called off; judges, newspapers, schools and other traditionally non-partisan umpires will be suborned to one side or the other; elections and legislative procedures will be corrupted through the tactical manipulation of the rules and even outright fraud, if that is what will work. After such a total breakdown of mutual trust, open violence will be just around the corner.

American politicians are sometimes denounced as corrupt and hypocritical for making compromising deals in the proverbial smoke-filled rooms. This may be true, too, to whatever extent the politicians have been lining their own pockets and advancing their personal careers at the expense of their constituents' actual interests. But in principle, there's nothing wrong with representatives bumping and pressing and adjusting to the competing interests of their political opponents, when the alternative is forcing their way on their opponents through temporary factional dominance or through the courts. In any case, this cooperative sort of playing by the rules and making gentlemen's agreements seems to be going by the boards in the last decade or two of increasingly bitter polarization in the United States, reaching a sour crescendo in the year 2000 when, for the first time in American presidential elections, a recount of votes was demanded after the returns had been certified by all the states, and state and federal courts were forced by lawsuits to intervene.[[301]](#footnote-301) Several subsequent statewide recount battles have resulted in protracted and nasty struggles in the courts, and this now seems to have become regular feature of our political campaigns.

I have not meant to suggest that Western-style liberal democracy, with or without jostling and gentlemen's agreements among factions, is the best system for everyone. My only universal suggestion is a program of political *change*, presumably in the direction of a liberal democracy like ours, through a balancing of traditional and progressive epistemic forces in whatever manner fits a given time and place. I like to imagine generally peaceful gradualism, as in the broad strokes of Britain’s political history (Cromwell excepted), from a minor shift in power from the king to barons, through the development and gradual empowerment of Parliament, and within Parliament from Lords to Commons, to the eventual abandonment, perhaps, of an increasingly embarrassing symbolic monarchy. But such moderation is not a necessary feature of political empiricism. There might well be times and places where revolution is required for any further progress (just as there has been in the development of science), or, if not strictly required, has better consequences overall than more peaceful but slower alternatives. The American Civil War may have been one such occasion. In any case, I do not mean to rule that sort of radical change out of bounds, especially in situations of outright tyranny. But short of that, especially once regimes have turned a certain corner on the way to peaceful reformability, they should ordinarily be left to continue on their own evolutionary paths, instead of being forced to create modern democratic constitutions out of whole cloth, something that rarely works.

The long-term consequences of the continued development of liberal democracy, in the West and elsewhere, are hopeful but not predictable. We know that our political system has promoted great scientific, economic, social, and military progress for at least a long while, so much so lately that some thinkers have predicted that our system will come to govern the whole world over the coming century.[[302]](#footnote-302) But we do not know whether our societies can maintain the complex balance of background beliefs that are necessary for stable progress into the indefinite future, whether the ideal of autonomous integrity will turn out to promote more powerful extremist factions of the left or right than we will be able to handle, or whether our own system will someday collapse from moral weakness in the face of a sufficiently determined external enemy. Considering our fairly narrow victories over fascism and communism in the century just ended, we might want to avoid promoting either unlimited autonomy of judgment in ourselves or an absolute refusal to pass judgment on others, and continue instead to deal with disagreement in a democratic spirit that is tempered by experience.

**9. GOOD SENSE**

Treat with respect the power you have to form an opinion. By it alone can the helmsman within you avoid forming opinions that are at variance with nature and the constitution of a reasonable being. From it, you may look to attain circumspection, good relations with your fellow-men, and conformity with the will of heaven.

- Marcus Aurelius, *Meditations*[[303]](#footnote-303)

9.1. Pure types of believer

I have argued that belief has the three aspects or dimensions of perception, opinion, and conviction, explained how these all fit together in most ordinary cases, and offered a general analysis of religious, philosophical, scientific, moral, and political disagreements in terms of this framework. But I have said relatively little about which sorts of beliefs are good ones to have. What, given that this is the general structure of people's beliefs, should people believe – or, rather, how should they go about believing in those cases where they have a choice? What principles or strategies for forming beliefs along the three dimensions make for the best sort of person? Epistemic purists try to apply the principles of rationality, autonomy, and integrity to first- and second-order evidence in some special order, either following one principle exclusively, or favoring one or two systematically over the others. This would simplify the issue of how and what to believe, which would be a good thing if the matter could be simplified without damaging things that we genuinely value in life. But I don't think it can; I don't think that a rigid formula for having the best beliefs is possible. In this chapter, I want to review six pure types of believer that I have discussed in different places in this book: the Pyrrhonian skeptic, the Bayesian probabilist, the Cartesian rationalist, the person of faith, the person of principle, and the independent thinker, and say what I think is wrong with each purist approach. Then I will say something about what it takes for ordinary, commonsense believers to be reasonable people in the face of disagreement, which is about the best that we can ever reasonably hope for in ourselves and other people. Finally, I'll make some suggestions for a certain type of imperfect believer liable to be reading this book, namely the seeker after wisdom.

*The Pyrrhonian skeptic*

I have mentioned repeatedly the pure perceiver, someone who observes and reasons but has no opinions and avoids all judgment. This type of person is by definition a skeptic of one sort or another. The ancient Greek and Roman skeptics aimed their thoughts at *epoche*, which means suspension of judgment, by way of maintaining a neutral attitude toward conflicting propositions about the external world. Since there is no clear criterion of truth between opposing possibilities and it is always possible to see things a different way, there can be no knowledge of the external world at all. For the most extreme such skeptics, called the Pyrrhonians because they followed the philosopher Pyrrho, this meant eschewing all judgments of probability as well as judgments of truth.

This is a rational position to the extent that it guarantees the skeptic to avoid having any false beliefs, by allowing no beliefs of any sort, about the external world. But can a Pyrrhonian skeptic be a *reasonable* person as well as a rational one? Can he be a useful or creative person? Can he be a prudent or morally good person? It is not entirely clear that he can be any of these things, either in practice or in principle. For along with guaranteeing no false beliefs comes a guarantee of having no true ones, either. Such a person can do very little to advance the cause of knowledge, since he can have no theories or opinions to share beyond whatever passes through his head at any moment. And he can offer no practical or moral leadership to others, either, since he has no convictions. Such a person can still speak and act in the world, but only out of habit or natural instinct, not according to an articulated model of the world that includes moral and prudential judgments.

The skeptics believed that this *epoche* would release them from worry, producing late antiquity's most sought-after state of *ataraxia*, meaning tranquility or peace of mind. This resembles the Zen Buddhist practice of contemplating paradoxical *koans* as a way of cultivating *sunyata* or emptiness, leading to *Zen* or inner peace.[[304]](#footnote-304) Other mystical and meditative practices aim at the same end in something of the same way – and not by accident, for inner peace is a great way to feel, and it is desirable to have as much control over such feelings as possible. It is also a great feeling to be eating ice cream, so it's good to keep some in the freezer. But you don't want to have it in your mouth *all the time*. Nor do most people want to feel inner peace all the time, because most of the time, or at least some of the time, almost everybody wants to engage in the world in a way that naturally excites the passions, and that results in good feelings other than inner peace, such as discovery, or victory, or the sense of having done something well or helped another person or achieved some other good thing in the outside world. It's wonderful to lie out on the beach over the summer, mindlessly soaking up the sun, or to get baked in other ways on weekends, or to do both on Spring Break – but that's vacation, not real life. For most of us, most of the time, inner peace is secondary to engagement and accomplishment, something you seek through mediation to keep yourself more peaceful than you would be otherwise, for sure, especially when your real-world activities are threatened by anxiety and stress. Robert Nozick writes about an imaginary Experience Machine that will provide you on request with any subjective experience you like, from inner peace to slaughtering your enemies, whenever you like, just for the asking.[[305]](#footnote-305) Raise your hands. How many people would use the machine some of the time? Just about everybody. How many would use it all the time? Practically nobody.[[306]](#footnote-306)

*The Bayesian probabilist*

More moderate skeptics, called the Academic Skeptics because they dominated Plato's Academy for a period long after Plato's death, also believed in avoiding categorical judgments, but did allow for judgments of probability. While they considered it irrational for us to ever believe that something was true categorically, because we can't ever be absolutely certain of anything, they saw no problem with believing on good evidence that one thing was more likelythan another to be true. This allowed the individual to develop probabilized models of the world that would support a certain rational approach to action: at any moment where he had to act, he could do the thing most likely to have results that were most likely to be good, though nothing would ever be fixed in place as a conviction. Having learned that it often seems better than otherwise for him to act, he acts, but only in the manner of a Bayesian gambler, never on principle. Right now, he will perform the action that appears to have the highest value according whatever ethical theory looks best right now. In another minute, he will reevaluate the situation and perform the action that looks probably best to him at that time. Nothing is ever fixed, except his general intention to do whatever looks like the right thing at the time, and even that default sort of principle can be revised in light of further evidence. This is a more reasonable theory than the pure perceptive rationality of the Pyrrhonian skeptic in that it allows for an intelligible and plausible criterion of momentary decision. But it still prevents the person from acting in a concerted or principled way over time. In settling for probabilized beliefs in place of categorical ones, the Bayesian skeptic limits his ability to *will* either a stable moral position or a long-term program of action. Even in purely intellectual matters, such a person is unlikely to make major progress of the sort that Plato or Galileo made, and in the practical sphere he will not be willful or decisive enough to lead.

I have known a number of people who took this view of themselves, that they were rational people entirely without fixed principles, and that this was the right way to live. Still, they had to have some kind of preliminary goal or criterion of rightness in mind, in order to be able to act at all. Some were self-styled rational egoists, supposedly concerned only with their own long-term well-being, and only ever good to others to avoid punishment, or to get something in return, or just because it pleased them at the moment to be nice. There are good philosophical grounds for thinking like this, too, since it is hard to see why as evolved biological beings we should be anything but self-interested, or have any choice in the matter. But I think that people who speak this way about themselves are probably deceiving themselves to some extent, unless they happen to be psychotic. They like the theory, which seems unexceptionable to them philosophically, and then they try to adjust their actions, or, more frequently, adjust their explanations of their actions, accordingly. In material terms, these people seem as principled as most of their neighbors, treating some actions as matters of dignity or duty, and some others as impossible for themselves, and not to be tolerated when done by others. They just don't talk that way. A particularly lovable fictional example was the professional gambler Bret Maverick on the Western television series *Maverick.*[[307]](#footnote-307) Maverick was a charming, calculating sort of character who made his living as a roving cardsharp in a dangerous world, always soberly playing the odds and always on the lookout for trouble. He took pride in his rational self-interest, and even boasted of his cowardice and lack of conviction. But once per episode or so, he would be called on to act bravely and decently, and always did so, shaking his head with regret, when he was just on the point of running away to save his own skin. Maverick would never have admitted to being a man of moral principle, but the audience knew him better than he knew himself. Though he never articulated his convictions out loud, they were plain to see: do not let decent people be abused by cheats and bullies, do not let women be abused by men, do whatever you have to do to save your reckless brother's life (a frequent necessity on the show), and always behave like a gentleman.

The deep reason for most people not to be unprincipled, however rational it may seem on Bayesian grounds, is that it is irrational, all evidence considered. When we are taught morality, we are taught to believe in principles of various sorts through categorically articulated testimony, initially from parents: you shouldn't steal, you shouldn't lie – that sort of thing.[[308]](#footnote-308) The rational person accepts this testimony when it is sufficiently unanimous and comes from rationally well-trusted sources. It may be rational for someone with no such upbringing to be unprincipled (even if still well-disposed toward others), but that isn't us. To *become* unprincipled after having been convinced that certain principles are true, just because it seems to you on first-order grounds alone that having principles doesn't make any sense, is on its face unreasonable. The thing that makes it possible for reasonable people to see themselves this way is only that they have been taught the epistemic principle that they must think for themselves and take the consequences. So, if it seems to them according to this principle that they should have no principles, then that is what they'll think that they should do. This won't actually wipe out their other principles, which still have plenty of their own inductive backing in normal cases. It will just make them contrary to their philosophical opinions.

*The Cartesian rationalist*

Descartes, you will remember, argued that the right way to believe in cases of conflicting authority is to ignore their testimony and start from scratch all on your own. If you start work from utterly indubitable foundations according to a method guaranteed to avoid error, and you are sufficiently careful about it, you can build a solid, totally reliable model of the world, or as much of the world as can be known, on an entirely first-order basis. You will have no problem deciding what to say and do if you *know* what to say and do, and you will know what to say and do if you follow his method conscientiously, only accepting what you have clearly and distinctly perceived to be true, and reasoning in only the most simple and obvious steps, so that you clearly and distinctly perceive all of your inferences to be correct as well. This ideal epistemic method will provide not just for your own error-free beliefs, but also for systematic progress in science that will benefit all humankind.

As I have noted before, though, there are at least three serious flaws in this project. First, the criterion for what counts as indubitable has to be either subjective or objective, and it is hard to see how it can work in either case. If the criterion is one that picks out certain propositions as *objectively* necessary, then you cannot be certain that you have applied it correctly in a given case. But if it is a purely subjective criterion, something like *obviousness-to-you*, then you cannot be certain that such perceptions guarantee truth. Second, this method relies too heavily on the highly your own fallible sensation, memory, and reasoning, and essentially ignores the second-order information that in real life constitutes most of your evidence about the world outside of your experience. Third, even if the method were guaranteed to lead inevitably to all the knowledge that can be had, and even if a single person's first-order resources were sufficient to provide all of the fuel that this machine required, it would be far too slow to do you much good in an ordinary human life-span. Indeed, looking at the tortuous development of science and philosophy among thousands and thousands of practitioners over many hundreds of years, it seems it would take a single person more time to complete than most of us would even *want* to spend alive.[[309]](#footnote-309)

Descartes realized that his project would take an indefinitely log time to complete, so he gave himself four maxims to live by in the meantime, which he called his "provisional moral code", and any rationalist would have to do the same just to get by until he figured things out properly. Descartes’ personal rules were: first, to be moderate and law-abiding and to fit in with the local customs; second, to be decisive in his actions, as though he knew for certain that what he did was right instead of just seeming to him like good guesses; third, to be satisfied with what he had in life, controlling his desires rather than trying to change the world; and fourth, to continue working at philosophy. These are all nice ideas, but hardly enough to guide you through your whole life. In my terms, they amount to saying little more than first, that you should practice second-order rationality in your daily life; second, that you should also make your judgments and form your own convictions; third, that you should be generally passive in your attitude toward life; and fourth, that you should use autonomous first-order reasoning to figure things out for yourself. I think it’s pretty easy to agree with all of that, but it does not have enough substance to tell me what to believe or what to do in any useful detail.

For one thing, in explaining the second maxim, Descartes says that when it comes to action, he intends to treat his current guesses as if they were certainties, even when he knows that he has chosen arbitrarily from among more or less equally probable positions. He compares himself to

…a traveler who, upon finding himself lost in a forest, should not wander about turning this way and that, and still less stay in one place, but should keep walking as straight as he can in one direction, never changing it for slight reasons even if mere chance made him choose it in the first place; for in this way, even if he does not go exactly where he wishes, he will at least end up in a place where he is likely to be better off than in the middle of a forest.[[310]](#footnote-310)

This passage has bothered me since the first time I read it. Walking in a randomly chosen straight line will indeed get you out of any finite woods, sooner or later.[[311]](#footnote-311) But does this arbitrary sort of procedure really get you "out of the woods" in real-life cases? Won't it depend on what really surrounds the woods, and which alternatives are actually better or worse than staying in the woods? If you absolutely commit to walking in an arbitrary straight line, maybe you'll end up stuck in a swamp instead, or falling off a cliff. Even supposing that nothing could be worse than being stuck in the woods, is it always better to walk than to stay still and wait for rescue? The latter thing is, after all, precisely what we are told we ought to do these days if we are actually lost in a forest. In the real-life situations Descartes is trying to illuminate, this might mean waiting for inspiration, or divine guidance, or the consensus of experts, any of which might turn out to be more effective than the random march of first-order reasoning he contemplates.[[312]](#footnote-312) Descartes plainly does not intend that he should *never* change his mind; as he puts it, he doesn't want to change his mind for "slight reasons". But which reasons are slight and which are not? If he were perfectly rational in the Bayesian sense he would evaluate his choices on a continuing basis, and be prepared to change his mind whenever the preponderance of probabilities opposed his present course. If he is willing to give up that perfect rationality, it is not clear what other criterion of choice he has available to tell him when to change his mind, beyond the observation that the most respected people that he knows are *reasonably* decisive people of *reasonably* firm principle, so it is probably better to be somebody more or less like that. In any case, the moment that Descartes commits himself to a plan action in a way that makes it at least temporarily immune from revision, he has already sacrificed some combination of the first- and second-order rationality that he commits to in his other maxims.

I don’t mean to pick on Descartes’ one example of a temporary moral code. My point here is that no merely provisional code of a few pages’ length could do much better. And it couldn’t be merely provisional in any case, since, to be at all realistic, you are bound to die long before you have discovered anything like a truly indubitable (that is, *indisputable*) foundation for belief and action. It is a reasonable idea in principle that you should work really hard on figuring out how best to live, while in the meantime just acting the way most people think you ought to act, and probably good advice for most young people if it isn’t taken too far. When you don’t know what to do, it is good to try to figure out what to do, and in the meantime it does make sense to be moderate in your decisions. But these aren’t really separable projects, where you act moderately up to a certain moment of enlightenment and then switch to doing what you know with certainty is right, foregoing moderation. In real life, we will never know what to do with absolute certainty, so we should always be trying to figure things out better than we have so far. But at the same time, we still have to make choices and to act in the world, and we shouldn’t be so dogmatic about moderation while waiting for certainty that we are unable to take action in a crisis. Our only reasonable choice, as we get quickly older and more slowly wiser, is to keep learning whatever we can, as we balance and rebalance the desire for ideal certainty against our pressing needs for action.

*The person of faith*

I have argued that instead of ignoring second-order evidence, the perfectly rational person humbly accepts that others are liable to know more about many things and takes their word on subjects that he is in no position to understand. I have said that quantum mechanics makes virtually no sense to me, for example, but I still believe in it to the extent that I can understand it, because I trust the experts who tell me that it is the truth. Many people, probably most people in human history, are in the same situation when it comes to the traditional religious beliefs of their societies. So, I have argued that it is rationally best for them to take such things on faith, even where these doctrines strike them as nonsense on first-order grounds alone, just as quantum mechanics strikes me. The person of faith follows the principle of rationality as a "consumer" of beliefs that are more likely, overall, to be correct than those that he produces by himself. Moreover, traditional systems of belief come with a certain guarantee of practical reliability, having already been shown to "work" well enough to stay in place over however many generations they have existed. But there are at least four problems with this narrowly rational approach to second-hand belief. First, beliefs are liable to become stale, inapplicable to new circumstances, and even potentially meaningless if they are held entirely on faith, i.e. with little or no understanding or first-order justification. Second, people are liable to get stuck in epistemic black holes if they take the word of their authorities as much to heart as momentary rationality seems to demand, for false as well as true beliefs will get passed along from one rational generation to another – and sometimes uncorrectably, when their recipients are not allowed to challenge them freely. Third, we know that some past traditional moral and religious systems, although they may be said to have "worked" well enough in their own day to have survived for a long time, are now known to have been full of crazy superstitions, brutal inequalities, and cruel punishments for things that we no longer even see as crimes. When intellectual, moral, and social progress has come, it has come at the hands of brave dissenters who rejected rationality to the extent of challenging the texts and experts of their own traditions.

The relation between faith and reason has been a core concern of Western Christianity since its beginnings, with some philosophers, like Origen (and, later, Aquinas) holding that the two can be totally reconciled, others like Tertullian (and later, Kierkegaard) that reason has no central place in religious belief, while still others like Augustine accept that the relationship will always be problematic, and that a running compromise of some sort between what I am calling first- and second-order rationality must always be sought. Still, as a matter of personal character, there is much to admire in the humble person of faith, who places no personal demands on his convictions that they make rational sense, as long as they makes sense in a spiritual sort of way, the practical test of which is something like the test of Pyrrhonian skepticism and Zen Buddhism, namely the peaceful feeling of a being reconciled to the conditions of his existence.

One of the great fictional exemplars of the humble person of faith is Dostoyevsky's hero Alyosha in *The Brothers Karamazov*.[[313]](#footnote-313) Alyosha is a young monk who gets involved in a great family struggle among his older brothers, the passionate and irresponsible but noble-minded man of action Dmitri and the brooding, nihilistic intellectual Ivan, and their father, the selfish, buffoonish, and soon to be murdered Fyodor Karamazov. At every turn, Alyosha treats his difficult relations with ideal Christian compassion, like a wingless angel whose absolute faith in goodness and God's benevolence motivates his every action. It is hard not to love this character, who offers so much to others while asking so little for himself. But Alyosha is also the *least* *realistic* person in the book, not just because it's hard for anyone to be so patient and giving, but also because he's never put under quite the same worldly stress as other characters. Though he is shown fleshly temptation by the wicked and beautiful Grushenka, which he resists, he is never put in a position where his Christian ideals would force him to be *unreasonable*. In particular, he is never challenged in his Christian aversion to violence, which is the true soft spot for anybody who believes in universal peace and love.[[314]](#footnote-314) What would he do if all the people close to him came under physical attack? At what point would he stop just being nice to everybody and pick up a gun? Never?

*The person of principle*

Some people believe that they should stick to their convictions no matter what. Often, these people understand that their principles come from fallible sources: themselves, their elders, or their peers. But what matters to them is not that their convictions are objective certainties, which they are not, but simply that they are *convictions*. To them, this means that they ought to behave as if they were absolutely certain of their truth. But it is hard for anyone to find a single principle that will not sometimes conflict with other principles or values they also hold, in which case absolute rigidity of action will seem wrong or foolish even to them. Nevertheless we like to think of ourselves as persons of principle, so what does this really mean? Perhaps only that we should *try* to be *fairly* rigid about this or that conviction, holding on to it up to some high level of pressure from other principles or goals. But then, of course, it won’t even be a *principle* any longer in the absolute sense we often think we ought to mean.

One of my oldest friends is a devoted pacifist, and I've always tried to talk him into hypothetical scenarios where he would feel obliged to violate his supposedly absolute commitment to non-violence. Here is a section of a fairly typical discussion.

Me: So look, how about if Genghis Khan came thundering down the street, chopping the heads off all your friends? Wouldn't you shoot him if you had the chance?

Him: No. Don’t you get it? Violence is wrong, *no matter what.*

Me: Okay. So, how about if I just put my finger on your nose and never take it off?

Him: No.

Me: Just like this.

Him: Hey, cut it out.

Me: Make me, you stupid hippie.

Him: Stop. This is really obnoxious.

Me: Nope. I am a man of principle, just like you.

Him: That's enough!

Me: Ha! Violence! You hit my hand!

Him: No, you ass, I…pushed it…gently away. And even if I did hit you, that only proves that flesh is weak and pacifists can lose their tempers. So what? I never said that I was perfect. If I hit you, which I didn't, but if I did, then I was wrong to do it, and I ought to apologize.

Me: Okay, accepted, if you give me one last try. Let’s say that Adolf Hitler has been resurrected, but he's got a really bad case of measles and he's going to die in about ten minutes anyway, so he's locked himself in a bunker with one little window and a Doomsday machine that's going to destroy the entire universe – the whole thing, forever – and you're the only person nearby and you have a gun and it's High Noon and Hitler is just about to press the button, so there's no time to tell him all about Jesus and Gandhi and the Buddha.

Him: Well…

Me: Here goes the Hitler-finger, down…down…

Him: All right, I'd pull the trigger...

Me: Great – that's all I ever wanted you to say.

Him: …but it would be *wrong*.

Me: Oh, come on! Saving the entire universe is *always* the right thing to do, *whatever* amount of violence it takes, not just killing Hitler, right? Just admit it, and we'll all be happy.

Him: Make me, you stupid troglodyte.

I don't think that my friend is a hypocrite, or hypothetical hypocrite, for saying that he would do something in extreme circumstances that he claims to know in advance is morally wrong. He calls himself, and is, an *idealist*. What this means, I think, is that he believes that he ought to speak and act according to a certain simple formula, despite its not being very reasonable if applied with absolute consistency, because if he does, then in the future more people will come closer to that ideal, without themselves needing to be unreasonable about it. Asserting that something is categorically right or wrong is, after all, an *action*, one that tends to make others more likely to believe that what is being said is true. His testimony in support of his semi-conviction is more forceful and influential if he perceives it and expresses it as an *absolute* conviction, provided he is still seen as a trustworthy person in general. And his own exemplary behavior shows an admirable man of principle taking a stand for his beliefs, not a fanatic, which makes him more effective as an advocate than either someone who qualifies everything he says or constantly admits that he is only voicing an opinion, or someone who behaves intolerantly and constantly denounces those who disagree. For these reasons, being an idealist entails holding convictions that are not absolute in fact, while refusing to admit on principle that they are anything less.

We know that in real life, Tibetan Buddhist monks put up a brave if pathetic effort to defend their country with a few old rifles when the Chinese People's Liberation Army conquered Tibet in 1950. Many were slaughtered as their country was reduced to subjection, and its peaceful Buddhist culture has largely been destroyed as a result. But at least they fought, and most people respect them for the desperate effort that they made, however much it went against the ideals of their faith. No reasonable person just abandons his friends and family and countrymen to ruthless invaders for the sole purpose of maintaining his own spiritual happiness, or even fidelity to his own articulated moral principles. Like Peter Singer with his mother, reasonable idealists in extreme conditions temper their own semi-convictions (semi-autonomous or not) with other reasonable people's common principles and expectations, that is, what we call common sense. To do otherwise is not idealism but fanaticism.

*The independent thinker*

The final kind of perfect believer is the sort of person we discussed at length in Chapter 7, someone who follows Emerson’s ideal of total self-reliance. This person places autonomy above all other values, and never defers in his judgments either to authorities or to unanimous peers. For the independent thinker, there is little difference between his own opinions and convictions. At the mild end of the scale, he is a mere eccentric, doing his own thing without harming other people, just because he doesn’t want to harm other people. At the harsh end, he is a self-imagined Nietzschean superman like Ted Kaczynski, or Charles Manson, acting on each of his casual opinions as if it were the word of God without acknowledging the possibility that other people have better ideas. When this violent self-reliant person’s moral notions disagree with ours, we will condemn him in the harshest terms, as we do Timothy McVeigh or Anders Breivik. When they agree with ours, though, we are somewhat more hesitant, and may even celebrate the violent person at a distance, as we sometimes do John Brown, or maybe Che Guevara.

In spite of its puzzling implications, the transcendentalist ideal of standing up for your autonomous beliefs seems to have been absorbed at every level of American culture, from Robert Frost’s “Road Not Taken” to Frank Sinatra’s “My Way”, to John Wayne and Jimmy Stewart in innumerable Westerns, to J. D. Salinger’s and Ken Kesey’s desperate anti-heroes who choose insanity and death over conformity, to Kesey himself as leader of the Merry Pranksters, Timothy Leary, Abbie Hoffman, and dozens of other joyfully rule-breaking figures of the Nineteen Sixties’ counter-culture. I don’t mean to say that this is a strictly American attitude. We see much the same ideal of autonomous integrity in the Bloomsbury group of British intellectuals, French existentialists, and of course Nietzsche himself in Germany. And all of it can be traced back to Luther, Descartes, and Galileo pleading for the primacy of individual judgment in religion, philosophy, and science. But in America, radical individualism is not seen as it seems mainly to have been in Europe, i.e. as the prerogative of aristocratic intellectuals, but as a core value for the entire nation. Thus Raymond Chandler’s classic punch-taking, wise-cracking hero Philip Marlowe begets today’s mainstream movie and television stereotype of the detective who throws away the rule book and does what his hunch tells him is right, despite his friend the world-weary police lieutenant’s warning that it will cost him dearly to go up against the “system”. The detective is no intellectual; just a clear-eyed student of the world who naturally thinks for himself and does what he thinks is right. And his autonomous perspective is always vindicated in the end. Even when corruption finally prevails, as in the movie *Chinatown*, the lone hero is still always right, in the sense that the powerful suspect really did commit the crime that the hero is convinced he did.[[315]](#footnote-315) Why is this? Why do we want so much to believe that the rebellious detective is in the right, rather than the stoical lieutenant? It’s certainly not the way things are in real life, where rebels with hunches are usually wrong and often do a lot of damage.

Thus, during the anti-Vietnam War movement in the US many of us spoke approvingly of revolutionary violence, but very few actually committed any. We approved of it in theory, but didn't have the confidence to act it out heroically in practice, with the exception of a small number of extremists like the Black Panthers and the Weathermen. Most of us didn't really respect people who murdered police and got their friends killed making bombs, though we had no very principled objection to their acting on their own convictions, which to some extent we shared. And it wasn't basically because we were afraid to get in serious trouble, though I am sure most of us were to some extent. Something else held most of us back, which is that somewhere in our minds we still kept track of probabilities and, consciously or not, *perceived* that we and they might well be wrong, whatever we believed we ought to think and say.

9.2. The reasonable person

There remains a kind of mixed ideal to which we all might aspire, namely reasonableness*.* On this imperfectionistic concept, we should always try to be as rational at all levels as possible, to keep our probabilistic calculations running in the backs of our minds as much as possible, consistently with thinking as independently and creatively as possible, consistently with making judgments as decisively as possible when necessary, consistently with acting as faithfully as possible to those convictions that our lives truly demand of us. As with the ideals of Christianity or Buddhism, there is no possibility of our ever matching this ideal of reasonableness or good sense perfectly. But we can always work to come a little closer. The key is finding the right *balance* among rationality, autonomy, and integrity, not just an ordering. We cannot just optimize them separately and add them up; some compromises have to be made. And there is no general formula for making these compromises, either. That will depend on all the circumstances of our lives: our families, friends, jobs, and other responsibilities, special duties that may spring from military or political emergencies, and all of the social and individual experiences that have formed our rational beliefs as individuals, including our beliefs about how rationality, autonomy, and integrity ought to be balanced. The only way in principle of guaranteeing that we won't turn out to be rigid, arrogant, destructive jerks is by restricting our autonomous integrity. We must be predisposed not to despise the testimony of our neighbors but to take them seriously as peers, which implies taking to heart the core values of our whole community. Otherwise, we will be missing the interpersonal epistemic connections that depend on our believing that our fellows are more or less as likely to be right as we are in a disagreement, and that constitute most of our social identity.

There is a less heroic conception of epistemic self-reliance than Nietzsche's or Emerson's that makes more sense. It proceeds not from radically autonomous opinion and judgment in all things, but from accepting a kind of autonomous *responsibility* for our convictions, regardless of where they came from in the first place. We have been *thrown*, to use the metaphor of existentialists, into an epistemic situation full of our own prior beliefs. Since it is possible that some of these beliefs are irrational or morally bad, despite their not seeming that way to us, we have an ongoing duty to rationalize them as well as we can. That is, given all of the first- and second-order evidence available to us, we need to find a maximally coherent model that includes moral perceptions as well as empirical ones. Nobody else can do this for us; in that sense, it's absolutely our autonomous responsibility. If our prior convictions survive this critical process, then we should continue to act on them with as much integrity as we can, consistent with our best understanding of our total moral and epistemic responsibilities, and consistent with accepting as many moral bumps from fellow passengers as we are giving them.

Such balancing and rationalizing efforts are usually better expressed and their complexities and subtleties better evoked in realistic novels than in didactic philosophy. The great master of reasonable balance in fiction is Jane Austen, whose young women protagonists find themselves in complex social situations, often in somewhat subordinate roles and under conflicting pressures from family, moral principle, and their own yearnings for happiness. Austen shows them to us figuring things out, changing their minds, adjusting their expectations, and learning more and more to see things from other points of view, all under strict constraints from moral principles inherited from others but strongly endorsed by themselves, rules of etiquette that they find stifling but also reassuring, and the rigid demands of dignity respecting social and financial status that Charles Dickens (not to mention Marx and Mill) was to attack a generation later.

Elinor Dashwood is the sensible young heroine of *Sense and Sensibility*, who saw to her fatherless family's needs and comforted her cruelly disappointed younger sister Marianne, while nursing her own silently broken heart over a quiet romance with her cousin Edward that had sputtered out for no apparent reason.[[316]](#footnote-316) Though otherwise tempted to denounce her sister's reckless romantic behavior, her deep love for Marianne and her sense of her own foolish thoughts and jealousies make her more tolerant and understanding than would be entailed by rectitude alone, so she makes every effort not to preach. She holds others up to very high standards on principle, but at the same time recognizes that these standards are not shared by everyone, and not easily lived up to by anyone, including herself. At the same time, Marianne grows and matures not just through suffering, but also through a dialectical relationship with Elinor, articulated partly in explicit moral arguments between the sisters but mostly expressed through Elinor's exemplary behavior under stress and Marianne's increasingly appreciative responses. The same sort of thing happens in Austen's later novels, where characters initially less reasonable than Elinor but more so than Marianne strive to understand themselves, their principles, their circumstances, and their fellow characters, so that they can make the most reasonable decisions that they can make. Elizabeth Bennet, for example, the protagonist in *Pride and Prejudice*, struggles mightily against her simultaneous attraction to and revulsion for the wealthy and arrogant Darcy, puzzles her way through to a tentative understanding of his motivations and allows herself gradually to be won over by a man of somewhat different, stricter principles from hers, as he allows himself to loosen them and share some of her warmer, equally honorable but more understanding nature.

Not all of us are sharp young English ladies of the late-Georgian lower gentry, and there are no hard and fast rules for being sensible that cover everybody all the time. The scholar, the lawyer, the farmer, the soldier, the prince, the prostitute, the parent, and the priest seek different equilibriums according to their duties. Even within these separate roles, different balances need to be struck depending on individual circumstances, temperament, and inclination, plus the simple need for some diversity. And even for an individual person, different situations, different times of life, and even different moods will shift the reasonable person's balance one way or another from cautious to bold, stubborn to flexible, obedient to independent. At all three levels, then, what's needed is the right mix of epistemic attitudes, not any single formula. And all that I can reasonably offer in place of a formula is a list of six suggestions – rules of thumb that I think people of good sense already follow for the most part, but that might be worth a little more direct attention with respect to disagreements. Here they are, in no useful order.

*Diversity.* Try to resist the natural (and often rationally justified) tendency to speak seriously only with people who agree with you. Be willing to discuss your views on public controversies with people on the other side who *seem* more-or-less reasonable, just in case they know something that you and your allies don't. It is frustrating to talk with enemies, but if you practice, you can learn the necessary patience. And it helps to have retreat-like situations where you can argue in an atmosphere of neutral analysis and re-evaluation. The best known such retreat is college, of course, but there are others, too, like drinks on Friday after work, holiday gatherings, backyard barbecues, family dinners, and other places where people are more relaxed than usual and more inclined toward friendly discussion. Take these occasions seriously as time out from whatever rigid thinking your convictions demand of you, and see what you can learn from those who disagree with you. As in Austen's novels, for example with Elizabeth and Darcy in *Pride and Prejudice*, mutual understanding sometimes requires awkward discussions between people not officially on speaking terms.

*Discrimination.* Bear in mind the complex structure of belief. People's beliefs all share a certain common skeleton, as I have argued in this book. They form as perceptions, opinions, and judgments; perceptions have first-and second-order sources; judgments vary from immediate decisions to firm convictions; and they all tend to coexist in the undifferentiated, largely unarticulated lumps that I've been calling our perceptive models. Make the effort to discern this hidden structure in your own beliefs and those of others. When someone says "I believe…" be ready to interpret this statement in all three dimensions, watching for feedback and interference among them and between the upper and lower levels of perception. In particular, remember that what you truly *know* only exists in the dimension of perception, not conviction or opinion, and even there it is usually vague, conditional, and probabilistic. If you wish to be or even to seem reasonable, remember that your perceptions are likely to be very different from those of others because of different second-order evidence, that your opinions are hypotheses in which you have no personal or moral stake, and that your convictions are derived from fallible epistemic sources rather than any direct connection between your brain and the objective truth. Reasonable people make these distinctions intuitively, of course, not according to a systematic theory that they've never heard of. But they use the same concepts frequently enough when they distinguish first-hand evidence from hearsay or expert analysis, facts from opinions, and negotiable interests from matters of principle.

*Subjectivity*. Do not expect to change most people's minds with arguments based solely on objective, public evidence. In ordinary arguments, opponents can only offer each other a portion of the reasons that they disagree, usually *public* reasons. Their other reasons, typically the more compelling ones for them, depend on who they trust and where they get their information. Your opponent's sources are not usually even present during the discussion (though it can be helpful to imagine them standing in ghostly form behind your interlocutor), so only a limited examination of the issue will ordinarily be possible. Therefore, the best response that you can usually hope for, even if you give a perfectly conclusive case against an opponent's position, is for the opponent to admit that this *seems* like a conclusive argument, and to say that they will look around among their sources for a counter-argument. It would be *irrational* for them to cave in right away, given the confidence they've gained from trusted sources that their own position is correct, just because your present argument looks plausible to them right now.

*Symmetry.* Refrain from diagnosing your opponents. Assume instead that they are intelligent people with good reasons to believe what they believe. Most people are quite rational in any case, so this will probably be true. But even where it isn't, it is better to err on the side of sympathy than of hostility. This means taking their proffered *reasons* seriously and conscientiously ignoring psychological and other non-rational *causes* for their beliefs, even if you can't help thinking they exist. Bear in mind that, because of the identical basic structure of everyone's beliefs, your opponent's view of the disagreement is likely to mirror yours. If you are tempted to see him as stupid or irrational or pig-headed, he's probably tempted to see you the same way. If you think he's been brainwashed by sinister forces through the media, he probably thinks the same way about you. If you can imagine yourself as a neutral observer or judge, it is easy to see how all these accusations cancel each other out; the remainder is what actually matters. If your opponent reciprocates, you may both learn a lot from the discussion.

*Discretion*. Determine whether or not your opponent is willing to be reasonable and act accordingly. If you are willing to consider his reasons without diagnosis or pre-judgment and he is not willing to reciprocate, there may be no point in continuing the argument, and to avoid frustration you should walk away. But there can also be good reasons to fight on. You may want to show third parties that there are good arguments on your side. You may want to show the other person that you won’t be bullied or insulted. You may just be fed up with people like your opponent and desire to vent your frustrations. Or you may want to impress your opponent or third parties with your very reasonableness by continuing to argue calmly while he loses his cool. In any case, be conscious of your motives and use discretion rather than just reacting to his provocations. Argue with unreasonable people only when you *decide* that it’s a good idea, then keep your cool except when you *decide* to lose it.

*Humility.* Convince yourself that you are pretty likely to be objectively wrong on issues that you care a lot about. This humility is rationally justified by history, including your own track record, the records of your peers and authorities, and the undeniable fact that the best, smartest, and wisest people in the world have all turned out to be wrong about important matters in the past. This is a difficult rule to follow because it means doubting your own convictions as well as your habitual perceptions, and convictions are by nature fixed beliefs, not subject to correction. But go ahead and doubt them anyway – cheat on your epistemic spouse; it's fun. One way of minimizing conflicts between rationality and integrity is to set aside some space in your life for serious reflection, totally independent of your active commitments and suspending even your deepest moral principles for re-evaluation. This will allow you to make mid-course corrections in your judgments and should give you added confidence in whatever convictions survive such re-examination. This is a bit like Descartes' practice of hypothetical doubt, but spread out over your life as an ongoing corrective, not a one-shot systematic reconstruction. Obviously, there are times when this sort of thing cannot be done, for example during hand-to- hand combat, but most of us spend more time on the internet goofing around than we do examining our principles and controversial judgments. A little intellectual discomfort, even a lot, is not an unreasonable price to pay for finding out that you are doing something foolish or even wicked. And if you feel incompetent to dig that deeply into your beliefs, at least bear in mind the probability that some of them are false, and act accordingly.

I have been saying that most people are mostly reasonable, and I believe that most of us follow most of these rules of thumb most of the time. There may well be good reasons for breaking some of them some of the time, and in a genuine crisis it may be the utterly *un*reasonable person who is strong and good, while the reasonable person is, simply by virtue of *being* reasonable, too weak – too humble and too hesitant – to do what’s necessary. Which times these are will not always be clear, of course, even in retrospect – and sometimes there may not even be a fact of the matter.[[317]](#footnote-317) In any case, it surely takes more than any brief list of rules of thumb to make you a completely reasonable person, if that is what you want to be. But if you break all of these rules persistently, and under normal as well as critical conditions, then you are bound to be an unreasonable person overall, and probably a bad one, too: ignorant, arrogant, and unreflective, just like your enemies.

9.3. The seeker after wisdom

Reasonableness isn't everything. It is essentially a negative capacity, a matter of avoiding intellectual mistakes rather than reaching a positive understanding of the world. So, an uninsightful, poorly informed sort of person can still be quite a reasonable one, provided he does not expect too much from other people or himself. If our goals in life are predominantly active ones, this may be all we need. The artist and the industrialist are focused on their work, not their beliefs, and it will usually satisfy for them if they stay out of the worst sorts of epistemic trouble outside of their practical lives. Beyond that, they don't care too much what they believe, except as it affects their work. Sometimes, adopting irrational beliefs can even make their jobs easier, and they may see this as a good deal overall. Thus, the actress who believes it is her destiny to be a star, the politician who believes that he is doing the will of God, or the warrior who believes he'll wake up in Valhalla if he dies courageously may be quite properly resentful of attempts to change their minds. When practical people get into arguments, it is typically because of practical conflicts, not because they see the arguments as good or useful things. If we are seekers after wisdom, though, then good sense is not just an instrument for getting through life, but also an end in itself. Seekers after wisdom must be reasonable all the time, reflecting on the deepest issues as a vocation, not just the occasional refreshing break. We get into arguments essentially in order to improve our understanding; not just to resolve practical problems, but to get to the very bottom of things. Changing people's minds, giving and gaining sympathy, and getting things off our chests are all beside the central point.

Different cultures, communities, and sets of parents set higher or lower values on wisdom as a goal in life, but only a few people are actually *raised* to be seekers after wisdom. The rest come to focus on wisdom later in life, often after spending a lot of time pursuing something else and finding only frustration. We may discover through experience that we find even small chances of being in error abhorrent, or that we find the sense of understanding something more satisfying than other achievements, or perhaps we are driven to seek wisdom by the same sort of vision or need that drives the painter or the politician. In any case, we cannot start from scratch. We've been thrown into seeking wisdom as imperfect people who already have a mixed lot of perceptions, opinions, and convictions. Our central project is to clean up the mess of beliefs that we already have – necessarily employing some of the same beliefs to run the cleaning process and to keep us alive and out of trouble in the meantime – not to replace it wholesale with a better, cleaner system. In Otto Neurath's well-traveled anti-Cartesian metaphor, we must rebuild the ship of our beliefs while we are still at sea.[[318]](#footnote-318)

Great, great Socrates is still the ideal seeker after wisdom. Or he would be, if he didn't sometimes foolishly advertise himself as both already wise and totally ignorant, the wisdom consisting in his knowledge and acceptance of the ignorance.[[319]](#footnote-319) A true seeker after wisdom is not nearly so ready to cash in his chips as this makes it sound like Socrates was. But it is hard to take his self-described skepticism altogether seriously. For he was far too actively engaged in inquiry his whole career to have believed that the whole enterprise was simply a lost cause. Unlike the ideal Pyrrhonist, Socrates *never* relaxed, never stopped thinking and arguing, right up to his death. And he couldn't really have believed that he had no knowledge whatsoever, while always insisting in his arguments that *this* follows from *that* so you cannot sensibly say *something else*. What he must have meant is simply that he had no absolute, categorical knowledge; conditional and probabilistic knowledge he had in abundance, and he never tried to hide it, though he didn't want to call it "knowledge".[[320]](#footnote-320) So, it is closer to the mark to see him as an *Academic* skeptic, as some of his distant followers were later called: there is no absolute knowledge, but conditional and probabilistic knowledge exists, which is enough to make some actions rationally preferable to others, though we can never be *certain* that we're in the right.

But Socrates was no pure Bayesian gambler, for he seems to have been a man of strong opinions, not just rationally flexible perceptions. In any case, Socrates says a lot of things that he appears to mean in dealing with outstanding philosophical problems, including some well-developed views like the elaborate theory of individual and social justice that he carefully constructs from first principles in the *Republic*.[[321]](#footnote-321) Yet Socrates is not a rationalist in the Cartesian sense, either, much as he loves and relies on reason, because he also often reminds us that his putative solutions to problems are all only hypothetical, given for purposes of argument. Though he feels free to argue for these personal opinions, and though Plato often represents him as convincing his interlocutors to agree, these conclusions are always tentative, and his criterion of even tentative success is never his own clear and distinct perceptions, but the very agreement with others that is the goal of all his arguments. Even in his highly idealized imaginary philosophical republic, Socrates would not permit a single philosopher king to rule according to first-order thinking alone, but insisted instead on a large group of equals, trained in dialectics to work problems out together.

Somewhat surprisingly, Socrates also sometimes puts himself forward as a man of faith. In the *Apology*, Plato reports Socrates' speeches at his trial on charges of impiety and corrupting the youth of Athens.[[322]](#footnote-322) In responding to the charge of religious impiety, he claims that in his philosophical pursuits he is only following a message from Apollo's oracle at Delphi. He makes a point also of his lifelong, conscientious observance of the rituals and prayers required of him. Moreover, he claims, he has been as faithful to the city of Athens and its laws as a child could be to his parents, defending it heroically in battle against the Spartans. After the trial he provides a further example of this piety by refusing to accept a comfortable exiled arranged by his friends in place of the execution formally mandated by the city authorities, arguing that as a faithful citizen he owes them unquestioning loyalty when their decision has been made, even if they made it wrongly, and even if it costs him his very life. But Socrates is not a *simple* man of faith like Alyosha Karamazov. He follows all the rituals and obeys all the laws that have been handed down to him without complaint, but he also insists on openly *questioning* his faith, and arguing publicly against some of its common notions, for example that the Gods undoubtedly are our moral betters, could really act as childishly as they do in the common myths.[[323]](#footnote-323) It is this examination, not any failure of proper observance, that gets Socrates into so much trouble and that, according to his prosecutors, the faith itself, properly understood, prohibits.[[324]](#footnote-324)

It is Socrates' actions, more than his statements, that show him to be not the humble, faithful citizen that he claimed to be, but instead a radical critic of the democratic government of Athens and an all-around pain in the neck for people in authority. This is made particularly clear in Socrates' behavior at his trial. In Athenian courts, the juries voted twice, once on guilt or innocence and, in the case of guilty verdicts, once again on punishment. The prosecution would suggest one penalty, the defense a second, lesser penalty, and the jury would then choose between the two by simple majority vote. After Socrates was pronounced guilty, his prosecutors proposed that he be given the death penalty, as expected. Standard procedure would then have been for Socrates to offer to go into exile instead, an alternative the jury would surely have accepted, since it would have gotten this annoying, vaguely dangerous but hardly violent character out of their hair just as effectively as killing him. Instead, Socrates insolently suggested a "punishment" of being given free meals at the city's expense for the rest of his life, then "moderated" his absurd proposal to a relatively paltry fine. Socrates' formally correct but impudent behavior so insulted the court that he was sentenced to death by a larger proportion of the jury than had voted him guilty in the first place. And even in accepting his punishment, when all involved understood that his friends would bribe the guards and take him safely off to retirement on a nice Aegean island, and few would have objected, he insisted on going through with the execution. To insist on being executed when his escaping out the back door was expected and would have been tolerated was obstinate and even arrogant of him in spirit, despite the show of pious obedience.[[325]](#footnote-325)

Though he was sometimes difficult to deal with and always seemed to do things his own way, Socrates was a far cry from Nietzsche's ideal of a superman. What he believed in was not self-assertion, not the will to dominance or Emersonian expression of the individual's subjective "truth", but the hard search for real, objective truth as an essentially communal undertaking. He believed in reason and in reasonableness, though he was not by nature very tractable or compromising, so he comes off as a rather flamboyant eccentric with a great affection for people who are psychologically more normal, but who share his love for argument and his drive always to probe beneath the common understanding.[[326]](#footnote-326) What Plato shows us is not a person with a batch of theories, or a method of proper reasoning, or a heroically detached or humble or self-confident approach to life, but as a complex, ambiguous but still compelling character that most of us admire and some of us love. He is full human being like the rest of us in his own epistemic situation, not an perfectionist but someone who acknowledges that he has his own first- and second-order perceptions, plenty of opinions, and even some beliefs much like convictions. All this is perfectly normal; we all have varied beliefs in a similar structure, including lots of opinions of our own, though typically less clever ones. What makes Socrates special, what makes him a seeker after wisdom, and a great one, is that he so steadfastly refuses to be *satisfied* with anything that he believes as long as other reasonable people *disagree*. Therefore, he dedicates his life to the pursuit of wisdom through discussion, never claiming that he has any absolute knowledge in substance, only the right sort of attitude toward seeking it – and even then, only an attitude for people dedicated to philosophy, not for everybody. Other individuals have other things to do, which Socrates respects. Good people have all kind of different goals and abilities, and it would be wrong for most of them to spend their lives searching obsessively for deeper and deeper understanding. But Socrates is also able to help these other people clarify their own opinions and convictions such as they are – to be a midwife, as he calls himself, to the birth of their *practical* wisdom. But for Socrates himself, and for his closest friends and followers, the search for understanding only ends when every reasonable person can agree.

There is no pat formula for the Socratic project, then, any more than there is one for avoiding unreasonableness in practical life. But the framework of this book, as well as the example of Socrates and other great philosophers, suggest another handful of somewhat more demanding rules of thumb, on top of the hard study and constructive effort that any philosopher would recommend. Here they are, again in no particular order:

*Optimism.* The history of ideas shows that so far, at least, there has always been a better theory waiting to be discovered; if even Newton's beautiful mechanics could be overthrown by Einstein, nothing can safely be considered permanently settled. People tend to think that two or three thousand years should be enough to solve those problems in philosophy that can be solved, and that another hundred years or so will be sufficient to complete the main projects of science. They are wrong. Almost certainly, most of what can be known remains to be discovered. So, try to imagine what people in two or three hundred years are going to think about the theories we are working with today, and this will help you avoid total commitment to any theory that has been made available so far. Some of the things that you believe may turn out to be quite correct, of course, but you are not in a position to know which ones they are, so it is best to view them all as hypothetical, and each of them as very possibly false, regardless of how plausible they seem to you. In the future, people will not just believe different things, but also think in different ways. Their new ideas will be ones that we would find strikingly weird or even paradoxical, as witness evolution, relativity, or existentialism. A thoughtful student of the world can always make some sort of progress, but the greatest insights seem to come from angles that no one had thought existed or that everyone had thought were plainly useless. So, if you want to find something really important, search where the light is bad. In particular, where there are long-standing controversies among schools of thought, as between behaviorist and Freudian psychology, this strongly suggests that the truth is not on the table yet, even as some kind of mixture or compromise. If you want to solve this sort of problem, don't take sides, but place yourself above the table and look down. In the meantime, take all of your thoughts seriously, including your reactions of the thoughts of others. Vague mental pictures, wispy analogies, and even jokes can turn into great theories. These stray thoughts are the raw materials of insight, and the more of them you look at carefully, the greater are the chances that you'll find something important.

*Uncertainty.* Most people don't like feeling confused, and tend to pull away from questions that provoke that feeling. But confusion is the father of understanding, unavoidable when people work to reconcile complex positions. If you resist confusion rather than trying to resolve it, you are resisting knowledge. So, have confidence that this feeling predicts success as long as you keep working through it. You can even learn to enjoy it, in the way that body-builders "feel the burn" as pleasure, not as pain. As Hume says, we naturally tend to judge as well as to perceive. But don't do it; don't ever set your subjective probabilities for empirical beliefs to either 0% or 100%, regardless of how things appear to you. After Einstein, you ought to accept that nothing is absolutely certain in the real world, so this advice makes sense inherently. But it is also important not to let our probabilities even get close tothese extremes for controversial propositions. As we approach Bayesian certainty on any question, we diminish our ability to perceive contrary evidence correctly, or to take it seriously, or to accept those who disagree as epistemic peers regardless of their other qualities. So, as long as an issue remains controversial among peers, you must continue to cultivate your own uncertainty, resisting the theories that you find most plausible right now, even if this means capping your subjective probabilities above or below what reason alone seems to demand. This sounds like an easy rule to follow, but it is not. Even the most successful thinkers fall into ruts, protecting their existing theories rather than moving past them into new discoveries. Talented students are easily drawn into situations where their cleverness earns them immediate, predictable rewards through accepting their professors' paradigms and adding something clever here or there. It is better to be both more autonomous and less: more independent of the local "boss" beliefs, and more humble in the face of all that you don't know. If you succeed, the academic world can be a very comfortable place to stop exploring the world deeply after a few good publications, tenure, and mastery of your own scholarly niche, so watch out. Take the job if you want it, but don't take the bait.

*Empathy*. There is a way of virtually guaranteeing measurable progress, if not directly towards the truth, then towards completing the systematic examination of a subject that should ultimately yield up the truth. Think of yourself as an explorer, not of the world of substance, but of the world of arguments and reasons. There is a gigantic network of facts and principles concerning what entails what, what is evidence for what, and what is analogous to what. Most of this network of rational connections – a whole universe of facts – is utterly unexplored. And people tend to agree on these connections: A being evidence for B, C being inconsistent with D, and so on, even when they disagree furiously over the truth of A, B, C, and D. Therefore, a kind of steady, cumulative progress can be made when people focus on these rational connections rather than material facts and judgments. Moreover, if we combine an agreed understanding of all these conditional relationships with even a partial agreement on categorical facts, say in the form of scientific data or just common knowledge, this partial agreement will tend to propagate throughout the network of implications. A good test of understanding reasons in this neutral way is your ability to argue for the other side as well as you can for your own. So, attend most seriously to the other side's best arguments, not the ones that are the easiest to counter. If your other intelligent people seem to be obviously wrong, then there is probably something attractive about their position that you don't understand and that they haven't fully articulated. Imagine planting your flag in their convictions, not your own. Look past their sentences into the perceptive models from which they're generated, and articulate better arguments yourself.

*Cooperation.* Part of the pleasure of a good discussion is the mutual tracing of opposed positions back to different trusted sources, root convictions, and subjective and objective evidence. If you keep that goal firmly in mind, making it clear to your opponents that your main interest lies in exploring networks of reasons rather than "winning" in the sense of gaining their agreement or impressing observers, they will often respond in the same spirit. It is a lot of fun to engage in arguments this way, and nothing is more useful to your intellectual development. In any case, seek out the very smartest and best-informed people who disagree with you, and do what it takes to make friends with them. The ideal interlocutor is someone with a very different view of the world, a similar desire for wisdom, and plenty of time on his hands – but you should take what you can get. Do not reject a person as a peer just because they disagree with you on an important issue. If you can manage to take *everybody* seriously, at least on matters of common interest, that will give you the broadest range of possibilities for self-correction and new understanding. A good, thorough discussion takes a lot of patience and a lot of time, perhaps a hundred hours for people of good will with serious disagreements to understand each other. Trying to settle things quickly just encourages the feeling that our opponents are stupid crazy selfish bastards. If your goal is understanding, then you must rigidly suppress the hostile feelings that come naturally in arguments and work hard to avoid giving avoidable offense while still expressing yourself clearly. When you find yourself getting angry, apologize to your opponent and calm down. When your opponent gets angry at you, turn the other cheek and wait until he regains his own composure. Such anger and frustration are the unwelcome but expected byproducts of serious discussion, like heat in a bakery. Cool yourself down and get back to work.

*Shamelessness*. Be prepared to feel embarrassed and rejected and alone. You will find yourself asking questions and offering speculations that others find silly or annoying or a waste of time. Socrates called himself a gadfly, a harmless but persistent pest, and he meant it seriously. It *bugs* people if you pester them with questions and doubts, and it doesn't feel very good when they ask you to shut up and leave them alone, or tell you seriously to get a real job. Even around other philosophers, proposing theories that turn out to be trivial or obviously flawed or someone else's idea and being brusquely corrected can be mighty painful. Too bad. One good new idea cancels a hundred embarrassing screw-ups, so scatter your opinions far and wide. Sometimes other people will get angry when you toss around suggestions or doubts that they find offensive, even though all you wanted was to get at the truth. You shouldn't be indifferent to other people's feelings, and you certainly shouldn't seek pleasure in provoking them, but sometimes you just have to let them be offended if you're going to get where you are committed to going, i.e. *ever deeper*. In the worst cases, people may come to see you as a wicked person and a bad influence on others. This can arise from a misunderstanding of your perspective when people see you as an enemy in some dispute simply because you press the argumentsthat their opponents sometimes use as ammunition. But sometimes other people understand your philosophical perspective perfectly well, and your neutrality itself is what they find immoral; they see your questions and speculations as harmful to a greater cause, or they see you as shirking your duty to commit to their side of the dispute. And they may be right. Do not be proud about your isolation. You work in the epistemic sewers underneath the common thoughts of decent and intelligent people, and if you take your job seriously, you will end up *smelling* a certain way. Your work is necessary for the intellectual equivalent of public health, but this doesn't mean that other people ought to find your company enjoyable. People who know you well will come to respect your serious purpose, and may even find you helpful when they run into confusions of their own; others will see you as hostile, irresponsible, indecent, or ridiculous. But remember this: whatever they might think of you, whoever argues with you seriously is an ally in your search for wisdom.

***Afterword***

I sometimes wonder what reasonable people in epistemic situations different from my own might make of this book. I have some evidence from having shown earlier drafts to students, friends, and colleagues, and parts of it to journal referees. The strongest praise I have received from students in my seminars has taken the form, "This is really good; you ought to show it to all the crazy morons who disagree with me about *X*," which I have found a bit disheartening. The main objection that I have received, from about half of the people who have seen this material in one form or another, is that regardless of my arguments to the contrary, no self-respecting human adult should just accept what they are told, at least about morality or politics, because each of us has a duty to think for himself and take responsibility for his beliefs; it simply *doesn't matter* what the neighbors think. I usually tell those people who say this that they have misunderstood me, that I *agree* with them about our having a duty to be independent, at least up to a point, but that I think it is a moral duty, not a strictly epistemic one. And while this moral duty may apply objectively to everybody in the world, it is not a duty that most people outside the modern West can rationally accept, since they have been raised by generally trustworthy people to think otherwise about their way of thinking. This is only a repetition of my central argument, of course. It just dismays me sometimes that it doesn't *work* that well with everybody, or even with most people.

It probably shouldn't, though. After all, it is extremely difficult to change rational people's mind about things that matter to them with mere objective reasoning of any sort. They will have far too much subjective first- and second-order evidence for their convictions than could be countered effectively by any book full of arguments, even if those arguments are absolutely sound. If, for example, readers are already convinced that their opponents are racists, elitists, infidels, etc., then even if they find the arguments about epistemology in this book persuasive, they are not likely to change their attitudes toward these enemies. Some of my most thoughtful friends have argued that even if I am right and their opponents have good reason to believe what they believe and do what they do, this doesn't mean they aren't *also* racists, elitists, etc., and that their actions aren't stilled *caused* by their racism, elitism, etc., not by their admittedly rational beliefs that what they do is morally right. And this is what my friends rationally *ought* to think instead of letting their beliefs be driven by a bunch of arguments that may seem credible, but that have not withstood thorough examination by them and their trusted sources. The most that I can really hope for, then, is that most readers will be *puzzled* by these arguments insofar as they can't figure out what's wrong with them, although they're pretty sure that something *must* be wrong with them since their conclusions are so plainly false.

The readers most open to my arguments are probably the ones most open to anything: young seekers after wisdom, in or out of college. Here is some parting advice for them: think about what you have read, but please don't *fall* for any of it. It is just one person's opinion, after all – a useful opinion, I hope, but only one statement among very many in a very long discussion about disagreement and belief – so don't let it *convince* you, even if it seems compelling at the moment. For what it's worth, I am determined not to fall for it myself, or even to get too involved in defending it, beyond maybe trying to restate my arguments more clearly. If I am actually right about this stuff, then I should take my own advice and get to work on something else.

1. e.g. Michael Shermer, *Why People Believe Weird Things* (revised and expanded edition,Owl, New York, 2002). [↑](#footnote-ref-1)
2. e.g. Jonathan Haidt, *The Righteous Mind: Why Good People are Divided by Politics and Religion* (Pantheon, New York, 2012) [↑](#footnote-ref-2)
3. e.g. Richard Dawkins, *The God Delusion* (Mariner, Boston, 2008); Thomas Frank, *What’s the Matter with Kansas? How Conservatives Won the Heart of America* (Holt, New York, 2005); Ann Coulter, *If Liberals had Any Brains, They’d Be Conservatives* (Crown Forum, New York, 2007). [↑](#footnote-ref-3)
4. Chapter IV.XVI.4, *Essay Concerning Human Understanding* (ed. P. H. Nidditch, Oxford, Clarendon Press, 1975), p. 660. [↑](#footnote-ref-4)
5. By “opponents” I simply mean people who disagree with me; they don’t have to be enemies in any way. [↑](#footnote-ref-5)
6. I will be using the term *epistemic peer* in a slightly broader sense than it is currently being used in the professional literature. In the current professional literature, the concept of epistemic peer is understood to include not just similar overall reliability, but possession of exactly the same evidence. This is convenient for theoretical arguments, but is too narrow for my purposes here. The real disagreements among actual people that I am addressing here rarely if ever involve perfect identity of evidence, especially the individual histories of testimonial evidence that I believe are essential to understanding most such disagreements. It is enough that we are roughly as likely to be right about the points in question, all things considered. [↑](#footnote-ref-6)
7. For this philosophical argument in full see Thomas Kelly, “The Epistemic Significance of Disagreement", in John Hawthorne and Tamar Gendler Szabo (eds.) *Oxford Studies in Epistemology, volume* 1 (Oxford University Press, 2005), pp. 167-196. [↑](#footnote-ref-7)
8. There will always be uncertainty in such inferences, of course, and there is no obvious standard of sufficient evidence. It might be better to say that counting someone else as a *probable* peer is a matter of having sufficient evidence that they are *probably* as likely as you are to be right. But this judgment of probability will have its own uncertainty, too. [↑](#footnote-ref-8)
9. Rene Descartes, *Discourse on the Method of Rightly Conducting One's Reason and Seeking Truth in the Sciences* (generally known as the *Discourse on Method*, or just the *Discourse*). John Cottingham, Robert Stoothoff, and Dugald Murdoch, trans., *The Philosophical Writings of Descartes* Vol. 1, Cambridge University Press (Cambridge, 1985), pp. 111-151. [↑](#footnote-ref-9)
10. Descartes himself recognizes that his project will take a very long time, and in Part IV of the *Discourse* provides a list of rules for getting through life in the uncertain meantime. I will discuss these rules at some length later on. [↑](#footnote-ref-10)
11. For example, however complete people may have thought the discussion of God's existence to have been, no one anticipated Nick Bostrum's "simulation argument" ("Are You Living in a Computer Simulation?" *Philosophical Quarterly*, 2003, Vol. 53, No. 211, pp. 243-255) which has been called (by the philosopher David Pearce) "perhaps the first interesting argument for the existence of a Creator in 2000 years." Perhaps the next one will take just as long. [↑](#footnote-ref-11)
12. Or, as Plato says in the *Theaetetus* (201c), it would make Protagoras no wiser than a pig. [↑](#footnote-ref-12)
13. Skeptics who took this Zen-like attitude were called Pyrrhonians after their founding philosopher Pyrrho. They were opposed by so-called Academic Skeptics, who dominated Plato’s academy long after his death and who argued that many things can be believed rationally, though none can be known. More on this later in chapters 5 and 9. [↑](#footnote-ref-13)
14. I am ignoring the philosophical view that statements about God or morality, or any other claims impossible to verify empirically, are literally meaningless. This would imply that genuine disagreements among such *apparent* beliefs are impossible. See for example A. J. Ayer, *Language, Truth, and Logic* (Gollancz, London, 1936). [↑](#footnote-ref-14)
15. Michael Shermer, *Why People Believe Weird Things* (revised and expanded edition,Owl, New York, 2002), p. 57. [↑](#footnote-ref-15)
16. See Michael Specter, *Denialism: How Irrational Thinking Harms the Planet and Threatens our Lives* (Penguin, London 2010). [↑](#footnote-ref-16)
17. Just to keep it fair: the name that people who are called denialists use for the people who call them denialists is “alarmist”, which is also pretty snappy. [↑](#footnote-ref-17)
18. Buckley routinely remarked that he would rather be ruled by the first two thousand names in the Boston telephone book than by the faculty of Harvard. [↑](#footnote-ref-18)
19. A chart purporting to show that citizens of Republican “red states” had substantially lower IQs than those of Democratic “blue states” that pinballed around the internet in 2004 proved to be a hoax; there is no noticeable such difference. The closest available real evidence is a new study by Gordon Hodson and Michael Busseri (“Bright Minds and Dark Attitudes: Lower Cognitive Ability Predicts Greater Prejudice Through Right-Wing Ideology and Low Intergroup Contact”, Psychological ScienceFebruary 2012 vol. 23 no. 2, 187-195 ) that is claimed to demonstrate a link between low childhood intelligence and racial prejudice in British adults. Hodson and Busseri speculate that conservative beliefs appeal more to the unintelligent because they offer comforting simplicity and structures of authority, not specifically because such people cannot reason soundly. [↑](#footnote-ref-19)
20. For two examples out of many, Michael Savage lays into progressives in *Liberalism is a Mental Disorder* (Thomas Nelson, New York, 2005), and Kate Zernike goes after conservatives in *Boiling Mad: Behind the Lines in Tea Party America* (St. Martin’s Griffin, New York, 2011). [↑](#footnote-ref-20)
21. The concept of homophobia in particular has been supported in some scholarly publications, but is not taken seriously as a clinical disorder by most working psychologists and has not yet earned a place in diagnostic manuals. Conservatives sometimes respond with pseudo-psychological terms like “Christophobia” against progressives, but it is hard to tell whether they mean them to be scientific or satirical. [↑](#footnote-ref-21)
22. For one among many recent examples, a study called “Political Conservatism as Motivated Social Cognition” (J. T. Jost, et al, *Psychological Bulletin* 129(3), 2003, pp.339-375) concludes that conservatism is primarily caused by “fear and aggression, dogmatism and the intolerance of ambiguity.” Spector argues (*op. cit.*) that “denialism” is also motivated by a need for security “…when an entire segment of society, often struggling with the trauma of change, turns away from reality in favor of a more comfortable lie” (p. 3). A seemingly somewhat contradictory report argues that white male conservatives are climate change denialists not because they are fearful, but rather because they are “cool dudes”, overly confident about the future due to their relatively untroubled experience of life so far. See also Y. Inbar *et al*, “Conservatives are more easily disgusted than liberals” (*Journal of Cognition and Emotion* 23, 2009, pp. 714-745). [↑](#footnote-ref-22)
23. # See for instance Lyle H. Rossiter M.D., *The Liberal Mind: The Psychological Causes of Political Madness* (Free World Books, 2006).

    [↑](#footnote-ref-23)
24. Technically, (2) could be counted either as true along with (1), or as false along with (3). This makes no difference to the argument. [↑](#footnote-ref-24)
25. In theory, I should probably say that all uses of "ought" are pragmatic, i.e. relative to this or that goal. Epistemic "oughts" have the goal of rationally justified belief, and moral "oughts" have the goal of people doing the right thing. But I will separate the categories of epistemic, moral, and pragmatic "oughts" as they are here, despite their not being absolutely basic, just because they are the ones that matter to my argument. [↑](#footnote-ref-25)
26. This single principle conflates John Locke’s principles of evidence, probability, and proportionality. My apologies to any experts still reading. [↑](#footnote-ref-26)
27. I still think it was a decent little joke – and also a perfectly rational argument, however inappropriate for serious papers. [↑](#footnote-ref-27)
28. Among the most prominent Oxfordians have been Sigmund Freud, John Gielgud, Derek Jacobi, and Orson Welles. In 1987, US Supreme Court Justices Stevens, Brennan, and Blackmun held a moot-court debate on this question, after which Blackmun and Stevens decided that Oxford probably wrote the plays. See William S. Niederkorn, "A Historic Whodunit: If Shakespeare Didn't, Who Did?", *New York Times* February 10, 2001. A recent movie called *Anonymous*, based on a particularly lurid version of the Oxford hypothesis, was generally panned by both experts on film and experts on Shakespeare. [↑](#footnote-ref-28)
29. So, Russell went to jail as a conscientious objector in Britain during World War One, while his close friend and protégé Ludwig Wittgenstein volunteered to fight for his own country in spite of a medical exemption, evidently not because he had concluded independently decided that the Austro-Hungarian Empire had an objectively just cause for declaring war on Serbia, but because he believed implicitly that it was the honorable thing to do. [↑](#footnote-ref-29)
30. I will discuss this idea at some length in later chapters. [↑](#footnote-ref-30)
31. Let me address two potential problems with my use of the word “perception” in this book. First, the word is sometimes used by philosophers and psychologists to refer only to sensory perception, excluding the products of memory and inference. In that restricted sense, perceptions would constitute only one ingredient in the construction of our mental models of the world, rather than the models themselves. But in ordinary speech, we also often claim to perceivecomplex inferential facts, such as that the Germans' war effort will be hopeless if they cannot win at Stalingrad; I will be using the term throughout this book in that broader sense. Second, “perception” is often used as a “success term”, meaning that we can only perceive a fact or a thing if that fact or thing really exists. In this second restricted sense, our perceptions could never be false or contradict anybody else’s perceptions. But again, common use also allows for people to differ in their perceptions, so that some of Hitler’s advisors perceived things differently from most of the General Staff, believing falsely that Stalingrad was only a temporary setback. And again, I am making the more inclusive choice. Thus, in my broad (but still ordinary) sense of this ambiguous word, our perceptions are no more or less than our *beliefs*, considered as parts of our mental models of the world that we experience. [↑](#footnote-ref-31)
32. So-called *representationalists* define belief more or less as I define perception. See for example Jerry Fodor, *The Language of Thought* (Cromwell, New York, 1975). See also Eric Schwitzgebel, "Belief", in the *Stanford Encyclopedia of Philosophy* for a detailed discussion of current theories of belief. *Dispositionalists* define beliefs as dispositions to behave in certain ways, especially to make assertions, more or less as I define opinion, or as more general dispositions to act on the basis of the thing believed, as I define judgment and conviction. See for example Gilbert Ryle, *The Concept of Mind* (Barnes and Noble, New York, 1949).

    [↑](#footnote-ref-32)
33. Meditations on First Philosophy, Fourth Meditation, *The Philosophical Writings of Descartes* Vol. 2, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge, Cambridge University Press 1984), p. 39. A number of recent philosophers have constructed other two-level typologies. See in particular Ronald De Sousa, "How to Give a Piece of Your Mind: or, the Logic of Belief and Assent", Review of Metaphysics 25, pp. 52-79, Daniel Dennett, "How to Change Your Mind" in his Brainstorms (Cambridge MA, MIT Press, 1981), pp. 300-309, and Keith Frankish, "A Matter of Opinion", Philosophical Psychology 11(4) 1998, pp. 423-442. These writers distinguish between the levels of "belief" and "opinion" (to use Dennett's terms), where "belief" is held to be the lower form, corresponding roughly to what I call perception, and "opinion" the high form, corresponding roughly to what I am calling judgment or conviction. It is a mistake to call the lower form "belief", I think, because all of these things are properly called belief in ordinary English. It is similarly misleading to call the higher form "opinion", because our ordinary concept of opinion is of something that typically requires little commitment, and that need not result in any action other than speech. [↑](#footnote-ref-33)
34. From what is presented in the news media, he seems clearly to be an excellent father, and he is evidently not much of a golfer, so they are probably talking about his role as president. [↑](#footnote-ref-34)
35. Alternatively, and more attractively to some philosophers, a single conception of belief could distinguish what I call conviction from opinion simply over whether the belief is held with 100% certainty or something less, while “perception” can refer just to the raw experience that we have *prior* to forming a belief. [↑](#footnote-ref-35)
36. The concept of reflective equilibrium was introduced by John Rawls in his *A Theory of Justice*, Cambridge, MA, Harvard University Press, 1971, p 17f. [↑](#footnote-ref-36)
37. cf. television’s Doctor House who, like a typical “rogue” detective, constantly violates established protocols in order to follow purely subjective hunches about his patients’ symptoms. [↑](#footnote-ref-37)
38. See for example the ABC news story at: <http://abcnews.go.com/TheLaw/article/exclusive-joran-van-der-sloots-mother-speaks-time/story?id=10935217>. [↑](#footnote-ref-38)
39. Jimmy Carter is the only recent president to send a child to public schools while he was in office. [↑](#footnote-ref-39)
40. To say “in my *opinion*, we have already seen this movie” would sound pretty odd, I think, suggesting as it would that the question involves more complex reasoning and a possibility of meaningful argument. [↑](#footnote-ref-40)
41. *Meditations on First Philosophy* I, in *The Philosophical Writings of Descartes* Vol. 2, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge, Cambridge University Press 1984), p. 13. [↑](#footnote-ref-41)
42. I will say more about this phenomenon of “seeing as” later on. The relation of prior perceptions to new perceptions is discussed in Chapter 4, and the role of judgment is discussed in Chapter 7. [↑](#footnote-ref-42)
43. The psychologist Elizabeth Loftus argues that many innocent defendants are convicted and imprisoned due to the systemic unreliability of memory in *Eyewitness Testimony* (Harvard University Press, Cambridge, MA, 1998). [↑](#footnote-ref-43)
44. *Republic* 509d.

    [↑](#footnote-ref-44)
45. *Theaetetus* 201c. [↑](#footnote-ref-45)
46. A famous exception is the *deductivist* Karl Popper, who argues that so-called inductive arguments are all plainly invalid, so we simply shouldn't use them. See *The Logic of Scientific Discovery* (Basic Books, New York, 1959). [↑](#footnote-ref-46)
47. Other standard inductive forms are probabilistic ones, with or without specific probabilities involved:

    All (or most) of the ravens that we have observed so far have been black.

    Therefore, probably, all (or most) ravens are black.

    Therefore, the next raven we observe will probably be black.

    96% of observed ravens are black.

    Therefore, probably, about 96% ravens are black.

    Therefore, there is about a 96% chance the next raven we observe will be black. [↑](#footnote-ref-47)
48. This scientific sort of inference “to the best explanation” is known sometimes as *ab*duction, to distinguish it from simple *in*duction. For an argument that this and other forms of non-deductive inference boil down to basic induction, see Richard Fumerton, "Induction and Reasoning to the Best Explanation", *Philosophy of Science* 47 (1980), pp. 589-600. [↑](#footnote-ref-48)
49. For a general analysis of induction as involving observations of such general facts, see my “Observation and Induction” in *Logos and Episteme* 1(2) 2010, pp. 303-324. [↑](#footnote-ref-49)
50. Over the past two decades, an enormous literature has developed in the new field of evolutionary psychology concerning human cognitive instincts. Steven Pinker provides an influential general treatment of the subject in *How the Mind Works* (Norton, New York, 1997). [↑](#footnote-ref-50)
51. For a quick example of such partly undigested perceptions, someone who knows nothing about cars might believe correctly that he needs a new “carburetor”, without having any idea what a carburetor actually is beyond its having something to do with his engine, just because he has been told by a reliable mechanic that this is the case. Hilary Putnam and other epistemological externalists infer from this that meanings “ain’t in the head” but are distributed throughout society. See Putnam, “The Meaning of Meaning” in Minnesota Studies in the Philosophy of Science 7, 1975, pp. 131-193. I prefer to say that individual people generally do know what they mean even if they don’t know what all their words refer to, as long as they can distinguish between statements and concepts that they have fully digested and those that they have not. Thus, a person can use a word like “carburetor” quite meaningfully without knowing what a carburetor is, if when he says something like “I need a new carburetor,” what he really means is something like “I need some part for my engine that is *called* a ‘carburetor’”. On this analysis, the speaker knows exactly what he’s saying, even though he has little notion of the referent of the word in quotes. I will deal with such undigested sentential perceptions at greater length in the next chapter. [↑](#footnote-ref-51)
52. Our perceptive beliefs appearing *introspectively* like holograms or sensory "clips" does not suffice to prove that they are not sets of sentences or propositions *objectively*. The relationship between symbolic and non-symbolic mental representation remains a very open issue within cognitive science, with the so called *classicists* or *computationalists*, who view the mind as much like a digital computer, on one side, and *connectionists*, who perceive the mind as a network of neural connections of varying strengths, on the other. The latter view seems friendlier to the sort of holistic mental representation that I’m talking about, and to Quine’s notion of the web of belief. Gary Marcus attempts to reconcile the two positions in *The Algebraic Mind: Integrating Connectionism and Cognitive Science* (Bradford Books, Cambridge, MA, 2003). [↑](#footnote-ref-52)
53. “Civil Disobedience”, in *Henry David Thoreau: Collected Essays and Poems* (Library of America, New York, 2001) p. 203. The more common attribution of this quote to Thomas Jefferson is evidently false, though he articulated the same sentiment in other ways. [↑](#footnote-ref-53)
54. I have also noticed that after a summer off or a sabbatical I sometimes find myself wondering how I can possibly lecture for 75 minutes without having everything written out beforehand. Just looking around in my memory for knowledge about this or that course of material, I don’t see much of anything to say; there's just this vague sense of the subject matter, plus a few fleeting images and other sensory cues. Over time, though, I have come to rely with better confidence on the other essential element in articulate memory, which is the magical sub-organ of the brain, entirely opaque to introspection, that actually generates sentences out of that hazy mass of stored perceptions. [↑](#footnote-ref-54)
55. In a very useful metaphor, W. V. O. Quine speaks of our beliefs as forming something like a web, with our most stable core beliefs in the center and a periphery of immediate perceptions that are much easier to alter or remove. See Quine and J. S. Ullian *The Web of Belief* (McGraw-Hill, New York, 1978). [↑](#footnote-ref-55)
56. The number of these worlds would be 2 to the *n*th power, where *n* is the number of propositions that could independently be either true or false. Thus, if there were exactly two things in the universe, and each of them might or might not have either of two different properties, there would be a total of 24 or 16 different possible worlds. If there were five things and each could be any of five ways, then there would be 225 or 33,554,432 different possible worlds. If there were millions of things and each could be millions of possible ways, then the number of possible worlds would be very large indeed. [↑](#footnote-ref-56)
57. I guess I just articulated that one. Please pick some other number of people for me to believe that there are more onetime hat-wearers than in China; anything under a billion will do. [↑](#footnote-ref-57)
58. In general, I will mean by categorical (or conditional, etc.) perceptions or beliefs those beliefs that are best articulated in statements of the relevant form, leaving it vague what constitutes the best articulation in one situation or another. [↑](#footnote-ref-58)
59. We also usually think that our background assumptions are true, or probably true, so it is usually no big deal that our beliefs contain them as conditions. In any case, what matters in informative discussion is not that these assumptions are true, but only that they are also being assumed by everybody else. [↑](#footnote-ref-59)
60. Subjective probabilities do not seem to be stored explicitly as numbers, but they can be evoked and measured, at least very roughly, in terms of bets that people would be willing to make, or the degree of surprise that they experience when what they take to be the most probable facts turn out to be false. The alternative concept is objective probability, which is sometimes defined in terms of actual frequencies of similar events. [↑](#footnote-ref-60)
61. In purely informative contexts, a simple statement "*p*" could usually be even more fully articulated in the very awkward form "if *a*, then *p* to *e* with *r*" where *e* stands for the extent to which *p* is supposed to be the case, *r* for the probability of *p*'s being true to extent *e* under background assumptions *a*. For still greater precision, we could include subjective probabilities and extents for *a* too, plus the subjective probability of the inference from *a* to *p* being correct. [↑](#footnote-ref-61)
62. This is a crucial flaw in the program of logical positivism (also called logical empiricism), which endeavored to analyze the vague and ambiguous statements of real empirical discourse in terms of an ideal, perfectly clear logical language. The ideal languages that positivists tried to construct were based on a simple “classical” logic where every basic statement is considered either true or false depending on whether the objects named by the nouns either belong or do not belong to the sets named by the predicates. Logical positivism as a movement ran out of steam long before its proponents had exhausted the possibilities of alternative, more sophisticated systems of logic. For a famous attack on the positivist’s unreasonable demands for perfect clarity, see W.V.O. Quine, “Two Dogmas of Empiricism” *Philosophical Review* 60 (1), 1951, 20-43. For a totally obscure response, see my “Analyticity without Synonymy in Simple Comparative Logic” *Synthese* 130 (2), 2002, 303-315. [↑](#footnote-ref-62)
63. By "degrees of truth" I don't mean anything fancy about logic; I just mean the degrees to which things are the ways they are. For example, saying that "it is cold out there" is true to a higher degree just means it's colder out there. There is a formal system called fuzzy logic which takes intermediate degrees of truth as real properties of propositions, with ordinary truth and falsehood at the extremes. In my view this is unnecessary. [↑](#footnote-ref-63)
64. For some details on how this standard-setting works, see my article "A Simple Logic for Comparisons and Vagueness", *Synthese* 123(2), 2000, pp.263-278; and Delia Graff, "Shifting Sands: An Interest-Relative Theory of Vagueness", *Philosophical Topics* 28(1), 2000, pp. 45–81. [↑](#footnote-ref-64)
65. Any direct attempt to do so leads to what is called the so-called *lottery paradox*. Suppose that we take any probability higher than, say, 99% as sufficient for believing something outright, i.e. treating it as certainly true. One consequence of this would be that you could not rationally participate in a lottery with more than 100 tickets. This is because the probability of any single ticket losing would be above 99%. Therefore, for any particular ticket that you bought, you would have to accept as certain truth the proposition that that ticket will lose. And it is plainly irrational to bet good money on something that one categorically believes will not happen. In real life, of course, it might well be rational for someone to invest in a lottery with a high enough pay-out to offset a less-than-1% chance of winning. Setting the uniform standard for categorical acceptance even higher, say to 99.999%, will not fix this problem, since we can always imagine a different lottery with a sufficiently large pay-out to make it rational for someone to bet in favor of whatever tiny chance of winning remains. See H. E. Kyburg, *Probability and the Logic of Rational Belief*, Middletown, CT: Wesleyan University Press (1961), p. 197. [↑](#footnote-ref-65)
66. To define this common sort of error (which might be called the *fallacy of categorical conjunction*, or just the *categorical fallacy*) schematically, just note that none of these argument forms is logically valid:

    if *A*, then *p*

    if *B*, then *q*

    if *A*, then *p* and *q*

    mostly *p*

    mostly *q*

    mostly *p* and *q*

    probably *p*

    probably *q*

    probably *p* and *q*

    But this one, which can be derived from any of those three by simplifying their qualified statements into categorical ones, *is* valid:

    *p*

    *­­q*

    *p*and*q*

    So, if we ever lose track of the information that is contained in the three qualified forms of argument, reducing them to their categorical "fragments" just for the sake of easy articulation, a bad piece of reasoning can be hidden under what appears to be a good one. [↑](#footnote-ref-66)
67. See Will, "Voters Prefer Divided Government", Seattle Post-Intelligencer, September 18, 2008. [↑](#footnote-ref-67)
68. See for example Richard S. Lindzen, "Climate Science: Is it currently designed to answer questions?” *Global Research* November 30, 2009.

    [↑](#footnote-ref-68)
69. Nassim Nicholas Taleb discusses such categorical neglect of unlikely disasters at length in *The Black Swan* (Random House, New York, 2007). [↑](#footnote-ref-69)
70. Bayes’s Theorem says that the new, adjusted probability of any hypothesis *H* being true, given some new piece of evidence *E*, is equal to the prior probability of *H* times the conditional probability of *E* given *H*, all divided by the probability of *E*, or:

    P(*H/E*) = P(*H*) x P(*E/H*) .

    P(*E*)

    Suppose, for example, that you discover that a sandwich you have just made for yourself has disappeared. Probably one of the children took it as a joke or out of thoughtless greed, but perhaps the culprit your dog. Next, you discover the dog cringing behind the couch. Does this new evidence increase the probability that it was the dog who stole your sandwich? Well, it does if your dog tends to hide behind the couch more when he is guilty than otherwise. Let your initial suspicion of the dog, before you found him cowering behind the couch, have given you a 20% subjective probability for the hypothesis (*H*) that he, not one of the children, stole the sandwich. And suppose that he hides behind the couch for one reason or another (*E*) about 40% the time, and further that *if* he stole the sandwich, then he would be hiding afterwards (*E/H*) with a high probability of 90%. The adjusted probability for the dog’s guilt is now P(*H/E*) = (0.2 x 0.9)/0.4 = 0.45. So, the proposition that it is the dog who stole your sandwich is not quite probably true yet, but your subjective probability for this belief has risen from 20% to 45% upon your finding him behind the couch. [↑](#footnote-ref-70)
71. How to make the right initial probability assignments is something of a mystery. To many Bayesians this doesn’t even matter, though, since enough additional evidence will eventually adjust these probabilities, however arbitrary their initial assignment, into an approximately correct relation to each other. If we set the initial probabilities to anything greater than 0 and less than 1, this will never result in an automatic yes-or-no answer to an empirical question, but only a higher or lower degree of rational belief as evidence for and against a proposition accumulates. However, assignments sufficiently *close* to 0 or 1 do make it difficult to perceive and adjust to contrary evidence. I will discuss this problem later on. [↑](#footnote-ref-71)
72. See Ernest Sosa, “Intellectual Virtue in Perspective,” in *Knowledge in Perspective: Selected Essays in Epistemology* (Cambridge: Cambridge University Press, 1991) p. 240f. [↑](#footnote-ref-72)
73. *Theaetetus* 201c-210a [↑](#footnote-ref-73)
74. This problem is named for its discoverer, Edmund Gettier. See "Is Justified True Belief Knowledge?” Analysis, v. 23. (1963), pp.121-123. [↑](#footnote-ref-74)
75. See Peter Unger, “In Defense of Skepticism”, *Philosophical Review* 80 (1971), pp. 198-211. [↑](#footnote-ref-75)
76. The actual full quote is, “You can’t do that! This patient is uninsured!” suggesting something like bureaucratic or financial impossibility. [↑](#footnote-ref-76)
77. It is odd that even moral principles will sometimes take a back seat to the conventions of discussion in this way. Of course, in some societies, or among some people in any society, they will not, and people will refuse to speak politely with those who do not share their moral assumptions. But in a cosmopolitan society where people must routinely deal with others who have different convictions, it seems we must remove some of our controversial beliefs to a more private space. I will discuss this sort of example at greater length in Chapter 7. [↑](#footnote-ref-77)
78. This sort of hypothesis is familiar these days to almost everyone from science-fiction movies like *Inception*, *Total Recall* and *The Matrix*. [↑](#footnote-ref-78)
79. *The Analysis of Mind*,London: George Allen and Unwin (1921), p.159. [↑](#footnote-ref-79)
80. Hume, *Enquiry Concerning Human Understanding* (P. H. Nidditch, ed., Oxford University Press, Oxford, 1975) , section VII, p. 60f. [↑](#footnote-ref-80)
81. This argument itself includes deductive reasoning, of course. We need to *use* at least a few basic deductive formulas even to make coherent sense about those very formulas. But even this does not entail that those formulas are always correct. [↑](#footnote-ref-81)
82. David Lewis speaks evocatively in this regard of the "elusiveness" of knowledge. See "Elusive Knowledge", *Australasian Journal of Philosophy* 74 (1996), pp. 549-567. [↑](#footnote-ref-82)
83. E.g. Unger, *op. cit.*, though he has modified his views in recent years. See *Philosophical Relativity* (University of Minnesota Press, Minneapolis, 2002). [↑](#footnote-ref-83)
84. For a full standard contextualist analysis of the skeptical problem, see Keith DeRose, "Solving the Skeptical Problem", *Philosophical Review* 104 (1995), pp.1-52. For a thorough account of my "conditionalist" alternative, see "Antiskeptical Conditionals", *Philosophy and Phenomenological Research* 73(3) (2006), pp. 505-536. [↑](#footnote-ref-84)
85. This must be clearly distinguished from saying that it really means "If I am not undetectably deceived, then I know that *p*". The latter is approximately Fred Dretske's well-known externalist analysis of knowledge claims. (See “Epistemic Operators", *Journal of Philosophy* 67 (1970), pp. 1007-1023.) Dretske argues that we do know many categorical propositions absolutely – for example, I know absolutely that I am holding a pen right now – but that this knowledge depends on the actual *truth* of the relevant background assumptions, including my not being undetectably deceived. An advantage of Dretske's externalist analysis is that we end up having outright knowledge of categorical facts, a result that satisfies a lot of people's epistemic aspirations and linguistic intuitions. The disadvantage is that we can never *know* that we have such knowledge, because we cannot actually tell whether our background assumptions are true. So, if we are *lucky enough* to live in a world that really matches our background assumptions, then we have absolute categorical knowledge, but we can never know whether we are lucky in that way, so we can never know whether we have any such knowledge or not. This puts us in the same basic situation as dogs and little children, who know such things as that there is food in front of them without having the capacity to know that they know anything at all. We might well ask what good this kind of knowledge, the stuff that Sosa calls "animal" knowledge, actually does us. But Dretske doesn't need to answer that question if he's correct and this is simply what we *mean* by the term "knowledge". [↑](#footnote-ref-85)
86. For readers familiar with formal logic, here is my theory of knowledge in a nutshell. I am ignoring probabilities, extents, and other possible qualifications here in order to get at what I think is essential: the implicit *conditionality* of all ordinary claims to knowledge.

    symbols:

    let a be any set of background assumptions

    let p be any proposition

    let s be any person

    let x → y mean “if x, then y”

    let x & y mean “x and y”

    let □ mean “this is absolutely necessary:”

    let "□" mean what we mean by "this is necessary:" in casual speech

    let B mean “believes absolutely”

    let "B" mean what we mean by "believes" in casual speech

    let K mean “knows absolutely”

    let "K" mean what we mean by "knows" in casual speech

    my central argument:

    (i) Ksp = Bsp & □(Bsp → p) premise

    (ii) "K"sp = "B"sp & "□"("B"sp → p) premise

    (iii) "B"sp = Bs(a → p) premise

    (iv) "□"p = □(a → p) premise

    (v) "K"sp = Ks(a → p) conclusion

    a proof of (v) from definitions (i) – (iv):

    (1) "K"sp = "B"sp & "□"("B"sp → p) premise (ii)

    (2) "K"sp = Bs(a→ p) & **□**((a → (Bs(a →p) → p)) (1) and premises (iii) and (iv)

    (3) "K"sp = Bs(a→ p) & □(Bs(a →p) → (a → p)) (2) and logical rules

    (4) "K"sp = Ks(a→ p) (3) and premise (i) [↑](#footnote-ref-86)
87. As an aside, let me note quickly two objections to this conditional analysis of knowledge claims. One is that my theory seemingly amounts to saying that knowledge claims really mean something empty like “if all my evidence is correct, I am right”, or just "if I am right, I am right" – too trivial to count as the knowledge that we think we have, or even as knowledge worthy of the name. But careful consideration will show that all of the *content* we express in ordinary knowledge claims exists in these conditionals as well. It is as if we took a person's total system of perceptive beliefs articulated categorically, and put one big "If I am not undetectably deceived" in front of it. Nothing else has to change. The other, more technical objection is that my analysis seems to produce the wrong results for people who actually happen to be brains in vats or otherwise undetectably deceived. In their case, any casual statement like "I know that I am holding a pear" would seem to become true automatically, because the conditional "If I am not undetectably deceived, then I am holding a pear" can be said to be true whenever *either* the consequent "I am holding a pear" is true or the antecedent "I am not undetectably deceived" is false. But this objection presupposes that the conditional above is a simple *material* conditional, equivalent to a disjunction with a negated antecedent: *p* → *q* = ~p v q. The logical form of what I say we really mean when making knowledge claims is the more complex (but arguably more intuitive) *subjunctive* conditional, for which that equivalence is false. (See Robert Stalnaker, "A theory of conditionals", in Nicholas Rescher, ed., Studies in Logical Theory, Oxford, Blackwell, 1968). These two objections among others are discussed at greater length in "Antiskeptical Conditionals", *op. cit*. [↑](#footnote-ref-87)
88. On this point see John Searle, "Literal Meaning", *Erkenntnis* 13 (1978), pp. 207-224. [↑](#footnote-ref-88)
89. Augustine, *Confessions* VI.5 (trans. R. S. Pine-Coffin, London: Penguin, 1961), p. 117. [↑](#footnote-ref-89)
90. Note that in my terms the same fact, for example that there is a mosquito on your neck, can be either a first or a second-order fact, depending on whether it is learned directly or through testimony. [↑](#footnote-ref-90)
91. *Enquiry Concerning Human Understanding*  X, (P. H. Nidditch, ed., Oxford University Press, Oxford, 1975)

    p.111. [↑](#footnote-ref-91)
92. 4 See Reid's "*Of the Analogy between Perception, and the Credit we give to Human Testimony"*, Section XXIV of *An Inquiry into the Human Mind: on the Principles of Common Sense* in Keith Lehrer and Ronald Beanblossom, eds., Thomas Reid, *Essays on the Intellectual Powers of Man* (Hackett, Indianapolis, 1983). [↑](#footnote-ref-92)
93. C. A. J. Coady, *Testimony: A Philosophical Study*, (Oxford: Oxford University Press, Oxford, 1992); also Tyler Burge, "Content Preservation", *Philosophical Review* 102(4) (1993), pp.457-488. [↑](#footnote-ref-93)
94. The first of these predispositions is encoded in what Noam Chomsky called abstractly universal grammar and which seems to be located in what contemporary neuroscience calls the language module of the brain. The latter two predispositions are what Reid called the principles of veracity and credulity (*op.cit*., p. 240). [↑](#footnote-ref-94)
95. By "bootstrapping" I mean nothing as specific as Clark Glymour’s theory of confirmation in *Theory and Evidence* (Princeton University Press 1980), just the general idea that from an initially small base of evidence we may develop provisional theories that enable us to obtain more helpful evidence, then better theories, then even better evidence, and so on. [↑](#footnote-ref-95)
96. Here is Plato in the *Theaetetus* 201c: “...suppose a jury has been justly persuaded of some matter which only an eye-witness could know, and which cannot otherwise be known; suppose they come to their decision upon hearsay, forming a true judgment: then they have decided the case without knowledge, but, granted they did their job well, being correctly persuaded...” (Trans. M. J. Levett, Indianapolis: Hackett, 1992, p. 80). See also M. F. Burnyeat, “Socrates and the Jury: Paradoxes in Plato’s Distinction between Knowledge and True Belief,” *Proceedings of the Aristotelian Society* Supp. 54, 1980, pp. 173-191. Locke agrees, using the term probability to cover all justified belief short of knowledge: “Probability then, being to supply the defect of our Knowledge, and to guide us where that fails, is always conversant about Propositions, whereof we have no certainty, but only some inducements to receive them for true. The grounds of it are, in short, these two following: First, the conformity of any thing with our own Knowledge, Observation, and Experience. Secondly, The Testimony of others, vouching their Observation and Experience.” *Essay Concerning Human Understanding* Book IV, Chapter XIV (Ed. Peter Nidditch, Oxford: Oxford University Press, 1975, pp. 655-656). [↑](#footnote-ref-96)
97. There are, of course, deeper problems as well, regarding which of indefinitely many definable properties one ought to project through induction, and whether there is any reason at all to rely on this form of reasoning in the first place. I must ignore these here. If we are to take the process seriously at all, enumerative induction must be viewed as providing only a defeasible, *prima facie* case for belief in its conclusions. [↑](#footnote-ref-97)
98. There is a long-standing dispute among empiricists over whether the fundamental objects of perception are ordinary physical objects and their properties, or "sense-data", which are more momentary and subjective. I do not mean to imply anything either way with respect to this dispute. My examples about apparent cats can, at the cost of some awkwardness, be replaced with ones about red spots, etc. [↑](#footnote-ref-98)
99. Ordinary "random" picking is not technically random, since it is generated by some kind of human choice, and we are never perfectly random agents. But such arbitrary choosing of examples might as well be random from the map’s designer’s point of view, which is what matters to my point. [↑](#footnote-ref-99)
100. If we took fifty-percent truthfulness as a minimum for general reliability, the chance of being misled randomly ten times in a row would seem to be 1/210, or .00098. [↑](#footnote-ref-100)
101. This argument would seem a lot more plausible if phrased in terms of probabilities, but again, I am trying to keep things simple so as not to distract from the main point. [↑](#footnote-ref-101)
102. To put it somewhat formally, second-order reasoning from testimony generally works like this:

     Source *S* made statement *p* on topic *T* under conditions *C*.

     Source *S* usually tells the truth on topic *T* under conditions *C*.

     Therefore, statement *p* is probably true.

     Therefore, probably *p*.

     As adults, most of us have learned inductively to temper our credulity according to *T* and *C* for every *S*.

     In our very early epistemic situations, though, where *S* is an as yet uncontroverted parent, *T* could be any topic at all, and *C* any circumstances. [↑](#footnote-ref-102)
103. Philosophers sometimes refer to any non-professional knowledge of human nature (rather demeaningly, I think) as "folk psychology". Several recent writers on testimony have urged that the best inferential path to testimonial reliability is through this kind of general knowledge. See Traiger 1993, Lyons 1997, and Faulkner 1998. Hume himself, as cited by Traiger, claims in the Treatise that there is not "any thing but our experience of the governing principles of human nature, which can give us any assurance of the veracity of men." (p. 113, quoted by Traiger p. 144). I see it as one, but only one, important feature of a complex inferential process. [↑](#footnote-ref-103)
104. The website *ratemyprofessors.com*, for example, is a good idea and probably a somewhat useful source of information, though its reliability is questioned because of its small sample of self-motivated participants, which is thought (by professors, at least) to over-represent students with axes to grind. [↑](#footnote-ref-104)
105. I should note that there are some philosophers, called direct realists, who claim the opposite, that we actually grasp external objects of perception and their properties, such as the rabbit and the white color of its fur, directly, not as constructs out of sensory information. Panayot Butchvarov uses direct realism to maintain an infallibilist definition of knowledge while rejecting its usual skeptical consequences in *Skepticism about the External World* (Oxford University Press, New York, 1998). A main problem for direct realists lies in accounting for apparent objects that are not really there, like a hallucinatory or a fictional white rabbit. Butchvarov discusses this problem at length in *Being Qua Being*: *A Theory of Identity, Existence, and Predication* (Indiana University Press, Bloomington, IN, 1979). [↑](#footnote-ref-105)
106. This abductive argument is fundamental to contemporary scientific realism, the theory that science studies an objective reality that is essentially independent of minds or linguistic conventions. See Richard Boyd, “The Current Status of Scientific Realism,” in J. Leplin (ed.), Scientific Realism, Berkeley CA: University of California Press (1984), pp. 41–82. [↑](#footnote-ref-106)
107. Mill makes this argument in Examination of Sir William Hamilton's Philosophy (Ch. XII, Appendix, Longman, *et al.* London 1865). Bertrand Russell approves, but denies that the argument properly inductive, preferring to reserve that term for arguments which extend only to potentially observable new instances. See “The argument from Analogy for Other Minds” in his *Human Knowledge: Its Scope and Limits* (New York: Simon and Schuster, 1948), p. 193. [↑](#footnote-ref-107)
108. Stuart Hampshire has an expanded version of the analogical argument, which is the closest thing I have found to my view in print. He says that one can gather evidence for other minds, not just from the correlations between one's own feelings and behavior, but also from the observed correctness of 'methods of inference' used by others to establish one's own states of mind. See Hampshire, "The Analogy of Feeling", *Mind* 61, no. 241 (January 1952), 1-12. [↑](#footnote-ref-108)
109. For further details, see my “Other Voices, Other Minds”, *Australasian Journal of Philosophy* 78(2), 2000, pp. 213-222. I should probably confess that this argument made its first appearance as a joke I told one day in graduate school. The professor, a young Richard Fumerton, was lecturing on the problem of other minds and had just asserted that inductive arguments for believing in other minds are worthless, since induction relies on making inferences from many cases, but there is only one case in which we can observe a person having a mind, namely our own, which yields us no substantial probability that other people also have minds. I interrupted him to boast that I could establish a probability of around 75% that he, Fumerton, had a mind. He politely asked how this could be, I responded, “Do you have a mind?”, he answered, “Yes”, and I said, “Well, in my experience you only lie about a quarter of the time.” As I remember it, we both agreed after class that this facetious argument deserved further investigation. [↑](#footnote-ref-109)
110. This is just an instance of Alfred Tarski's famous definition of truth as disquotation: "*p*" is true if and only if *p*. See "The Semantic Conception of Truth and the Foundations of Semantics", *Philosophy and Phenomenological Research* 4 (1944), pp. 341-375. [↑](#footnote-ref-110)
111. As with other minds, there may also be some good first-order arguments for our belief in the physical world, e. g. the argument that physical causes best explain the regularities among our sense impressions. [↑](#footnote-ref-111)
112. Fifth Meditation, *op. cit.* [↑](#footnote-ref-112)
113. I am ignoring many distinctions among varieties of egoism, moral skepticism, and related views that do not matter to the argument here; this basic view dates back through Hobbes and Machiavelli at least to Plato's *Republic*. For a thorough treatment of rational egoism see Richard Fumerton, *Reason and Morality* (Cornell University Press, Ithaca, NY, 1990). [↑](#footnote-ref-113)
114. Hume argues for the insufficiency of belief in causing action in his *Treatise of Human Nature* II.3, ed. L.A. Selby-Bigge, 2nd ed., revised by P.H. Nidditch (Oxford: Clarendon Press, 1978), p. 414. [↑](#footnote-ref-114)
115. My mother does not mean to imply here that "it's wrong" means literally the same thing as "don't do it". That theory, called prescriptivism, belongs instead to R. M. Hare, in *The Language of Morals* (Oxford: Clarendon Press, 1952). [↑](#footnote-ref-115)
116. *The New World,* Volume 5 (June, 1896), pp. 327-347. [↑](#footnote-ref-116)
117. John Hardwig argues along similar lines for a somewhat more restricted principle of deference, to the effect that laymen can have good reason to defer to experts, especially in matters of science. He declines to count this reason as a kind of evidence. See his “Epistemic Dependence,” *Journal of Philosophy* LXXXII, 7 (1985): 335-349. [↑](#footnote-ref-117)
118. Assuming one of us is right in this disagreement, the chance of its being me is .2423 and of its being you is .7577, which makes you 3.13 times as likely to be right as me. [↑](#footnote-ref-118)
119. The probability of my being right while both of you are wrong is only .1536, whether you agree with each other or not. There is actually not enough information here to calculate the probability on the further hypothesis that you agree with each other, but it will have to be even lower. [↑](#footnote-ref-119)
120. See Solomon Asch, “Opinions and Social Pressure”, *Scientific American* 193 (1955), 31-35. [↑](#footnote-ref-120)
121. Stanley Milgram, "Behavioral Study of Obedience", *Journal of Abnormal and Social Psychology* 67/4 (1963), pp. 371–8. [↑](#footnote-ref-121)
122. The outlaw Sundance, not mentally disabled but a humble person intellectually, relies in much the same way on his partner in the movie *Butch Cassidy and the Sundance Kid*: “You just keep thinking, Butch. That's what you're good at.” [↑](#footnote-ref-122)
123. Sorry – what I called the planet Pluto has just been reclassified by astronomical authorities as a mere "dwarf planet", of which there are evidently quite a few out in the Kuiper belt. That’s what I now believe, because the experts say so. [↑](#footnote-ref-123)
124. New York, Norton (1961), p. 11. [↑](#footnote-ref-124)
125. *The Future of an Illusion*, trans. James Strachey (New York, Norton, 1961), p. 60. [↑](#footnote-ref-125)
126. *Why I am Not a Christian and other essays on religion and related subjects* (New York, Simon and Schuster, 1957), p. 22. [↑](#footnote-ref-126)
127. "Is Science a Religion?", *The Humanist* 57(January-February 1997), pp. 26 [↑](#footnote-ref-127)
128. # Reprinted in *Early Modern Philosophy: Essential Reading*s with Commentary, eds. Aloysius Martinich, Fritz Allhoff, Anand Vaidya (Malden, MA, Blackwell, 2007), p. 145.

     [↑](#footnote-ref-128)
129. The philosopher William Alston discusses a similar case in *Epistemic Justification: Essays in the Theory of Knowledge* (Ithaca, Cornell University Press, 1989), pp. 95 and 145-146, and argues that a person who grows up inside a culturally “isolated community” cannot be *blamed* for adopting the peculiar traditions of his elders, in part because he will never have considered questioning them. In this weak sense, Alston allows that such people can be “deontologically justified” in their beliefs (his overall point is that such a conception of justification is faulty, precisely because it leads to such cases). The position that I am suggesting here is much stronger. I say that ordinarily, a culturally isolated individual is rationally *obligated* to believe in the local traditions, even if he has considered them quite thoroughly and he directly perceives what is wrong with them. The only possible exception is if the individual has sufficiently good reason to believe that he is more reliable on the subject than the sum of his elders. [↑](#footnote-ref-129)
130. *Enquiry* X, *op cit.* pp. 108-131. [↑](#footnote-ref-130)
131. Indeed, if he wanted to, Dawkins could easily say: “Yes, the brain viruses I'm talking about are transmitted precisely through valid second-order reasoning. This is how healthy people *catch* such diseases.” That sort of approach might allow Dawkins to develop his core theory of memetics without disparaging the rationality of most other human beings. [↑](#footnote-ref-131)
132. This is not a matter of the surface form of the statements, of course, but a matter of the total propositional content that is being passed on, whether through words, gestures, or anything else. [↑](#footnote-ref-132)
133. I suppose I should say multiply rather than add. Other things being equal, the probability assigned to probabilized testimony would presumably be the product of the source’s reliability (probability of being right about things) and the probability the source assigns to the proposition in question. So, if a source I have found to be right 75% of the time tells me there is an 80% chance of rain tomorrow, I should base my own decisions on a 60% chance of rain. [↑](#footnote-ref-133)
134. Unsurprisingly, Descartes experienced all three as a young man before deciding that he had to rethink everything from scratch all on his own, as he describes in the *Discourse on Method* Part One (op. cit.). [↑](#footnote-ref-134)
135. The recent case of arch anti-socialist Andres Breivik's murder spree in Oslo provides a clear-cut exception. [↑](#footnote-ref-135)
136. Learning to draw is more crucially a matter of developing this ability to see without interpretation than of learning to handle a pencil. See Betty Edwards, *Drawing on the Right Side of the Brain*, New York, J. P. Tarcher, 1989. [↑](#footnote-ref-136)
137. Sixteenth Century Poetry: An Annotated Anthology, *Wiley Online Library*, 2007: http://onlinelibrary.wiley.com/doi/10.1002/9780470996317.fmatter/pdf [↑](#footnote-ref-137)
138. J. Thomas Looney, *"Shakespeare" Identified In Edward De Vere the Seventeenth Earl of Oxford (*Frederick A. Stokes Company, New York: 1920.), pp. 246-7. [↑](#footnote-ref-138)
139. Bill Bryson, *Shakespeare: the world as stage* (HarperCollins, New York, 2007), p.188. [↑](#footnote-ref-139)
140. *The Structure of Scientific Revolutions* (second ed., Chicago, University of Chicago Press, 1970). [↑](#footnote-ref-140)
141. See C.J. Brainerd, *The Science of False Memory* (Oxford University Press, Oxford 2005). A friend who was among the 400,000 or so people who attended Woodstock jokes that at least a million of them remember being there. [↑](#footnote-ref-141)
142. See Dorothy Rabinowitz, *No Crueler Tyrannies: Accusation, False Witness, and Other Terrors of Our Times* (New York, Free Press, 2003). [↑](#footnote-ref-142)
143. See Stephan Jay Gould, *The Mismeasure of Man* (revised and expanded)*, New York, Norton, 1996.*  Gould expanded the book in order to critique Richard Herrnstein and Charles Murray, *The Bell Curve*, New York, Free Press 1994. [↑](#footnote-ref-143)
144. The Bureau of Justice Statistics is the usual source of numbers for both sides, with many tables online at: http://bjs.ojp.usdoj.gov/index.cfm?ty=pbse&sid=1. [↑](#footnote-ref-144)
145. The feminist leader Gloria Steinem coined the term "click experience" for the same sort of gestalt-like perceptive shift that happens when women whose consciousnesses have just been raised abruptly stop perceiving themselves as caged rabbits (or “bunnies”!) of traditional femininity and become free-flying ducks of liberated womanhood. [↑](#footnote-ref-145)
146. For an essential collection of essays on the topic by ex-communist writers see Arthur Koestler, *et al*, *The God that Failed* (Harper, New York, 1950). [↑](#footnote-ref-146)
147. Remember Bayes's Theorem: P(*H/E*) = P(*H*) x P(*E/H*). The closer P(*H*) gets to 1, the less P(*E*) matters in the

     P(*E*)

     equation, since P(*E/H*) must approach 1 simultaneously, and therefore P(*H/E*) must do the same. If we assign a

     P*(E)* P(*H*)

     prior probability of exactly 1to the hypothesis *H*, then the posterior probability P(*H/E*) must be 1as well for any value over 0 that we assigned to evidence *E*, since whenever P*(H)*= 1, P(*E/H*) = P*(E)*. [↑](#footnote-ref-147)
148. In case it isn’t adequately clear, let me explain the distinction between an epistemic black hole and a closed epistemic community. The former is an epistemic situation for an individual, in which it is rationally impossible for that person to perceive evidence against a given belief. The latter is a social group whose where membership is restricted to people who agree with a given doctrine. It is plainly easier for black holes to occur for individuals in closed epistemic communities, but it is quite possible for either exist without the other. Thus, a single person who has become (for whatever reason) totally certain that everybody else is part of an extremely powerful conspiracy, so that all evidence of other people’s good will is necessarily construed as deceptive, could be in an epistemic black hole while inside a perfectly open epistemic community. Alternatively, somebody in a closed epistemic community may nevertheless rationally conclude on the basis of unsuppressed evidence that the community is wrong on basic matters of doctrine and thereby *leave* that community, as some ex-Scientologists are said to have done. [↑](#footnote-ref-148)
149. It might seem a reasonable rule for Bayesians to restrict all prior probabilities to the range strictly in-between 0 and 1 to avoid utterly unrevisable empirical beliefs. But under the circumstances I am discussing, it seems that authoritative testimony would force us to break this rule in any case. [↑](#footnote-ref-149)
150. At the same time, I cannot *forget* that this belief is contrary to my first-order perceptions, and the combination will produce a kind of epistemic ambivalence that I’ll call *uncertain certainty.* I will discuss this problem at length in chapter 7. [↑](#footnote-ref-150)
151. I know next to nothing about the religions of ancient Egypt or Peru, but I imagine that they had less need for a similar doctrine of faith, since it was comparatively rare for their adherents to get into arguments with unbelievers. [↑](#footnote-ref-151)
152. For convenience as opposed to a delight in paradox, I will continue using the term *rational irrationality* for this learned attitude toward justifying beliefs. Regardless of the label, it is still clearly distinguishable from the *irrational* sort of irrationality that Linus applies to his Great Pumpkin. In his case, we see how absurd it is for someone to treat “pure faith” as a reason for belief when it is without foundation in authoritative testimony. [↑](#footnote-ref-152)
153. As I write this, a Pakistani woman who converted to Christianity has been sitting in prison for several months

     awaiting hanging for the crime of blasphemy, and a provincial governor who argued for clemency in her case has been assassinated. See Salman Masood and Carlotta Gall, "Killing of Governor Deepens Crisis in Pakistan", *New York Times*, January 11, 2011. [↑](#footnote-ref-153)
154. "It is certain, because it is impossible." [↑](#footnote-ref-154)
155. See Gary Gutting, *French Philosophy in the Twentieth Century*, (Cambridge, Cambridge University Press, 2001),

     p. 295f. The variant that has realist philosophers most often rolling their eyes is the feminist critique of *phallogocentrism*, which is the idea that men's domination over women includes "phallic" or male-centered logic as one of its tools. The argument here is that men insist on defining reason in black-and-white, deterministic ways, while rejecting feminine forms of thinking that are more holistic and empathetic, as a means of enforcing patriarchal authority by making women seem irrational, hence incapable of using power responsibly. [↑](#footnote-ref-155)
156. Kuhn himself showed deep ambivalence about this interpretation in his postscript to *The Structure of Scientific Revolutions* (second ed., Chicago, University of Chicago Press, 1970, pp. 174-210), sometimes almost explicitly endorsing it and other times seeming to pull back. [↑](#footnote-ref-156)
157. *On Liberty and Other Essays*, ed. John Gray (New York, Oxford University Press, USA, 2008) p.59. [↑](#footnote-ref-157)
158. *The Philosophical Writings of Descartes*, vol. 1., ed. John Cottingham, Robert Stoothoff, and Dugald Murdoch (New York: Cambridge 1988), p. 115f. [↑](#footnote-ref-158)
159. As I suggested earlier, Descartes' actual project is rather more expansive; he really wishes to develop the most rational long-term intellectual procedure, not just to calculate the best guesses available at every moment. [↑](#footnote-ref-159)
160. This is not the same thing as the minority behaving more assertively, of course. I just mean the degree of perceptive confidence that our evidence tells us they actually possess. [↑](#footnote-ref-160)
161. I am thinking here of ordinary informative discourse. There may be circumstances where more apparent confidence actually implies a lesser probability of truth. In political discussions, for example, the people most confident in their views might be the ones who pay least attention to their opponents’ arguments. People we count as fanatics are probably less reliable on most things than moderately skeptical thinkers. [↑](#footnote-ref-161)
162. It is not certain even then, of course, given the possibility of a consensus belief being false – something plainly evident in the history of *changes* of consensus among experts in all fields. Unless he has been authoritatively taught otherwise, the rational perceiver must always account for the probability of his authorities being mistaken even when they all agree. [↑](#footnote-ref-162)
163. For example, in the case of Augusto & Michaela Odone, whose amateur medical investigations aimed at finding a treatment for their dying son Lorenzo inspired the 1992 film *Lorenzo's Oil*. [↑](#footnote-ref-163)
164. Just to keep track of things, I also argued that alternative, conditional articulations of the same beliefs *can* sometimes be absolutely known, and that when we say we know things, these are often what we really mean. [↑](#footnote-ref-164)
165. Training in philosophy is seen as an especially good preparation for law school. According to a recent study, philosophy majors (combined with theology majors) are tied for second place in LSAT (the law school entrance examination) scores with economics majors, behind physics and math majors taken together. See Michael Nieswiadomy, *LSAT Scores of Economics Majors: The 2008-2009 Class Update* (June 25, 2009). Available at SSRN: http://ssrn.com/abstract=1430654. [↑](#footnote-ref-165)
166. The process of digestion is an apt and common metaphor for such degrees of understanding: some ideas you take in have been fully digested, some less so, and some you are still chewing over, as we say. Some ideas you might find palatable, others hard to swallow; others you don’t digest at all but just regurgitate for an examination. (This analogy can be extended more extravagantly; I’ll leave this as an exercise for younger readers.) [↑](#footnote-ref-166)
167. Actually, there are several competing theories of modal logic, which differ on axioms that don’t strike all philosophers as equally intuitive. Going into it, you’d expect logic *itself* to be simple and obvious, right? And some parts of it are. But at the expert level it’s as much a mess as everything else. [↑](#footnote-ref-167)
168. See for example Olivia Wu, "Rice goes digital cooked the fuzzy logic way", *San Francisco Chronicle*, December 10, 2003. [↑](#footnote-ref-168)
169. Yes, decades or centuries after you’ve been buried in threadbare obscurity, despised and ridiculed by all the so-called authorities, posterity will sing your praises! Now, get back to work! [↑](#footnote-ref-169)
170. There used to be a big, bluff, old-school reporter named Jack Germond who appeared every week on a political talk show called *The McLaughlin Group*. As part of the show’s routine, the host would frequently demand predictions from his panel: “Prediction! Who will win the presidential election this November?” And Germond would sometimes get a laugh from his fellow panelists by answering, “How the hell would I know?” This was a good sort of running joke, not because Germond really did know what was going to happen – none of them did, of course – but because his basic *job* on the show was to state plausible opinions and argue for them, not to assert what he actually thought he knew. [↑](#footnote-ref-170)
171. Even the über-Communist Mao Zedong agreed with this point during his evanescent “Hundred Flowers” campaign: “Let a hundred flowers blossom and let a hundred schools of thought contend.” I’ll discuss the more political forms of argumentation in some detail later on. [↑](#footnote-ref-171)
172. I don’t mean this as a statement of ideology, just a general observation. Competition can also sometimes produce inferior or even absurd results, as with crude television shows or breakfast cereal made out of cookies. And there are certainly exceptional cases where geniuses create wonderful things right off the bat (e.g. the Ford Model T or the Apple II computer) or where top-down leadership is required for a consolidated one-time effort (e.g. the Manhattan Project). [↑](#footnote-ref-172)
173. *op. cit.*, Part II (p. 20f.). [↑](#footnote-ref-173)
174. This also explains why the human forces typically defeat more powerful robot armies in science fiction: we understand what's going on, and they just follow rules. It is also why the modern American military depends more and more on the tactical insights of low-level personnel (the so-called *strategic* *corporal*) rather than simply on drilled maneuvers and obedience – as the German forces did to great effect in World War Two. [↑](#footnote-ref-174)
175. *Republic I*, 331c-d. Don't quote my casual rendering. [↑](#footnote-ref-175)
176. *Republic, Book I* 336b – 354b. This pre-modern Sophist theory is, as I noted earlier, echoed in numerous modern and postmodern versions of politically grounded moral relativism. [↑](#footnote-ref-176)
177. Socrates himself refused on principle to charge for lessons, though he wouldn’t say no to a nice meal or a few cups of wine. [↑](#footnote-ref-177)
178. *op. cit.*, p. 112. [↑](#footnote-ref-178)
179. Walter Soffer argues for this position in *From Science to Subjectivity: An Interpretation of Descartes' Meditations* (Westport, CT: Greenwood, 1987). [↑](#footnote-ref-179)
180. For what it’s worth, I think that I am making the same sort of effort here by saying how it *seems to me* that disagreement and belief arise and function (after sitting for years in warm rooms trying to figure these things out for myself, inspired by the efforts of Descartes) without kidding myself that I’ve found the whole truth, or expecting every reasonable person to agree with me. [↑](#footnote-ref-180)
181. Hegel argues that the universe itself is structured in this way in *Phenomenology of Spirit*, trans. by A. V. Miller

     with analysis of the text and foreword by J. N. Findlay (Oxford: Clarendon Press, 1977). [↑](#footnote-ref-181)
182. Ah, yes, but where is home? The science-is-better theory is generally called *naturalism*. See Daniel Dennett, *Darwin’s Dangerous Idea* (Touchstone, New York, 1995) for one prominent source. Give-up philosophies come in a range of types, from ancient Pyrrhonian skepticism to Ludwig Wittgenstein’s therapeutic *linguisticism* (*Philosophical Investigations*, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 1973) to Colin McGinn’s *mysterianism* about consciousness (“Can We Solve the Mind-Body Problem?” in Brian Cooney, *The Place of Mind* , Wadsworth, Belmont, CA , 1999). [↑](#footnote-ref-182)
183. It is actually a crime in some Western European countries to deny the Holocaust. In 2006, the British historian David Irving was sentenced to three years in prison for this crime by an Austrian court, and served thirteen months before being deported and banned from Austria for life. [↑](#footnote-ref-183)
184. See Richard Posner, *How Judges Think,* Cambridge, MA: Harvard University Press, 2008. [↑](#footnote-ref-184)
185. In a recent case, the Chief Justice of the Alabama Supreme Court, Roy Moore, was removed from office by a special panel for defying a federal judge's order to remove a plaque containing the Ten Commandments from the wall of his courtroom. See Jeffrey Gettelman, "Court Orders Alabama’s Chief Justice Removed from Bench", *New York Times*, November 13, 2003. [↑](#footnote-ref-185)
186. All right, I'll tell you. What has always bothered me is that certain forms of argument strike me as valid in a very basic way that aren't included in the standard formal logic that you find in logic books. The "classical" system works quite well to handle arguments like this:

     Jack is thirsty.

     If Jack is thirsty, then Jill is thirsty.

     Therefore, Jill is thirsty.

     But there are other arguments, just as good and of a generally similar nature, that are *not* valid in classical logic, such as:

     Jack is thirsty.

     Jill is more thirsty than Jack.

     Therefore, Jill is thirsty.

     I have a marvelous solution to this problem, but this footnote is too small to contain it. If you're interested, see "A Simple Logic for Comparisons and Vagueness", *Synthese* 123(2), May 2000, 263-278. [↑](#footnote-ref-186)
187. Socrates and Plato disliked democracy because they thought that most people knew nothing about the technical aspects of governance and were too easily swayed by oratory and self-interest. In the *Republic* (488b) Socrates compares democracy to a ship run by an ignorant crew instead of a captain who understands the principles of navigation, and argues for rule by a class of “philosopher-kings” instead. These days, we all seem to prefer some form of democracy over the alternatives, but the tension between popular will and expert authority remains a major systemic problem. I’ll say more on this issue in Chapter 8. [↑](#footnote-ref-187)
188. *Apology* 30e. [↑](#footnote-ref-188)
189. *Republic* 377e. But again, the more dogmatic Socrates becomes in Plato’s dialogues, the more often he is seen as a mouthpiece for Plato himself, who did not pretend to be as skeptical in general as was his mentor. [↑](#footnote-ref-189)
190. In Athenian justice, after a conviction both the prosecution and the defense suggest a punishment, and the jury then decides between the two suggestions. After his joking bid for free food, Plato and other friends of Socrates offered to pay a huge fine for him instead, but the damage had been done. Now that I think of it, though, Socrates actually was provided free meals for the rest of his life – on death row. We should be careful what we wish for. [↑](#footnote-ref-190)
191. The journalist I. F. Stone develops this sort of argument at length in his *The Trial of Socrates* (Anchor, New York 1989). I don’t mean to take a position here on the historical Socrates’s actual guilt or innocence, but only to illustrate a point. [↑](#footnote-ref-191)
192. For an influential popular account of Galileo’s martyrdom, see Bertolt Brecht's play *Life of Galileo*, New York: Penguin Classics (2008). [↑](#footnote-ref-192)
193. For the version of the Galileo case that I am following see Arthur Koestler, *The Sleepwalkers: A History of Man's Changing Vision of the Universe*, London, Penguin (1959), pp. 471-503. As with Socrates, the whole truth about Galileo is probably more ambiguous than I am representing here, and I am not trying to pass judgment here on Galileo as an actual person. I just want to give a plausible example of the usefulness of arrogance in science, and the *story* of Galileo is an obvious choice. [↑](#footnote-ref-193)
194. For an engaging account of the medieval commentators see Richard E. Rubenstein, *Aristotle’s Children: How Christians, Muslims, and Jews rediscovered ancient wisdom and illuminated the Middle Ages* (Harcourt, New York, 2003). [↑](#footnote-ref-194)
195. The great Scholastic philosopher Thomas Aquinas framed his masterwork, the *Summa Theologica*, as one of these formally analyzed disputations. [↑](#footnote-ref-195)
196. I mean, it’s understandable that they condemned him in some way. Burning him at the stake was probably excessive. [↑](#footnote-ref-196)
197. New York: Modern Library, 2001. [↑](#footnote-ref-197)
198. It might still be argued that Galileo had sufficient reason to believe that he was so extraordinarily perceptive that he had no peers at all, or none who disagreed with him. In that case, contrary testimony wouldn't count strongly as evidence for him against his own first-order astronomical conclusions, making it rational for him to rely exclusively on his own authority. If so, this would make him an exception to the general rule that I am using his case to illustrate. [↑](#footnote-ref-198)
199. This isn't generally true in social sciences, though, where elegant and comprehensive theories often miss the mark because the subject matter is not actually simple. [↑](#footnote-ref-199)
200. I don't mean to imply that he was *sincere* when he recanted. Legend has it that while being led away afterwards he muttered "*Eppur si mouve*" (still, it moves), meaning the Earth. Koestler states flatly that he did not say this at his trial. (*Sleepwalkers*, p. 502-503.) [↑](#footnote-ref-200)
201. In a speech in 1992, Pope John Paul II stated publicly for the first time that the Inquisition was mistaken in its judgment of Galileo, though the mistake was made, he said, in good faith. L'Osservatore Romano 44 (1264), November 4, 1992. [↑](#footnote-ref-201)
202. It is controversial among philosophers, psychologists, and biologists whether moral altruism, understood as totally non-self-interested action, is truly possible. Some, for example Elliott Sober and David Sloan Wilson, believe (following Darwin) that altruism can evolve through the mechanism of group selection, where groups of cooperatively self-sacrificing organisms are able to outperform other groups comprising individually fitter, but less mutually cooperative ones. See their *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge: Harvard, 1999). This is an inviting model for the notion of objective rationality that I am suggesting is the basis of science. But my point does not depend on this. I am only claiming that science tends to progress against the purely *epistemic* self-interest of individual participants, in favor of the epistemic interests of the group. As with other species, it is not required that this type of altruistic behavior be recognized as such by human agents. In any case, present-day scientists (and even philosophers) are routinely compensated in financial and many other non-epistemic ways for their efforts, so it is arguably still in their *overall* self-interest to do what they do. [↑](#footnote-ref-202)
203. Koestler, *Sleepwalkers*, pp. 244-248. [↑](#footnote-ref-203)
204. # See Natalie Angier, "Moonlighting as a Conjurer of Chemicals", *New York Times*, October 11, 2010.

     [↑](#footnote-ref-204)
205. # The famous "Unabomber", Ted Kaczynski, is a striking recent example. See Alston Chase, *A Mind for Murder*: *The Education of the Unabomber and the Origins of Modern Terrorism* (New York, Norton 2004).

     [↑](#footnote-ref-205)
206. By successful society I do not mean one that makes us *happier* in the sense of finding greater contentment. It is more a matter of getting what we seek, including wealth, power, and security, and holding on to it in a competitive world. [↑](#footnote-ref-206)
207. College teachers can be amazingly touchy on this point. Once, when I had just started teaching as a graduate assistant, I received a paper in a freshman course in ethics that I now think I reacted to unreasonably. The assignment must have been something like: “Should abortion be legal on demand? Why or why not? Explain your reasons in detail.” And I got this paper from a student that began with the words, “I am a child of God,” laid out the Catholic doctrine on abortion, and explained that she accepted that position simply because she was a good Catholic who obeyed the Church in which she had been raised. After showing it around to my friends, I gave this openly “conformist” paper quite a low grade and some pretty hostile comments, despite its being well-written and directly responsive to the question. You’re in college now and you’re supposed to *think critically* about issues, I told her, not justrepeat what you’ve been told! But I didn’t realize that she was nevertheless thinking and writing quite rationally about the issue at hand. So, though I think I gave her more or less correct instructions for producing acceptable papers in philosophy courses – making it rational for her to write more critically from then on, which she did – I’m sure I failed to explain what was wrong with her initial heartfelt and intelligent response, because I didn’t really understand what was wrong it myself. (I got paid back, though, later, by the way, when, at the end of a paper of my own that criticized some point in Aristotle’s *Metaphysics*, I added a little footnote arguing that Aristotle was still probably right in spite of everything I’d said, on that grounds that he was almost certainly a better philosopher at the height of his career than I was as a mere grad student. This was only a *joke*, of course, but the professor really chewed me out for making such an inappropriate remark.) [rep on Ari §1.3!] [↑](#footnote-ref-207)
208. The *locus classicus* for this notion of conformism is the famous balcony scene in Monty Python’s 1979 film *The Life of Brian*, where the reluctant savior Brian commands his clueless followers to think for themselves, and they all shout their agreement in unison. [↑](#footnote-ref-208)
209. There are occasional partial exceptions like the great early-Twentieth-Century Indian mathematician Srinivasa Ramanujan who, with only a basic education in the field, was able to produce extraordinary new ideas on top of recapitulating much that was already known in the profession at the time. [↑](#footnote-ref-209)
210. *The Structure of Scientific Revolutions*, p. 5. [↑](#footnote-ref-210)
211. This story is told with rather brutal frankness by Watson in his brief memoir *The* *Double Helix* (Atheneum, New York, 1968). [↑](#footnote-ref-211)
212. To some extent, prominent Russian scientists like Andrei Sakharov were even allowed the Western privilege of criticizing government policy. Alexander Solzhenitsyn writes piercingly about the moral struggles of such privileged intellectuals in his novel *The First Circle* (Harper and Row, New York, 1968). [↑](#footnote-ref-212)
213. For example, in his book *Black Holes and Time Warps: Einstein's Outrageous Legacy* (New York: Norton, 1994), Kip S. Thorne writes that he sometime makes bets with Stephen Hawking over their untested hypotheses in physics, but does not bet automatically in favor of his own ideas. (p. 521). [↑](#footnote-ref-213)
214. Their publicity-seeking behavior wouldn’t have mattered much if their experiments had been successfully replicated, but despite occasional equivocal results, this hasn’t happened. For more on this case, and a far more detailed discussion of how authority actually works in science, see Philip Kitcher, "Authority, Deference, and the Role of Individual Reasoning in Science,” in McMullin, E., *The Social Dimensions of Science* (South Bend: Notre Dame, 1992). [↑](#footnote-ref-214)
215. John Hardwig (*op. cit.*, p 346)gives an example of a scientific paper with 99 co-authors: K Abe *et al*, “Charm Photoproduction Cross Section at 20GeV”, *Physical Review Letters* 51 (1984), pp. 156-160..

     [↑](#footnote-ref-215)
216. The tendency of science to get stuck in consensus is well known. Helen Longino argues in *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry* (Princeton University Press, Princeton, NY, 1990) that the scientific community can only retain if it remains responsive to critics from social groups with different interests. Kathleen Okruhlik sees a particular problem with unconscious social discrimination preventing mainstream scientists from even acknowledging the different points of view of excluded groups, in particular women, in “Gender and the Biological Sciences”, *Biology and Society, Canadian Journal of Philosophy*, Supplementary vol. 20 (1994) pp.21-42. Thus, what would appear to be inherently scientific problems may require partly political solutions. [↑](#footnote-ref-216)
217. Linus Pauling, for example, lost much of his gigantic reputation as a Nobel Prize-winning biochemist by committing himself to the radical theory that cancer could be cured with high doses Vitamin C, a theory that has not panned out, though it still has some adherents. See Richard Severo, "Linus Pauling Dies at 93; Chemist and Voice for Peace" *New York Times*, August 21, 1994. [↑](#footnote-ref-217)
218. In philosophy, much of the most revolutionary work is done by people with backgrounds in mathematics, including Pythagoras, Descartes, Russell, Wittgenstein, and more recently, Saul Kripke. [↑](#footnote-ref-218)
219. Darwin only found out about Wallace’s near-scoop because Wallace had sent his work to Darwin for advice on publication. For an engrossing account of Darwin's struggle with his conscience over taking credit for the theory, see Robert Wright, *The Moral Animal: Why We Are the Way We Are* (Vintage, New York, 1995), pp. 287-312. [↑](#footnote-ref-219)
220. See Wilson's own account in his *Naturalist*, Island Press, Washington DC (1994), p. 307. Wilson’s colleagues, the biologists Stephen Jay Gould and Richard Lewontin, both colleagues of Wilson (plus several other co-authors in the Sociobiology Study Group) denounced him and his work in a famous letter to *The* *New York Review of Books* (November 13, 1975) titled “Against Sociobiology”. [↑](#footnote-ref-220)
221. The increasingly frank discussion among psychologists of sex differences in cognitive abilities has not yet spread to ethnic and racial differences, perhaps because such differences simply don’t exist. Richard Herrnstein and Charles Murray’s effort to distinguish social (including racial) groups by average intelligence statistically in *The Bell Curve: Intelligence and Class Structure in American Life* (Free Press, New York, 1994) provoked immediate hostile responses from a long list of prominent scientists in Steven Fraser, ed. *The Bell Curve Wars: Race, Intelligence, and the Future of America* (Basic Books, New York, 1995), and the question has not been taken up again since, at least not in a way that’s visible to the general public. [↑](#footnote-ref-221)
222. # For a sharp response to this forced resignation, see the editorial "Prejudice Wins" in *The Washington Post*, February 22, 2006.

     [↑](#footnote-ref-222)
223. See for example the MIT meteorologist Richard Lindzen, "The Climate Science Isn't Settled", *Wall Street Journal*, November 30, 2009. [↑](#footnote-ref-223)
224. Buck v. Bell, 274 U.S. 200, Holmes, J., Opinion of the Court, Decided May 2, 1927.

     [↑](#footnote-ref-224)
225. Some medical scientists now argue that millions of (mostly children’s) have been lost unnecessarily over the past four decades due to the ban on DDT. See Amir Attaran, Donald R. Roberts, Chris F. Curtis, & Wenceslaus L. Kilama, "Balancing risks on the backs of the poor", *Nature Medicine* 6, 729-731 (2000). [↑](#footnote-ref-225)
226. Here is a quote from the opening of this alarming book (New York, Sierra Club/ Ballantine Books, 1968), which sold over three million copies:

     *The battle to feed all of humanity is over. In the 1970s and 1980s hundreds of millions of people will starve to death in spite of any crash programs embarked upon now. At this late date nothing can prevent a substantial increase in the world death rate..."*  (p. 11)

     In fact, the world death rate has fallen substantially, from 13.2 deaths per 1000 people in the years 1965-1970 to 8.23 in 2011, according to the CIA’s World Factbook, (Washington, DC, 2009). [↑](#footnote-ref-226)
227. Most scientists (including friends of mine) are naturally outraged when their objectivity is questioned in cases like this. The evidential problem is that they’d be outraged either way. [↑](#footnote-ref-227)
228. *The Skeptical Environmentalist: Measuring the Real State of the World* (Cambridge, Cambridge University Press, 2001). For example, All this material is linked under "Skepticism toward the Skeptical Environmentalist" (April 15, 2002) at http://www.scientificamerican.com/article.cfm?id=skepticism-toward-the-ske. [↑](#footnote-ref-228)
229. All this material is now linked under "Skepticism toward the Skeptical Environmentalist" (April 15, 2002) at http://www.scientificamerican.com/article.cfm?id=skepticism-toward-the-ske. [↑](#footnote-ref-229)
230. In 2009, a thousand stolen e-mail messages in the files of the Climate Research Unit of the University of East Anglia, a prominent center for mainstream climate information, were published on the internet, some of which suggested that scientists there were manipulating or suppressing data to exaggerate the evidence for catastrophic global warming. Several formal investigations into this “Climategate scandal” were held, none of which concluded that the e-mails constituted serious evidence against the consensus theory, and none of which concluded that the people involved had committed outright scientific fraud, but some of which, particularly the Muir Russell Report (Sir Muir Russell, *The Independent Climate Change E-mails Review* of July 2010), perceived a “lack of openness” on the part of the scientists at CRU that risked “the credibility of UK climate science”. Predictably, mainstream experts generally views these reports as definitive exonerations of the scientists involved, while dissenting experts, noting that these investigations were all conducted by mainstream scientists and other like-minded academics, perceived a whitewash. All this left uncommitted non-experts with essentially no net second-order evidence for either side. [↑](#footnote-ref-230)
231. *Discover*, October 1989, pp. 45-48. [↑](#footnote-ref-231)
232. Schneider argues on his website that this statement has been misrepresented by opponents. See "'Mediarology': The Roles of Citizens, Journalists, and Scientists in Debunking Climate Change Myths", at http://stephenschneider.stanford.edu/Mediarology/Mediarology.html. [↑](#footnote-ref-232)
233. *Treatise of Human Nature* (London: Oxford University Press, 1967) I.IV.I, p. 183. [↑](#footnote-ref-233)
234. An engaging recent best-seller on this topic is Malcolm Gladwell’s *Blink: The Power of Thinking Without Thinking* (Little, Brown, New York, 2005). [↑](#footnote-ref-234)
235. This is what economists mean by the term *satisficing*, which is opposed to *optimizing or maximizing* as a strategy for making decisions, to be used when the time and trouble of making a decision is counted as a cost in the possible outcomes. See Michael Byron, ed*., Satisficing and Maximizing: Moral Theorists on Practical Reason* (Cambridge University Press 2004). [↑](#footnote-ref-235)
236. To demonstrate the extent to which convictions really define your personal identity, just ask yourself: If forced to choose, would you rather give up your good looks, your fond memories, your charming personality, your renowned professional expertise, or the mere fact that you are not a Nazi? [↑](#footnote-ref-236)
237. The slogan "failure is not an option" is in wide use nowadays after being attributed in the movie *Apollo 13* to Gene Kranz, the director of a determined (and successful) effort at NASA's mission control center to save three astronauts aboard a damaged spacecraft close to the moon. [↑](#footnote-ref-237)
238. US v. Seeger, 380 US, 163 (1965). This hasn’t been much of an issue in the United States since the military draft was ended in 1973. [↑](#footnote-ref-238)
239. Feel free to substitute “Stop abortion!”, or whatever else you’d like to stop. [↑](#footnote-ref-239)
240. Mahatma Gandhi similarly put conviction ahead of calculation in pushing for Indian independence from Great Britain, despite the long-predicted civil war that followed when his newly independent nation split into hostile Hindu and Muslim states, costing a million lives. And Abraham Lincoln held the wrongness of slavery and the indivisibility of the United States as convictions firm enough to spend one Union or Confederate soldier's life for every eight slaves ultimately freed, and at an economic cost equivalent to having purchased all the slaves for $1500 each and emancipated them with no blood being shed. (Patricia L. Faust, ed., *Historical Times Illustrated Encyclopedia of the Civil War*, New York, Harper and Row, 1986). Both men are hugely admired for their convictions in spite of the horrific costs borne by so many others (as well as themselves, of course) for their unwillingness to bend. [↑](#footnote-ref-240)
241. *Hamlet*, Act 3, Scenes 2-4, author unknown. [↑](#footnote-ref-241)
242. *Op. cit.* p. 329. [↑](#footnote-ref-242)
243. It is a wonderful irony of Homer's that Achilles feels he has to make this *argument* to Hector, rather than just attacking him as a real wild animal would do. [↑](#footnote-ref-243)
244. I don't mean to suggest that mutual assured destruction is always the wisest approach to conflict, especially when the two sides have unequal resources. A similar strategy certainly backfired on the inhabitants of Melos during the Peloponnesian War. When the Athenians desired to occupy this neutral island as an outpost, they gave the Melians a friendly ultimatum: submit to our rule or we will wipe you out. The Melians replied that this demand was patently unjust, and that for them to surrender, even to the Athenians’ vastly superior forces, would bring dishonor on their country. The Athenians explained to them that this position was irrational, since the Melians had no realistic hope of winning any battle, and subordination was a whole lot better than certain destruction. The Melians acknowledged this fact, but insisted that they would fight for their island to the death in any case, in hopes of getting the Athenians to leave them alone: "Yes, we really are irrational, and you will probably wipe us out, but we will kill a lot of you before we die. Therefore, it’s still not in your overall interest to conquer us.” As it happened, the Athenians attacked anyway and, as everyone expected, quickly won the battle. All of the Melian men were then slaughtered as promised by the victorious Athenians, and the women and children all sold into slavery. (Thucydides, *History of the Peloponnesian War* 5.84-5.116) [↑](#footnote-ref-244)
245. Thus, in the O.J. Simpson murder case, Simpson's team of lawyers somehow managed to avoid drawing the conclusions that most casual observers had been driven to accept, and loudly celebrated his acquittal as a triumph of justice. [↑](#footnote-ref-245)
246. For one notorious example, Walter Duranty, the reporter for the New York Times won the Pulitzer Prize for articles praising the Soviet Union that later proved to have willfully ignored Stalin's deliberate and tyrannical policy of using famine to destroy traditional farming in the Ukraine, at the expense of millions of lives. Other, previously sympathetic observers like George Orwell were revolted by such cruelties and quickly withdrew support for the ongoing Soviet "experiment". See S. J. Taylor, *Stalin's Apologist: Walter Duranty: The New York Times's Man in Moscow* (Oxford University Press, 1990) [↑](#footnote-ref-246)
247. *Pragmatist* philosophers, William James notably included, go so far as to claim that this feature of beliefs – their “working” in the larger context of practical life – is all there is to truth itself. See James, *Pragmatism, a new name for some old ways of thinking* (Longmans, Green, and Co., New York, 1907). [↑](#footnote-ref-247)
248. If this sort of thing needs to be said, I am not talking here about fanatics who blow themselves up, but about traditional working people of faith who take care of their families, treat others honestly in business, give to the poor, and offer hospitality to strangers. [↑](#footnote-ref-248)
249. They Might be Giants discuss this issue in “Your Racist Friend” (the album *Flood*,Elektra, 1990):

     This is where the party ends

     I just sit here wondering how you  
     Can stand by your racist friend  
     I know politics bore you  
     But I feel like a hypocrite talking to you  
     You and your racist friend

     Local opinion is divided on whether this song is intended as a critique of people with racist friends or people with third-order convictions. [↑](#footnote-ref-249)
250. At a recent party at our college president's house, I quietly re-raised the possibility of holding an authorship debate to one of my other friends from English, who loudly responded, “Oh, God, Ted! Not this again!” This piqued the interest of our betters, who soon joined in the hilarity at my tender expense. Since then, I’ve been trying to keep my mouth shut on this topic in public. (Eppur si mouve. Maybe.) [↑](#footnote-ref-250)
251. This is plain to see on popular partisan internet message boards like Free Republic on the right and Democratic Underground on the left, where anyone expressing any sympathy at all for people or principles on the other side is instantly "flamed" as a "troll". At least they can agree on jargon. [↑](#footnote-ref-251)
252. Thus, in the present broad polarization of American politics, we have books like Chris Mooney’s *The Republican Brain: The Science of Why They Deny Science – and Reality* (Wiley, Hoboken, NJ, 2012) and Michelle Malkin’s *Unhinged: Exposing Liberals Gone Wild* (Regnery, Washington DC, 2005), the purpose (or at least the effect) of which is precisely to rule out large classes of opponents as intellectually, emotionally, or morally deficient interlocutors. [↑](#footnote-ref-252)
253. Perhaps in Sparta with Menelaus during the school year, and at the beach in Troy with Paris over the summers plus one weekend per month; Christmas to alternate. Once she comes of age, she must of course choose for herself. [↑](#footnote-ref-253)
254. We shouldn't blame Nietzsche personally for the horrors of totalitarianism, any more than we should blame Darwin for Social-Darwinian eugenics in the Third Reich and elsewhere. Neither would have endorsed the wicked excesses of some of the people they inspired. It is okay, however, to blame the philosopher Martin Heidegger, an enthusiastic Nazi all through World War II. For details see Victor Farias, *Heidegger and Nazism* (Temple University Press, Philadelphia, 1991). [↑](#footnote-ref-254)
255. *The Limits of State Action* (Cambridge: Cambridge University Press, 1969), pp.20-21. [↑](#footnote-ref-255)
256. *Op. cit.*, p. 58. [↑](#footnote-ref-256)
257. "Self-Reliance", in Riverside Edition, Vol. II, of *Emerson's Complete Works*.  (Boston, Houghton Mifflin, 1883), p. 47. [↑](#footnote-ref-257)
258. *Civil Disobedience* (Amherst, NY: Prometheus, 1998) p. 15. [↑](#footnote-ref-258)
259. “A Plea for Captain John Brown”, in *Thoreau: Essays and Poems* (Library of America, New York, 2001), p. 416. [↑](#footnote-ref-259)
260. Dan Amira, "Anders Breivik Plagiarized Ted Kaczynski", *New York Magazine*, July 25, 2011. [↑](#footnote-ref-260)
261. A particularly galling recent case was that of Republican Congressman Larry Craig, an opponent of gay rights who was arrested for soliciting sex in a public men's room (though convicted only of disorderly conduct, whatever that means under the circumstances. See Adam Cohen, "Larry Craig's Great Adventure: Suddenly, He's a Civil Libertarian", *New York Times*, September 24, 2007.) People on the left are correct when they argue that only conservatives can be hypocrites (in the malign sense) about traditional values, even though progressives formally subscribe to most of the same standards, e.g. about adultery, because they aren’t constantly pushing them on other people, but subscribe to them more in the manner of opinions. But this distinction sometimes comes back to haunt progressives who trespass against their own more doctrinaire convictions, like Bill Clinton with sexual harassment. [↑](#footnote-ref-261)
262. *Animal Liberation: The Definitive Classic of the Animal Movement* (New York: Harper Perennial Modern Classics, 2009), p. xiii. Another bold claim of Singer's that strikes some people as unintuitive is that it's morally permissible for humans to have sex with dogs, as long as the dogs don't mind and nobody gets hurt. ("Heavy Petting", *Nerve* 2001). Age-of-consent laws would presumably have to be changed for this, since most dogs don't live to be eighteen, and those who do are typically no longer very attractive. [↑](#footnote-ref-262)
263. See Peter Berkowitz, "Other People's Mothers", *New Republic* January 10, 2000, 27-37. [↑](#footnote-ref-263)
264. “The Pursuit of Happiness: Peter Singer interviewed by Ronald Bailey”, *Reason* 32(7), December, 2000, pp. 28-35. [↑](#footnote-ref-264)
265. He comes encouragingly close in his recent *A Darwinian Left: Politics, Evolution, and Cooperation* (New Haven, CT: Yale University Press, 2000), in which he admits that recent work in evolutionary psychology supports his political positions only vaguely, and seems contrary to them in some ways, and that much remains to be worked out. I don't mean to pick on Singer for his politics, here. A parallel analysis could be given for the views of conservative and libertarian philosophers who favor equal opportunity for everyone, but don’t (and in my view shouldn’t) apply this idea to their own children’s education. [↑](#footnote-ref-265)
266. *On Liberty and Other Essays*, ed. John Gray (New York, Oxford University Press, USA, 2008) p.17. [↑](#footnote-ref-266)
267. It is a mark of the preeminence of individual autonomy in present-day America, I think, that doctors more and more refrain from making medical decisions for their patients, forcing the patients (who are not only medically ignorant, but also often distraught about their situations) to choose among technical options that they can hardly understand. This may have arisen largely from fear of lawsuits, but it is generally regarded by doctors and medical ethicists as a principled "respect for autonomy". See any current textbook in biomedical ethics for details, for example Tom Beauchamp and James Childress, *Principles of Biomedical Ethics* (5th edition, Oxford University Press, Oxford, 2008), p. 22.

     [↑](#footnote-ref-267)
268. Machiavelli makes a big point of this in his advice to conquerors: whatever else you do, don't raise the taxes. See *The Prince*, Chapter 3, in Peter Constantine, trans., *The Essential Writings of Machiavelli* (New York: Modern Library, 2007), p. 10. [↑](#footnote-ref-268)
269. Marx is particularly harsh on Christianity, the ideology he saw as a chief obstacle to social progress in his day. He famously condemns it in the *Critique of Hegel's Philosophy of Right* as "the opium of the people" for promising heavenly rewards to everyone who humbly works himself to death for someone else's profits here on earth (Joseph O’Malley, ed., Cambridge: Cambridge University Press, 1977), p.131. [↑](#footnote-ref-269)
270. This was the standing policy of BBC radio during World War II and of Radio Free Europe during the Cold War, as I understand it: be more honest than the local authorities on questions where the listeners can test what you are saying, e.g. reports of shortages or bombing damage. If you can gain more trust than those authorities by proving yourself more trustworthy, then listeners will have that much more reason to believe what you are telling them on other matters. This was surely not a major factor in the Western victories in World War II and the Cold War, and the people we wanted to sway were largely Western anyway, with varying amounts of experience in running their own affairs, but it was probably a factor. It also makes sense in these terms that our soldiers overseas engage as they do in community projects such as building schools and water systems, where the local people learn not just that we are well-intentioned people, but that we are knowledgeable, reasonable, and reliable people. Abrupt abandonments and other changes of policy quickly destroy these minor community-based epistemic victories, so they do more harm than good in short-term occupations. [↑](#footnote-ref-270)
271. For a generally sympathetic discussion see Gail Warshofsky Lapidus, *Women in Soviet Society: Equality, Development, and Social Change* (University of California Press, Berkeley, 1978). [↑](#footnote-ref-271)
272. Thus, one of the mythic heroes of early Soviet propaganda was Pavel Morozov, a 13-year-old boy who supposedly reported his father to the authorities for hoarding grain and was allegedly murdered in revenge by other relatives (Yuri Druzhnikov, *Informer 001: The Myth of Pavlik Morozov*, Transaction, Piscataway, New Jersey, 1996). Since the fall of the Soviet Union, his many statues have been removed and Pavlik Morozov Children’s Park in Moscow has been renamed. [↑](#footnote-ref-272)
273. Matthew 10: 34-37 [↑](#footnote-ref-273)
274. For details on Soviet biology see David Joravsky, *The Lysenko Affair* (University of Chicago Press, Chicago, 1986). [↑](#footnote-ref-274)
275. Meters and liters are no huge improvement over yards and quarts in daily life, but the metric system works extremely well in science. On the other hand, the metricized calendar (12 months of 30 days apiece, all with new names and each including three 10-day weeks, plus a five-day holiday at the end) strikes people now as pretty silly. In a heroic effort to exaggerate the excesses of rationalistic revolutionary governments, Woody Allen has the Castro-like dictator in his film *Bananas* (1971) declare in his inaugural speech: "[A]ll citizens will be required to change their underwear every half-hour. Underwear will be worn on the outside, so we can check. Furthermore, all children under sixteen years old are now…sixteen years old." [↑](#footnote-ref-275)
276. Even serious alcoholism became chic in the 1930s and retained some cachet for a long time thereafter. In the "Thin Man" movies, wealthy amateur detectives Nick and Nora Charles (played by William Powell and Myrna Loy) were portrayed as three-quarters in the bag the whole time from their breakfast of Bloody Mary's to their sizeable nightcaps after the crime had been solved. (The series begins with W.S. Van Dyke's film of Dashiell Hammett’s The *Thin Man*, 1932.) Three decades later members of the Rat Pack, a group of sharply-dressed entertainers led by Frank Sinatra, were rarely photographed without a drink in one hand and a cigarette in the other, and Dean Martin in particular was (or pretended to be) constantly *soused* as host of his popular variety show on television. (NBC, 1965-1974). [↑](#footnote-ref-276)
277. In an analysis of this strategy, sometimes called Entertainment Education or EE, we are told that “…EE efforts within the United States have experienced their greatest success when integrated into existing commercially‐produced television programs with familiar characters that audiences already know and love. Such a feat is managed by sensitive outreach to scriptwriters, coordinated by a handful of academic, non‐profit, and advocacy groups, as well as through cooperative agreements from the federal government and other funding sources.” Huang, G. C., Murphy, S. T., de Castro Buffington, S., Nahm, S., Singer, & Felt, L. J., *Impact of global health topics in US television and transnational audiences.* (Los Angeles: California: University of Southern California, Annenberg Norman Lear Center, Hollywood, Health & Society Program, 2010). [↑](#footnote-ref-277)
278. For example Stanley Fish, in *Save the World on Your Own Time* (Oxford University Press, 2008). [↑](#footnote-ref-278)
279. For one example, Bill Ayers was a main figure in the violent revolutionary group Weather Underground, but switched to non-violent activism once released from jail trouble and has had a successful career as a professor of education who emphasizes “teaching for social justice”. Ayers is allegedly close to the current US President – how close is a matter of dispute on internet message boards – and is the author of at least one best-selling memoir, *Fugitive Days* (New York: Penguin, 2003). [↑](#footnote-ref-279)
280. This may seem too convenient an excuse for Mill’s own involvement in imperialism as an employee of the British East India Company, but I think the same view was almost universally accepted by the progressive Europeans of his day. [↑](#footnote-ref-280)
281. *Republic* 488b-e. [↑](#footnote-ref-281)
282. Speech to the House of Commons, November 11, 1947. [↑](#footnote-ref-282)
283. Factions have been accused of breaking these agreements in the United States ever since the republic was founded, for instance when Franklin Roosevelt for third and fourth terms as President, defying the informal limit set by George Washington’s example, or when he attempted to expand and “pack” the Supreme Court in order to guarantee favorable rulings. More recently, and still controversially, a longstanding gentlemen’s agreement was said to have been broken when a partisan majority in the US Senate refused to confirm Robert Bork to the Supreme Court in 1987 despite his obvious high competence as a constitutional scholar. Still more recently and controversially, when the Court ended recounts in the Bush/Gore election of 2000 with a 5-4 decision widely seen as political, leaving the legitimacy of Bush’s presidency in question. Perhaps even more damaging to democratic harmony will be the ongoing partisan struggle over an unpopular health care bill, alleged by conservatives and libertarians to involve multiple violations of established procedure, bribery of Senators and Congressmen, and intimidation of the courts. Some conservative activists, upset by a recent Supreme Court ruling accepting this Affordable Care Act as constitutional on questionable grounds, are claiming that the end of the republic has been reached. As Ron Meyer of Young America’s Foundation put it, “Our Constitution is dead, and you can thank the chief justice for that.” (Caroline May, “After Obamacare ruling, conservatives turn on Chief Justice Roberts”, *The Daily Caller*, June 28, 2012). [↑](#footnote-ref-283)
284. The most pessimistic current writer is probably the conservative Mark Steyn, author of the bestsellers *America Alone: The End of the World as We Know It* (Regnery, Washington DC, 2006) and *After America: Get Ready for Armageddon* (Regnery, Washington DC, 2011), who argues that decades of Western economic and cultural self-indulgence have made us incapable of either solving internal problems or even perceiving external threats. But James Burnham took a similar view of the Cold War in *Suicide of the West: an Essay on the Meaning and Destiny of Liberalism* (John Day, New York, 1964), and we managed to win that one, which suggests that we might muddle through again despite our great moral ambivalence about our own prosperity and power. [↑](#footnote-ref-284)
285. Here for simplicity I am ignoring possible three- or more-sided disagreements, such as the complex conflict in the Middle East since World War Two between Westernizers, Marxists aligned with the Soviet Union, and Islamic fundamentalists, and the similar if milder struggle currently emerging in Europe among Christian Democrats, Socialists, and Muslim immigrants. [↑](#footnote-ref-285)
286. Either “socialism” or “communism” would be a more natural label to use than the awkward “communalism”, except that those words have taken on partisan meanings, especially in the United States, that I do not intend. In fact, they’d be particularly misleading given what I am saying here. [↑](#footnote-ref-286)
287. Just as with “socialism” and “communism”, the philosophical term “liberalism” has become attached to a partisan position which I don’t want to confuse with liberalism in the classical sense, especially since in current American usage it refers to something like social progressivism rather than the classical demand for individual liberties. Sadly, the usual alternative term “libertarianism” is lately also taking on a partisan and somewhat narrower meaning, though it is still the best available choice. [↑](#footnote-ref-287)
288. Individual liberties can and do exist outside of liberal democracies – though this is somewhat exceptional, tyranny being what it is. For example, even when the English monarch was powerful, even the lowest of his subjects had some acknowledged rights, especially inside their homes, as the common law holds: “The wind may enter, the rain may enter, the sun may enter, but the king may not enter.” [↑](#footnote-ref-288)
289. *On Liberty*, p. 10. [↑](#footnote-ref-289)
290. I don’t say this as a coolly detached observer. The same thing still happens to me sometimes – I still find myself abruptly getting angry at perfectly good friends in political discussions and calling them names, at least in my head – although less frequently now than it did when I was younger and had more convictions. [↑](#footnote-ref-290)
291. In a renowned, grimly beautiful story by Lawrence Sargent Hall called "The Ledge" (*Hudson Review*, 1959) a man is stranded with his son and nephew on an underwater ledge after their boat drifts away on the rising tide, and chooses similarly. [↑](#footnote-ref-291)
292. But not exactly the same, perhaps. Just yesterday as I write this, my family and I came across a kitten in the road that had been hit by a car but not killed. The cat was a stray, and I ended up spending hundreds of dollars on emergency veterinary care that failed to save the kitten’s life, and that could more easily have been spent on saving human lives in Africa. Many good people would have chosen differently in this case, but again, I don’t think I did what I did because I’m a hateful or callous sort of person. [↑](#footnote-ref-292)
293. Plato admits that organic moral relations will inevitably destroy his ideal philosophical republic. The best we can do to prevent this, he says, is to raise children communally and try to keep the rulers from finding out which children are their own, but he admits that system this is bound to break down sooner or later, and nepotism will eventually corrupt even an initially perfect government. There has never been a perfect government, of course, but we have certainly seen better governments replaced by worse ones for this reason, including the Roman Empire under the philosopher-Emperor Marcus Aurelius, who unaccountably allowed his actual son Commodus to replace him on the throne, rather than a person of proven worth adopted for the purpose, as had been done under the previous four "Good Emperors". Plato himself, ironically, left his own nephew Speusippus in charge of the Academy upon his death, making his renowned colleague Aristotle angry enough to quit and found a rival school at the Lyceum. [↑](#footnote-ref-293)
294. As L. P. Hartley famously said at the beginning of his novel *The Go-Between*: “The past is a foreign country: they do things differently there.” Given the obvious essential truth of this statement, I am surprised by how fiercely most proponents of cultural relativism in contemporary politics and education reject its analogue regarding past stages of our own culture. [↑](#footnote-ref-294)
295. "Observations on a Late Publication, Intituled, The Present State of the Nation", in *The Works of Edmund Burke*,Volume 1 (New York: George Dearborn, 1834), p. 134. [↑](#footnote-ref-295)
296. *Politics* II.8, (trans. T. A. Sinclair, Penguin, Baltimore 1962), pp.81-83. [↑](#footnote-ref-296)
297. *On Liberty*, p. 50. [↑](#footnote-ref-297)
298. Some people have argued that ordinary traffic lights are actually dangerous, for the reason that they make us less attentive to each other by giving us the *right* to go certain directions regardless of what others are doing. See Martin Cassini, "Traffic Lights: Weapons of Mass Distraction, Danger, and Delay", *Economic Affairs* 30(2), 2010:79-80. [↑](#footnote-ref-298)
299. See in particular F.A. Hayek libertarian classic *The Fatal Conceit: The Errors of Socialism*, (University of Chicago Press, Chicago, 1991). [↑](#footnote-ref-299)
300. *In Defense of Politics* (fourth edition, Chicago: University of Chicago Press, 1992), p. 22-23. [↑](#footnote-ref-300)
301. The Electoral College had not yet formally voted, so those results were still unofficial at the national level. [↑](#footnote-ref-301)
302. Francis Fukuyama famously made this argument in "The End of History?" (*The National Interest*, Summer 1989) as Soviet power was crumbling in Eastern Europe, then backed away from this neoconservative position after the Iraq War seemed to be failing during the Bush administration, and recently he seems to have become ambivalent about Western political values in general ("US democracy has little to teach China", *Financial Times*, January 17, 2011). Samuel Huntington responded to Fukuyama's initial argument with gloomy predictions of conflict in "The Clash of Civilizations?” (*Foreign Affairs*, Summer 1993). [↑](#footnote-ref-302)
303. *Meditations* III.9, (London: Penguin, 1964), pp. 58-59. [↑](#footnote-ref-303)
304. Some have argued that the two philosophies are essentially the very same thing; that Pyrrho himself learned Buddhism in India and then simply translated it into Greek philosophical idiom. See Adrian Kuzminski, *Pyrrhonism: How the Ancient Greeks Reinvented Buddhism*, (Lexington Books, Lanham, Maryland, 2010). [↑](#footnote-ref-304)
305. “Philosophy and the Meaning of Life” in Nozick, *Philosophical Explanations* (Harvard University Press, Cambridge, 1981), p.595f. [↑](#footnote-ref-305)
306. One cinematic exception is the little ratty guy played by Joe Pantoliano in the movie *The Matrix* (dir. Andy and Lana Wachowski, 1999), who voluntarily returns to the pleasant brain-in-a-vat existence provided by humanity's mechanical masters, after discovering just how crummy life is in the real world. All the other liberated humans fight like hell to stay in reality, defeat the machines, and destroy the Matrix. [↑](#footnote-ref-306)
307. ABC, 1957-1962. [↑](#footnote-ref-307)
308. See Chapter 3, section 3.5. [↑](#footnote-ref-308)
309. Once he hits the singularity and makes himself immortal, maybe Ray Kurzweil can take a crack at it. [↑](#footnote-ref-309)
310. *Discourse,* p. 123. [↑](#footnote-ref-310)
311. For myself, I have always suspected that a spiral path would be the best way to escape from a strange forest, since it would give you a much better chance of finding a road out, while still guaranteeing ultimate success if no road exists. [↑](#footnote-ref-311)
312. For example, Dante is only able to escape from his strange woods in the *Inferno* with considerable coaching, both infernal and divine. Why Descartes rejects this avenue, especially if he is the committed Christian that he claims to be, he doesn't say. [↑](#footnote-ref-312)
313. New York: Modern Library, 1996. [↑](#footnote-ref-313)
314. It's a safe bet that Alyosha would have been more seriously challenged to take action in the intended sequel on Alyosha's future life, but Dostoyevsky died before it could be written. [↑](#footnote-ref-314)
315. Directed by Roman Polanski in 1974. Polanski himself is both a victim of autonomous integrity and a bit of a rebel in his own right. In 1969, his pregnant wife, the actress Sharon Tate, was brutally murdered in the morally autonomous Manson killings. Then, in 1977, Polanski boldly threw away the rule book on drugging and raping thirteen-year-old-girls. After pleading guilty to a reduced charge of unlawful sex with a minor, he escaped to Europe and has never been extradited. He remains a free man in Paris with many supporters in the international artistic community. [↑](#footnote-ref-315)
316. New York: Norton, 2001. [↑](#footnote-ref-316)
317. Sophocles confronts us with a compelling case of such indeterminacy in his tragedy *Antigone*, where the two daughters of the late king Oedipus, headstrong Antigone and cautious Ismene, struggle over Antigone’s determination to perform burial rites for their dead older brother, rites that are required by their religion but forbidden by their uncle the new king. Unlike Jane Austen’s Dashwood sisters, Antigone and Ismene are in a moral crisis that cannot be resolved by long talks and careful ruminations, but that demands a more or less immediate decision. And Sophocles does not provide us with a clear answer as to who we ought to think was right, as the play ends with Antigone dead and Ismene heartbroken; only the king has been taught any useful lesson. For what it’s worth, among my students the Antigonists outnumber the Ismeneacs by about two to one, though I suspect they’re voting for the person whom they find more admirable at a literary distance, not the person they would really want themselves (or their big sisters) to be. [↑](#footnote-ref-317)
318. Otto Neurath, "Protocol Sentences," in A. J. Ayer, *Logical Positivism* (NY: Free Press, 1959), p. 201. [↑](#footnote-ref-318)
319. *Apology*, 21d. [↑](#footnote-ref-319)
320. *Theaetetus*, 200d-201c. [↑](#footnote-ref-320)
321. It is famously hard to tell how much of what he says here and in other dialogues is Socrates' own ideas, and how much is Plato's. [↑](#footnote-ref-321)
322. Probably pretty accurately, since this was a recent, major political trial in Athens, attended by hundreds of jurors and other citizens. [↑](#footnote-ref-322)
323. *Republic*, 382b-d. [↑](#footnote-ref-323)
324. *Apology*, 36a-38c. [↑](#footnote-ref-324)
325. Don’t let me suggest that it was *only* obstinacy that kept him from escaping. In the *Crito*, Socrates makes an extensive case on several different grounds that this was the morally right thing, and even the prudent thing, for him to do. Still, this is the kind of action that Socrates was evidently *temperamentally* inclined to take, and later justify. [↑](#footnote-ref-325)
326. Nietzsche himself really despises Socrates. In a brilliantly sarcastic passage, he attacks him as a sick man, dwarfish and ugly, questioning life because he is unable to succeed by life's own terms. Socratic argument he sees as ignoble sophistry, just a way for nature's losers to change the rules of combat so that they can finally win. He says, at least Socrates got something right when he insisted on drinking the hemlock, because he secretly realized how sick he was. When Socrates asks his friends, just before he dies, to sacrifice a rooster to the god of health, he clearly means that he is being cured of something by his death. Traditionally, this has been read as Socrates expressing an expectation that after he dies he will be freed from his body to live in a realm of pure spirit and understanding. On Nietzsche's view, Socrates just wants to be *dead*, which is the only cure for the disease of being Socrates. (See "The Problem of Socrates" in *Twilight of the Idols*, reprinted in Walter Kaufmann, ed., *The Portable Nietzsche*, Viking, New York 1954, 473-479). [↑](#footnote-ref-326)