1

What Should We Believe About Belief?

Beliefs, as philosophers use the term, are mental states that represent the world. We have lots and lots of beliefs, most of them very boring. I believe *that it is hot today*; *that I am in Australia*; *that it is mid-afternoon.* We tend not to notice these beliefs. We notice our beliefs more when they are controversial (*that abortion is morally permissible*; *that facemasks save lives*) and when they are identity-forming for us (*that Jesus died for our sins*; *that the US is the greatest country in the world*). The contents of these beliefs are very different from one another. A major focus of the recent philosophical literature on beliefs has been whether these different beliefs are all in fact the same kind of mental state. A belief like *the cat is on the mat* differs from the belief *the banana is unripe* in its content. The first is about cats and mats; the second about bananas and ripeness. But perhaps *the cat is on the mat* differs from the belief *Manchester United is the best team* not only in content, but also in the kind of state it is.

One reason to think that there might be different kinds of beliefs is that some beliefs are bizarre. It is hard to see how anyone could genuinely believe *that Trump is a decent person* if by "belief" we mean *take the world to be a certain way*. Moreover, beliefs have the function of guiding behavior: we represent the world as being a certain way because we need to know how to act on and in it. But some beliefs don't guide behavior. For example, *delusions* have a notoriously loose relationship with behavior. Someone who suffers from the Capgras delusion believes that some familiar individual (usually a close family member) has been replaced by an impostor. But sufferers are typically little concerned by this startling fact (Bayne & Pacherie 2005). The man who believes, or says he believes, that his wife has been replaced by a robot or an alien typically doesn't report her missing or avoid the impostor. Capgras sufferers have had a brain injury; as we'll see, healthy people are also unmotivated by some of their most fervently expressed beliefs.

I won't be tackling these issues directly here. I'm not especially concerned with the nature or structure of belief. I care about beliefs at all because of the role they play in explaining and causing behavior, and there's no real doubt that most of the beliefs that people profess are apt to cause significant behavior (even if there are interesting differences between them in this regard). A state can qualify as a belief, on the relaxed standard I'm using here, even if it falls well short of what Stephen Stich (1978) calls *inferential promiscuity* (that is, aptness to enter into inferential relations with an indefinitely broad range of propositions) or broad guidance of behavior, just so long as it drives a sufficient amount of our sufficiently consequential behavior.

In this chapter, I'll survey some of the rich philosophical and psychological literature on how beliefs are acquired and updated. I survey this literature not only for its own sake, but to set out the background for my own account. I aim to show that the existing literature, illuminating though it is, doesn't adequately explain how and why we come to believe what we do and act as we do. In setting it out, I'll also be setting out the standard and highly influential picture of ourselves that has emerged from psychology. It is now part of educated lay wisdom that human beings are far from fully rational animals, and the psychological literature on beliefs is an important part of the motivation for this piece of wisdom. I'll be arguing we're a lot more rational than the standard view holds.

Before I turn to the psychology, I'll provide some essential background on why belief matters and why we should therefore be concerned with bad beliefs. Beliefs matter because they guide behavior; accurate belief is required for successful navigation of the world. I'll also show how high the stakes are: some of the most significant political challenges of our time are, in part, battles over belief. In the second and third sections of this chapter, I'll examine accounts that hold that some of our controversial beliefs don't guide behavior in the manner characteristic of mundane beliefs; if these accounts are correct, some of our (apparent) beliefs may not matter all that much. I'll argue that we can't take comfort from these accounts: many controversial beliefs play enough of the explanatory and causal role we expect of beliefs to be extremely consequential. I'll then turn to a discussion of the family of accounts that explain bad belief as a result of

our rationality deficits. I'll give some preliminary reasons for thinking that these accounts are inadequate. The rest of the book will constitute a much fuller case for an alternative, on which we're rational animals after all.

Belief and Behavior

Belief *matters*. Many philosophers identify beliefs with functional states. On this view, to believe that *p* is to possess a representation that plays a characteristic role in behavior.[1] It *explains* our behavior (why did Anya go to the fridge? She was thirsty and believed there was beer in there) and it *predicts* behavior (will Bindi take her umbrella? She will, if she believes it might rain). Even those philosophers who reject functionalism accept that beliefs at very least typically play these explanatory roles in behavior. Beliefs are the kind of things that show up in behavior (especially, but not only, verbal behavior: assertion).[2] Attention to behavior therefore enables us to attribute beliefs. In a popular metaphor, beliefs typically function as *maps*. They enable us to navigate the world.

The (characteristic) functional role of beliefs provides us with a heuristic for belief attribution: beliefs are those states that make best sense of agents' behavior (given plausible assumptions about their desires). This fact also brings into relief why it is so important to understand the factors that explain belief acquisition and update. *Because* beliefs play a pivotal role in behavior, accurate beliefs are essential for appropriate behavior. "Appropriate," here, may mean many things. Accurate beliefs are typically required for *adaptive* behavior. Agents with accurate beliefs

---

[1] The focus on behavior comes easily to philosophers, but it is worth saying a few words about it for those less steeped in the philosophical literature. Focusing on behavior doesn't entail accepting behaviorism: rather, it is a way of getting at the very internal states that behaviorism ignored (or denied). A belief, on this view, is a functional state: it is the internal state of the agent that makes best sense of her behavior, given her desires. This "interpretivist" picture of belief attribution should be familiar upon reflection: if your friend says she is liberal, but votes for Donald Trump and believes that taxes should be lower, you're probably going to conclude she's lying (or self-deceived). Attributing the belief she claims to have doesn't make sense of her behavior. Some other functional state must underlie it, not the belief she professes.

[2] Rose, Buckwalter, & Turri (2014) present experimental evidence that ordinary people take assertion to be the single most powerful cue for belief ascription.

tend to do very much better at setting and achieving goals than those with inaccurate beliefs. There may be exceptions: more than a century ago, William James observed that sometimes having an inflated confidence in our abilities may help us achieve our goals. Work on "depressive realism" (Ackermann & DeRubeis 1991; Moore & Fresco 2012) also suggests that overconfidence is normal and adaptive (though the reliability of this research is now in doubt). But it's uncontroversial that accurate beliefs are generally required for effective goal pursuit. The man who succeeds in leaping across a gulf only because he has an inflated assessment of his own athletic prowess had better retain an accurate assessment of where the edge is, lest he step off it inadvertently.

Accurate beliefs are also required for pro-social behavior. Again, there are surely exceptions. Perhaps a belief in free will is false, but nevertheless required for the maintenance of social order (as Vohs & Schooler (2008) suggest).[3] Again, though, accurate beliefs are generally required for pro-social behavior. We need to have accurate beliefs about who is sentient if we are to avoid hurting others inadvertently; about what benefits them and what harms them; about where they are and where they will be, and so on. Driving a car, closing a door, simply walking across a room: all of these are actions that could cause harm, but usually do not, because we have roughly accurate beliefs about a host of banal topics.

Beliefs, and especially bad beliefs, matter not only for our banal behaviors; they're also at the heart of our major political challenges. Those who seek to manipulate our behavior have been quick to grasp this fact. They aim to manipulate us by targeting our beliefs, rather than by getting us to act directly.

In the wake of recent scandals involving Cambridge Analytica and fake news on Facebook, journalists and ordinary people sometimes express a worry that ordinary people can be easily manipulated into believing more or less anything. Actually, shifting opinions is often quite difficult: when people have pre-existing beliefs, when the topic is one that they take to concern them directly, and when they get reliable confirming

---

[3] This paper has been very influential. It has been cited 970 times, according to Google Scholar. Its influence indicates the difficulty of correcting false beliefs: later work finds little to no effect of free will belief on prosocial behavior (Crone & Levy 2019; Monroe et al. 2017; Nadelhoffer et al. 2020).

or disconfirming feedback, changing minds is hard and effects tend to be small and often short-lived (Mercier 2020). But when they have few pre-existing beliefs or when the topic is one on which they are not able to access direct evidence, manipulation can succeed. Unsurprisingly, we're more vigilant and careful concerning questions that touch us directly than those which seem distant and of little significance to how our lives go. The selectivity of epistemic vigilance leaves us vulnerable to manipulation on important questions when they don't make a direct difference to our lives, or we can't see how they make a difference. The obvious example here is climate change. On this topic, manipulation of public opinion has been spectacularly successful.

Most or all of my readers live in democracies, however imperfect. Democracies are responsive to the priorities of voters. Decades ago, fossil fuel interests took action to forestall voters agitating strongly for effective action on climate change. In *Merchants of Doubt* (2011), Oreskes and Conway document how they went about it, and how successful they were.

Oreskes and Conway's title stems from a now infamous 1969 memo by a tobacco executive. Big tobacco fought a long and effective battle against governmental regulation and bad publicity by casting doubt on the science linking smoking and cancer. "Doubt is our product," the unnamed author of the memo wrote, "since it is the best means of competing with the 'body of fact' that exists in the mind of the general public." The tactics the tobacco industry pioneered to engineer doubt (often, in fact, promoted by the very same individuals) came to be adopted by fossil fuel interests. The merchants of doubt don't aim to convince the public that climate change isn't genuine or isn't a serious challenge. Rather, they aim to convince us that the science isn't settled: that there is ongoing debate about these claims and that there are reasonable views on both sides. They aim in their own words, to convince the public that "significant uncertainties exist in climate science;" sufficient to cast doubt on the central claims of the IPCC (Readfearn 2015). Success for them consists in putting sufficient doubt in people's minds to ensure that calls to action are not seen as priorities, rather than in bringing people to reject the science. The tactic has been successful in bringing about higher rates of skepticism about the need for action (Capstick et al.

2015). Given the probable effects of climate change on the world, their success may be the biggest tragedy of the past half century.

The tactic of sowing doubt (rather than trying to convince people) has been adopted by campaign after campaign: the battle against the regulation of DDT, of chlorofluorocarbons, of the emissions that cause acid rain, as well as tobacco and fossil fuels. In its most recent manifestation, the tactic has taken on an even more cynical guise. The aim now is no longer to bring people to think that there may be legitimate debate about an issue, but to create confusion. Disorientation, rather than doubt, is now the product.

Allegedly, this tactic is widely used in Putin's Russia, where it is aimed at Russians themselves by state-linked agencies. Russian state media put forward a dizzying array of incompatible explanations for the downing of Malaysia Airlines Flight 17, for example, with no attempt made to reconcile them. Flight 17 was brought down by the Americans to frame Russia, *and* it was brought down by accident in a failed attack on Vladimir Putin's private jet, *and* it was brought down by Russian forces who knew that it was not a scheduled passenger plane at all. Similarly, Russian-linked bots on Twitter have not supported a particular line on vaccines; rather, they have played both sides (Broniatowski et al. 2018). They don't aim to convince viewers of some state-sanctioned narrative, "but rather to leave them confused, paranoid, and passive—living in a Kremlin-controlled virtual reality that can no longer be mediated or debated by any appeal to 'truth'" (Pomerantsev 2014). In their book-length study of the right-wing media, Benkler, Faris, & Roberts (2018) argue that the promotion of bizarre conspiracy theories by Alex Jones and the like should be understood in the same kind of way. It's not because they seriously entertain the thought that the Sandy Hook shooting was a false flag operation that the right-wing media ecosystem has devoted attention to the claim; rather, the aim (or the explanation; people need not always explicitly aim at an end to effectively pursue it) is disorientation. They aim "simply to create a profound disorientation and disconnect from any sense that there is anyone who actually 'knows the truth.' Left with nothing but this anomic disorientation, audiences can no longer tell truth from fiction, even if they want to" (Benkler et al. 2018: 37).

Should We Believe in Belief?

The apparent success of merchants of doubt in pursuing their goals by targeting our beliefs suggests that beliefs matter: what we believe explains how we act. Perhaps there's another explanation of their success, however; one that doesn't turn on belief. In this section, I will discuss accounts (due to the cognitive scientist Hugo Mercier and to the philosopher Neil Van Leeuwen), which suggest, or could be taken to suggest, that apparent beliefs matter much less than we tend to think. If they're right, there must be some other explanation of the success of the merchants of doubt.

Mercier, Van Leeuwen, and others who have pursued this kind of line draw our attention to an important fact. They're right that people do not (straightforwardly) believe many of the things they profess to believe. Still, I'll argue, differences in our beliefs (in states that are beliefy enough to count for our purposes) are at the heart of many of the central political issues of our day, such as disputes over climate change and vaccination. What people say they believe tends to explain and predict how they will act. Let's begin with the case for thinking that many of the beliefs people express are relatively inert or epiphenomenal.

Reports of startling ignorance or of bizarre beliefs are staple fare for the media. Consider some recent headlines:

One-third of Americans think alien UFOs have visited Earth (Whalen 2019).

One-third of Americans don't believe 6 million Jews were murdered during the Holocaust (D. Brennan 2018).

One in four Americans think Obama may be the antichrist, survey says (P. Harris 2013).

Nearly half of Republicans see truth in "Pizzagate" theory: Poll (Blake 2016).

Polls like these suggest that a large proportion of our fellow citizens are fundamentally disconnected from reality.

Consider the "Pizzagate" conspiracy mentioned in the last headline for an egregious example. In March 2016, John Podesta, the chairman of Hilary Clinton's election campaign, had his email hacked. Several months later, a number of these emails were published by Wikileaks. Apparently as a joke (Kunzru 2020), some people posting on social media and discussion boards began to suggest that the emails contained code words referring to pedophilia, and that senior figures in the Democratic party were involved in the trafficking of children for sex. Allegedly, food related words in the emails referred to sex with children (for instance, "cheese pizza" meant "child pornography," since they share initial letters).

As the conspiracy theory transmuted, it came to be associated with the Comet Ping Pong restaurant, a Washington D.C. pizzeria that had hosted fundraising events for Obama. Comet Pizza was alleged to host parties involving sexual abuse of children. At least some of the people who fabricated ever wilder versions of these stories or who helped spread them were trolls or pranksters. Some, in fact, probably thought that the stories were harmless jokes because they were so implausible. But others took them seriously. On December 4, 2016, Edgar Maddison Welch drove to Comet Pizza with an assault rifle. He was there, he said, to investigate the claims and free any child slaves he found. He fired several shots before he was arrested.

As Pizzagate illustrates, not everyone who promotes bizarre conspiracy theories genuinely believes them. In the face of the sheer wildness of Pizzagate and some other conspiracy theories (the moon landings were fake; the government is hiding the bodies of aliens who crashed at Area 51 in Nevada; the British Royal family is heavily involved in trafficking heroin, and so on), we may wonder whether the Welches of this world, those who take these theories at face value, are in fact a tiny minority. Perhaps most of the people who report belief in these conspiracies are pranksters or trolls, or aiming to fit in with others who they take to believe them. Perhaps belief reports are explained by something like the phenomenon called *pluralistic ignorance*, whereby most or all people conform to a norm they don't themselves take to be justified because they think a majority of others take it to be justified (D. T. Miller & McFarland 1991).

Psychologists—if not pollsters—have long recognized that it can be very difficult to probe people's *real* beliefs. On many topics, we lack

settled beliefs: asked what we believe, we may take a side rather than admit ignorance, because admitting ignorance can be threatening to our self-image (see Motta et al. 2019 for evidence that giving people the option of skipping a question—giving them a "soft don't know" option— has a significant impact on responses compared to requiring them to *report* they don't know). Many responses are produced simply for the purpose of answering the question asked and not to report a pre-existing mental state. We therefore ought to take reported support for conspiracy theories and other bizarre beliefs with several grains of salt. Perhaps people are more rational than these surveys suggest, or at any rate have more realistic beliefs.

There are a variety of reasons why people may not report what they genuinely believe: they may be deceptive in their reports, or they might be engaging in some kind of social signaling, rather than reporting (Funkhouser forthcoming; Mercier 2020; Sterelny 2015, 2018). Equally, they may not really believe what they sincerely report: they may be mistaken about what their beliefs are. These are all possibilities that I take very seriously: in fact, I think each of them explains some belief reports. Even taken together, however, they don't give us reason to think that our exemplars of bad belief are typically not genuine beliefs (that is, beliefy enough to guide a wide class of significant behavior).

Let's begin with the cognitive scientist Hugo Mercier. Following Sperber (Sperber 1997), Mercier (2020) distinguishes between intuitive and reflective beliefs; only the former govern behavior broadly, he suggests. He argues that we are liable to (reflectively) accept bizarre conspiracy theories and rumors *because* they do not have "serious practical consequences" (159) for our behavior. One of his primary examples is the rumor that swept the town of Orleans in 1969, that Jewish shopkeepers were kidnapping local girls and selling them into slavery. Mercier notes that people didn't take the kinds of steps that full belief in such stories would seem to warrant (demanding police action or raiding the shops, for example). Like the Pizzagate conspiracy theory, for most people these assertions were causally insipid: they motivated only low cost and relatively trivial kinds of behavior (Mercier reports that at the height of the rumor, some people stopped and stared at the shops alleged to be involved; hardly a proportionate response to slave trading). Perhaps

there are odd examples of people who go on to act consistently with their theories, but their behavior is probably best explained by some quirk peculiar to them, not the apparent belief they share with many others.

Why do people profess bizarre beliefs, if these states don't govern very much of their behavior? Mercier gives three reasons. First, professions of belief may function as a kind of commitment device. Because these beliefs are bizarre and at odds with the respectable consensus, professing to hold them serves as a reliable signal of allegiance to a group. Talk is usually cheap, but willingness to engage in talk that paints one as a kook is not cheap and is therefore a good signal of group belonging. Second, people may be willing to accept certain claims (say, "that Obama is a secret Muslim") because these beliefs have little effect on their behavior: when we don't expect to be able to act on an assertion, we have little incentive to test its plausibility. Third, people accept and repeat rumors to justify how they wanted to act in any case, rather than to acquire new information that is consequential for further behavior. Voltaire is often paraphrased as saying that "those who can make you believe absurdities can make you commit atrocities." Mercier claims this is backwards: "it is wanting to commit atrocities that makes you believe absurdities" (202).

Mercier gives several striking illustrations to back up his claims of the relative inertness of the beliefs people express. One comes from the history of medicine. On a naïve view, he suggests, the millennia-long persistence of bloodletting in the Western world as a medical treatment could be attributed to the influence of the ancient Greek Hippocratic writers and those who developed their ideas, in particular the hugely influential Galen. But that would be wrong, Mercier claims: bloodletting is practiced in at least one fourth of the world's cultures, including cultures very far removed from Greek influence. It is practiced in small-scale societies in Central and South America, in Africa, and in South East Asia. This indicates that it was because bloodletting is intuitive, not because of the influence of the Galenic manuscripts, that it was practiced. Another illustration Mercier provides is the blood libel (i.e., that Jews kill Christian children for ritual purposes). The blood libel was a common accompaniment of pogroms, but these rumors sometimes spread without precipitating violence and pogroms were sometimes accompanied by much milder rumors. Both examples suggest that people look to

apparently factual claims to justify their behavior, rather than using them as a map to steer by.

But Mercier's examples don't show that these professed beliefs were causally inert or insipid. That many societies with no exposure to the Galenic manuscripts practice bloodletting shows that the influence of Hippocratic medicine wasn't necessary for such practices to develop. It doesn't show that bloodletting would have been practiced in medieval Europe in the absence of Galenic influence. After all, most societies *don't* practice bloodletting. Moreover, even if the Galenic influence wasn't necessary for the development of bloodletting in Europe, that influence *wasn't* causally inert, on Mercier's own telling: bloodletting in Europe had distinctive features (for example, drawing much more blood than seen elsewhere) due to the influence of Galen.

Mercier is right to emphasize that beliefs take hold only under certain conditions. The inhabitants of Orleans, or of Kishinev where an infamous pogrom resulted in the death of 49 Jews in 1903, accepted outlandish stories because they were already disposed to feel ill-will toward Jews. They embraced a handy justification for their ill-will and (in the latter case) of the atrocities they went on to commit. But the justifications don't seem to have been inert: without them, the acts may not have taken place, or may have been less widespread or less serious. For some people they functioned as an excuse, but for those on the fringes they may have functioned as a reason. In fact, it seems impossible to explain the events at Orleans or Kishinev except by citing the kinds of rumors and misinformation Mercier claims to be inconsequential. We can't explain these events by citing envy or worries about competition alone, because that explanation leaves certain important facts mysterious. Why the Jews and not other successful shopkeepers, or government officials? Handy scapegoats are handy because they are stigmatized, and stigmatization involves belief. The rumors may have taken hold only because people were already disposed to hate and despise Jews, but they were disposed to hate and despise Jews in very important part due to centuries of previous rumors and propaganda.

At the same time, Mercier is persuasive that it is much harder to shift people's firm beliefs than is usually thought. Propaganda, advertising, and rumor have very limited power to move people to reject a belief that

is entrenched to any significant degree (on the other hand, as we will see, it is often trivially easy to shift people from one belief to an opposed belief when the first is *not* entrenched, even when people take themselves to be fervently committed to it). He is also persuasive that we are very much less likely to accept rumors and conspiracy theories when they are seen by us as having "serious practical consequences" (159) for our lives (as opposed to the lives of others for whom we have little sympathy).[4] But the limited power of propagandists is sometimes enough to bring about serious consequences.

Propagandists can play to our prejudices and make our beliefs more extreme and us more likely to act on them. They can also sometimes take advantage of decreased vigilance with regard to beliefs we see as inconsequential for ourselves by dissimulating the consequences, and they can have spectacular success when the consequences are distant and abstract. Were the consequences of our beliefs about climate change more easily perceptible and more personal, we might be less apt to accept conspiracy theories about it. Because the consequences are far removed from individual behavior, however, we aren't vigilant. Our relative credulity on this topic helps to explain why we face a climate crisis with little political will to address it.

Neil Van Leeuwen's (Van Leeuwen 2014, 2018, forthcoming) focus is especially on religious beliefs. He argues that religious beliefs are not factual beliefs: they belong to a different class of representations, which play a different role in behavior. They play an identity-constituting role; any role in guiding behavior is more constrained and secondary to this primary role. His case is built around evidence from cognitive science, which seems to indicate that religious beliefs don't consistently guide behavior. Van Leeuwen maintains that ideological credences typically belong to the same class of representation as religious beliefs; they're identity-constituting, not action-guiding. His view is therefore a

---

[4] The COVID-19 pandemic has provided some telling examples of how people may come to accept conspiracy theories with tragic consequences. While in most Western countries, these conspiracy theories have (at the time of writing) caused little consequential behavior beyond the apparent targeting of 5G mobile phone towers, the rumor that alcohol can prevent or cure illness has led to hundreds of deaths in Iran (Associated Press 2020). Sometimes, fake news kills.

challenge to one (like mine) that holds that bad beliefs matter because of the relatively broad role they play in behavior.[5]

Evidence of mismatches between professed religious beliefs and behavior is plentiful.[6] Christians, Hindus, and Muslims seem more likely to behave consistently with the tenets of their religion when it's salient to them than at other times. For example, Christians are more charitable on Sundays than on other days of the week (Malhotra 2010); they also appear to consume less pornography on Sundays (in highly religious US states, consumption of internet porn is lower than in less religious states only on Sundays (Edelman 2009)); Malhotra dubs this propensity of Christians to behave more consistently with their professed beliefs on one day of the week the *Sunday Effect*. Analogous effects have been reported for other religions. Xygalatas (2012) found that Hindus who played an economic game in the context of a temple withdrew significantly less money from a pot that would otherwise benefit the entire group of players than did those who played the game in a secular setting. Similarly, Duhaime (2015) found that Muslim shopkeepers in Morocco gave significantly more to charity within 20 minutes of hearing the call to prayer than at other times. Of course, the faithful claim to accept ethical norms that govern behavior at all times, not just in certain settings, but their behavior suggests that these norms have a much weaker grip on them outside contexts that remind them of their commitments.

Other instances of apparent mismatch in the religious domain don't involve pro-social norms. Consider the phenomenon that has come to

---

[5] Van Leeuwen isn't committed to denying that bad beliefs explain a lot of consequential behavior. He accepts that there may be some ideological beliefs (and some religious beliefs) that are held as factual beliefs, though he thinks they are atypical, at least today. He also accepts that the "secondary attitudes" govern *some* behavior, and that might be enough to make them troublesome. Given that (as we shall see) he predicts that secondary attitudes tend not to guide behavior in high-stakes situations (except when the stake is the agent's identity), though, he seems committed to holding that secondary attitudes are less directly and significantly troublesome than I take bad beliefs to be.

[6] Alas, plentiful data is not always reliable data. The sample sizes for most of this work are small and little of it was preregistered. I have more confidence in the data underlying the Sunday effect (if not necessarily the explanation offered) than in the other reported results. It is intrinsically difficult to gather large samples for experimental work that aims for ecological validity. The Sunday effect is based on correlational data, which is part of the reason the sample is much bigger.

be called *theological incorrectness*. Theological incorrectness occurs when people who profess to believe the official tenets of their religion appear to utilize conflicting representations in interpreting religious stimuli; implicitly attributing limited knowledge or a limited capacity for attending to events to God, for example (Barrett 1999). Similarly, there is evidence that afterlife beliefs are context dependent. Harris & Giménez (2005) and Astuti & Harris (2008) found that Spanish children and Vezo children and adults (respectively) were more likely to attribute continuing mental life to the recently dead when the person was described as dying in a context that featured primes for religion, such as an attending priest, rather than in a more secular context.

Mismatches like these suggest that people don't factually believe the tenets of their religion. Factual beliefs predict and make sense of our behavior across all contexts to which they're relevant; religious beliefs appear to lack this property. People act consistently with their religious beliefs only when they're salient to them. Van Leeuwen takes this functional difference to be good evidence that they belong to a different class to factual beliefs.

According to him, religious beliefs are closely akin to (perhaps even identical to) imaginings. Imagination, too, only guides our behavior in certain contexts (only when Wendy is playing *fire engines* does she cover her ears in response to the noise of the siren she imagines; outside the game, her truck is silent), whereas belief guides our behavior all the time (even while she's playing fire engines, Wendy doesn't worry that her tree house will catch fire, and she still takes care to stay away from the edge). Imagination requires effort to be sustained: imagined states tend to fade quickly and reality—ordinary belief—takes over. The anthropologist Tanya Luhrmann (2012) has proposed a similar account. She argues that the relationship to God her informants—members of the Vineyard Evangelical Church—cultivate is half recognized by them as an effortfully sustained imagining.

I'm deeply skeptical of much of the evidence for the claim that religious beliefs function differently to factual beliefs. Much of it comes from the (too recent) bad old days of psychology, and reports surprisingly large effects that seem more likely due to chance than to the detection of an underlying reality. Nevertheless, I'll proceed under the assumption that

the data is reliable. Even on that assumption, we should still think that religious beliefs are (typically) beliefy enough to count for our purposes: they still guide a great deal of consequential behavior.

Van Leeuwen argues that religious beliefs tend not to guide highly consequential behavior, because religious "believers" operate with what he calls a "two-map cognitive structure" (Van Leeuwen 2018, forthcoming). One map represents the world as it is factually taken to be; the other map represents the world as they religiously represent it. The faithful never lose sight of the real world in their religious imaginings. They (implicitly or explicitly) track the difference between the world as they religiously imagine it and the way the world really is (just as Wendy tracks her tree house, even while she's playing fire engines). They are thereby able to monitor the gap between the two maps. This monitoring enables them to steer away from putting their religious beliefs to too rigorous a test. For example, believers sometimes call on God to perform miracles, but they're careful to ensure that they ask only for things that might happen anyway. They pray for rain, not showers of money; they pray that someone's cancer goes into remission, but not that a leg grows back (Barrett 2001). This two-map cognitive structure also ensures that religious believers avoid paying the high costs that might be associated with factually believing the tenets of their religion. When stakes rise, Van Leeuwen claims, factual reality rushes back. The devotee may claim that God looks out for her, but continue to take out health insurance.

But believers often *do* engage in behavior that seems to make sense only if they really—factually—believe much of what they claim. They make large donations to churches. They pay for masses to be said for them after their deaths (when they can no longer reap any ancillary benefits). They refuse conventional health care in favor of an exorcism or a faithhealer. They refuse lifesaving blood transfusions. They may do so on behalf of their children as well as themselves. Medieval trials by ordeal provide a dramatic illustration of how religious beliefs may shape behavior in high-stakes contexts.

Trial by ordeal was usually used when evidence was sparse or contradictory. It gave the accused the opportunity to clear their name by performing a painful ritual. Trial by fire (walking a set distance over red-hot

ploughshares or clutching a red-hot iron) or trial by water (plunging one's hand into boiling water to retrieve a stone) were the most common ordeals. Trial by ordeal was held to be determinative of guilt on the basis that God would protect the innocent, by preventing injury or by speeding the healing process. To contemporary eyes, trial by ordeal seems not only barbaric but useless: we'd expect everyone who underwent it to be found guilty. The available evidence tells a different story: a surprisingly high proportion of those who underwent trial by ordeal were acquitted.

Peter Leeson (2012) points out that the rituals surrounding trial by ordeal gave the attending priests ample opportunities to manipulate the results. They could, for example, exert a lot of control over the temperature of the irons. Why would they engage in such manipulation? Leeson suggests that manipulation might have occurred in response to the demeanor of the accused. Those who faced the trial with relative equanimity thereby gave evidence that they believed God would protect them, and unintentionally signaled innocence to the priests. Conversely, those who were more fearful thereby provided evidence of their guilt and suffered accordingly. The contemporary evidence indicates that such manipulation must have been selective: faced with trial by ordeal, many people pleaded guilty rather than undergo it (and face the harsher penalties associated with being found guilty in this way to boot).

If the behavior of those who faced trial by ordeal really provided evidence in this kind of way, however, then they must have really—factually—believed that their innocence was protective. There's no sign here of a two-map cognitive structure: rather their religious beliefs seem to function as factual beliefs and govern behavior, in this very high-stakes context, accordingly. Of course, this leaves the behavior of the priests unexplained. If they factually believe that God will protect the innocent, why do they intervene? Leeson suggests that the priests might come to see themselves as serving as God's instruments in manipulating the trials in this kind of way.

Van Leeuwen always acknowledged that some religious beliefs might be held as factual beliefs. In his most recent work he goes further, suggesting that religious credences, as he calls them, may not be seen in all cultures or all religious communities. Perhaps some communities

accept most or all their religious beliefs factually.[7] He claims that religious credences are typical today, not universal. Especially given this restriction on the scope of his view, it's difficult to produce decisive evidence against it. There are many reasons why people may behave inconsistently with their factual beliefs, after all, and many why they might behave consistently with their imaginings. Our implicit attitudes may conflict with our genuine beliefs, and these attitudes sometimes cause behavior (Levy 2014a, 2015). We may have inconsistent beliefs (Brownstein et al. 2019; Mandelbaum 2016). We may have in-between beliefs that conflict with our genuine beliefs (Schwitzgebel 2010). We may be mistaken about what we believe, because we haven't reflected deeply, and because our beliefs are not open to introspection (Carruthers 2013). Motivated cognition (Kunda 1990) may make our beliefs insensitive to evidence and to the context we find ourselves in. All these phenomena are consistent with Van Leeuwen's account and could explain divergences between the behavior he predicts and what we actually observe.

This dizzying variety of alternative explanations of departures from behavior consistent with belief and with imaginings renders a search for a decisive test of Van Leeuwen's account futile. What matters for our purposes is whether our paradigms of bad belief (concerning climate

---

[7] In his new book, Van Leeuwen emphasizes ethnographic work on the Vineyard movement, an evangelical Christian church that began in the United States and has since spread to many other (developed) countries. To what extent the Vineyard movement is representative or typical of religion worldwide is, however, a difficult question. Members of the church live in secular societies, in which science has a great deal of prestige and in which many scientific claims are more or less universally known. Perhaps the Christian who lives in the United States doesn't take a factual attitude to the proposition that *God created the Earth in 6 days less than 10,000 years ago*, given she knows that this proposition is rejected by scientists, and she knows that science is widely held to be the most reliable way of discovering the truth of propositions like this. But Christians who asserted that same proposition in medieval Europe were in nothing like that position. Theistic creation may have seemed (perhaps *been*) the most plausible explanation of the origin of the world available. That medieval Christians took a factual attitude to propositions like this one seems to be overwhelmingly plausible; so, perhaps, do religious believers in other parts of the world, in which there is less access to good education and/or science is held in lower regard. A more difficult question is whether people at such times and in such places took and take a factual attitude to propositions like *there are ghosts and other spirits who cause good and bad events in the world*. They have and had access to naturalistic explanations of (many) such events, which might have militated against such beliefs. But given the lower prestige of such explanations in these times and places, they may have given them less credence and been more inclined to factual belief in non-naturalistic propositions.

change or vaccinations, for instance) are in fact sufficiently belief-like to help explain and predict a great deal of consequential behavior. In the original paper introducing his theory, Van Leeuwen (2014) suggested that those who reject the science of climate change might hold an essentially religious attitude toward certain factual propositions. He's right there are good reasons to think that climate change skeptics often don't have very determinate beliefs. They seem to oscillate between believing *that climate change isn't happening*, *that it's happening but we're not causing it* and *it's happening and we're causing it, but it'd be too expensive to fix*, depending on which is handiest. The fact that they move between incompatible propositions suggests that they don't have a very determinate or stable belief (they're not unusual in that; Levy 2018). Nonetheless, their behavior is best explained by something beliefy. Climate change skeptics aren't merely ignorant: they don't just fail to know that climate science is true. Most qualify as skeptics at least in part in virtue of holding a distinctive belief; and it's a belief with a fairly precise content, even if it shades into imprecision when they attempt to flesh it out. The settled content is the content that's common to all the inconsistent propositions they oscillate between. They believe something along the lines of *climate change isn't a problem we need to address.*

Unless we attribute to climate denialists a belief with a content along those lines, we can't begin to explain their behavior. This belief makes best sense of their verbal behavior (they *say* they reject the science of climate change), and what people say is a good, though far from infallible, guide to what they believe. But more than that, this belief makes best sense of their voting behavior, their consumption habits, and the ways in which they invest their time and energies. If you have a stake in the future—if you're young enough to expect to live to see the dramatic changes that climate change will bring even to wealthy individuals in wealthy countries, or if you have children—then failures to support action to rapidly reduce emissions and seek to induce other nations to do the same is good evidence that you don't actually accept the science.[8]

---

[8] It's not conclusive evidence, of course. If you are completely fatalistic about climate change, you might accept the science but lack any motivation to address it. So called "doomers" are a real phenomenon, but doomers don't exhibit the combination of attitudes seen with skeptics: verbal denial of the science of climate change (remember, "denial" here encompasses saying that you believe it's real but not a priority) together with lack of motivation to do anything about it.

The science very clearly entails that significant action is urgent, if we and our children are to have decent lives (IPCC 2018). Beliefs motivate behavior in conjunction with desires: given that we can safely attribute the desire to maintain a habitable planet and a comfortable standard of living for their children to the majority of skeptics, we can only explain their behavior by attributing genuine—factual—disbelief to them, and therefore a belief *that there isn't a problem we need to address*.

We'll return to the apparent inconsistency in behavior exhibited by religious believers (and plenty of other people) in a later chapter. Let's turn, now, from the domain of religion to that of ideology, and examine evidence that partisans may not believe some or much of what they assert.

Expressive Responding

For the past few years, a favorite sport of many US liberals has been mocking the credulousness of the supporters of Donald Trump. The single most shared election-related story on Facebook in the three months leading up to the 2016 election was literally unbelievable: *Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement* (O'Connor & Weatherall 2019). But that's far from the most bizarre thing that Republicans have apparently believed. In one survey, 20 percent of Republicans reported believing that Obama is the Antichrist (P. Harris 2013). Consumption and sharing of fake news online is disproportionately common on the political right (Benkler et al. 2018), but the left is by no means immune. Fake stories congenial to the right were shared around 30 million times during the three months before the 2016 election, and those congenial to the left were shared around 8 million times (Allcott & Gentzkow 2017). Fake news appears to spread further and faster than real news (Vosoughi et al. 2018).

How could anyone possibly be taken in by these often-bizarre claims? There is some evidence that older people are considerably more likely to share fake news than younger: one study of online behavior during the 2016 election found that Republicans over 65 were seven times more likely to share fake news on Facebook than people of any political

leaning aged 18-29. Overall, around 11 percent of over 65s shared fake news (Guess et al. 2019). One possible explanation is that this group is less internet savvy or more trusting (having grown up in an age with fewer and arguably more trustworthy media outlets) than younger people, and therefore more easily taken in. But there are other plausible explanation for the sharing of fake news and the expression of support for partisan falsehoods, including some that suggest that people aren't all that credulous after all.

Around a third of Americans report believing the "Birther" conspiracy theory, according to which Barack Obama was not born in the United States (Uscinski & Parent 2014). However, there's good reason to think that the surveys overestimate the true extent of belief. On these highly politicized questions, agents may engage in *expressive responding* (Berinsky 2018; Bullock et al. 2015). Expressive responding occurs when people report beliefs to express their support for their "side," rather than because they genuinely hold them. Expressive responding may help explain some mismatches between people's reported beliefs and their behavior. Surveys have long documented large partisan gaps in attitudes, with each side perceiving the world in a way that seems to conform to their normative views (Lerman, Sadin, & Trachtman 2017). Republicans and Democrats report divergent beliefs about factual claims (e.g., the effects of economic policies). But there is some—albeit mixed—evidence that they don't go on to act in ways that are consistent with these professed beliefs (see Bullock & Lenz, 2019 for review).

There's persuasive experimental evidence that people sometimes engage in expressive responding. Schaffner and Luks (2018) took advantage of the controversy over Trump's inauguration crowds to probe its extent. The Trump administration notoriously claimed that the crowd was the largest ever to witness an inauguration, claims that flew in the face of the photographic evidence (it was this incident that led Kellyanne Conway to introduce the phrase "alternative facts" to describe administration claims). On the two days immediately following the controversy, Schaffner and Luks showed participants photographs of the Trump and the Obama inaugurations (without identifying them), and asked which depicted a larger crowd. Given how widely reported the story was, they

knew that many participants would recognize the photos and be aware of their sources. A very small proportion of non-voters and Clinton voters identified the photo of the Trump inauguration as depicting a larger crowd (3 and 2 percent respectively). In contrast, 15 percent Trump voters identified the image of his inauguration as depicting the larger crowd. It's hard to believe they were reporting a genuine belief: the photographic evidence was clear. Instead, it seems that many people are willing to report a belief they don't hold in order to express support for their preferred party or candidate. Schaffner and Luks note that some Trump supporters were probably unaware of the controversy or failed to recognize the photos, and therefore didn't see the task as presenting them with an opportunity for expressive responding, so it may be that the percentage of people willing to respond expressively is higher than 15 percent.[9]

Other studies have used different methodologies in an attempt to measure the prevalence of expressive responding. Incentives seem to be effective in reducing the partisan gap in responses. For instance, Prior et al. (2015) found (relatively) small monetary rewards for correct responses halved partisan bias (from 12 percent to 6 percent). Bullock et al. (2015) report similar results, and an apparent dose-dependence of reduction: the larger the incentive, the bigger the reduction in bias. A combination of treatments apparently succeeded in eliminating bias altogether. In contrast, Berinsky (2018) found little or no evidence of expressive responding on the questions he probed, despite offering an incentive (albeit one of a different type: a reduction in time spent on the survey). Taken together, and despite some failures to narrow the partisan gap via the provision of incentives, the evidence suggests that a substantial number of survey respondents knowingly and deliberately misrepresent their true beliefs for expressive purposes.

---

[9] As Michael Brownstein pointed out to me, this suggestion is supported by the fact that better educated Trump voters were much more likely to say the smaller crowd was bigger, presumably because they were more likely to recognize the photo. When Schaffner and Luks told participants of the source of the photos (in the second experiment), the effect of education vanished.

In fact, the studies to date may underestimate the extent to which people fail to report their prior true beliefs. First, incentives for accuracy might have perverse effects: they provide an opportunity for more powerful expressive responsive. The bigger the monetary reward forgone, the stronger the signal a belief report sends. For this reason, we should expect the most partisan participants to be difficult to shift by monetary reward. Further, if participants count support for a person, policy, or stance as a sacred value, they'll likely spurn the opportunity for financial reward for accuracy: sacred values are usually held to be incommensurable with and tainted by financial reward (Tetlock 2003).

A second reason why incentivization may not result in people reporting their true beliefs is that people may often lack any prior belief at all. Political scientists have long recognized that a substantial proportion of survey respondents construct their responses on the spot (J. Zaller & Feldman 1992). Instead of responding expressively, some participants may use partisan heuristics, biased sampling methods or motivated inference to generate a response, in the absence of a prior belief. To the extent to which this occurs, surveys of public opinion play a role in producing the responses they aim to probe. Someone who reports believing *that Obama is a secret Muslim* or *that Hilary Clinton gave uranium to Russia in exchange for donations* may not believe these things prior to being asked. Rather, they engage in biased memory search or biased inference procedures, or apply heuristics, to construct a belief report. In part, this may be explained by an aversion to admitting ignorance: Bullock et al. (2015) found that while only 15 percent of respondents gave a "don't know" response in the absence of incentives for accuracy, provision of incentives dramatically increased selection of the option. While the responses given *may* persist as beliefs for some time after the survey, surveys that construct such attitudes overreport their prevalence in the population.

Taking expressive responding and overreporting as a consequence of belief (or attitude) construction into account, there's little doubt that surveys often exaggerate the extent to which people genuinely believe fake news and the like. Still, there's also little doubt that substantial numbers of people do accept some fake news some of the time sufficiently to have an influence on their consequential behavior. In 2017 for example, a

false story that the founder of Ethereum had died in a car accident caused the market value of the company to drop by $4 billion (Dunning 2019). More disturbingly, a number of conspiracy theorists have escalated their harassment of parents who lost children in school shootings beyond online trolling to confronting and threatening them in person (Raphelson 2018; Robles 2016). Similarly, people who deny the existence of Covid have invaded UK hospitals and attempted to remove patients from ventilators (Quinn & Campbell 2021). While these conspiracy theorists are, no doubt, outliers, other differences in behavior are large enough for us to be confident that they aren't driven by outliers. For instance, Lerman, Sadin, & Trachtman (2017) report that Republicans don't merely *say* that they distrust Obamacare; they are also less likely to enrol in it (see Bullock & Lenz 2019 for further examples).

It's likely that estimates of the prevalence of anti-vaxx beliefs are inflated by expressive responding and attitude construction (for the record, I predict that the COVID-19 vaccine will be taken by significant numbers of people who, on surveys, report that they won't), but assertion of anti-vaxx sentiment correlates with behavior: anti-vaxxers vaccinate their children at lower rates, leading the World Health Organization to list vaccines hesitancy as one of the top ten threats to health in 2019 (WHO 2019). It's hard to explain willingness to put one's children at risk expressively; nor are the majority of people who refuse to vaccinate those who have previously participated in a poll. Similarly, even if reported skepticism about climate change is exaggerated by expressive responding or a desire to "own the libs," the fact that those on the right tolerate and encourage inaction on climate is evidence they genuinely believe it isn't a significant problem. While expressive responding might explain a large proportion of those who espouse really bizarre claims, with only a few outliers really buying into Pizzagate or the Sandy Hook conspiracy theories, it can't explain away some of the most consequential cases.

To this point in the chapter, we've been considering evidence that bad beliefs are not a pressing problem. I've argued that the evidence suggests that reports of bad belief are exaggerated: people are less credulous than we often think. But that's little comfort: the most consequential cases continue to center, very significantly, on bad belief. Your neighbor may not believe the QAnon guff he posts on Facebook, but he genuinely

doesn't believe that climate change is a problem requiring urgent action. Bad beliefs should continue to be our focus. With that in mind, let's turn to consider theories of bad belief formation.

Deficit Accounts

On the political right, acceptance of the science of climate change is much lower than in the center or the left. While 66 percent of Democrats support policies aimed at reducing or mitigating global warming, only 27 percent of Republicans express support for such policies (Funk & Kennedy 2019). Republican skepticism is not motivated by doubts about the efficacy of these policies as a means to address the problem: it is motivated by skepticism about the *existence* of the problem. Only a minority of Republicans, and a small minority of those who identify as conservative Republicans, report believing that climate change constitutes a serious threat (B. Kennedy & Hefferon 2019). The gap between Republicans and others is likely inflated by expressive responding and the like, but it's too large to be eliminated by these considerations and there are good reasons to think that on this topic, the skepticism expressed is largely genuine. Climate change is a central, and crucially important, illustration of bad belief, but it's by no means the only one. Around half the population of the United States rejects the theory of evolution, for example (Newport 2014).

In their broad outlines, neither evolution nor climate science is remotely controversial among scientists. Not only is there a scientific consensus on the reality and the urgency of anthropogenic climate change; there is also a scientific consensus on the existence of the consensus (Cook et al. 2016). Moreover, neither are matters on which the general public is well positioned to dissent from expert opinion. It is not as though the kind of evidence on which the theory of evolution is based is easily available to the general public, or that rival accounts could easily be tested without an enormous amount of specialized knowledge and tools. Why do so many people have beliefs at odds with the scientific consensus? In the light of the evidence surveyed in *Merchants of Doubt*, some kind of disinformation program—by fossil fuel interests

in the first case, and from the pulpit in the second—seems a plausible explanation.

On this kind of account, bad beliefs might be due to an *information* deficit. People accept false claims because they haven't been exposed to better information. Information deficit accounts are popular, but *rationality* deficit accounts are perhaps even more common. On these accounts, bad beliefs are not due to bad information, but some kind problem in *processing* information. Accounts of this sort are typically motivated by work in the cognitive sciences, allegedly demonstrating widespread irrationality. A third, related, kind of account is inspired by work in philosophy rather than in the mind sciences. *Virtue* deficit accounts explain bad belief formation as arising from a lack of one or more intellectual virtues. Of course, these accounts needn't be exclusive: perhaps bad beliefs arise from a variety of causes.

In the rest of this chapter, I'll discuss the first two kinds of deficit account. I'll leave virtue deficits for a later chapter. My aim is to show that these accounts face problems sufficiently large to motivate the development of an alternative. While all three may well explain many cases of bad belief formation, even in combination they fall well short of explaining the kinds of cases I'm focusing on here. We need some alternative. Developing that alternative is the goal of the rest of the book.

(a) Information Deficits

Information deficit accounts are familiar. People often blame bad beliefs on a failing school system, or on exposure to Fox News, or on echo chambers on Facebook. Books like *Merchants of Doubt* provide evidence that supports these theories: if tampering with people's information didn't work, presumably companies like Exxon wouldn't have spent so much money doing it (despite knowing that the narrative they were promoting was false (Banerjee et al. 2015)).

Surely there are people who don't accept the science of climate change because (for whatever reason) they lack access to good information. But there's good evidence that this isn't the principal explanation. Disbelief in climate change is predicted by political allegiance, not misinformation.

Being on the political right, and especially on the pro-market right, is the single best predictor of climate change denial (Lewandowsky, Gignac, & Vaughan 2013; S. van der Linden et al. 2021; McCright & Dunlap 2011; Schuldt et al. 2011). Rejecting the scientific consensus may correlate with misinformation about one important topic: skeptics are apt to underestimate the degree of scientific consensus (Cook et al. 2016; Leiserowitz et al. 2014; S. L. van der Linden et al. 2015). But this apparent fact aside, climate change skeptics don't seem worse informed than those who accept the science.

The Ordinary Science Intelligence test measures people's basic scientific literacy. For non-politicized topics, OSI scores predict accuracy in belief. Those with higher scores are more likely to answer correctly when asked whether electrons are smaller than atoms or to identify which gas is most plentiful in the Earth's atmosphere. But the neat correlations dissolve when people are asked about politicized topics. While the positive correlation between OSI scores and accurate beliefs about climate change and evolution holds for (so-called) liberals, it fails for those on the political right and those higher in religiosity (respectively). Being well informed about the mechanisms of natural selection doesn't predict accepting the theory of evolution among those higher in religiosity (Lawson & Worsnop 1992). Equally, knowing what scientists say about climate change doesn't predict accepting the scientific consensus for those on the political right (Funk & Kennedy 2019; Kahan 2015). While there may be a correlation between lack of information and bad belief formation, those who reject the science don't do so *because* they lack information. If anything, the causal arrow probably points in the opposite direction; people are unmotivated to seek better information because they take it to be irrelevant or misleading.

Of course, evolution and climate change are unusual issues. On many other topics, we do find a strong correlation between lack of accurate information and false beliefs, or lack of any belief. That should be obvious. If my only source about the weather in Chicago is misleading, it's likely that I'll acquire false beliefs about the weather in Chicago. If pressed, I'll have to confess that I have no idea when the next train from Milan to Verona departs (or even whether there is such a train). My lack of information about the timetabling of Italian trains explains my lack of

belief here. Some common false beliefs (including many pseudoscientific claims—e.g., that people are left or right brained, or that we only use 10 percent of our brains) are surely explained by misinformation. But there's little reason to believe that lack of, or bad, information is the principal explanation of bad beliefs about highly contentious topics like climate change.

There's more evidence against information deficit accounts. But that evidence is probably best discussed in the next section, which focuses on a lack of rationality rather than of information.

(b) Rationality Deficit Accounts

Rationality deficit accounts explain bad beliefs by reference to how information is processed (see Bardon 2019 for a recent book-length defence of this kind of account). Folk psychology and scientific psychology both recognize the (apparent) existence of *motivated reasoning*. When we find a conclusion unpalatable, we engage in intellectual contortions to reject it. We seem to take mixed evidence to support our prior views (Lord et al. 1979), for example, because we engage in *biased assimilation*, applying more lenient standards to evidence we take to support us than to evidence against our views (Ditto & Lopez 1992; Lord & Taylor 2009; Mercier & Sperber 2017). Motivated cognition may even saturate perception. In a classic study, Hastorf & Cantril (1954) had Princeton and Dartmouth students watch film of a football match between the schools. Princeton students perceived twice as many (and more severe) fouls committed by Dartmouth players than did Dartmouth students. More recently, Dan Kahan and colleagues updated this study. They showed their participants video of a political demonstration. One group was told that the demonstration was against the provision of abortion, while the other group was told the demonstration was against the military's "don't ask don't tell" policy on recruitment of gay people, and in favor of a more inclusive recruitment policy. Perceptions of the actions of the demonstrators (for example, whether they obstructed or threatened passers-by) were predicted by the perceivers' prior political outlook, with egalitarians perceiving more aggressive

behavior from anti-abortion protesters than from those protesting the recruitment policy, and vice versa for those with more conservative views (Kahan et al. 2011).

Data like these are often interpreted through the lens of a dual process account of cognition (J. St. B. T. Evans 2008; J. St. B. T. Evans & Stanovich 2013). Type 1 cognition is (always or typically, depending on the version) fast, mandatory in its operation, effortless (both in its phenomenology and insofar as it is not dependent on the availability of cognitive resources), and typically unconscious. Type 2 cognition has the opposite profile: it's slow, must be engaged and sustained effortfully, degrades under cognitive load and is conscious, in the sense that agents know when they're engaged in it. Type 1 cognition is typically assumed to be evolutionarily ancient, whereas Type 2 cognition is a more recent evolutionary adaptation. Type 2 cognition is the kind of cognition we typically associate with intelligence. It's the kind of cognition required for science, math, and philosophy (though Type 1 cognition plays a role in all thought).

Dual process accounts were made famous through the work of Daniel Kahneman and Amos Tversky (Kahneman 2011; Kahneman et al. 1982). Kahneman was awarded the 2002 Nobel Prize in economics; sadly, his co-author died 6 years earlier. Kahneman and Tversky studied *heuristics* and *biases*; rules of thumbs, mental shortcuts and dispositions to weigh information in a variety of ways, all of which make adaptive sense as responses to challenges we faced in the environment in which we evolved, but which may mislead us in contemporary environments. For example, the *salience bias* arises from the way in which information that is easily accessed or emotionally colored is given heavier weight in decision-making than information that is relatively pallid (Bordalo et al. 2010). The salience bias explains why terrorist attacks or mass shootings, which are very salient for us because they're emotionally charged, are apt to be given far more weight than they deserve in our decision making compared to higher probability risks which are much more common and much less salient (Sunstein & Zeckhauser 2011).

Dual process theories provide a neat explanation for our rationality deficits. If the rational response to a problem consists in the response we should reach after careful deliberation, then we should expect departures from rationality to be common. Type 2 cognition is a scarce resource,

one we must use sparingly. Most of our responses will be due to, or heavily influenced by, Type 1 processes, and they will routinely lead us astray. Type 1 cognition may be reliable in the environment in which we evolved, but the contemporary world routinely throws up challenges utterly unlike those it's designed for. We live in groups that are orders of magnitude bigger than those we're adapted for; we're required to make decisions (about retirement savings; about our health; about government policy) that require us to take into account probabilities and effects years or decades into the future, and so on. Type 1 cognition is ill suited to these kinds of challenges, but it remains an important determinant of how we decide. Guided by unreliable cues, we choose badly. We may, for instance, trade away our civil liberties for a tiny increase in protection against terrorist attacks, while ignoring much higher probability (and much more easily addressed) risks like heart attacks and strokes caused by urban pollution.

Bad belief formation might plausibly arise from reliance on Type 1 cognition. It could arise from biased assimilation, for instance. As mentioned above, support for free markets is the best predictor of climate change skepticism. For those who support unfettered markets, climate change is threatening (Bardon 2019; Keller 2015). Because an adequate response seems to require interference with the market, strongly pro-market individuals have a strong motivation to be skeptical. This motivation could help explain their beliefs. They might discount evidence that conflicts with the belief they want to accept, directly or by casting doubt on its source, while being relatively credulous with regard to evidence that supports their views.

Dan Kahan's novel variant on this kind of view has been particularly influential.[10] He explains bad belief formation as a product of *cultural*

---

[10] Kahan might object to his account being described as a rationality deficit account. On his view, when we ask survey respondents questions like "is climate change real?" we're asking them about their identities, not their views on science: they respond by telling us who they are (this is not expressive responding, as usually understood, because on Kahan's account people report their genuine beliefs, rather than beliefs they don't hold; belief is shaped by identity on this view). But if the task is really protecting, or reporting, our identities, it's far from clear that we act irrationally at all. Nevertheless, it seems appropriate to treat Kahan's account under this heading, both because motivated cognition is one of the mechanisms people have foremost to mind when they postulate rationality deficit accounts of bad belief, and because the mechanism Kahan identifies is supposed to lead individuals to accept false beliefs due to a selective inability to process information as Type 2 cognition would mandate (see Williams, 2021, for further discussion).

cognition. Cultural cognition is motivated reasoning made social: for Kahan, it's our cultural and social identity that shapes how we perceive the world and how we make inferences about it. Kahan sees cultural identities as tending to come in more or less coherent packages. Those who value free markets also tend to be highly individualistic, for example. In the contemporary United States, cultural identities are organized around two principal axes: a hierarchical/egalitarian axis, on the one hand, and an individualistic/communitarian axis, on the other. Hierarchical individualists are the most supportive of unfettered markets, while communitarian egalitarians the least. When our cultural identities are threatened, Kahan suggests, we engage in *identity protective cognition* to defend it (Kahan 2008, 2017; Kahan et al. 2010).

Kahan has produced an impressive array of evidence in support of the cultural cognition hypothesis. One lovely experiment involved participants assessing the efficacy either of a skin cream as a treatment for a rash or of a ban on carrying concealed weapons as a response to crime (Kahan, Peters, Dawson, & Slovic 2017). Participants had to make this assessment on the basis of a 2×2 contingency table. The numbers used were identical across conditions: it was therefore possible to assess the extent to which the topic alone made a difference to participants' capacities to engage in numerical reasoning, independently of the mathematical challenge the tasks involved.

As Kahan and his colleagues predicted, topic made a significant difference to the results. In the two skin treatment conditions, better numeracy correlated with a higher probability of choosing the right answer. That's unsurprising, since the task is moderately difficult and getting the right answer requires comparing ratios to detect covariance between treatment or its absence and improvement. But in the (mathematically identical) gun control conditions, better numeracy didn't predict a higher probability of picking the right response. Just the reverse: more numerate participants exhibited *greater* polarization than those with lower numeracy. This depressing finding—that higher ability may make motivated individuals *less* accurate, not more—is supported by a range of other data. Above, we cited evidence from Kahan himself that greater science literacy and higher levels of education correlate with greater

skepticism about climate change among Republicans (Kahan 2015). This basic finding is supported by other work (Drummond & Fischhoff 2017; McCright et al. 2016). Nurse & Grant (2020) have recently demonstrated the existence of motivated numeracy with regard to climate change specifically and Connor et al. forthcoming replicated the basic finding in a European sample, though they did not find evidence of increased polarization in high numeracy participants.

Kahan suggests that the greater polarization seen among more capable and informed participants is due to their greater capacity. This capacity gives them an ability less capable participants don't possess: to clearly recognize how threatening the correct response is to their worldview, or identity. They are therefore motivated to selectively inhibit Type 2 cognition. Other dual process theorists have suggested different explanations. Perhaps bad beliefs among the cognitively sophisticated arise from the selective *deployment*, rather than inhibition, of Type 2 cognition. Taber & Lodge (2006) suggest something along these lines: having greater reasoning skill and more information available may give motivated reasoners more tools to defend the position they are motivated to accept against unfavorable evidence. Either kind of explanation might explain not only why Republicans who score higher in Ordinary Scientific Intelligence are less likely to accept climate change, but also higher levels of belief that Barack Obama is a secret Muslim among better educated Republicans (Lewandowsky et al. 2012), or why philosophers who specialize in ethics don't behave better than those in other areas of philosophy (Rust and Schwitzgebel 2009; Schwitzgebel 2009a).

Just as information deficits surely explain some cases of bad belief, rationality deficits probably really play a role in explaining some of the kinds of cases we're interested in. But they fall well short of a comprehensive explanation. I'll focus on Kahan's sophisticated version of a dual process account, since he has explicitly developed his theory with an eye toward explaining bad beliefs.

One reason for some skepticism about Kahan's account is the apparent fact that bad belief formation—while systematic in a way that calls for explanation—doesn't always correlate with identity. Skepticism about

climate science and about evolution may be unusual in this regard. In other cases, there is little or no correlation with markers of group belonging. As Kahan himself has noted, anti-vaxx sentiment doesn't seem to be predicted by group identity (Kahan 2014).[11] The same seems to be true for opposition to genetically modified organisms (GMOs) (Lewandowsky, Gignac, & Oberauer 2013). According to Kahan, we engage in motivated reasoning most powerfully when a query is perceived as probing our identities. He seems therefore to be committed to thinking that on GMOs and vaccination sentiment, we should see a close correlation between scientific literacy and accuracy. But that's not the case: anti-vaxx sentiment doesn't correlate with ordinary scientific intelligence.

Further, there are grounds for skepticism about a central plank of accounts that turn on motivated cognition. On these views, we are motivated to reject some hypothesis because it is threatening to our group identity or our self-esteem. Evolution is a clear case in which this might be true: while there are theological views entirely consistent with evolution, theists may be passionately committed to a creation story on which God created the world in seven days less than 10,000 years ago. Evolution is therefore intrinsically threatening to this identity-constitutive commitment. But other cases are much less clear.

An analogous story is often told about climate change: it's inherently threatening to those who support free markets, since any adequate response to it would involve heavy regulation of the market (Bardon 2019; Keller 2015). But it is far from obvious there's any inherent conflict between climate change and the ideological commitments of the vast majority of those who reject it. Ideologies are usually too indistinct to entail or even imply positions on policy. For instance, most people who call themselves fiscal conservatives express as much support for government spending as those who don't think of themselves as fiscally conservative (Barber & Pope 2019; Merkley & Stecula 2018). The gulf

---

[11] There is, however, some evidence that anti-vaxx sentiment is currently coming to be correlated with right-wing political views (Quintana et al. n.d.). This may reflect the politicization of the COVID-19 pandemic: as a consequence of Trump's opposition to lockdowns, acceptance of public health messaging came to be associated with Democrats and opposition to them with Republicans. Of course, partisan polarization on vaccination beliefs won't save Kahan's account, because he needs to explain why anti-vaxx sentiments were common prior to the polarization.

between our ideological commitments and policy suggests that we rarely reject a proposition due to conflict between them.

Even setting this issue aside, explanations that explain the rejection of a proposition by citing this kind of conflict vastly exaggerate the degree to which there is a fit between particular policies and broad ideological orientations. Adrian Bardon (Bardon 2019), for example, cites the right's commitment to the status quo as a driver. Of course, conservatism is part of the right's brand. But for the past century or more, the right has also been the home of people fervently dedicated to shaking up the status quo. Across the global North, parties on the right are strongly supportive of relatively unfettered capitalism, and capitalism is the most dynamic economic system the world has ever seen and the most corrosive of established practices and institutions. As Marx and Engels (in what might rightly be seen as at least as much a paean to capitalism as a condemnation of it) wrote:

> Constant revolutionising of production, uninterrupted disturbance of all social conditions, everlasting uncertainty and agitation distinguish the bourgeois epoch from all earlier ones. All fixed, fast-frozen relations, with their train of ancient and venerable prejudices and opinions, are swept away, all new-formed ones become antiquated before they can ossify. All that is solid melts into air, all that is holy is profaned.
>
> (Marx & Engels 2012).

In the intervening century and three-quarters, the revolutionizing power of capitalism has only intensified.

The dual commitment of the right to dynamism and to stability ensures that being a Republican has no determinate policy implications. From a contradiction, anything follows. Until fairly recently, in fact, the conservative strain dominated the capitalist strain within the Republican party when it came to the environment. Conservatives might as easily be conservationists as capitalists. Indeed, conservative conservationism has historically been a strong current within conservative thought, and up until recently there was no partisan divide on the environment, either within Congress or among the general public in the United States (McCright et al. 2014). Conservative distrust of environmentalism appears to arise from, rather than cause, the partisan split.

Even now, it's very hard to see right-wing opposition to climate change as genuinely due to ideological commitments. The same right that frets about market interference strongly supports at least $20 billion dollars in annual subsidies to fossil fuel interests within the United States (Environmental and Energy Study Institute 2019). In Australia, in the face of growing competition from renewables the conservative government has floated the idea of subsidizing coal-fired power stations to keep them competitive (Murphy 2020). Absurdly, while the Australian government rejects the idea of a price on carbon emissions due to its negative effects on business, many of the businesses that would allegedly suffer have called for its introduction (Toscano 2018). While the rhetoric might be that climate change is threatening because it interferes with the market, those who engage in this rhetoric are quite willing to interfere with the market to see off the threat.

In a 1988 campaign speech, George H. W. Bush pledged to tackle the greenhouse effect through the "White House effect" (Hudson 2018). While concerns about climate may be very uncongenial to the contemporary Republican party, that fact is probably as much due to the role that opposition to environmentalism played in shaping the ideology of the party over recent decades as to anything intrinsic in protecting the environment. The party might easily have instead followed a path more closely akin to the European center-right, which recognizes the need for constraints on capital in the name of protecting the environment and social institutions.[12] Of course, the left is equally vague in its ideological commitments. The fact that environmental concerns and support for the science of climate change currently find a more congenial environment on the left is a historical accident, not a reflection of the nature of either ideology or the people who hold them.

Defenders of motivated cognition of bad belief formation might shrug their shoulders in response. It's an interesting question, they may

---

[12] That said, Republican ties to big business certainly played a role in slowing down and even reversing gains on the environment on more than one occasion. Bush rapidly lost enthusiasm for tackling climate change once elected. Oreskes & Conway (2011) tell the story of how people close to the Reagan administration deliberately distorted the science to prevent action on CFCs to tackle ozone depletion. Whether these problems are best seen as arising from partisan ideology or the influence of big business over policy is an open question, however, in light of the failure of Democrat administrations to tackle climate change effectively.

say, *why* people are passionately motivated to defend certain hypotheses, but that's a question for historians, or political scientists, or sociologists. We can help ourselves to the fact that they *are* passionately motivated to defend them, without needing to know why. In later chapters, I'll suggest this response won't do. In fact, people *aren't* passionately committed to defending much of anything. Given the right nudges, they'll passionately defend positions diametrically opposed to those they formerly espoused.

I don't take any of these points to be decisive. For a start, I haven't attempted anything like a comprehensive survey of even the major highlights of the literature explaining bad beliefs within a dual process framework (for instance, I've said nothing about the important work of Gordon Pennycook and David Rand (2019), who explain some cases of bad belief by reference to cognitive laziness and are compelling critics of Kahan's work (Tappin et al. 2020)). Theory choice is a comparative affair: sometimes the best reason to reject a theory is because there's a better one available. Developing such a theory is my task across the remainder of this book. On the account I will develop, people are far more responsive to genuine evidence than the dual process hypothesis concedes. Rather than deploy mechanisms that respond to cues *rather than* evidence, I will argue we respond to cues *as* evidence, and we do so rationally. We are trying to *find out*, not just to fit in.[13]

I won't immediately present the main lines of the alternative account I'll defend. There's more background we need, and this background (too) comes from cognitive science. We are rational creatures and we are social creatures: these two aspects of our being are not independent but intertwined. In responding to social cues, we respond to reasons. The nature of these cues and our responses to them is the topic of the next chapter.

---

[13] De Cruz (2020) argues for a view she presents as midway between Kahan's and mine. She suggests that while Kahan does not give sufficient weight to epistemic factors, I don't give sufficient weight to the motivation to belong. Some of the evidence she cites does testify to the power of this motivation. For instance, in Asch-style paradigms, people often conform behaviorally without changing their beliefs (see Mercier (2020) for further discussion). However, other considerations she takes to be non-epistemic (e.g., markers of group belonging like accent) I will argue are actually evidential: they turn on higher-order evidence.