

What should educators do about bad beliefs?

Neil Levy ^{1,2}

¹Department of Philosophy, Macquarie University, NSW 2109, Australia

²Uehiro Oxford Institute, University of Oxford, Suite 8, Littlegate House, 16/17 St Ebbes St, OX1 1PT, UK

Corresponding author. E-mail: neil.levy@uehiro.ox.ac.uk

Abstract

Bad beliefs are beliefs that are contrary to settled expert opinion; beliefs like ‘global warming is a myth’ or ‘vaccines cause autism’. Non-experts should defer to experts when they are not themselves able to assess the evidence, so we should reject bad beliefs. Notoriously, though, large numbers of people report believing these things, and even more bizarre assertions. Many philosophers think we can avoid bad beliefs by teaching critical thinking or inculcating virtues. While these strategies have some effect, I am sceptical that they achieve very much. In fact, most people—including those who accept bad beliefs—are already broadly rational and believe what they have most reason to believe. I argue that educators do better to tackle the problem indirectly: by acknowledging the rationality of bad believers and by being realistic about what we can know and about how knowledge is produced.

Keywords: belief, conspiracy theories, rationality, education

Surveys and psychological research alike find that people believe all sorts of odd things. They believe that the moon landing was faked (Hamilton 2022), that the Earth is flat (Weill 2022), or that lizard people rule the world (Bump 2013). Many of these beliefs are—apparently, at any rate—harmless. But people also report believing things that are far from harmless.

Few people with credentials in immunology doubt *that vaccines do far more good than harm* and few relevant scientists doubt *that climate change is anthropogenic in origin and a very serious challenge to human well-being*. Given the strength of the expert consensus, we would expect ordinary people (who cannot assess these issues for themselves) to defer to the experts. But significant numbers do not, or so it seems (Marlon et al. 2025; Yam et al. 2025).

Other harmful beliefs are not merely contrary to the scientific consensus, but downright bizarre. Consider the QAnon conspiracy, or Pizzagate (which was a precursor to QAnon). According to the Pizzagate conspiracy, senior Democrats were holding children hostage in the basement of a Washington DC pizza restaurant, and there subjecting them to Satanic abuse. At its height, around a third of Americans reported believing that the theory was ‘definitely’ or ‘probably’ true (Blake 2016). Not only is the belief itself extremely implausible, the evidence cited for it is too weak to

Received: 1 April 2026. Revised: 16 May 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of Philosophy of Education Society of Great Britain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

be described as flimsy, consisting in the fact that leaked emails from the Democratic National Committee mentioned ‘cheese pizza’. This was allegedly code for ‘child pornography’, on the grounds that they share the same initial letters.

I call beliefs like these *bad beliefs* (Levy 2021). A bad belief is a belief held despite the fact that the overwhelming preponderance of evidence is inconsistent with the belief, and the believer is in an excellent position to know this fact. A bad belief may concern any topic on which (1) there is sufficient evidence to establish a consensus, and (2) this fact is very widely known. I am sympathetic to the claim that normative beliefs can and have been established sufficiently well to count as bad beliefs. However, to avoid begging the question against meta-ethical views on which we cannot know moral facts, I will limit discussion to bad beliefs whose badness consists in being inconsistent with *empirical* evidence (note, though, that this restriction allows some normative beliefs to count as bad beliefs: for example, racist and sexist beliefs which rest, at least in part, on empirical claims about differences between human groups). Since it is (more or less) uncontroversial that these beliefs are bad, in my sense, they make a much cleaner test case for rival accounts that aim to explain bad belief formation.

How can people come to hold such bad beliefs?

There are two standard explanations in the literature. The first, which is very common in the psychological literature and widely accepted by philosophers, cites irrationality. The second, which is the province almost exclusively of philosophers, cites intellectual vices.

The irrationality account most often cites motivated reasoning: people do not assess their evidence well when they see it as supporting a view they find congenial. That is why we find *systematic* biases; biases that are correlated with partisan outlooks. The best predictor of being a climate change sceptic in the USA is being a Republican. That fact is explained, on the irrationality account, by the fact that rejecting climate change is congenial to Republicans (either because accepting it would require regulation of business, or because rejecting it is central to right-wing identity). Because it is congenial to Republicans, they are motivated to reject the science of climate change. Theories like this have been advanced by many psychologists (e.g. Kahan 2012, 2015) and echoed by philosophers (Keller 2015; Bardon 2019).

On the intellectual vice view, the problem is not a lack of rationality. It is a failure akin to a moral failure. Virtue ethicists hold that morality should be assessed, and moral behaviour cultivated, by reference to moral virtues. Bad actions are caused by bad character traits and good actions by good character traits. Similarly, virtue and vice epistemologists explain bad beliefs by reference to bad traits of intellectual character (Cassam 2018). People cling on to beliefs in the face of evidence *because they are close-minded*. They dismiss the opinions of people better informed than they are *because they are intellectually arrogant*, they accept conspiracy theories because they are *gullible*, and so on.

Irrationality and intellectual vice may have some part to play in explaining bad beliefs. But I deny that they are particularly important factors. I argue that bad beliefs arise, for the most part, from people doing their best to believe what is true in an intellectually polluted environment.

Which beliefs?

I will present the case for thinking that bad beliefs arise rationally in a moment. As a first step, though, I will make my task easier. There are a lot of frankly bizarre bad beliefs out there. I have already mentioned some examples, and it is easy to add more. According to one survey, one in four Americans believed that Barack Obama is the literal Antichrist (P. Harris 2013). According to another, 15 per cent believe that ‘media or government adds secret mind-controlling technology to television broadcast signals’ and 4 per cent believe that ‘shape-shifting reptilian people control our world by taking on human form and gaining political power to manipulate our societies’ (Jensen 2013). These apparent beliefs are hard for *anyone* to explain. No doubt there are many

gullible people out there, and we have all met people who are prone to bias, but c'mon! Shape-shifting reptilians? True, only 4 per cent of Americans report believing that, but 4 per cent of Americans is a lot of people. That 4 per cent of US adults is over 12 million people.

There may be some people gullible or irrational enough to believe things like that. But 10 million adults is a *lot* of highly dysfunctional people. Their dysfunction should be visible. Beliefs are *maps* of the world; they represent how we take things to be (Ramsey 2013). Like maps, they are used for navigation: my beliefs about where the supermarket is guides me when I need food, and my belief about how much money is in my account guides me in how much I spend. People who are so gullible or so irrational as to believe these bizarre conspiracy theories do not seem likely to have just one or two bad beliefs; they are going to have highly distorted maps of the world. They are very likely to struggle in everyday life. I would be surprised if many of them even have the wherewithal to respond to surveys.

I do not believe that most of these people believe what they report believing. So what is going on?

There are at least two reasons why people might report believing things they do not really believe. One is what has come to be called expressive responding. The other is sheer trolling.

Expressive responding is the intentional reporting of beliefs one does not in fact hold in order to express an attitude. Consider a football fan asked *will Spurs win the Premier League next year?* With all due respect to Spurs, I do not think they are genuine contenders and I doubt many people disagree, even among their fans. But we would not be surprised if some of their fans reported believing they will win. We would probably interpret that report as the expression of an attitude, not the report of a belief. They are taking the opportunity to say, in effect, *yay Spurs!*

Expressive responding has largely been studied in the context of politically partisan attitudes. Much of the literature on this question is open to other interpretations. Consider the partisan gap in perceptions of the state of the economy as a function of who is in power (i.e. people are more positive about the state of the economy when the party they support is in the White House). It has been found that offering monetary incentives for accuracy reduces the partisan gap (Bullock et al. 2015; Prior et al. 2015). The success of this manipulation has been taken to be evidence that people were insincere in their reports; now that money is on the line, they report what they *really* believe. That is a plausible interpretation of the data. But another interpretation is also plausible: perhaps participants are reporting what they think the *experimenters* believe in order to win the money (Levy and Ross 2021).

The most compelling evidence that expressive responding is a genuine and important phenomenon comes from an influential study by Brian Schaffner and Samantha Luks. Schaffner and Luks (2018) took advantage of the then current controversy over the size of the crowd at Trump's first inauguration, sparked by Sean Spicer's false claim that it was the biggest ever. Schaffner and Luks gave their participants unlabelled photos of the Trump inauguration and the first Obama inauguration and asked, simply, which photo had more people. The photos they used were starkly different; it is not plausible that people looked at the photo at the Trump inauguration and thought it showed more people. Yet a substantial minority of Trump voters (but not Clinton voters) picked that photo as showing more people. That is expressive responding: reporting a belief to show one's support for one's side (see Ross and Levy 2023 for replication).

Expressive responding is reasonably well studied (see Graham 2025 for an overview of existing research). Trolling, reporting beliefs one does not hold because one thinks it is funny or upsetting to others to do so, is much less well studied. There is, however, compelling evidence it is a genuine phenomenon.

Consider another Trump-era controversy. After some typically ill-judged remarks by Trump, it was widely reported that Trump supporters were drinking bleach to prevent or combat COVID-19 (Reimann 2020; Reuters staff 2020). But were they really doing that? A Center for Disease Control (CDC) study seemed to show they were; it found that 4 per cent of respondents reported drinking or gargling with bleach or other cleaning products for these purposes (Gharpure et al. 2020). Assuming the sample is representative (as it was designed to be), that is a *lot* of people

ingesting cleaning products: we would expect emergency departments to be overwhelmed. Trolling is a more plausible explanation.

Litman et al. (2023) replicated the CDC study with a larger sample and found that 8 per cent of respondents reported using bleach or other cleaning products to fight COVID. A very different picture emerged, however, when they eliminated ‘problematic respondents’. Problematic respondents are people who show signs of insincerity or inattention. Attention checks (e.g. asking participants to answer ‘strongly disagree’ to a particular question) are already common in psychological studies. Litman et al. added sincerity checks: can you name every US senator ever from memory? Do you know what the (made-up) word ‘wuttlet’ means? Have you previously died from a heart attack? Have you ever used the internet? And so on. Eliminating problematic responders almost entirely eliminated the effect. It left twelve people who were not problematic responders but answered ‘yes’ to the ‘ingesting household cleaners’ question. On follow up, nine of these people reported that they had clicked ‘yes’ by accident and two reported they had accidentally ingested cleaning products. That left one individual. It is unlikely that this person was sincere, given that they also reported being 20 years old and having four children, having a weight of 1,900 pounds, and a height of ‘100’.

In our own research, we used a different method to probe for insincerity. We recruited 1,044 Australian participants, matched for population demographics, and asked them their degree of support (on a five-point Likert scale) for seven conspiracy theories. The first six we drew from the literature:

1. ‘The Labour Party is planning to engage in widespread voter fraud to steal the 2022 Australian federal election’
2. ‘The Liberal Party is planning to engage in widespread voter fraud to steal the 2022 Australian federal election’
3. ‘The idea of man-made global warming is a hoax that was invented to deceive people’
4. ‘A secret group of Satan-worshipping paedophiles has taken control of parts of the Australian federal government and mainstream Australian media’
5. ‘Governments are covering up the fact that 5G mobile networks spread Coronavirus’
6. ‘Coronavirus is a myth created by some powerful forces, and the virus does not really exist’

We also added a seventh conspiracy theory; one that we did not find in the existing literature:

7. ‘The Canadian Armed Forces have been secretly developing an elite army of genetically engineered, super intelligent, giant raccoons to invade nearby countries’

We coded ‘definitely’ or ‘probably’ true as endorsing the theory. We found that endorsing (7) was a stronger predictor of endorsing the other conspiracy theories. The mean number of theories other than (7) endorsed by those who endorsed (7) was 4.08; among those who *rejected* (7) it was 0.50 (Ross et al. 2026; see Williams et al. 2026 for a preregistered replication).

We think this is strong evidence of insincerity. Our raccoons theory is intended to function like the photo of Trump’s inauguration: it is so obviously false that the most plausible explanation of endorsement is not belief but insincere reporting of belief. Of course you might deny that—you might hold, for example, that this shows how gullible conspiracy theorists are. We have two responses.

First, we think that if conspiracy theorists were really *that* gullible, we should see that in their daily lives. They would have difficulty functioning in the contemporary world. They should be perennial victims of scams, for example. As far as we can tell, this is not the case. The literature on the victims of scams shows that those who are emotionally vulnerable are the most often victimized (Levy forthcoming), and emotional vulnerability does not seem to correlate with endorsement of conspiracy theories.

Our second response is to point out that two of our six pre-existing conspiracy theories—(5) and (6)—are contradictory. One cannot simultaneously believe *that Coronavirus is a myth* and *that 5G*

towers are spreading Coronavirus. It is important to point out, therefore, that 56 per cent of participants who endorsed ‘racoons’ endorsed both (5) and (6), compared to less than 2 per cent of those who rejected the conspiracy. We think this makes the case for thinking that insincerity is a major driver of endorsement of conspiracy theories, especially the more bizarre ones.

Pointing to evidence of insincerity is important, because it enables me to respond to an obvious objection to my claim that people embrace their beliefs, whatever they are, rationally. The objection, as we have already seen, is that some of their beliefs are utterly bizarre, and it is hard to see how so many people could rationally believe things like *that*. Once we recognize that fewer people than is thought really believe these bizarre theories, my view becomes more palatable. We can explain the genuine belief of the remainder by reference to their unusual epistemic position—the isolated epistemic communities they belong to, or the edifice of evidence they have built for themselves. They are true believers. But the more bizarre the theory, the fewer true believers there are (I think there are people who are not true believers but who mistake themselves for being believers, but I set that aside for reasons of space; interested readers are directed to [Levy 2022, 2024](#)).

Rational believers

The itinerary I have been following is an odd one: I have been developing a response to an objection to my view without presenting my view. Here it is. It is utterly banal.

People believe the things they do largely on the basis of testimony from others they trust. It is rational to do so.

Let me try to establish, first, it is rational to rely on testimony. It is important to address this issue, because the Enlightenment ethos opposes this view. In his famous essay on the Enlightenment, Kant argued that it is characterized by daring to use our ‘own understanding’ ([Kant 1991](#): 54; emphasis in original), and in emerging from the ‘immaturity’ that consists in ‘the inability to use one’s own understanding without the guidance of others’.¹ Or think of the motto of the Royal Society: *Nullius in verba* (Take no one’s word for it). I think this sort of advice is utterly mistaken. Reliance on testimony is inevitable; moreover, it is entirely rational to rely on testimony so long as we remain sensitive to evidence of unreliability (as we ordinarily do).

We are deeply reliant on testimony for almost all our knowledge. This is most obvious when it comes to assertions that depend on special expertise: *that climate change is anthropogenic in origin and caused by burning fossil fuels; that the MMR vaccine prevents much more morbidity and mortality than it causes; that life on Earth is the product of billions of years of evolution*. None of these are things that laypeople can verify for themselves. They cannot, for example, gather the enormous amounts of data required to track the comparative benefits and costs of vaccines, not to mention the (magnitudes larger) amount of data needed to show that climate change is occurring and that it is largely

¹ While Kant is widely regarded as rejecting reliance on testimony as a source of knowledge, his account is actually more subtle than that suggests. As [Gelfert \(2006\)](#) has shown, Kant sees testimony from others as on an epistemic par with sense experience: both are routinely sources of knowledge. Indeed, Kant says it is a moral and rational failing to be suspicious of such testimony without good reason. However, he appears to limit the scope of knowledge via testimony to propositions *we could have known* via sense perception were we appropriately placed. Gelfert recognizes that this restricts the scope of knowledge via testimony, but he does not recognize how far-reaching this restriction is: it is not only ‘propositions of reason’ we cannot know on the basis of experience—and, therefore, on the basis of testimony—but also empirical propositions that require expertise or extensive inference to be known (note, however, that even with regard to these propositions, Kant recognizes the need for engagement with others: communication of these propositions and finding them to be valid for everyone is a ‘touchstone’ of their truth). Were our capacity to know on the basis of testimony really restricted in this way, we would know very much less than we actually do. Scientific propositions could not be known, nor even historical propositions like *World War Two ended in Europe with the defeat of Germany by the allies*, since this kind of fact is not something we can or ever could observe for ourselves. I am grateful to David Bakhurst for bringing me to consider Gelfert’s interpretation of Kant.

caused by the burning of fossil fuels. Less obviously, *scientists* cannot verify things for themselves (Gerken 2022). They, too, are deeply dependent on testimony, even in their own areas of expertise.

Science is a collective enterprise, both diachronically and synchronically. No scientist works alone. They inherit tools and problems and concepts from previous generations, and they work on them in concert with many others. This is true at every level of scientific research. The basic unit of scientific research is the lab, which consists in people with different tasks and skills. Often, they are not intersubstitutable: the person who conducts one task (say statistical analysis) typically could not do the work of another (say preparing samples). Even when they can do the work in principle, they cannot do it in practice: there simply is not time. Forty years ago, Hardwig (1985) gave the example of a physics paper with ninety-nine authors who divided tasks between themselves and between subgroups. One subgroup spent *sixty person years* on their task; obviously, this is not work that the other authors could also do. Multi-author papers are very much the norm now, and the trend is towards larger numbers of people involved (indeed, good research in psychology and in medicine often brings together multiple laboratories).

The distribution of intellectual labour is pervasive in science, at every level of enquiry. In fact, science owes its reliability to that fact. Research groups output papers that they submit to journals. These papers go through peer review. Peer review serves as a (flawed) check on research quality and value by exposing the work to different sets of people with different sets of skills and different sets of biases. All going well, the review process improves the resulting work. If it is published, it is then subject to post-publication peer review, which is increasingly important in science. That process potentially exposes the research to a much bigger group of people with a much wider set of skills and biases. It is normal for this process to identify flaws the authors may not have seen. In the best cases, these flaws—and the strengths of the paper, too—spur further work, aimed at building on it or refuting it.

The papers produced by research groups and subjected to this process rarely—very, very rarely—aim at answering what we might think of as big scientific questions. You will not find papers in *Nature* or *Science* with titles like ‘Evidence that Evolution is True’. In the vast majority of cases, they address much, much smaller questions, questions the answer to which helps to constitute the case for evolution (or the reality of climate change, or whatever). This is another, and most significant, reason why scientists are reliant on testimony. Working scientists rarely have genuinely deep expertise in *climate change*. Rather, they are a palaeoclimatologist, or an atmospheric physicist, or a glaciologist ... They either work on narrow problems within (say) atmospheric physics, or they work on bigger problems in groups that distribute work between specialists in their discipline and specialists in (say) boundary layers, or in computer modelling. Again, they are not in a position to verify the big claims for themselves; it is only by relying on each other that scientists produce scientific knowledge.

So we are all heavily dependent on testimony when it comes to expert knowledge. Perhaps surprisingly, though, we are also dependent on testimony for ordinary knowledge. Every fact you know about places you have never visited and times at which you did not exist is a fact you know on the basis of testimony. This is true even for facts about *you*: where and when you were born, who your parents are, and so on. Even facts about your own city you know largely on the basis of testimony (how can you verify that *this is Birmingham* or that *Birmingham is in England* without testimony?).

Our political knowledge and partisan allegiances are also owed to testimony. Policy is hard. You might support a political party because they *pledge to cut immigration*, and because you believe *that immigration is responsible for the housing crisis*, but all the beliefs implicated—that *they promise to cut immigration*, *that cutting immigration will help with the housing crisis*, *that there is a housing crisis* in the first place—are owed to testimony. There is, in fact, very extensive evidence that partisan political beliefs are the product of deference to political elites (Zaller 1992; Bullock 2011; Slothuus and Bisgaard 2021). That is rational: since political elites are the (apparent) experts, it is rational to defer to them.

Of course, we should be careful in how we defer to testimony. Some people are unreliable and some people are unscrupulous. Rational agents are sensitive to evidence of these kinds of failings. Indeed, there is extensive evidence that we weigh testimony by reference to this sort of evidence (Mascaro and Sperber 2009; Sperber et al. 2010; Harris 2012; Harris et al. 2018). Children and adults alike prefer testimony from competent agents over testimony from incompetent agents. That is obviously rational. They also prefer testimony from benevolent agents. That is also rational, for slightly less obvious reasons: benevolent agents—those who share our values and are well disposed towards us and what we value—are less likely to seek to exploit us. Of course, children and adults are also sensitive to plausibility, preferring plausible to implausible testimony. Partisan deference, I claim, is just what we would expect of rational agents: political elites share my values and their elite status is evidence of their competence (Levy 2019, 2021).

There is an obvious objection to this perspective. Is there not extensive evidence of pervasive human irrationality (especially, but not only, in politics)? Do people not reject climate change (say) in the face of clear evidence because they are biased? I do not have space to address the concern at length, but I think the evidence for irrationality in the processing of information is hugely overblown. While it may seem obvious to you that the case for the reality of climate change is overwhelming, this is not because you possess the expertise to assess it for yourself (*no one* has that). It is because you trust the scientists and therefore the digests of the scientific research they make available to you. If you did not trust them—but you did trust those who rejected the science—things would look very different to you. The extensive empirical evidence supposedly demonstrating irrationality in information processing is flawed because it rarely takes these background relations of trust into account. In slightly technical jargon, it fails to account for participants' prior probabilities (Tappin and Gadsby 2019; Tappin et al. 2020, 2021). Take them into account, and the supposed irrationality appears to go away. The issue remains highly controversial, but I think it is fair to say, with Alexander Coppock (2023), that right now we have no decisive way to distinguish between motivated accounts of belief formation and those that explain it as Bayesian rational.

So what should we do about bad beliefs?

With this picture of how belief formation, including bad belief formation, works, we can finally turn to the titular question.

If the picture I have presented is on the right track, certain initially tempting responses lose much of their allure. A number of thinkers have suggested that we address bad beliefs by teaching critical thinking. Traditional critical thinking courses focus on (so-called) argumentative fallacies—the *ad hominem* fallacy, appeal to authority, and the like—and species-typical biases, like the confirmation bias or the availability heuristic. While the skills taught in these courses may have their value, there is limited evidence that they are effective either in improving reasoning outside the classroom (Mercier et al. 2017) or in enabling us to identify misinformation (Mercier 2020). In the literature, the failure of critical thinking to improve reasoning is usually ascribed to the so-called 'problem of transfer' (Swartz 1987), but I think the problem is more basic than that. The critical thinking model rests on a mistaken individualism. Those claims we are able to assess on our own are very easy to verify without any special critical thinking skills, and those that are hard to verify we cannot verify on our own. That said, reasoning *can* be improved; as we will soon see, the route to such improvement requires rejection of the individualist model.

Another approach has many adherents among philosophers: inculcating the intellectual virtues. I am sceptical that the intellectual virtues do a great deal of explanatory work in thinking in any case (Levy and Alfano 2020; Levy 2023), but you need not accept my case for that view to see the problem. People accept unreliable information for exactly the same reason as other people accept reliable information: because they trust the source. The real problem is that unreliable sources can seem trustworthy to intelligent, rational, and virtuous agents.

To address the root problem, we must address the sources of trust and mistrust. This is no easy task. While educators have a role to play, it is important to recognize the extent to which it is

out of our hands. We need to address the forces that impel people away from actually reliable sources and those that attract them to the reliable, and doing so requires social, political, and institutional changes.

It is widely noted that conspiracy theories tend to be especially attractive to those who see themselves (often, but not always, rightly) as losing out. Political conspiracy theories, for example, are more widespread among partisans of the party that lost the last election. But the phenomenon is more widespread than that: conspiratorial ideation is more common among the disenfranchised in a wider sense—those who lack a voice in decision-making. As [Uscinski and Parent \(2014\)](#) put it, ‘conspiracy theories are for losers’, in a purely descriptive sense of losers.

A correlation between being a loser and a readiness to accept theories that are at odds with those of the epistemic authorities (with that phrase understood sociologically) is what we ought to expect, on the picture of rational belief acceptance. The epistemic authorities, in this sociological sense, are those institutions tasked with disseminating reliable information and who are held to be credible by mainstream institutions. All going well, these authorities, in this sociological sense of the term, are also genuinely reliable epistemic authorities. In good cases, those people recognized by the press and the state as being able to speak authoritatively about vaccines or about carcinogens or about inflation are also those people who possess the expertise to gather and assess this information. Even in good cases, though, we all know that there is a lot of slack between epistemic and sociological authority.

In the best cases, governments take advice from advisory bodies that are independent of them, but it would be naïve to expect that these bodies are not politicized to some degree and not influenced by extra-epistemic pressures (like lobbying). We expect that media will be somewhat independent of government and able to identify institutions that are reliable, but it would be naïve to expect these institutions, and the media who report on them, not to be influenced by commercial considerations and not to have their own agendas to some degree. Moreover, some domains of knowledge may be intrinsically politicized, such that neutral advice is not possible (e.g. macroeconomics). The media, the epistemic authority that is—or at any rate seems—most influential when it comes to elections will always fall *at least* somewhat short of actual reliability (it is another question, of course, whether perceptions of the degree and direction of its biases are accurate).

Even in the good cases, then, we should expect—and most of us know to expect—that the advice offered by epistemic authorities will weigh the interests of some people more than others. Losers may therefore reasonably see themselves as belonging to a class of people the state has shown itself willing to deprioritize, perhaps because they do not vote for the party in power or because their vote is highly concentrated. This perception will often be true to some degree; in bad cases, it may be true to a very large degree (right now, the USA is in a bad case, with central epistemic authorities in the hands of people who place point-scoring above health outcomes). It is unsurprising, therefore, that distrust of official health advice is highest among those groups with the worst health outcomes, and who have been the victims of a range of medical harms from neglect to illegal and unethical experimentation ([Savoia et al. 2024](#)).

A central way to address bad beliefs, therefore, is to address the determinants of distrust, recognizing that distrust is often a rational response to evidence. In a slogan, if we want to be trusted we need to become trustworthy. Addressing the crisis of distrust requires showing that the epistemic authorities are reasonably taken to value the well-being of those we want to take their advice, and the best way to do that is to *actually* value their well-being. That is no easy, and no quick, task, of course. It is also not a task that philosophers are well placed to advise on.

The second urgent task is the depoliticization of science. In our highly polarized environment, taking a position on scientific issues is often seen as taking a position on a range of unrelated issues. Someone who asserts that *climate change is real and an urgent problem*, for example, is heard as saying not only that *we need to cut emissions* but also *government must interfere with the market* and even that *people should be denied the freedom to buy the car of their choice*. It is also heard, and not without all reason, as expressing support for the view that *trans rights are human rights*, and so on. Some degree of politicization is inevitable and defensible, but we should roll it back to the extent we can.

It is easy to see how politicization occurs. The same epistemic authorities who tell us that climate change is an urgent problem also tell us that decarbonization is a very important part of the solution. But decarbonization requires government action, of a kind that is uncongenial to conservatives. It is a small step from there to seeing support for climate change as a partisan view. It is only a slightly bigger step from there to support for climate change becoming a signal for the whole cluster of views on which the Left and Right diverge.

Because climate change does have fairly direct implications for policy, it is somewhat difficult to unwind the process of politicization, and there are limits on how far we can hope to go. But it is not impossible. The right does not oppose interference with the market on all occasions; on the contrary, there are strands of conservative thought that are very open to such interference to accomplish other goals. And as the Global Financial Collapse and COVID-19 illustrate, each in different ways, both sides will suspend commitments when the challenge is urgent. Indeed, there are grounds for suggesting that the politicization we see over climate change is not, as it were, organic, but was produced by actors who wanted to block change (Oreskes and Conway 2011).

How might we unwind politicization? One way is by stressing the separability of the issues when they are genuinely independent. Another is by having the message delivered by those who share the audience's values. There is evidence that political partisans who testify against the party line—Republicans who accept the science of climate change, for example—are especially persuasive (Nyhan and Reifler 2013; Berinsky 2017), perhaps because their testimony is evidence that the evidence is strong enough to cut through partisan biases. It is increasingly difficult to find such people, sadly, and we can expect their audience to find their credentials suspect in virtue of their testimony. The next best thing is to seek testimony from those with other markers of credibility to the audience. Evangelical pastors who accept evolution, or CEOs who urge action on climate change can be expected to be more persuasive than scientists to certain audiences.

The responses to the problem of bad beliefs I have just sketched are social and political. They require institutional change and large-scale changes in political and social discourse. While we can each play some small role in bringing them about, they are largely out of the hands of each of us. What can we, and in particular educators, do?

Above, I claimed that reasoning can be improved if we reject the individualism typically assumed by critical thinking courses. Without any special training, people are already (perhaps surprisingly) good at evaluating arguments and at discovering solutions to difficult questions. We best do these things *socially*, rather than alone. There is a marked asymmetry in how good we are at evaluating arguments. When we agree with a claim, we are relatively poor at evaluating arguments against it. Even when the problem with an argument is, apparently, glaring, we find it difficult to spot when the premises and the conclusion are all obviously true (Evans et al. 1983). For example, most people assess an argument like this as valid (which, of course, it is not):

All animals need water;
Dogs need water;
Therefore, dogs are animals.

Performance improves significantly when people find a conclusion *unbelievable*. They do well at identifying genuine problems with arguments they are motivated to reject. As a consequence, people are very much better at reasoning together than when they are alone, so long as individuals within the group are motivated to defend different conclusions (Mercier 2011).

People deliberating in groups are also very much better at identifying solutions to problems than when they are alone. The Wason Selection Task is a classic reasoning problem, which only 10 per cent of people get right (despite it being the kind of logical task that is well within ordinary competence). Success in groups, however, rises to 70 per cent (Mercier et al. 2015). This is by no means an isolated finding: across a range of tasks, groups outperform individuals

very significantly. They often outperform even the best individuals within the groups (the ‘assembly bonus effect’). Groups of expert deliberators even do better when *non-experts* are added to deliberation (Sunstein 2002).

Attention to these facts provides educators with an opportunity. We can do two things to improve reasoning. First, there is some evidence that, unlike traditional critical thinking courses, teaching reasoning by *doing*—arguing in groups—has persisting effects on individual reasoning, including their reasoning away from the groups (Kuhn and Crowell 2011). It is unsurprising, of course, that learning by doing is the most effective way of learning: the only surprise is what the doing consists in: not individual identification of argumentative fallacies, and not even group discussion, but group argument (of course, learning will not consist *only* in arguing; discussion will be an important part of identifying which arguments worked and why).

Even in this case, however, the effect sizes are relatively small. The second thing educators can do is to teach the importance of group deliberation and argument (and that argument is not something to be avoided, but to be welcomed). Educators can emphasize not only that we do best in thinking things over when we talk with others, but that good deliberation requires a sufficiently diverse group, with different perspectives and biases.² Educators can play an important role in helping to overcome the individualist bias that is common in our culture, stressing how ‘thinking for oneself’ is never something one does alone.

Another thing educators can do is to present a far more realistic picture of science. The account of science many of us acquire is idealized and individualized in ways that render trust in it fragile (Contessa 2023). Even when it is working well, science is far messier than the folk conception of it suggests. Scientists must make contestable decisions about what questions to address, what data to gather to address them and how much, how to analyse that data and about what alternative explanations are credible, all in ways that are, inevitably, partly reflective of their values (Douglas et al. 2019). They do all this in the service of a variety of motives, from prestige to funding to the hunt for jobs, and under a variety of institutional and social pressures. These facts mean that scientists are not and cannot be the disinterested truth seekers of the myth (John 2018). Moreover, the folk picture of science also presents it as an enterprise driven by lone geniuses (Mercier 2020), rather than by relatively ordinary people working together. When people discover, as they inevitably do, that science is never value free and that scientists are deeply dependent on testimony from others, they may conclude that it is not reliable at all, or that the ‘real science’ is on the side of the fringe theorists.

We might avoid this by teaching a far more realistic picture of science; one that acknowledges the messiness, the mixed motives, the missteps and sometimes the misdeeds, but that also stresses that science remains our best way of generating knowledge about the natural world. Our teaching should also acknowledge the limits of scientific claims today, even when scientists restrict themselves to making claims that fall within their rightful domain. Because science is a collective enterprise and because it takes time, some scientific claims are much more secure than others: those on which a large and diverse group of people have been working for a long time. *That life on Earth is the product of evolution* and *that climate change is anthropogenic in origin* are as well established as facts like *World War One ended in 1918*; there is no serious dispute about them. Other claims are very likely true, but do not enjoy that level of certitude (e.g. *that an asteroid impact is responsible for the extinction of the non-avian dinosaurs*). Others still are subject to lively and ongoing dispute: for example, there is no consensus on what explains, or even causes, consciousness.

² There is a common view that we ought to ‘reach across the aisle’, ensuring that we talk to people with very different political and moral views to us. While such discussions are potentially valuable, fundamental disagreement on issues that are central to us is corrosive of debate quality. Disagreement is productive when we respect one another as epistemic peers, and a sense that our interlocutor holds views we find repugnant blocks such respect. Discussion is most productive when it takes place against the background of a body of values and beliefs that are shared.

Science is not unusual in any of these properties. *Every* domain of human knowledge is the product of fallible and biased people working together, and sometimes in conflict with one another. Science is continuous with them in that respect. We can and should teach how we are reliant on one another for knowledge of so much: history, geography, sociology, but also ordinary facts about the world around us. Human culture is the product of epistemic labour across generations (Richerson and Boyd 2008; Henrich 2015). There are good epistemic reasons to teach the extent of our interdependence, but there are also *moral* reasons to teach how much we do and must rely on one another. Trust is absolutely central to every aspect of human lives, and educators should emphasize this.

Educators working with socio-economically disadvantaged populations face special challenges. As we have seen, such populations are more susceptible to bad beliefs, not because they are less likely to be well educated but because they have lower trust in epistemic authorities. This lower level of trust is explained, in turn, by their recognition that their welfare is often not a high priority for those in power. One route forward is to stress that authorities are not a monolith. While—once again—it would be naïve to think that peak bodies, health authorities, universities, and other sources of information are not influenced by political agendas, they typically make an honest attempt to minimize this influence.³ This does not make them perfectly reliable, by any means; instead, we should expect them to be *more* reliable than fringe groups, and especially when these groups are endorsed or boosted by those with obvious political agendas.

Finally, we can and should teach not only what sorts of cues we should rely on in deciding who to trust, but also why these are the relevant cues. Why should we trust experts? It is not because they are better than us (they are not), or even because they are smarter than us (they often are, but not by all that much). It is because (and when) they are representatives of a discipline that is structured so that truth emerges from a process of discussion and criticism, and when they are reporting claims that are well established in that discipline. We believe them not because they can show us the evidence in a form we can understand, but because they have the right credentials and the right education and a track record of publication, or they are in a position to convey the consensus of their discipline. Certain sorts of challenges are irrelevant: that they have a good response to a ‘gotcha’ question, for example. They are not reliable because they are good debaters or can trot out facts and figures; they are reliable because of all the work they and so many other people have done.

Educating for trust is difficult, in very important part, because educators themselves may not be trusted. They may be seen (and not *entirely* inaccurately) as representing a system that does not value students. Trust must be bootstrapped, not simply taught. There is no easy way to do this, but the surest route is by modelling trust; both being trusting (that the student did not cheat, for example), to the extent possible, and by being trustworthy. The need to model the epistemic virtues is, of course, a central aspect of teaching that has been emphasized by a number of philosophers of education. In this domain, as is so often the case, what is most important ‘could not be explicitly taught but must be *shown*’ (Bakhurst 2013: 201; italics added).

Conclusion

In this article, I have stressed that people are much more rational than many narratives would have us think. They accept truly bizarre claims less often than they say they do, and when they accept conspiracy theories and the like, they do so for the same general reasons as others accept the mainstream views: because they trust their sources. I have gone on to suggest some ways in which educators might make people more likely to trust genuinely reliable sources: by presenting a more realistic and less idealized picture of science, by stressing how trust is

³ It is not experts’ virtue that explains why they seek to minimize political influence; rather, doing so reflects their incentives. So long as the primary rewards of academic life, like job security, promotions, and prestige, are tied to collective recognition, we should expect this to continue. Teaching a realistic view of science might include teaching how scientists are motivated to produce reliable research.

central to knowledge in every domain, and by explaining why the markers of expertise are good indicators of trustworthiness.

It is worth mentioning one more point. The picture I have painted here, of ordinary people—including conspiracy theorists (and those who say they are conspiracy theorists)—as rational agents is itself something worth stressing in education. It might help to reduce the temperature of debate and reach those students who might already be tempted by fringe theories. Labelling them as irrational or stupid is not only false, it is also counterproductive. Education should always aim at bringing people to respect everyone else with whom they share a polity; emphasizing our common rationality is at the heart of stressing our common humanity.

Funding

This research was funded in whole, or in part, by the Arts and Humanities Research Council (AH/W005077/1). The funder had no role in the preparation of this manuscript or the decision to submit for publication. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

- Bakhurst, David (2013) 'Learning from Others', *Journal of Philosophy of Education*, 47: 187–203.
- Bardon, Adrian (2019) *The Truth about Denial: Bias and Self-Deception in Science, Politics, and Religion*. Oxford: Oxford University Press.
- Berinsky, Adam J. (2017) 'Rumors and Health Care Reform: Experiments in Political Misinformation', *British Journal of Political Science*, 47: 241–62.
- Blake, Andrew (2016) "Pizzagate" Theory Believed by Nearly Half of Republicans: Poll', *The Washington Times*, 28 December. <https://www.washingtontimes.com/news/2016/dec/28/pizzagate-theory-believed-by-nearly-half-of-republ/>
- Bullock, John G. (2011) 'Elite Influence on Public Opinion in an Informed Electorate', *American Political Science Review*, 105: 496–515.
- Bullock, John G. et al. (2015) 'Partisan Bias in Factual Beliefs about Politics', *Quarterly Journal of Political Science*, 10: 519–78.
- Bump, Philip (2013) '12 Million Americans Believe Lizard People Run Our Country', *The Atlantic*, 2 April. <https://www.theatlantic.com/national/2013/04/12-million-americans-believe-lizard-people-run-our-country/316706/>
- Cassam, Quassim (2018) *Vices of the Mind: From the Intellectual to the Political*. Oxford: Oxford University Press.
- Contessa, Gabriele (2023) 'It Takes a Village to Trust Science: Towards a (Thoroughly) Social Approach to Public Trust in Science', *Erkenntnis*, 88: 2,941–66.
- Coppock, Alexander (2023) *Persuasion in Politics: How Information Changes Minds About Politics*. Chicago: University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/P/bo181475008.html>.
- Douglas, Karen, Sutton, Robbie M., and Cichocka, Aleksandra (2019) 'Belief in Conspiracy Theories: Looking beyond Gullibility', in J. Forgas and R. Baumeister (eds) *The Social Psychology of Gullibility: Conspiracy Theories, Fake News and Irrational Beliefs*. pp. 61–76, Abingdon, Oxon, Routledge. <https://www.routledge.com/The-Social-Psychology-of-Gullibility-Conspiracy-Theories-Fake-News-and/Forgas-Baumeister/p/book/9780367187934>
- Evans, J. St. B. T., Barston, Julie L., and Pollard, Paul (1983) 'On the Conflict between Logic and Belief in Syllogistic Reasoning', *Memory & Cognition*, 11: 295–306.
- Gelfert, Axel (2006) 'Kant on Testimony', *British Journal for the History of Philosophy*, 14: 627–52.

- Mascaro, Olivier and Sperber, Dan (2009) 'The Moral, Epistemic, and Mindreading Components of Children's Vigilance towards Deception', *Cognition*, 112: 367–80.
- Mercier, Hugo (2011) 'When Experts Argue: Explaining the Best and the Worst of Reasoning', *Argumentation*, 25: 313–27.
- Mercier, Hugo et al. (2015) 'Experts and Laymen Grossly Underestimate the Benefits of Argumentation for Reasoning', *Thinking & Reasoning (United Kingdom)*, 21: 341–55.
- Mercier, Hugo et al. (2017) 'Natural-Born Arguers: Teaching How to Make the Best of Our Reasoning Abilities', *Educational Psychologist*, 52: 1–16.
- Mercier, Hugo (2020) *Not Born Yesterday: The Science of Who We Trust and What We Believe*. Princeton: Princeton University Press.
- Nyhan, Brendan and Reifler, Jason (2013) *Which Corrections Work?* Washington DC: New America Foundation, Research Paper.
- Oreskes, Naomi and Conway, Erik M. (2011) *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. London: A&C Black.
- Prior, Markus, Sood, Gaurav, and Khanna, Kabir (2015) 'You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions', *Quarterly Journal of Political Science*, 10: 489–518.
- Ramsey, Frank Plumpton (2013) 'General Propositions and Causality', in R. B. Braithwaite (ed.) *The Foundations of Mathematics and Other Logical Essays*, pp. 237–55. Mansfield Center, CT: Martino Fine Books.
- Reimann, Nicholas (2020) 'Some Americans Are Tragically Still Drinking Bleach as a Coronavirus "Cure"', *Forbes*, 24 April. <https://www.forbes.com/sites/nicholasreimann/2020/08/24/some-americans-are-tragically-still-drinking-bleach-as-a-coronavirus-cure/?sh=1add6ed36748>
- Reuters staff (2020) 'Gargling with Bleach? Americans Misusing Disinfectants to Prevent Coronavirus, Survey Finds', *Reuters*, 6 June. <https://www.reuters.com/article/idUSKBN23D06H/>
- Richerson, Peter J. and Boyd, Robert (2008) *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.
- Ross, Robert M. et al. (2026) 'Do People Sincerely Believe Conspiracy Theories that They Endorse?', *Collabra: Psychology*, 12: 159253.
- Ross, Robert M. and Levy, Neil (2023) 'Expressive Responding in Support of Donald Trump: An Extended Replication of Schaffner and Luks (2018)', *Collabra: Psychology*, 9: 68054.
- Savoia, Elena et al. (2024) 'Determinants of Vaccine Hesitancy among African American and Black Individuals in the United States of America: A Systematic Literature Review', *Vaccines*, 12: 277.
- Schaffner, Brian F. and Luks, Samantha (2018) 'Misinformation or Expressive Responding? What an Inauguration Crowd Can Tell Us about the Source of Political Misinformation in Surveys', *Political Opinion Quarterly*, 82: 135–47.
- Slothuus, Rune and Bisgaard, Martin (2021) 'How Political Parties Shape Public Opinion in the Real World', *American Journal of Political Science*, 65: 896–911.
- Sperber, Dan et al. (2010) 'Epistemic Vigilance', *Mind & Language*, 25: 359–93.
- Sunstein, Cass R. (2002) 'The Law of Group Polarization', *Journal of Political Philosophy*, 10: 175–95.
- Swartz, Robert J. (1987) 'Critical Thinking, the Curriculum, and the Problem of Transfer', in D. N. Perkins, J. Lochhead, and J. C. Bishop (eds) *Thinking*, pp. 261–84. London: Routledge.
- Tappin, Ben M. and Gadsby, Stephen (2019) 'Biased Belief in the Bayesian Brain: A Deeper Look at the Evidence', *Consciousness and Cognition*, 68: 107–14.
- Tappin, Ben M., Pennycook, Gordon, and Rand, David G. (2020) 'Bayesian or Biased? Analytic Thinking and Political Belief Updating', *Cognition*, 204: 104375.

- Tappin, B. M., Pennycook, G., and Rand, D. G. (2021) 'Rethinking the Link between Cognitive Sophistication and Politically Motivated Reasoning', *Journal of Experimental Psychology: General*, 150: 1,095–114.
- Uscinski, Joseph E. and Parent, Joseph M. (2014) *American Conspiracy Theories*. Oxford: Oxford University Press. <https://global.oup.com/academic/product/american-conspiracy-theories-9780199351800?cc=gb&lang=en&>
- Weill, Kelly (2022) *Off the Edge: Flat Earthers, Conspiracy Culture, and Why People Will Believe Anything*. New York: Algonquin Books.
- Williams, Matt et al. (2026) 'Do You Really Believe That? Examining the Prevalence and Predictors of Belief in Conspiracy Theories When Accounting for Insincerity', *Royal Society Open Science*, 13: 260,163.
- Yam, Eileen, Kikuchi, Emma, and Pula, Isabelle (2025) 'How Do Americans View Childhood Vaccines, Vaccine Research and Policy?', Pew Research Center, 18 November. <https://www.pewresearch.org/science/2025/11/18/how-do-americans-view-childhood-vaccines-vaccine-research-and-policy/>
- Zaller, John R. (1992) *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.