Is Conspiracy Theorising Irrational?

Neil Levy, Macquarie University and University of Oxford, neil.levy@philosophy.ox.ac.uk

Abstract

Conspiratorial ideation—as I will call the disposition to be accepting of unwarranted conspiracy theories—is widely regarded as a product of irrationality or epistemic vice. I argue that it is not: the dispositions that underlie it are not rationally criticisable. Some of the dispositions underlying such ideation are the product of mistrust and heightened vigilance, and these dispositions are warranted as responses to (usually real) inequality and exploitation. Other dispositions are warranted as adaptations for filtering testimony. While those who accept unwarranted conspiracy theories are being led astray epistemically, the solution to this problem is not to alter their dispositions but instead to change the conditions that make their mistrust appropriate.

Whether we should regard conspiracy theorising as irrational depends in part on what notion of ‘rationality’ we have in mind. We sometimes use ‘rationality’ in a way that implicitly presuppose an objectivist standard. On this way of thinking, someone is irrational if they fail to use the procedure that has the best chance of reaching the truth. But we can also use ‘rationality’ in a more subjectivist fashion. According to this standard, someone is rational if they use the procedure that is by their lights best (when we model behavior as Bayesian rational, we have such a standard in mind: is the agent updating on their priors?). Obviously these two standards can diverge. Someone who consults tea leaves before acting may be acting rationally by their lights, while being extremely irrational by an objective standard.

When we ask whether a conspiracy theory (or conspiracy theorising) is rational, we might ask whether it is objectively or subjectively rational. The answer to the first is not as obvious as it might seem. Whether it is or not depends, inter alia, on how common conspiracy theories actually are (Dentith and Keeley 2018). The knee jerk rejection of such theories is justified only if real conspiracies are rare or usually ineffectual. A number of philosophers have argued that conspiracies are in fact relatively common and relatively successful, such that it is irrational to rule them out without serious investigation (Coady 2007; Pigden 2007). It is noteworthy, in fact, that we reserve the term ‘conspiracy theory’ for theories that are not widely accepted (those that conflict with what Coady (2003) calls the ‘official story’). The official story for the events of 9/11 centrally feature a conspiracy (coordinated by Al Qaeda) to hijack planes. Even those who reject many hypotheses as ‘just conspiracy theories’ therefore accept that conspiratorial explanations are often appropriate. Of course, we cannot be certain about the actual base rates of conspiracies, since some conspiracies—perhaps the more successful, on average—go undetected.

However, there is no doubt that a great many of the theories that get smeared as ‘just conspiracy theories’ are objectively irrational. They are accepted despite conclusive evidence that they are false, or at least in the face of evidence that strongly supports a rival hypothesis. Familiar examples include the claim that the moon landing was faked, allegations that the Sandy Hook shootings was a false flag operation and that British intelligence was involved in the death of Princess Diana. These conspiratorial explanations are sometimes harmless and
sometimes highly pernicious: plausibly, false flag theories have played a role in resistance to effective gun control legislation and conspiratorial explanations of the scientific consensus on climate change may have played a role in preventing effective action on the most serious challenge of our time (see Douglas et al. 2019 for a review of, among other things, the behavioral consequences of accepting conspiracy theories). When I ask whether conspiracy theorising is rational, it is this subset of conspiratorial theories I have in mind: theories that are objectively irrational and subjectively irrational for most of us. From now on, I will use ‘conspiracy theories’ to refer to this subset of unwarranted theories.

Just as the question ‘are conspiracy theories rational?’ divides into two subquestions, so the question ‘are objectively irrational conspiracy theories nevertheless subjectively rational for conspiracy theorists?’ itself subdivides in two. Almost all the research so far has focused on existing conspiracy theories and the causes and correlates of accepting them. Following this large body of work, we might ask whether acceptance of conspiracy theories is subjectively rational. Alternatively, we might ask whether the generation of such theories is rational. Conspiracy theories sometimes manifest a great deal of ingenuity and creativity. Thus, they display marks of intelligence. Of course, that does not entail they are generated rationally: the capacities required to generate conspiracies may be admirable yet deployed irrationally.

Alternatively, they might be deployed rationally: because conspiracy theorists have unusual Bayesian priors or perhaps because they do not generate them with the intention of being believed but rather for a joke or to troll those who are easily taken in (the ‘pizzagate’ conspiracy—alleging that Comet Pizza in Washington D.C. was a centre of a child sexual exploitation ring involving the highest echelons of the ‘Democrat’ party—appears to have begun as a joke or as trolling on 4chan). A further possibility is that they owe their intelligence to distributed cognition: by the time they come to public and academic attention, conspiracy theories have been elaborated by multiple—perhaps thousands—of hands, and may display a corresponding degree of complexity and creativity in virtue of that fact.

While questions about the rationality in the generation of conspiracy theories are fascinating, they are harder to study than questions about their acceptance. It is far easier to identify the causes and correlates of the acceptance of conspiracies, but lab-based studies of generation face an obvious problem with lack of ecological validity. Perhaps the availability of large data sets from subreddits devoted to conspiracy theories may allow for this obstacle to be overcome (see Klein, Clutton, and Polito 2018 for the use of such data). In this paper, I will focus on the easier question: are those who accept conspiracy theories irrational to do so? I will refer to such acceptance as ‘conspiratorial ideation’ for ease of reference.

I will argue that accepting conspiracy theories is, for many people, subjectively rational. That is, these agents apply their epistemic standards appropriately in accepting such theories. In addition and more surprisingly, however, I will argue that these epistemic standards themselves are in some sense rational, perhaps even objectively rational. While there is no reasonable doubt that these epistemic standards are leading these agents very badly astray, that fact is not itself a reason to indict the standards themselves, or the agents who deploy them. Thus, I go further than previous defenders of the rationality of conspiracy theories. Charles Pigden
(2017), in what is probably the most wholehearted defence of such theories, limits his defence to the subset of theories that are not objectively likely, and argues that those who accept the theories that are considered here are “epistemically vicious.” I will argue against this view; in fact, they are agents deploying defensible epistemic dispositions in defensible ways, in environments that might be thought of epistemically vicious.

The Psychological Causes and Correlates of Conspiratorial Ideation

As mentioned above, there is a large literature on the correlates of acceptance of conspiracy theories (see Douglas et al. 2019 for review). This literature provides the basis of an assessment of the rationality of those who accept such theories. Van Prooijen (2019) argues that if conspiracy theorists are rational, we ought to see group differences between those who accept such differences and those who reject them suggestive of higher rationality in the former. But we see the reverse: individuals more likely to accept conspiracy theories also tend to be more reliant on irrational cognitive processes and have other unjustified beliefs. Moreover, and worse for their rationality, these irrational dispositions seem to be causally involved in conspiratorial ideation.

There is extensive evidence that those who accept conspiracy theories are also more likely to accept supernatural and paranormal claims (e.g. Drinkwater, Dagnall, and Parker 2012; Darwin, Neave, and Holmes 2011). Perhaps part of the explanation for being higher in supernatural and paranormal belief is a greater disposition to perceive agency where there is none (Douglas et al. 2016; van der Tempel and Alcock, 2015) and a disposition to detect patterns in random noise (van Prooijen et al. 2018). There is also a link between acceptance of conspiracy theories and a tendency to see meaning in pseudo-profound bullshit (Pennycook et al. 2015). The correlation between conspiratorial thinking and other unfounded beliefs suggests that those who accept the former suffer from some kind of deficit in rational thinking.

The deficit model receives additional and more direct support from evidence of a reliance on type 1 cognition and on heuristics and biases. For example, believers in conspiracy theories are more likely to commit the conjunction error (thinking that a conjunct is more likely than its individual conjuncts; Brotherton and French 2014). Most directly of all, conspiratorial ideation is associated with lower levels of analytic thinking, and the association seems to be causal (Swami et al. 2014). Finally, conspiracy theories are more attractive to those lower in education (Douglas et al. 2016) and those lower in intelligence (Stieger et al. 2013). All of this seems to add up to a strong case for the claim that acceptance of conspiracy theories is the product of irrationality.

A partial, but only partial, rebuttal might be found in the work of those people who have looked to social, rather than psychological, explanations for the acceptance of conspiracy theories. Just as conspiratorial ideation has its psychological correlates, it has its social correlates. These social correlates can be summed up in Uscinski and Parent’s (2014) evocative phrase: “conspiracy theories are for losers.” This is true in multiple ways. As they document, conspiracy theories are more likely to be accepted by those on the losing side of
elections (perhaps the Trump era presents an exception to this pattern, as Rosenblum and Muirhead (2019) suggest, though perhaps this reflects a deeper sense among many Trump supporters that they have lost out despite the election victory). Conspiracy theories are more likely to be accepted by those who feel they lack control (Bruder et al. 2013) and affirming a sense of control reduces conspiratorial ideation (van Prooijen and Acker 2015). These theories are therefore more attractive to those of low status, due to income or ethnicity, for example (Uscinski and Parent 2014).

On the basis of this evidence, we might argue that conspiratorial ideation arises from dispositions that are psychologically, and perhaps socially, adaptive. If one is a member of a low status group or is otherwise at the mercy of forces that one cannot control and which cannot be expected to have your best interests at heart, hypervigilance with regard to threats might be adaptive. Whether people are intentionally conspiring against one or one’s group or not, institutions may be structured, and individuals may act, in ways that reinforce inequalities, and it may be adaptive to be alert for these possibilities (Bost 2018). Evidence that people who report having personally experienced group-based victimization are also more likely to believe in conspiracies targeting their group (Parsons et al. 1999; Simmons and Parsons 2005) provides additional support for the hypothesis that conspiratorial ideation may be an adaptive response to threat.

This hypothesis not only explains why it may be adaptive for members of some groups to take the hypothesis that there are people conspiring against them seriously and to seek evidence for that hypothesis, it also explains some of the supposedly irrational biases involved in conspiratorial ideation. It explains, for example, a bias toward pattern seeking and hypersensitivity of agency detection. Perceiving oneself to be in a dangerous environment alters the relative costs of false positives and false negatives: if people are conspiring against one, false positives may be a small price to pay for protection against genuine plots. The same dispositions toward pattern seeking and agency detection might explain the higher observed rate of acceptance of supernatural and paranormal beliefs, which may arise from the perception of hidden agency or forces at work. It might even explain the illusion of meaning in pseudo-profound bullshit: perhaps that, too, arises from a disposition to see meaningful patterns in noise.

This explanation not only explains conspiratorial ideation; it rationalizes it. Accepting bizarre conspiracies is the price agents pay for being alert to real dangers. Compare the way in which ordinary human agents are disposed to be oversensitive to cues for the presence of danger, like the jogger who startles when a stick falls near her path. Detection of movement by a snake-shaped object might trigger subpersonal mechanisms dedicated to threat detection, in a way that would be adaptive in an environment that actually contained snakes. While it would be better to have mechanisms that are more reliable detectors of genuine threats, such mechanisms may not be possible. It might, for instance, take too long for such a mechanism to accumulate the evidence needed for a higher rate of accuracy: too long compared to the speed at which a snake may strike. Thus, reliance on such a mechanism is not something for which the agent can be criticised.
However, this explanation only goes some way toward rationalizing conspiratorial ideation. While it explains some of the correlates and causes of such ideation as adaptive, it leaves others unexplained and unjustified. Why is such ideation associated with lower levels of analytic thinking and with the conjunction fallacy, and so on?

We might explain conspiratorial ideation in the following way: there is a sense in which such ideation is subjectively rational inasmuch as it is the best that very limited agents, who find themselves in potentially threatening environments, can achieve. If one is in a threatening environment, a disposition to seek evidence of threats is adaptive. But we see the acceptance of bizarre conspiracies only from those agents who in addition to finding themselves in such environments do not have the cognitive capacity to reject these theories when strong evidence against them accumulates. Acceptance of these theories is the price that those who must be on the lookout for danger must pay, given that they lack the capacities to do better.

I think there are grounds for going further: for holding that there is a deeper sense in which conspiratorial ideation is rational. The dispositions that underlie it are not merely rational in the sense that the agents whose dispositions they are cannot be expected to do better (because they lack the capacity), but in the deeper sense that these dispositions are cognitive adaptations. We can rationalize, and not merely excuse, the causes and correlates of conspiratorial ideation,

**Testimony**

At a high enough level of abstraction, those who accept moon landing conspiracies or the claim that scientists are conspiring to make people fear climate change accept these theories for precisely the same reasons as I accept the official story about the moon landing and the science of climate change: on the basis of testimony. In that respect, and again at that level of abstraction, we are all behaving rationally. We cannot coherently forgo this route to belief acquisition, because it provides us with the very tools we need for epistemic assessment; abandon testimony and we will find ourselves without the means for engaging in complex cognition at all. Nowhere is the depth of our dependence on testimony more apparent than within science.

To an extent that is larger than laypeople recognize, scientists must trust one another. They must trust that their tools (intellectual and physical: algorithms, methods of analysing data, software, scanning equipment, and so on), which were developed by other scientists, work as advertised, because they typically lack the skills (and almost always lack the time) to calibrate them, or even to understand the technical details of their implementation. They must trust that scientists in different areas conduct their research responsibly and produce data and theories that are reliable. Think of the way in which in which psychologists use evolutionary theories as constraints in their own work, despite lacking the expertise of evolutionary biologists. Even within a single lab or for the purposes of a single paper, scientists rely on one another to play their part, while lacking the expertise and the time to check each other’s work. In this age of big science, in which research is routinely conducted by multiple labs and across disciplines, the need for such trust increases further.
Science is unusual in the way it institutionalises the social distribution of cognition, harnessing sometimes conflictual (peer review and post-publication review) and sometimes cooperative processes, but we are all reliant on one another in this kind of way. This reliance is in fact distinctive of human beings as a species (Levy and Alfano 2019). It is not something we can dispense with when it comes to the explanation of complex events: it is impossible for any of us to assess a conspiracy theory, or a scientific theory, all by ourselves.\(^1\) Even scientists who test theories in the domain of their expertise rely on each other for their epistemic priors in the light of which they process their data and develop their hypotheses. Thus, the fact that those who accept conspiracy theories do so on the basis of testimony is no basis for criticism of them.

What differentiates those who accept the official account of the 9/11 terrorist attacks (which is, of course, a conspiracy theory) from those who reject it in favour of a theory turning on other conspirators is not that they accept testimony, but from whom they accept it. We can go a long way toward explaining receptivity to conspiracy theories by explaining this differential receptivity to sources of testimony.

We know a great deal about receptivity to testimony, and what we know throws light on this differential receptivity. We know that agents filter testimony by reference to cues of reliability and benevolence. It is obviously adaptive to prefer competent testifiers to incompetent, and both children and adults exhibit this preference (Harris 2012). It is also adaptive to prefer the testimony of those who have our best interests at heart to those who might seek to exploit us or be indifferent to our welfare, and we also exhibit a preference for testimony from those who have track records of behaving benevolently (Sperber et al., 2010). To some degree, we have no choice but to assess both competence and benevolence by our own lights: someone is incompetent insofar as she manifests belief that we assess as false, and someone is malevolent insofar as she exhibits behaviors that we judge immoral.\(^2\) In the light of all these facts, it is unsurprising that we prefer testimony from people who resemble ourselves: from those who share our partisan leanings (which are evidence, _inter alia_, of prioritising values we share) and those who share our background beliefs.

By itself, this receptivity to different sources of testimony helps to explain why some people are more receptive to official stories than others. Consider, for instance, the correlation between lower education and lower socio-economic status and conspiratorial ideation (Douglas et al. 2019). Since the official narrative are promulgated by elites, we should expect elite audiences—those who can identify with these sources—to be more receptive to them than those who cannot identify with them. Couple that with the heightened vigilance to threats that can be expected from groups that, due to the ideology dominant in their society or to lack of resources, are more vulnerable to exploitation, and a heightened receptivity to alternative narratives is easily explicable.

---

\(^{1}\) In fact, a disposition to check official stories for ourselves may leave us vulnerable conspiratorial ideation, because it cuts us off from the social sources of reliable information. See (Levy 2007).

\(^{2}\) We are not, of course, incorrigible: we will change our own beliefs in the light not only of evidence, but of disagreement by those we take to be better placed than we are or are simply more numerous.
But we still haven’t explained differences in analytic thinking or reliance on heuristics and biases. So far as it goes, the account just sketched predicts that we should see differences in acceptance of conspiracy theories correlated with socio-economic status and other social differences, but it does not explain why we also see differences in reliance on intuition and so on. I will now argue that these differences, too, should be understood as differences in receptivity to testimony.

Analytic Thinking and Testimony

We commonly associate analytic thinking with rational thought. There can be no doubt that analytic thought—or type two cognition more generally—is required for engagement in our paradigmatically epistemic practices. Mathematics, science and philosophy can only be conducted by way of slow, effortful, serial processes. While even in these domains, trained intuitions have an important role to play, they must be systematized and probed by analytic thought to be truth conducive. It does not follow, however, that analytic thinking can simply be identified with rational thought and intuitive response with irrational. Rather, which kind of cognition is better depends on the kind of challenge the person faces and the context in which they face it.

Of course, it is widely recognized that type two cognition is a limited resource and that given the costs of deploying it, agents often do better to rely on the “adaptive toolbox” bequeathed to us by evolution in the form of heuristics and biases (Gigerenzer and Selten 2002). The claim I am making goes beyond this point however. Heuristics and biases are not merely adaptive, given our time and resource constraints: for many challenges, they perform better than type two cognition even if we are able to deploy them. Given the right context, we ought to rely on them even if we have the time and opportunity to rely on type two cognition instead. Most importantly for us here, they typically work by responding appropriately to implicit testimony (Levy 2019).

Consider, for illustration, framing effects. Such effects are often cited as paradigmatically irrational (e.g. Shafir and LeBoeuf 2002), on the grounds that the way in which options are framed may lead agents to prefer one or another: since nothing has changed except the frame, it is supposedly irrational to change one’s preferences. But if frames constitute implicit testimony, that’s a mistake: there is obviously nothing irrational in choosing option A over B on the grounds that A was recommended by someone, and that remains true even if I would have chosen B instead given a different recommendation. That’s just how recommendations are supposed to work: they lead us to rational changes in our preferences without (necessarily) adding any additional information about the options themselves. So if framing effects arise out of implicit recommendations, they are rational.

There is, in fact, evidence that framing effects work in just this kind of way (Sher and McKenzie 2006). The choice of frames reflects choosers’ attitudes to the options and frames are implicitly understood as conveying this information. Agents who lack grounds on which to decide between options for themselves do well to take the information conveyed by framing as decisive and act accordingly. Relying on this ‘bias’ is in fact rational, and
overriding it and instead relying on effortful processing (in the absence of any grounds for type two choice) would lead to worse choice on average.

Framing effects are by no means unusual in conveying implicit information. Our (putatively) “mindless” (Thaler and Sunstein 2008) tendency to be guided by options presented as the default works in precisely the same way: the fact that an option has been selected as a default is evidence that those who selected it considered it choiceworthy (Carlin et al. 2013; McKenzie et al. 2006). Since—once again—being guided by implicit testimony is rational, it is rational to rely on this apparent bias. Elsewhere, I have argued that the ballot order effect (the tendency of low-information voters to prefer candidates listed earlier on the ballot to those listed later), which serves as John Doris’s (2018) principal example of a deeply unintelligent process, functions in precisely the same way: it provides implicit testimony, and in the absence of reasons to override this testimony, it is rational to be guided by it (Levy 2019).

Ballot order effects look irrational because as a matter of fact, candidate quality does not correlate with candidate order. Similarly, framing effects, or preference for the default option, might look irrational when (as in the lab), frames are chosen without regard to option quality, or when the options do not differ in quality. But the fact that a particular mechanism may fail to lead us to choose well is not evidence that it is not a mechanism that has the function of leading us to choose well. If these mechanisms all function by providing us with implicit testimony, then we should expect them to fail when they highlight bad options: that’s how testimony works. It is no objection to reliance on testimony that testimony can be used to deceive. Or rather, it is no objection to the reliance on testimony that it can be used to deceive if it is not usually used to deceive, or (more carefully) the costs of being misled by testimony are lower than the benefits of being appropriately guided by it.

The costs and benefits of testimony are dependent on facts external to the agent: the proportion of deceptive testifiers in their environment and facts about the reliability of sincere testimony (for instance, in a highly unstable environment, testimony is likely to have low reliability, because the experiences of those who testify are not good guides to the circumstances the recipients of testimony will encounter). Since heuristics and biases are apparently adaptations, to the extent they function by the provision of implicit testimony we can be confident that the payoffs were, in the environment of evolutionary adaptation, sufficiently large to constitute significant selection pressures.

These considerations provide us with a basis for assessing the costs and benefits of reliance on analytic versus non-analytic cognitive strategies. While all agents filter testimony in light of cues carrying information about the reliability and benevolence of testifiers, their filters will be more or less sensitive depending on how they weight the costs of false positives versus false negatives. Agents will differ, too, in the degree to which they check testimony for plausibility, by analysing it. More analytic strategies are adaptations to lower trust environments; less analytic strategies are adaptations to higher trust environments.
Note that the relevant environment here is local: it consists of the agents with whom the person can be expected to interact most (or rather of the kind of agents with whom the relevant mechanism ‘expects’ to interact; presumably, sensitivity levels will be set by cues that were, in the past, reliably correlated with agents of different kinds). We should expect agents who live in smaller and more homogenous communities (for instance) to employ less analytic strategies while those in larger and more heterogeneous communities (particularly communities which are heterogeneous in terms of values) to employ more analytic strategies.

The Paradox of Conspiratorial Ideation

If the forgoing is correct, we ought to see higher levels of conspiratorial ideation in individuals who are at once high trust and low trust. They have low levels of trust toward outgroups, especially to those they perceive as powerful and potentially malevolent. Conspiracy theories are for losers: those who see themselves as worse off, in virtue of their group belonging or the fortunes of the political party they identify with, are likely to exhibit comparatively low levels of trust toward those they see as representing the powerful; accordingly, they will be less likely to accept the official story. But being low trust in official sources isn’t sufficient for being a conspiracy theorist. In addition, you need to have relatively high levels of trust toward unofficial sources: toward in-group members and toward others who are assessed a working against or in opposition to the elites. This latter disposition toward (selective) high trust is manifested in a reliance on heuristics and biases—on type one cognition more generally—in the absence of cues for rejection of testimony.

Interestingly, there is evidence that those higher in conspiratorial ideation are also more prone to the illusion of explanatory depth (Vitriol and Marsh 2018). The illusion of explanatory depth may best be understood as an index of the extent to which agents, unknowingly, offload cognition on others (Levy 2007; Rabb et al. 2019; Wilson 2004). That is, the illusion arises when people take themselves to be in individual possession of knowledge that is available to them only inasmuch as they are embedded in a community of knowers. Thus, their vulnerability to the illusion is indirect evidence that those who are prone to accepting conspiracy theories also exhibit a high degree of something like trust in others, insofar as these others do not trigger their filters for rejecting testimony.

If the forgoing account is correct, agents more prone to conspiratorial ideation do not manifest cognitive dispositions that are rationally criticisable. Pace Meyer (2019; see also Cassam 2018), they do not exhibit intellectual vices (while I will not argue the point here, I suspect the point generalizes: what vice epistemologists call vices are typically adaptations mismatched to the agent’s environment). While the fact that conspiratorial ideation does not manifest intellectual vice does not entail that the problems it gives rise to are not best addressed by attempts to change agents’ cognitive dispositions (any more than the fact that headaches are not caused by an absence of aspirin entails that they cannot be treated by its administration), it does suggest that an alternative approach might be in order.

This is not the place to develop such an approach. However, a recognition of how deeply we are dependent for our knowledge on our epistemic environment—material, intellectual and
social—suggests that a prime target of intervention should be that environment. We might ensure that conspiracy theories do not circulate. Or—recognizing the dangers of censorship and of overreach by those who take it upon themselves to identify such theories—we might take steps to increase and to broaden trust in the epistemic authorities. Those who are most prone to conspiratorial ideation often have good grounds for their mistrust, even if it overgeneralizes: they really have good reason not to see themselves as having a stake in the wellbeing of society and really have good reason to worry that elites do not place much weight on their welfare. Trust can be increased if it is earned: the best way to counter their pernicious effects is to ensure that glaring inequalities and inequities are addressed.

References


— Recognition of these dangers is central to Coady (2019). It should be clear, however, that I reject his solution to the problem. Self-conscious vigilance over what we are told is the cause of the problem, not its solution (of course, we all ought to exercise vigilance, but by playing our role in distributed epistemic networks: we ought to exercise vigilance mainly at a local area or within our special area of expertise, defined narrowly. Outside these areas, we should mainly take things on trust.


