Review of Philosophy and Psychology
https://doi.org/10.1007/s13164-024-00765-6

Sub-Category Generalism About Conspiracy Theories

Paul Noordhof[1]

Accepted: 23 December 2024

Abstract
I argue for a kind of sub-category generalism about conspiracy theories. I identify four features that negatively contribute to the evaluation of any conspiracy theory that has them. I argue that particularism about conspiracy theories is in conflict even with this moderate position. I explain the implications of sub-category generalism for questions about the epistemic environment in which epistemic agents operate. I argue that even if there are pragmatic reasons for being vigilant about the possibility of malign conspiracies, this does not support rejecting the claim that it is a prima facie epistemically wrong to believe a conspiracy with any of the four features I identify. I discuss how the favoured sub-category generalism may help to refine discussion about the relationship between the psychology of those who believe conspiracy theories and subjects who have delusions. There are grounds for thinking that believers in conspiracy theories display irrationality at least as bad as those suffering from delusions. This is not to pathologise believers in conspiracy theories but recognise the continuities between those who have delusions and believers who are not the subject of clinical attention.

Keywords Conspiracy theory · Delusion · Generalism · Particularism

A central philosophical debate about conspiracy theories concerns whether they should be assessed solely upon their evidential merits, as Particularists urge, or whether there are features which negatively contribute to their epistemic evaluation independent of the evidence in favour or against their content, as Generalists claim. Although this debate can be taken in isolation, it has implications for the connection between the attribution of epistemic properties to subjects and how collectives should be organised to maintain appropriate vigilance with regard to ways in which their democratic organisation may be undermined, for the psychology of those who believe in conspiracy theories, and how they relate to subjects with delusions.

✉ Paul Noordhof
paul.noordhof@york.ac.uk

[1] Department of Philosophy, University of York, York, UK

Published online: 08 January 2025

The paper will proceed as follows. In the first section, I characterise the Particularism vs. Generalism debate more precisely. We shall see that particularism also denies that there is even a sub-category of conspiracy theories characterised in terms of features that negatively contribute to their epistemic evaluation *independent of evidence* for or against the theory. In the second section, I will identify four features that negatively contribute to the evaluation of any conspiracy theory that has them. These features cover quite a large class of conspiracy theories including ones that some say that we should take seriously. In defending the significance of these features, I will be arguing for a Sub-Category Generalism. In the third section, I will explain the implications of this position for questions about the epistemic environment in which epistemic agents operate and how pragmatic benefits for collectives of certain epistemic behaviour relate to the assignment of epistemic properties to individuals regarding what they believe. In the fourth and final section, I will discuss how the favoured sub-category generalism may help to refine discussion about the relationship between the psychology of those who believe conspiracy theories and subjects who have delusions. There are grounds for thinking that believers in conspiracy theories display irrationality at least as bad as those suffering from delusions. This is not to pathologise believers in conspiracy theories but recognise the continuities between those who have delusions and believers who are not the subject of clinical attention.

1 Particularism-Generalism Debate

The recent consensus of philosophical thinking about conspiracy theories is that one should be a particularist. According to the particularist, we should assess individual conspiracy theories on their evidential merits (*the positive thesis*), we should not make generalisations about the class of things labelled conspiracy theories concerning whether or not they should be believed (*the negative thesis*) (Buenting and Taylor 2010, pp. 568-9; Dentith 2016, p. 582; Keeley 2019, p. 424; Dentith 2021, p. 9900-1).

An important part of the support for the particularist position is their championing of a minimal neutral understanding of conspiracy theories according to which conspiracy theories are simply theories that explain a certain phenomenon by appealing to a conspiracy. A conspiracy is understood to be a plan developed between two or more people directed toward a certain end where they keep their intentions, and some activities related to the plan, hidden as necessary in order for the plan to be successful (cf. Keeley 1999, p. 116; Pigden 1995, p. 9, Dentith 2016, p. 577; Dentith 2017, p. 3; Dentith 2018; Dentith and Orr 2018 for why secrecy may be necessary in many circumstances but not essential to any conspiracy; Keeley 2019, p. 423). The phenomena explained do not need to be the target of a conspiracy. The reason why they occurred may be due to a conspiracy that was botched (Pigden 1995, pp. 12 – 4). Particularists urge that there is no more plausible definition of a conspiracy theory, with a typically negative connotation, that applies to what people generally have in mind by conspiracy theories – for example, that covid vaccinations are to implant a micro-chip to get us under

control – and not to well-established historical conspiracies. They conclude that the only way to assess any conspiracy theory, in this sense, as a result is on the basis of evidence.

As the negative thesis of particularism indicates, it is defined in opposition to Generalism. Although particularists can sometimes define generalism very loosely as the thesis that it is irrational to believe in conspiracy theories, in their more circumspect moments, they identify a weaker thesis more plausibly attributed to generalists that the particularists judge to be mistaken:

> There is, at least, something (a) prima facie or (b) by default, epistemically wrong with believing anything which falls in the class of conspiracy theories in virtue of certain of features that *all* of their class share (Buenting and Taylor 2010), pp. 568-9; Dentith 2017, p. 5; Dentith 2016, p. 581, talks about believers being thought 'typically irrational', Dentith 2021, pp. 9900-1, some plausibly understood moderate generalists agree with this characterisation of their position, e.g. (Harris 2018, p. 240).

Here we may take 'prima facie' to be a contribution to the issue of whether having such a belief is wrong (cf. Ross's notion of prima facie duty, Ross 1930, pp. 19–20). If the 'default' alternative adds something, it is that the contribution may be defeated if the condition, which is required for the contribution to the issue to hold, fails to hold. The wrongness in question is epistemic. When particularists assert that conspiracy theories should be assessed by their evidential merits they are not taking a stand against those who claim there might be pragmatic reasons in favour or against beliefs (for example, you are not taken seriously by the public). The debate is over whether there are features relevant to the truth of a theory, or epistemic rationality in believing it, distinct from evidence for or against it. Here *relevant to* should be understood to be compatible with, in fact, the presence of the features sometimes being misleading. Some particularists characterise generalism in terms of conspiracy theories being prima facie false (e.g. Basham 2016, p. 4; Dentith 2021, p. 9906). My preferred characterisation explains the idea of something being prima facie false, although perhaps, in fact true, in terms of there being an epistemic mistake in believing it.

Without the negative thesis, there is no immediate clash between particularism and generalism. If there is something prima facie, or by default, wrong with believing something that falls within the class of conspiracy theories, it doesn't follow that they should not be assessed on a case by case basis. If the prima facie wrongness may be outweighed by evidence in favour of the theory (e.g. highly fortuitous events that would have to be written off as chance by the official story but receive satisfying explanation by a conspiracy theory), then conspiracy theories still need to be assessed on a case by case basis (for illustrations, see Buenting and Taylor 2010, pp. 573-7).

The clash with the positive thesis of particularism only arises if it is taken as the claim that, on a case by case basis, conspiracy theories should *only* be assessed on their evidential merits. Refining the positive thesis of particularism in this way brings out its connection to the negative thesis. Any generalisation

about a conspiracy theory that presents a hurdle for evidence (or candidate evidence) in favour of the theory to overcome would imply that the theory should not be assessed on its evidential merits alone. Let Particularism be the thesis that whether or not it is epistemically wrong for a subject to believe *any* conspiracy theory, T, is settled by the evidence that the subject has in favour or against the conspiracy theory alone (Harris characterises this as *weak particularism* and claims that it is unobjectionable and no philosopher has objected to it (Harris 2022, p. 13). Tightening the characterisation of particularism in this way brings out the conflict with the weaker position that I intend to support.

Sub-Category Generalism Certain types of conspiracy theories have features such that it is prima facie, or default, epistemically wrong to believe any of that category.

If there is a characterisation of a sub-category of conspiracy theories in which theories in this sub-category have features that present a hurdle for the evidence (or candidate evidence) to overcome, then it is not true that *all* conspiracy theories should be assessed on the basis of evidence as to their truth or falsity *alone*. Particularism is not just a denial of generalism but of any form of sub-category generalism (contrary to what Clarke, and Boudry and Napolitano, seem to suppose, Clarke 2023 p. 58; Boudry and Napolitano 2023, pp. 22 – 3).

Recently, some particularists have accepted that some conspiracy theories may have 'suspicious-making' features, which brings them much closer to the sub-category generalist position (e.g. Dentith 2022). There is an important difference of emphasis though. They take these features to indicate that a theory with them is not currently warranted. Nevertheless, they insist that we only discover that a theory *should not* be believed if the theory is investigated to see whether the suspicion is merited and the theory is shown to be false (Dentith 2022, pp. 6–10). The suspicious-making features prioritise which theories we should take more or less seriously but the features, by their lights, do not license the conclusion that we should not believe a theory that has suspicious-making features. So they retain the position that all conspiracy theories should be assessed on the basis of evidence as to their truth and falsity alone, which sub-category generalism denies.

My position is no heroic defence of the prima facie or default epistemic wrongness of believing any theory of the kind specified by the minimal neutral definition of conspiracy theory. I claim that there is a sub-category of conspiracy theories for which it is prima facie, or default, epistemically wrong to believe (hereafter, I'll just say wrongness rather than epistemic wrongness). We can set aside the question of whether this sub-category is what people always have in mind when they talk of 'conspiracy theories'. Although there is some linguistic evidence that 'conspiracy theory' has a negative connotation, we can focus on the interesting epistemic question regarding conspiracy theories of a certain type without settling the general semantic question (for linguistic evidence in favour of a negative connotation, see Napolitano and Reuter 2023). In this respect, I concur with Boudry and Napolitano's separation of the epistemic issue from the semantic issue as the basis for discussion to progress (Boudry and Napolitano 2023, pp. 23 – 4).

In the next section, I will identify four features drawing upon the conspiracy theory literature. The failure to appreciate that these features support a version of

sub-category generalism derives from certain epistemic mistakes and a failure to focus on the accurate characterisation of the generalist position. Obviously, sub-category generalism is in danger of being trivially true if the feature in question is the epistemic one of being unjustified or its negation being more justified. So the features in question have the interesting feature of having epistemic implications while, at the same time, being characterised either in non-epistemic terms or, although they are, at least partly, characterised in broadly epistemic terminology – for example, by talking about treatment of evidence – the characterisation does not prejudge that there is something epistemically wrong with believing a theory which has that feature. The claim that there is something epistemically wrong remains informative.

This raises the question of how there can be features that contribute to a prima facie epistemic evaluation of a belief in a theory which are not evidence bearing on the theory. One example is that of epistemic entitlement. Perception, memory, inference and testimony are candidate processes that entitle a subject to believe a proposition without the nature of these processes figuring as evidence for, or justification in, the proposition in question (e.g. Burge 1993, p. 458). That is not to deny that what is perceived or given in testimony may be evidence for the belief. For example, if I see the actions of some conspirators related to the conspiracy, I have evidence in favour of the conspiracy. The claim is that the belief is not supported by evidence in favour of, or a justified belief concerning, the process of perception, memory, inference, or testimony being, for example, reliable or having the function, in normal circumstances, of conveying reliable information or being truth preserving. This is compatible with the outcome of these processes being evidence for a belief.

A second example is higher-order evidence concerning the nature of the process that resulted in a particular belief, for example, evidence that the kind of process that typically results in epistemic entitlements is questionable in the circumstances or that the justification that one has offered may involve a mistake. In the case of higher-order evidence of either sort, the evidence is not in favour or against the content of the outcome except in so far as it is the outcome of a process and the higher-order evidence concerns that process. The higher-order evidence is that, whatever the outcome, there is something epistemically suspect about that outcome. For example, evidence that one is very tired and tiredness results in overconfidence, throws into question the confidence you have in the outcome of a process, undertaken while tired, that typically results in an epistemic entitlement. Potentially more relevant to the case of conspiracy theories, the existence of peer, or indeed expert, disagreement is evidence that the process at which one arrived at a particular belief is suspect. Some of the features below might plausibly be thought of as *higher-order evidence* concerning the quality of the process that gave rise to a subject's belief in a conspiracy theory. Others involve general criticism of presumed evidential relations relied upon by conspiracy theorists.

The debate between particularists and generalists is either a debate over whether there are such features or whether these features should be taken into account in the overall evaluation of our belief practices. We shall look at the second option in Sect. 3. It faces the challenge of shifting an epistemic discussion in the direction of something more pragmatic. Understanding the debate in the first way presents particularists with a different challenge. Is their resistance to the existence of such

features for a sub-category of conspiracy theories based upon a general resistance to relations of entitlement, taking higher-order evidence, or candidate evidence-undermining features, to bear upon whether a subject should believe a certain proposition, or does their resistance just apply in the particular case of conspiracy theories? If so, why? It has been noted that appeals to such features can be the basis for a pessimistic meta-induction against conspiracy theories derived from the past falsity of theories with these features (Harris 2022, pp. 14 – 7). The considerations canvassed in the next section are independent of any such argument.

One final qualification is needed. Consider the case of the subject who is told by somebody they trust as a medical practitioner that there is something wrong with vaccines but governments' health programmes/big pharma cover this up. Is there something prima facie wrong with the subject believing this no matter what features, unbeknownst to the subject, the theory may have? Arguably not. They are arriving at a belief on the basis of testimony from an expert they may reasonably trust.

It is better to see the relationship between the putative features of a sub-category of conspiracy theories to be detailed below, and the prima facie or default wrongness of belief, in the following way. If the subject is, at least implicitly, aware of these features of conspiracy theories and they form a belief in the theory, then the belief is prima facie or default wrong to have been formed. I talk in terms of implicit awareness to capture the idea that they might not have quite conceptualised these features in the suggested way but seem to recognise them in some way. They would be even more mistaken if, not only were they aware of these features but this awareness was either appropriately conceptualised or even part of the reason they had for forming the belief in the theory. Hereafter, the implicit awareness condition should be read into what follows.

The implicit awareness condition highlights the fact that the proposed features identify when epistemic blame is appropriately attributed. Epistemic blame is typically understood as the guiding idea behind internalist accounts of knowledge and justification. I remain neutral about what the appropriate character of knowledge and justification is because there will be a question of epistemic blame even if externalism is the appropriate approach for knowledge and justification. It is to this that talk of prima facie or default wrongness of a belief in a conspiracy theory adverts. One consequence of this point is that a subject might be epistemically blameworthy even if, in fact, the feature identified fails to bear on the support for the conspiracy theory so long as it is reasonable to think that it does.

2 Prima Facie or Default Wrong-Making Features of Conspiracy Theories

2.1 Counter to a Well-Supported Official Story

The first feature, of the sub-category of conspiracy theories in which we are interested, is that they are counter to some official story of what has happened that is well supported by relevant expertise from socially accredited epistemic authorities,

that is those authorities that, for a particular subject matter, are taken by the society of which they are a part to provide support for belief contents on that subject matter. Call this feature the ***Counter to Well-Supported Official Story*** feature. This is a complex idea that needs a bit of unwrapping.

If a conspiracy theory is counter to an official story, that, in itself, is no reason to think that there is prima facie something wrong with believing it (Keeley mentions this feature, Keeley 1999, pp. 116-7). Official stories can be poorly supported and, indeed, official stories can be deliberate misrepresentations of some situation, as particularists are keen to point out (e.g. Coady 2003, p. 206; Harris 2018, pp. 236-7). The latter may even be widely known. For example, the 'dodgy dossier' that is sometimes alleged to be the basis for the UK Parliatment's support for the Iraq War was widely reported to be dodgy at the time with, it was alleged, some parts copied from a 12 year old PhD thesis (see Guardian, Tuesday 19th August, Timeline; the evolution of the Iraq dossier, https://www.theguardian.com/media/2003/aug/19/bbc.iraqdossier). The interesting case is where there is an official story that is supported by relevant sources of expertise upon which the official story may even draw.

A related idea is that a conspiracy theory of the relevant type is counter to an official story which is the *obvious* account (Keeley 1999, pp. 116-7). If an account is obvious in the relevant sense, then it will be well supported by socially accredited epistemic authorities. But there may be official stories that are not obvious but well supported. It would be a prima facie error to believe a conspiracy theory that conflicts with the official story in this case.

A number of points should be noted about the socially accredited epistemic authorities. The first is that these need not correspond to those endorsed by currently dominant political authorities. Populist parties in government can define themselves against socially accredited epistemic authorities. The second is that, although they are socially accredited epistemic authorities, it does not follow that they are reliable. Social accreditation can be either mistaken or making the best of a bad job because, for that particular subject matter, no source of knowledge is reliable (e.g. early weather forecasting). Some have argued that reliability is a necessary condition. This is a mistake. The question is not whether they *are reliable* but rather whether they are *more reliable* than other sources of information on a matter. If they provide the best grounds a subject has for beliefs based upon testimony upon these matters, a subject is criticisable for not drawing upon them.

If governments have too much control over the social accreditation of epistemic authorities relevant to a particular conspiracy theory – such as limiting them to the promotion of certain views as in authoritarian or totalitarian states – then this feature ceases to be a prima facie or default wrong-making feature concerning a conspiracy theory in that subject area. The qualification does not apply to states in which there is a pretty good spread of opinion allowed in the subject area that may figure in public debate. There can be reasonable disagreement over how much control a government has compatible with this condition being met. When it is met, that is sufficient to support the first feature indicating something prima facie or default wrong with believing a conspiracy theory in conflict with a well-supported official story.

Some particularists argue that the question of whether governments have too much control over the social accreditation of epistemic authorities can't be answered

independently of the question of the extent to which there is evidence that a government conspires to control its citizens in various ways including social accreditation (e.g. Basham (2002), pp. 273-4). Presumption that the first feature shows something prima facie or default wrong with believing a conspiracy theory is question-begging.

There may be reasonable disagreement about precisely how much control governments have but it is hard to hide whether we are under a totalitarian or authoritarian regime. It is not question-begging to dismiss this possibility if there is no evidence that we are. Even if it were question-begging, this would not be sufficient to undermine application of the first feature. The fact that you can't rule out the possibility of being under a totalitarian or authoritarian regime doesn't imply that you fail to be aware of the first feature. By being aware that a conspiracy theory has the first feature, without having reason to believe that you are not under such a regime, you are still prima facie or default wrong if you believe the theory. You only avoid epistemic fault if your awareness of the feature can be set aside because you have reason to believe that you *are* living under a totalitarian or authoritarian regime.

Although being against what socially accredited epistemic authorities conclude can be itself a mark against a conspiracy theory, talk of an official story is not redundant. An official story can be based upon additional information available to governments but not to other institutions. Although governments may have an advantage in the information available to them, that does not mean they cannot be subject to correction by socially accredited epistemic authorities because the latter can often detect whether there is genuine additional information or the government is spinning a line. Governments also face additional responsibilities in the development of an official story because they take it to be an official story for which they will be held responsible and so must be robust.

Concerns about the character of the society you are in, and the motives about the actors in government, are potentially good grounds for *failing to believe* the official story but they are much more dubious grounds for *believing* a conspiracy theory. Journalists, and investigators more generally, espousing a conspiracy theory running against the official story have to establish their epistemic credentials in the mind of the consumer to make their considerations sufficient to be a basis for believing a theory as opposed to making the theory one hypothesis among potentially many concerning the target subject matter. The fact that some well-supported official stories in the past have been shown to be wrong doesn't change the picture. The question isn't whether a conspiracy theory might be true but rather what are the conditions in which a subject should believe a conspiracy theory. The suggestion is that there is something prima facie or default wrong in believing it if it has the first feature we have identified and, thus, substantial evidence in favour of it is required.

2.2 Contrary to Preponderance of Appropriately Independent Relevant Experts

The second feature is that the preponderance of socially accredited sources of expertise, or experts, hold a theory that contradicts the conspiracy theory or, at best, couldn't be true if the conspiracy theory is true. The qualification of 'appropriately independent' to the talk of preponderance reflects the fact that the feature doesn't

involve a straightforward appeal to numbers. The experts must be appropriately independent in the sense that each of their beliefs in a theory T, one that contradicts the conspiracy theory, is relatively causally independent of each others' for each expert to count towards the number of experts. The proper characterisation of this idea requires some further work that cannot be undertaken here but the preliminary idea is as follows. While there is most likely to be, and entirely properly, *some* causal connection between each expert's beliefs as they draw upon each others' work, it is not the case that a belief having the content of the theory in question in one expert is part of the most precise causal explanation of a belief having the content of the theory in question in another expert, if they are both to count towards the number of experts holding the theory in question (see Noordhof (1999), Noordhof 2006, Noordhof 2020, pp. 274-5, for further discussion of related issues, see Goldman 2001, pp. 150-6). In less abstract terms, if one expert believes the theory only because another does, and they defer to them, both of the experts shouldn't both count towards the number of experts. It may not be easy to detect how independent the experts are from each other in this sense. Nevertheless we are used to receiving information about group think, the influence of one expert on another, and so on, which qualifies the extent to which we take expert consensus seriously.

We'll call this second feature the ***Contrary to Preponderance of Experts***. It is of additional importance if the government of a particular country is not to be trusted and, thus, any official story may be a misrepresentation.

Subjects who are led to distrust socially accredited epistemic authorities find it difficult to go it alone. We are all dependent on many other people to form beliefs about the world required for leading our lives successfully. Such subjects are led to put their trust in more discreditable sources of which one of the more likely is the source of the conspiracy theories they started to believe in the first place. Believing a theory on the basis of some other source than a socially epistemically accredited one, or at best one that is in a considerable minority, already displays a way in which the subjects are making a cognitive mistake. The point is not that the preponderance of socially epistemically accredited ones are reliable, or cannot fail to be true, but rather that in most cases they are still the best source we have (Harris 2022, p. 7). The credentials of the preponderance of expert opinion are better and there is readily available information about how the sources of conspiracy theories tend to be unreliable (Harris 2018, p. 255).

It is not enough to argue that appeal to typical sources of expertise may be mistaken because these have proven to be unreliable. It must also be argued that these sources of expertise are *less* reliable, and worthy to be the basis of belief, than what a subject knows about the proponents of conspiracy theories. Without the latter claim, there would still be something prima facie wrong with believing a conspiracy theory rather than the view expressed by the preponderance of experts with which it conflicts, even if the view expressed by the preponderance of experts is, in fact, wrong or, as it turns out, unreliable in this instance. While a conspiracy theory rightly may be given some credence, and the contents of beliefs formed from a socially accredited epistemic authority lose some credence, as Hagen suggests, any more radical response than this displays an error (Hagen 2022, p. 212-4). If somebody believes, as opposed to give some credence to a conspiracy theory, or withhold belief in the

negation of the conspiracy theory, over what is proposed by the preponderance of socially accredited epistemic authorities, then they are doing something prima facie, or by default, mistaken.

2.3 Counterevidence Neutralisation

The third feature is a certain characteristic way of dealing with counterevidence as follows. Suppose that a conspiracy theory is a conjunction of statements, H, concerning a particular conspiracy. H together with some auxiliary hypotheses, A, which are assumptions about the way the world works that are not part of the conspiracy theory, have implications about what has happened: O (for occurrence). In fact, there is evidence that not-O is the case. The conspiracy theory is retained with some adjustment either to the conjunction of statements that constitute it or to the auxiliary hypotheses. Call these adjustments H* or A* respectively. H* and A or H and A* or H* and A*, depending on the adjustments made, either fail to imply that O or they imply that not-O. The adjustments have neutralised the counterevidence while retaining something which is a plausible successor of the conspiracy theory. The adjustments to the conspiracy theory, if any, don't change its central claims. Treatment of evidence in this way is the invariable response to counterevidence and, when the adjustments are made, no implications about what has happened or will happen are identified which would be suitable tests of whether the adjustments are appropriate. The adjustments are just made for the sake of dealing with the evidence (Clarke 2002, notes a feature like this). For example, evidence against the conspiracy theory is taken as evidence that the conspirators are highly effective at covering up traces without acknowledging that there would be likely traces of the cover up which could then be investigated (Boudry 2023, pp. 619 – 22). Call this feature ***Counterevidence Neutralisation***.

As with the preceding two features, the third feature is a relational feature characterising a particular type of theory, this time in virtue of how its proponents respond to evidence against it. There may be two theories making the same claims about the world, one of which has the feature concerned, the other of which does not and, as a result, the second theory is unproblematically rejected because of counterevidence without further theoretical development. Partly for this reason, the significance of the feature does not suffer from the general problems raised with criteria for the demarcation of the sciences from the non-sciences. The aim is not to identify a general feature of the content of some theory, or to distinguish one practice, scientific practice, from another, but characterise a relational feature that, if a theory possesses it, renders it prima facie wrong to believe it.

The feature is not properly thought of as *evidential insulation* (proposed by Napolitano 2021, pp. 88–95). The feature does not describe a theory insulated from evidence. The evidence against a conspiracy is taken to imply an adjustment to the theory, for example, to involve, in addition, competent covering of tracks that might have revealed a conspiracy of that type. Adjustments in a theory to explain away the evidence apparently against it is not necessarily a flaw. It is standard that a scientific hypothesis H plus auxiliary hypotheses A relating to how there may be evidence

relating to the truth of the hypothesis allows retention of high credence in H by adjusting to low one's credence in A, or adopting A*. In Bayesian terms, a situation in which evidence E is made less likely by H and A than not-(H and A) allows retention of a high credence in H at the expense of lowering credence in A, or rejecting A, with raised credence in a substitute proposition A*, without irrationality (Poth and Dolega 2023, pp. 7–14).

When scientists are engaged in legitimate evidence neutralisation, they recognise the dangers in protecting a favoured hypothesis by adjusting other claims, say substituting A* for A, and look for the possibility of straightforward supporting evidence for H together with A*. An illustration of this characterisation of legitimate scientific practice is Imre Lakatos' idea that, when scientists deal with counterevidence to theories in this way, we don't have a case of a degenerating scientific research programme if there is the ability to make *novel* predictions as a result of doing it (Lakatos 1978, pp. 31–37; Harris 2018, p. 246). In the case of conspiracy theories with the damaging feature identified, the aim is evidential neutralisation for its own sake. There is no interest in considering the implications of how one has neutralised the evidence, making new predictions on the basis of the evidence-neutralising adjustment and searching for evidence that these predictions are born out. As Clarke notes with regard to the controlled demolition theory of the World Trade Centre Towers' destruction, sometimes the conspiracy theory is not even developed in sufficient detail to have testable predictions but rather persists simply by pointing out failures in the official theory which are consistent with an alternative theory of that general type (Clarke 2007, pp. 172-5). If this is a persistent feature, then we have a way in which counterevidence is neutralised at the start rather than as an upshot of the adjustment I mentioned earlier.

Probably one of the main reasons why conspiracy theorists neutralise evidence for its own sake without misgivings is that it seems a natural explanatory move to suggest that conspirators suppress evidence for the conspiracy and create what may be easily taken as evidence to show that there is no conspiracy (Keeley 1999, pp. 122-5; Harris 2018, p. 247). Lee Basham argues that competent conspiracies would be ones in which the absence of evidence for the conspiracy is to be expected and can indicate the presence of a competent conspiracy. Since such a conspiracy cannot be ruled out, the rational response to the possibility of competent conspiracy theories explaining a sequence of events is agnosticism (Basham 2011, pp. 8–9). If Basham is right, it might be thought my candidate third feature cannot be a feature showing the prima facie, or default, wrongness in believing a conspiracy theory.

There are a number of problems with this line of thought. First, a feature that shows the prima facie, or default, wrongness in believing a conspiracy theory of the relevant kind doesn't touch on the question of whether a subject should believe that a conspiracy theory of the relevant kind is false. The existence of the feature is even compatible with the evidence favouring a credence in the conspiracy theory being greater than credence in its negation so long as the credence falls short of that required for belief.

Second, even if a competent conspiracy cannot be ruled out to explain the absence of evidence, if there is a much more probable explanation, that may be enough to believe that there is no conspiracy behind the events. Given what we know about

human incompetence, indiscretion, and the potential benefits of defecting on a conspiracy, then, although we cannot rule it out, a more probable explanation of the absence of evidence may well be that there is no conspiracy at work (for discussion of the possibility of defection, see Pidgen 2016, pp. 209 – 15). Taking the absence of evidence to be evidence of particularly competent conspirators covering their tracks is not properly constrained by evaluation of how effective people are in conspiratorial behaviour. That's why we should be interested in evidence of cover-up.

Sometimes friends of conspiracy theories claim that earlier conspiracies of which we now have evidence provide grounds for taking the possibility of a competent conspiracy to explain the target events seriously. However, evidence of prior conspiracies is evidence of how conspiracies give themselves away. In which case, it is less plausible as a basis for allowing that it is probable that there is a conspiracy that hasn't given itself away for the events in question. Any preliminary support derived from the idea that there have been conspiracies that didn't initially give themselves away dwindles because of the evidence that even competent conspiracies are eventually detected.

The assumption behind the envisaged objection to the third feature seems to be that the only options are that either we believe that there is no conspiracy at work or remain agnostic on the subject. A natural way of understanding remaining agnostic is withholding belief that there is no conspiracy at work and withholding belief that there is. In fact, something more general still seems to be needed. We are not agnostic that there is no conspiracy at work if we give more credence (short of belief) to there being no conspiracy at work than to there being a conspiracy at work. Being agnostic in this more general sense is to hold that either no credences are to be attached or that the credence in the proposition that there is a conspiracy at work is identical to the credence that there is no conspiracy at work so that the subject is neutral on whether there is. Once we appreciate what is involved in agnosticism understood either way, a third option is available. While we attach some credence to there being a conspiracy at work, we attach more that there isn't a conspiracy. None of the considerations offered count against this option.

2.4 Lack of Explanatory Constraint

The feature of evidential neutralisation for its own sake interacts with the last feature of the sub-category of conspiracy theories I want to identify: ***Lack of Explanatory Constraint***. It is the result of three factors. A fundamental difficulty with conspiracy theories involves the resources upon which their candidate explanations draw. Psychology is still, unfortunately, if not in its infancy, in its early adolescence with regard to its application to individual people. We have no detailed understanding of how a particular environment interacts with the history of particular individuals, and their psychological make-up, to result in certain behaviour. There are very few (if any laws) at the requisite level of detail for application in these situations. Let's call the first factor *Current Lack of Explanatory Detail in Relevant Psychological Explanation*. At the same time, we have an interpretative theory – sometimes called a folk theory – the application of which enables us to *make sense* of how agents behave

in the past, the consequence of their behaviour, and make limited predictions about how they will behave (for details, see Davidson 1980, Dennett 1987). This is not a matter of having a detailed causal explanation, because we don't have that, but rather we attribute reasons to understand the behaviour. Although such a theory may give us a strong sense that we understand what is going, this understanding is often partly an illusion. The illusion stems from the way in which the attribution of reasons can give us the impression that we understand the agent's perspective, in the same way we think we understand ourselves, and partly stems from the holistic character of our attribution of attitudes – governed by constraints of rationality or more general considerations of what seems comprehensible as part of human nature – which is relatively unconstrained and indefinitely adjustable to result in an attribution that makes sense of the agent's behaviour and consequences of it. We may characterise the second factor as *Interpretation Provides Potentially Illusory Understanding*. This brings me to the last point. Whatever our position on the question of free will, when we think about how agents may behave, we often take them to, within their abilities, be able to act to give rise to a multiplicity of different kinds of events relatively unconstrained by the environment they are in and how they interact with each other (for discussion of the various ways this has been understood, see Noordhof 2020, pp. 391–400). This is more of an assumption than an empirically well-grounded observation about fellow agents. Call this factor the *Openness of Agency*.

These factors taken together should, and often do, make us hesitant about any explanation of, or prediction about, someone's behaviour. The hesitancy is reduced when we know someone very well, there is a mass of observation and supporting testimony. Call these the mitigating conditions. The amount of observation and supporting testimony required may be somewhat less if the explanation is highly general and so compatible with many different resolutions of detail. The fourth feature is the *lack of explanatory constraint in the absence of mitigating conditions*.

The feature to which these factors contribute has implications for some conspiracy theories. These conspiracy theories often start by offering explanations of errant data that a more limited theory may fail to explain (Keeley 1999; Keeley 2019, p. 425; Boudry 2023, p. 623). A conspiracy theory is generated because it enables us to see a pattern in the events, including the events the official theory explains, but extends the pattern further. However, the explanatory virtues of the conspiracy theory are overstated by this ability to see a pattern because of the nature and limitations of psychological explanations of individuals I have described above. We have relatively unconstrained explanatory development. By that, I don't mean that there are no principles to which somebody developing a conspiracy theory may appeal. The point is rather that for any piece of evidence, there will always be available some explanation of it just by working back from it to a possible attribution of motivational states and beliefs of two or more individuals to fit it into a pattern.

The levels of competence and discretion attributed in support of a conspiracy are a case in point. As we saw, it has been argued that some degree of counter-evidence neutralisation is to be expected given that a conspiracy theory is postulating actors capable of assessing the evidence in favour of their presence, and trying to create evidence that throws people off the scent while at the same time not giving away what they are up to (Keeley 2019, pp. 428-9). Although this point has some force,

the danger that we have identified is that it is easy for a kind of theorising lacking explanatory constraint to deal with any form of prima facie evidence against a theory of that type. The failure to give too much consideration to the extent to which it is more advantageous for one or more of the co-conspirators to abandon the conspiracy and reveal what is going on, to defect as opposed to sticking with the plan, is another example (Pigden 2016, for a discussion of defection).

The feature is related to, but importantly distinct from, Steve Clarke's idea that conspiracy theorists display the *fundamental attribution error* privileging dispositional over situational explanations (Clarke 2002). Cases of the fundamental attribution error involve explaining behaviour by specific dispositional properties of a person's character. You explain why somebody failed to help a stranger in apparent need by taking them to be callous and self-absorbed rather than taking them to be no worse in that respect than anybody else but, as things were, they were in a rush and other people were around who might assist (for empirical evidence in favour of our bias against context, see Darley and Batson 1973). By contrast, the factors I have identified emphasise our relative lack of knowledge of how a situation may affect a particular individual and the way in which agents, have a number of options open to them that we take not to be closed off by their character. While we think that a particular individual may be subject to temptation, we presume they can fight against it and, for example, resist their desire to boast about, or express their anxieties concerning, a conspiracy of which they are a part. If there were a focus on the attribution of character dispositions in the case of conspiracy theories, there would be far greater concern about the possibility of indiscretion, defection, the strain of keeping a conspiracy going, and the like. Although attributions of beliefs and desires explaining an agent's response are attributions of, at least partly, dispositional states, the resulting explanation does not draw upon the specific dispositional properties of a person's character at work in cases of fundamental attribution error.

While the lack of explanatory constraint has its basis in something distinct from the fundamental attribution error, the explanation I gave of the lack of explanatory constraint may also help to explain why we are inclined to fundamental attribution errors. It will always seem that agents can respond in many ways to the context in which they operate unless we attribute to them particular character dispositions. In that eventuality, the attributed character dispositions will seem particularly salient to us, and general truths about how agents tend to operate under time constraints seem commensurately inadequate. For example, although rushed, it might seem that an agent could have acted more humanely unless we take them to have the character disposition of being self-absorbed and callous.

2.5 Errant Data: Lack of Explanatory Detail vs. the Fallacy of Probabilistic Modus Tollens

It has been argued that the problem with conspiracy theories' treatment of errant data is that it involves a problematic form of reasoning: *probabilistic modus tollens*.

> If if p then q is *highly improbable*, and not-q, then it is improbable that not-p (Harris 2018, p. 250).

The fourth feature I have identified provides a more effective explanation.

Here is one illustration of the problem with probabilistic modus tollens. Suppose that if p (a lottery is fair) then not-q (where q is 'I win') is highly probable. However, I win. Therefore, the mistaken reasoning would go, it is probable that the lottery is unfair (Harris 2018, p. 250, Harris references Sober 2002). Similarly, if a non-conspiratorial hypothesis gives low probability to certain errant data which, in fact, obtain, it doesn't follow that we should attribute low probability to the non-conspiratorial hypothesis (Harris 2018, pp. 250-1).

There are, at least, two related problems with the identification of probabilistic modus tollens as the fault behind conspiracy theories' explanation of errant data. The first is that probabilistic modus tollens is an implausible principle of reasoning and there is little reason to suppose that, if conspiracy theorists are making an error, it captures the error they are making. If an hypothesis attributes a high, but below 1, probability to a certain outcome and its negation occurs, this does not have an immediate implication for the probability of the hypothesis. The occurrence of the negation is compatible with, indeed a consequence of, the hypothesis being true and the hypothesis antecedently may be highly probable when compared with the alternatives.

Second, the rejection of probabilistic modus tollens is compatible with the recognition that one theory can be favoured over another by comparing the likelihoods that they give a piece of evidence which, prima facie, can seem to favour a conspiracy theory. Thus, to quote from Sober Evidence E favours H_1 over H_2 iff $P(E/H_1) > P(E/H_2)$.

Recognition of the validity of this relationship is compatible with the verdict we arrived at in the lottery case. Suppose we are comparing the hypothesis that the lottery is fair with one in which it is unfair. My winning doesn't favour the hypothesis that the lottery is unfair because there is no reason to suppose that the lottery being unfair will make my winning any more probable than any other person winning. If the evidence to be explained is just somebody winning, then this doesn't favour an unfair lottery over a fair one. By contrast, although my winning may favour the hypothesis that the lottery is unfair in that it favours me, the antecedent improbability of the lottery being unfair in this way means that the hypothesis remains improbable.

The explanation for why evidence about errant data doesn't favour conspiracy theories is drawn from the fourth negative feature I have identified. The putative high probability that a conspiracy gives a piece of errant data is inappropriately constrained since the capacity of human agents to carry out a particular conspiracy is not appropriately taken into account (cf. Sober's point about intelligent design, Sober 2002, pp. 72 – 8). Human agents are just viewed as free competent agents. The apparent probabilities that a conspiracy explanation gives errant data are in danger of being either inflated or empty (as Harris 2018 also notes, pp. 251-2). Equally, the improbability of a conspiracy hypothesis involving highly competent conspirators is underestimated as result of the absence of explanatory constraint, in particular due to the openness of agency. Of course, that doesn't mean that we shouldn't appeal to human agency at all to explain why things happened. The issue is rather that we should be circumspect in using this style of explanation's ability to mop up

errant data to overthrow an otherwise well-evidenced theory supported by the preponderance of experts, for example.

Taken together, these reasons explain why the fourth feature I have identified is a more plausible and deeper account of what is wrong with conspiracy theories than attributing the fallacy of probabilistic modus tollens.

2.6 Application of the Features

My suggestion is that it is possession of these four features, individually or taken together, that makes a sub-category of conspiracy theories prima facie or default wrong to believe. The more of the features possessed by a theory, the more concerned we should be. The first two concern the conflict between the theory and other theories that have certain sorts of support, for example, the fact they are taken up by a pre-ponderance of experts. The latter two concern how evidence that appears to support the theory may, in fact, be deceptive.

Perhaps the last feature – lack of explanatory constraint – may seem to need the presence of one or more of the others to make it prima facie wrong to believe a theory of this kind. This is questionable. Consider a familiar experience in daily life, gossiping with a friend about the motivations of other people you know given how they behave. We certainly seem to take ourselves to be sharing insights even though, by the lights of the last feature, there is a lack of explanatory constraint unless there are mitigating conditions. Even here, I think the verdict is correct. We might find certain attributions of beliefs and desires to other people plausible, and it can make sense of a pattern of behaviour, but we are circumspect about believing that that is what was going on in most cases. It is an intriguing hypothesis rather than something we believe. That's because we recognise that there is something prima facie wrong with believing it due to the lack of explanatory constraint. If this holds in cases in which we are very familiar, it should hold so much more in the attribution of conspiracies involving shadowy actors.

A theory with one or more of these features doesn't have to be false. It is quite possible that the theory is true and, indeed, that a subject may be aware of evidence sufficient for it to be correct for them to believe the theory. It is important to distinguish between mistakenly believing that a conspiracy theory is false and correctly failing to believe that it is true. Sceptics about conspiracy theorists with one or more of the features I have identified claim that we should not believe the theory unless evidence overcomes the significance of the features in a particular case. That is why there is something prima facie or default wrong with believing a conspiracy theory. They can accept that there is evidence that conspiracies occur, and allow a certain probability that a conspiracy is true (just as Basham argues Basham 2002, pp. 271-2). It doesn't follow that there are sufficient grounds for believing a particular conspiracy theory.

The difference between the generalist position I am recommending and the kind of particularist who allows that there may be suspicious-making features of some conspiracy theories is this. They deny that the suspicious-making features are sufficient to make a theory prima facie wrong to believe. Instead, the suspicious-making

features characterise which theories should be given a lower priority in investigations. We only discover that a theory should not be believed if it is investigated and shown to be false. By contrast, the position recommended here is that the features I have identified make it prima facie, or default, wrong to believe a theory with one or more of them. The occurrence of 'prima facie, or default' allows for the possibility that there is sufficient evidence in favour of the conspiracy theory to overcome this. However, in the absence of the evidence, it is straightforwardly wrong to believe the conspiracy theory. We do not need to discover, by investigation, that the theory should not be believed.

The first two features I have identified might be thought to be evidence against the theory that has them. In which case, the objection runs, taking them into account is neutral ground between the particularist and the generalist. It would be evaluating a theory on its evidential merits alone, which is all that the particularist insists upon.

If the pre-ponderance of experts, say, deny that p (or assert that q which is evidence against p), a favoured conspiracy theory, then that is certainly evidence against p. It is neutral ground that this should be taken into account and is one dimension of the contribution that an instantiation of this feature makes. On the other side, there is all the evidence in favour of the conspiracy theory which, those who are concerned about a conspiracy theory being dismissed, may say demonstrates that the denial is mistaken, or that q is outweighed. The pre-ponderance of experts denying that p makes another contribution. It is higher-order evidence that the process by which those who arrived at p may be flawed in some way. This second contribution is evidence about the process, whatever the output of the process, that a reasoner picks up by noticing the disagreement. In this sense, it is not evidence against p but rather raises a question over the process by which the proponent of the conspiracy theory arrived at p. To set this evidence aside, the proponent of p might have to be extra-careful in thinking about whether the evidence they have is enough to support that p, whether they may have additional evidence that experts may not have, whether biases may be in operation, and so on. While these may ultimately bear upon the overall case for believing that p or not p, they are not evidence for, or against, p, but rather evidence bearing on the status of the case that a proponent of the conspiracy theory has for p. The particularist position – motivated by the desire to take the possibility of conspiracy theories seriously – does not cover this evidence when they say that we should consider the evidence for or against p alone.

A third contribution that the pre-ponderance of experts may make is an assessment of the process by which those in favour of p arrived at the conclusion that they did which, those who favour p, may disagree with. Those who favour p may also disagree about the weight that should be given to q (the evidence favoured by the experts). The requirement that we should assess p in terms of the evidence for and against it alone does not make clear what a thinker should do in situations where there is disagreement over whether something should be counted as evidence, or the weight it should receive. A proponent of a conspiracy theory may reject the putative higher-order evidence concerning the process that led them to p as a result of the bias of so-called experts. The claim I am defending is that, regardless of whether a thinker takes the features to be evidence, although of course they are aware of the features, it is prima facie wrong for the thinker to believe a conspiracy theory with

one or both of the first two features. It is in terms of the second and third contributions of these features that there is no common ground between the particularist and sub-category generalist.

Opponents of generalism tend to view the generalist position as requiring the identification of features of conspiracy theories that would explain why a conspiracy theory is a bad theory and then explain the mistake a subject is making in believing a conspiracy theory in terms of the badness of the theory. As a result, they take it to be an objection to generalism to point out that, while ordinary epistemic agents may not have a reason to believe a conspiracy theory in the context in which they find themselves, it doesn't follow that a conspiracy theory is a bad theory (e.g. Dentith 2018, p. 200). As I have already pointed out, the features in question are relational properties of a particular theory. There is nothing in the content of the conspiracy theory itself that makes it a bad theory. The generalist is not concerned with the question of whether there is something wrong with the theory in itself. The sub-category generalist isn't committed to the claim that there are no circumstances in which it would be rational to raise one's credence in response to evidence in favour of a theory with any or all of the four features identified. The question is whether, if a subject believes it, they are prima facie, or by default, making an epistemic mistake. I have argued that this is the case.

3 Collective Attributions of Epistemic Properties and the Implications for Sub-Category Generalism

Sub-category generalism raises questions about the relationship between epistemically relevant properties of groups and the epistemic properties attributed to the beliefs of the individuals that make them up.

3.1 Epistemically Polluted Environments

We are epistemically dependent upon our social context for our ability to form well-grounded beliefs on a large range of topics. The proper exercise of our epistemic abilities in isolation is insufficient to ensure that the beliefs that we form are likely to be well grounded or true. Individuals with bad beliefs are in an epistemically polluted environment (i.e. one that contains other agents that shape it in a way that makes the epistemic situation of other individuals worse off than it might have been) and it is in virtue of the latter that they arrive at the beliefs they do (Levy 2022, pp. 110, 122-4). Because the collective has this negative epistemic property, the individuals who are part of this collective may be epistemically blameless even, presumably, in the circumstances I envisage involving one or more of the features identified above.

One way to put the challenge from this line of thought is that we mistake an environmental failure for an epistemic failure of the individual. There is nothing prima facie or default wrong with an individual believing a conspiracy theory if that is taken to imply that they are displaying some kind of irrationality. The proper

characterisation of what is prima facie or default wrong can only refer to the connection between the individual and the social environment with no attribution of fault to the former.

My response to the challenge is that the four features plausibly identify features of a theory which, if an individual subject is aware of the features, makes their belief subject to criticism. It is not the case that they are victims of misinformation with no insight into its character. They show a failure to deal with high-order evidence in an appropriate way or recognise failings in how they take evidence to bear upon a theory. Given that we are socially dependent for our beliefs, we should expect that our higher-order epistemic mistakes are some of the most significant. Some subjects in similar social situations fall prey to conspiracy theory beliefs while others don't. Many middle class well-educated people fall for anti-vaxxer theories and yet there are others from the same background who, presumably, are better able to identify the sources of information on which they should rely.

When it is hard to differentiate between different sources of expertise, the right approach is to fail to have confidence about whether one should believe something. For example, suppose that somebody has been interacting with lots of people who are sceptical about socially accredited sources of expertise. This may be sufficient to mean that a subject fail to have a belief that they should have as a result of the expert testimony of these sources. But there is a step from here to believing a conspiracy theory. There is still something prima facie, or default, wrong with that given the epistemic situation in which they are. Observations about epistemic pollution are more effective in explaining failures of belief rather than belief in the opposite.

Subjects' inconsistent treatment of theories with the features in question can reinforce this point. For example, if they tend to believe official stories when their favoured party is in power, generally rely upon the preponderance of experts for their information, or are critical of theories which display the features of evidence neutralisation or lack of explanatory constraint in other areas, then, despite the social dependence of beliefs a subject has upon the environment, another explanation is needed to explain what is going on in the case of belief in conspiracy theories (for further discussion, see Williams 2023, pp. 824 – 30). So while it might be natural for people to make mistakes in this area, and their sense of group identity may give rise to them having a particular belief to express their identity, this does not touch the point about prima facie, or default, wrongness.

3.2 Democratic Dangers of Sub-Category Generalism

Particularists often argue that if it is prima facie or default wrong for individuals to believe conspiracy theories of any type (independent of evidence against them), then societies made up of those individuals will fail to be vigilant against conspiracies that undermine the interests of individuals making up those societies, for example, conspiracies that undermine the democratic structures and decision-making of the societies (e.g. Basham 2011, p. 9–11; Dentith 2021, p. 9901). Pragmatic considerations drawn from the need to be vigilant show that there are no features of the kind that I have identified. If groups can be assigned beliefs, then, for a society to

be vigilant, it should not be prima facie or default wrong for a society to believe a conspiracy theory with any of the features identified. If groups can't be assigned beliefs, then social vigilance requires that it can't be prima facie or default wrong for individuals of that society to believe conspiracy theories with any of the features identified.

We have seen how sub-category generalism is compatible with considering the evidence for a particular conspiracy theory and, indeed, giving some credence to a conspiracy theory short of belief. There is no reason to suppose that this is insufficient to support the requisite degree of vigilance. If that proves to be a mistake, it shouldn't be taken to be grounds for dismissing epistemic properties relating to believing a conspiracy theory. Differences of opinion about the nature of the threat – for example, the number of true conspiracy theories with the features I have identified – does not bear upon the question of whether it is prima facie wrong to believe a theory with one or more of the features I have identified. Instead, it is grounds for organising a society so that, given the fact that there are these epistemic properties relating to beliefs in conspiracy theories, there are other mechanisms to maintain vigilance. This is an instance of a familiar problem when what may be rational for individuals can, collectively, be self-defeating. For example, we task certain institutions or media with the role of ventilating conspiracy theories to which some credibility attaches and subjecting them to further evaluation. Some particularists recognise that this is a possibility but they conceive of it as providing a substitute for sub-category generalism. They identify suspicious-making features of conspiracy theories and rank conspiracy theories, for investigation, to the extent to which they lack these suspicious-making features (e.g. Dentith 2022; esp pp. 9, 11). It is more plausible to see this kind of task as a means of attributing to subjects investigative roles *despite* the epistemic properties attaching to beliefs about conspiracy theories.

The alternative approach recommended here is supported by two further considerations. The first is that the extent to which subjects should be tasked with the investigative role should be determined by the importance of devoting social resources into finding out whether there has been a conspiracy in operation. We might think that investigating conspiracy theories with one or more of the features I have identified is not justified by the resources it would involve. Investigation of conspiracy theories is by nature often retrospective, identifying blame. It might be much better to focus on whether political actors are pursuing effective policies rather than working out whether they conspired, or were just clueless, in pursuing more malevolent policies earlier.

The second is that sub-category generalism of the type I have identified reflects an important epistemological cost that subjects who believe in conspiracy theories with one or more of the features identified face. They are isolated from probably the most generally reliable sources of knowledge there are. This is particularly important politically for populist politicians who have no actual solutions to the problems of society – by that I mean none that will work – but, by discrediting socially accredited authorities, disguise this from the voters. The existence of a conspiracy will explain why the solutions aren't working and, at the same time, direct people away from solutions that might work (Harris 2023, pp. 26–30).

4 Psychological Implications: Belief in Conspiracy Theories and Delusion

The features I have identified are not meant to be exhaustive. Other features may be identified. Nevertheless, the features do relate to, and deepen, our understanding of why some subjects are prone to believe in conspiracy theories and allow us to arrive at preliminary conclusions about how the epistemic mistakes involved in belief in conspiracy theories with these features relate to subjects with delusions.

Individuals who believe a conspiracy theory with the first two features are isolated from important social sources of knowledge, relevant to the phenomena the conspiracy theory concerns, in societies broadly committed to democratic ideals, against the manipulation of its citizens, and with properly accredited sources of expertise. Subjects drawn to such theories are often poorly integrated with sources of expertise in society and often feel themselves disadvantaged by the society in which they live and, consequently, the government that is in operation. Commitment to conspiracy theories with these features can further this isolation (Douglas, Sutton and Cichocka 2017, p. 540, Bowes, Costello and Tasmi 2023, pp. 263-4). The poor integration of these conspiracy prone subjects partly explains why, even when aware of the features, such subjects feel them to be less of a negative and, as they believe more such theories, are inclined to see the presence of the features as less and less of a negative.

The second two features compound the epistemic isolation while providing subjects who believe a theory with the apparent assurance that what they believe, and what they develop in response to evidence, is worthy of belief. If a theory seems capable of development to deal with errant data without being prone to being undermined by further evidence and features of the theory facilitate your own ability to develop it, then this allows you to close down interaction with other epistemic agents who do not share but challenge your beliefs. Developing a sequence of theories within an explanatory scheme, that is relatively unconstrained but with which one is familiar (the attribution of beliefs, desires and intentions to subjects), may be done independently of relevant expertise. The fact that it gives one the illusion of insight reinforces the relative isolation. While we should not ignore the way in which belief in a conspiracy theory may bring you closer to co-believers, and a reduction in isolation from them, it remains the case that a believer in a conspiracy theory with these features is epistemically isolated from everybody else while having the illusion that they do not suffer from epistemic isolation. The apparent explanatory success of their framework makes alternative approaches to the phenomena they are seeking to explain not real options for them.

A belief in one conspiracy theory of the relevant kind may indicate little more than an epistemic mistake and a correspondingly small amount of epistemic isolation. Empirical findings in social psychology are more plausibly related to those who *tend* to favour theories with these negative features. It has been found that subjects who are more inclined to anthropomorphic explanations in general are more prone to believe in conspiracy theories (Bruder et al. 2013, pp. 9–10).

Subjects who are more inclined to perceive patterns in randomness, look for patterns and meaning, or overestimate their capacity to understand complex causal phenomena, by appealing to explanations attributing agency, are also more prone to conspiracy beliefs (Vitriol and Marsh 2018). These tendencies naturally relate to the third and fourth features. Evidence is accommodated by a conspiracy without testing for further predictions due to overestimating the capacity to understand complex phenomena which is, in part, based on the illusion of understanding supplied by explanations involving rational agency. If a subject has a need for cognitive closure – for example, an answer to a salient and potentially puzzling phenomenon for which an official explanation is unavailable – then a conspiracy theory is attractive (Marchlewska, Cichocka and Kossowska 2017). As the last point emphasises, subjects drawn to the latter two features of the sub-category of conspiracy theories in question need not be drawn to theories with the first two elements but when a theory combines the features, the resulting tendency to believe these conspiracy theories in some subjects will be stronger, for example, those combining lack of social integration with a tendency towards anthropomorphic explanations.

A background assumption behind work on conspiracy theories and delusions is that, typically, subjects with delusions display greater failures of rationality. In fact, the situation is much more complex. In the case of conspiracy theories of the kind identified, there is often a clash of testimonies concerning what is going on to explain certain phenomena. On the one hand, there is that provided by official or socially accredited epistemic authorities. On the other, there is that provided by the proponents of conspiracy theories with the features I have identified. Even taking into account claims about epistemically polluted environments and relative social isolation, it is hard to see how those who are the source of conspiracy theories are established as more reliable or authoritative with regard to the theories they favour. So believing in conspiracy theories of this kind typically involves either a departure from what a subject has epistemic reason to believe (*epistemic irrationality*) or improper functioning of one process of belief formation – testimony – (*procedural irrationality*), probably both (see Noordhof 2024, p. 266). Obviously, irrationality of either kind is not invariable because overriding evidence may make belief in the conspiracy theory rational.

By contrast, those with monothematic delusions, at least, are usually seeking to make sense of anomalous experiences (for further discussion and references, see Noordhof and Sullivan Bissett 2021). It is true that there is testimony from friends, family and medical practitioners that clashes with what they take their experience to support – for example, that a loved one has been replaced by an imposter – and, it might be argued, this places their situation on a par with those who tend to believe conspiracy theories. This is to neglect the fact that it is quite standard for us to privilege experience above testimony, whatever epistemic credentials we give to the latter. However much we trust interlocuters when they describe what went on in a particular situation, we would have much more confidence if we could have seen it for ourselves. As a result, the epistemic or procedural irrationality of delusions is less obvious. The grip a subject's delusions have on them is due to the salience of their anomalous experiences, often concerning a central aspect of their identity or

personal relations, rather than because of any greater irrationality that subjects with delusions display.

Some delusions have a comparable social dimension. In the case of folie à deux delusions, while the primary subject has a delusory belief typically in response to anomalous experiences, the secondary subject comes to share the belief of the primary subject because they often stand in an intimate relationship to the primary subject (e.g. partner, child, sibling etc.), trust the individual's testimony and are, themselves more passive and/or in a felt position of inferiority (Bourgeois et al. 1992, pp. 709 – 11; Langdon 2013, p. 75; Rahman et al. 2013, pp. 1110–1111). This may lead them to interpret their own experiences as supporting the delusory belief in question or even have similar experiences (Kim et al. 2003, p. 463). On the face of it, there is less reason to doubt the rationality of secondary subjects in folie à deux situations, setting aside the question of how bizarre the delusory content may be, than those who believe conspiracy theories. For example, the first two features, we identified relating to conspiracy theories, indicate ways in which believers should not trust the source of the conspiracy theories. By contrast, the relationship of trust in folie à deux cases provides a basis for belief. The picture becomes more complex where we consider subjects who believe conspiracy theories with the features in question because they trust friends and family who didn't develop the conspiracy theory for themselves but who believe it. In this situation, it is relevant to consider the extent to which the subjects have considered the source of the beliefs of their friends and family.

Folie à deux cases are a highly distinctive and limited example of social delusions in general. The interest of these more general phenomena is, in part, how subjects arrive at quite strange beliefs (which may in turn distort their experience) when the same relations of trust and felt inferiority are not in place. One example is someone with anorexia who believes that they are fat, sees themselves as fat in the mirror, invests considerable importance in being thin, and, in fact, is thin (Radden 2011, pp. 87–90). It is plausible that in such cases there is no more substantial failure of epistemic or procedural irrationality but rather the difference is to be found in the disproportionate value placed upon being a certain way.

In some cases, the practical significance of believing a conspiracy theory of the requisite kind is low. This may mean that the confidence threshold for the belief is correspondingly lower. By contrast, a subject's delusory beliefs are often highly practically significant with mixed implications for the confidence threshold for belief. The confidence threshold may be lower as a precautionary measure if the delusory hypothesis involves a threat. On the other hand, the impact that a delusory hypothesis has on one's life is significant and, in that case, the confidence threshold may be high. These comments qualify the observation made above concerning the potentially greater irrationality involved in belief in conspiracy theories. Conversely, if a subject believes a range of conspiracy theories of the relevant type, the origin of the irrationality they display is plausibly attributed to the character dispositions we have described above. This is a second way in which the irrationality of believing in conspiracy theories may be deep seated, and thus more substantial, deriving from the presence of a systematic explanation of the beliefs in terms of character dispositions.

Other work has suggested that, contrary to what might be thought, self-deception often involves a more significant irrationality than monothematic delusions, at least, apart from those cases of monothematic delusion, such as patients with anosognosia, that have a self-deceptive element (Noordhof and Sullivan-Bissett 2023, pp. 96–106; Noordhof 2024, pp. 294–301; Ramachandran 1996, pp. 348, 355). It is plausible that some conspiracy theory beliefs are due to self-deception, for example, those in which individual or group narcissistic tendencies result in belief in conspiracies to explain why their situation is a difficult one. Beliefs in some conspiracy theories have also been argued to be a signal of group commitment and, thus the product of motivated irrationalities one sub-category of which plausibly involves self-deception (Mercier 2020, pp. 191-8; Williams 2022, pp. 1015-6). A key component of both types of cases is that the conspiracy theories in question will have one or more of the features I have identified. For example, in the case of group commitment, belief in a conspiracy theory is taken to be a sign of group commitment if there are social costs in believing such a conspiracy theory because the beliefs are seen to be absurd, with consequent reputational damage. The fact that you have committed yourself to their content makes it hard for you to defect from your group. The epistemic isolation I identified earlier is strategic.

The observations of the present section suggest that conspiracy beliefs may often involve more significant irrationality than monothematic delusions when the subject is aware of a theory's possession of one or more of the features I have identified. This is not to pathologise believers in such conspiracy theories but recognise the continuities, within the general population, of those who are the subject of clinical attention already – subjects with delusions – and the complexity of characterising what is involved with those who receive clinical attention from those that don't (for further discussion emphasising the non-pathological continuities, see Ichino and Sullivan-Bissett (forthcoming)).

5 Concluding Remarks

I have identified four features of conspiracy theories that would make them prima facie, or default, wrong to believe. My observations apply to the sub-categories of conspiracy theories that have one or more of these features. Some of the features are quite inclusive. The fourth, Lack of Explanatory Constraint, might in theory apply to any conspiracy theory with the key qualification that the lack of explanatory constraint can be mitigated when we are attributing a conspiracy to actors we know very well, there is a mass of observation and supporting testimony and/or the explanation is highly general indifferent to many different resolutions of detail. For most cases, lack of explanatory constraint points to prima facie wrongness of moving to full belief rather than some credence. I have discussed how these four bases for prima facie, or default, epistemic properties hold in epistemically polluted environments, are relatively independent of pragmatic considerations concerning appropriately levels of social vigilance regarding conspiracy theories, and how they relate to some

of the findings of social psychology, and the insight they provide concerning the relationship between those who have delusions and those who believe conspiracy theories.[1]

Funding I wrote this paper while being employed by the Department of Philosophy, University of York and in receipt of funding from AHRC *Deluded by Experience* Project (grant no. AH/T013486/1). Neither the author, the university or the AHRC have financial interests related to the content of the article as far as I can discern.

Declarations

Conflict of Interest I have no conflict of interest.

References

Basham, Lee. 2002. Living the conspiracy. *The Philosophical Forum* 32 (3): 265–280.

Basham, Lee. 2011. Joining the Conspiracy. *Argumenta* 3 (2): 271–290.

Basham, Lee. 2016. Between two generalisms: A reply to Stokes. *Social Epistemology Review and Reply Collective* 5 (12):4–12.

Boudry, Maarten. 2023. Why We Should Be Suspicious of Conspiracy Theories: A Novel Demarcation Problem. *Episteme* 20:611–631.

Boudry, Maarten, and M. Guilia Napolitano. 2023. Why we should stop talking about Generalism and Particularism: Moving the debate on Conspiracy theories Forward. *Social Epistemology Review and Reply Collective* 12 (9): 22–26.

Bourgeois, Marc L., Pascale Duhamel, and Helene Verdoux. 1992. Delusional parasitosis: Folie à Deux and attempted murder of a family doctor. *British Journal of Psychiatry* 161:709–711.

Bowes, Shauna M., Costello Thomas, and Arber Tasimi. 2023. The Conspiratorial Mind: A Meta-Analytic Review of Motivational and Personological Correlates. *Psychological Bulletin* 149 (5–6): 259–293.

Bruder, Martin, Peter Haffke, Nick Neave, Nina Nouripanah, and Roland Imhoff. 2013. Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy Mentality Questionnaire. *Frontiers in Psychology* 4:1–15.

Buenting, Joel, and Jason Taylor. 2010. Conspiracy theories and fortuitous data. *Philosophy of the Social Sciences* 40 (4): 567–578.

Burge, Tyler. 1993. Content preservation. *The Philosophical Review* 102 (4): 457–488.

[1] Many thanks to three referees, the audiences of two workshops which were part of the Conspiratorial Ideation and Pathological Belief Project, and Ema Sullivan-Bissett and Anna Ichino, for comments on the material in this paper. I would also like to recognise the funding support of the Deluded by Experience AHRC project (grant no. AH/T013486/10) for the writing of this paper.

Clarke, Stephen. 2002. Conspiracy Theories and Conspiracy Theorising. *Philosophy of the Social Sciences* 32 (2): 131–150.
Clarke, Stephen. 2007. Conspiracy Theories and the Internet: Controlled Demolition and Arrested Development. *Episteme* 4 (2): 167–180.
Clarke, Stephen. 2023. Two problems with the generalist-particularist distinction in the philosophy of Conspiracy Theory and why I am not a generalist. *Social Epistemology Review and Reply Collective* 12:54–60.
Coady, David. 2003. Conspiracy Theories and Official Stories. *International Journal of Applied Philosophy* 17 (2): 197–209.
Darley, John M., and C. Daniel Batson. 1973. From Jerusalem to Jericho: A study of situational and dispositional variables in helping Behaviour. *Journal of Personality and Social Psychology* 27 (1): 100–108.
Davidson, Donald. 1980. *Essays on actions and events*. Oxford: Oxford University Press.
Dennett, Daniel. 1987. *The intentional stance*. Cambridge, Massachusetts: The MIT Press.
Dentith, Matthew R. X. 2016. When inferring to a conspiracy might be the best explanation. *Social Epistemology* 30:5–6.
Dentith, Matthew R. X. 2017. The Problem of Conspiricism. *Argumenta*: 1–17.
Dentith, M. R. X. 2018. Expertise and conspiracy theories. *Social Epistemology* 32 (3): 196–208.
Dentith, M.R.X. 2021. Debunking Conspiracy Theories. *Synthese* 198:9897–9911.
Dentith, M.R.X. 2022. Suspicious conspiracy theories. *Synthese* 200 (243): 1–14.
Dentith, Matthew R.X., and Martin Orr. 2018. Secrecy and Conspiracy. *Episteme* 15 (4): 433–450.
Douglas, Karen M., Robbie M. Sutton, and Aleksandra Cichocka. 2017. The psychology of conspiracy theories. *Current Directions in Psychological Science* 26 (6): 538–542.
Goldman, Alvin I. 2001. *'Experts: Which ones should you trust?', Philosophy and Phenomenological Research, 63, reprinted in his (2002), Pathways to Knowledge*, 139–163. Oxford: Oxford University Press.
Hagen, Kurtis. 2022. Is Conspiracy Theorising Really Epistemically Problematic. *Episteme* 19:197–219.
Harris, Keith. 2018. What's epistemically wrong with Conspiracy Theorising. *Royal Institute of Philosophy, Supplementary* 84:235–257.
Harris, Keith Raymond. 2022. Some problems with particularism. *Synthese* 200 (447): 1–25.
Harris, Keith. 2023. Conspiracy theories, populism and epistemic autonomy. *Journal of the American Philosophical Association* 9 (1): 21–36.
Ichino, Anna, and Ema Sullivan-Bissett. forthcoming. Conspiracy beliefs and Monothematic delusions: Case for De-pathologizing. *Erkenntnis.*
Keeley, Brian L. 1999. Of conspiracy theories. *Journal of Philosophy* 93:109–126.
Keeley, Brian. 2019. The credulity of conspiracy theorists. In *Conspiracy theories and the people who believe them*, ed. Joseph E. Uscinski, 422–431. Oxford: Oxford University Press.
Kim, Chuleung, Jinmi Kim, Mounghoon Lee, and Minhee Kang. 2003. Delusional parasitosis as 'Folie à Deux'. *Journal of Korean Medical Science* 18:462–465.
Lakatos, Imre. 1978. *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.
Langdon, Robyn. 2013. Folie à deux and its lessons for two factor theorists. *Mind and Language* 28 (1): 72–82.
Levy, Neil. 2022. *Bad beliefs*. Oxford: Oxford University Press).
Marchlewska, Marta, Aleksandra Cichocka Małgorzata, and Kossowska. 2017. Addicted to answers: Need for cognitive closure and the endorsement of conspiracy beliefs. *European Journal of Social Psychology* 48:109–117.
Mercier, Hugo. 2020. *Not born yesterday*. Princeton: Princeton University Press.
Napolitano, M. Giulia. 2021. Conspiracy theories and evidential self-insulation. In *The epistemology of fake news*, ed. Sven Bernecker, Amy K. Flowerree, and Thomas Grundmann, 82–105. Oxford: Oxford University Press.
Napolitano, M. Giulia, and Kevin Reuter. 2023. What is a Conspiracy Theory? *Erkenntnis* 88:2035–2062.
Noordhof, Paul. 1999. Causation by Content? *Mind and Language* 14 (3): 291–320.
Noordhof, Paul. 2006. Environment-Dependent Content and the Virtues of Causal Explanation. *Synthese* 149 (3): 551–575.
Noordhof, Paul. 2020. *A Variety of causes*. Oxford: Oxford University Press.
Noordhof, Paul. 2024. Irrationality and the Failures of Consciousness. In *Belief, Imagination and Delusion*, ed. Ema Sullivan-Bissett. Oxford: Oxford University Press.

Noordhof, Paul, and Ema Sullivan-Bissett. 2021. The clinical significance of anomalous experience in the explanation of monothematic delusion. *Synthese* 199:10277–10309.

Noordhof, Paul, and Ema Sullivan-Bissett. 2023. The Everyday Irrationality of Monothematic Delusion. In *Experimental Advances in Philosophy of Action*, ed. S. Murray and P. Henne, 87–112. London, Bloomsbury.

Pigden, Charles. 1995. Popper Revisited, or what is wrong with conspiracy theories? *Philosophy of Social Sciences* 25 (1): 3–34.

Pigden, Charles. 2016. Conspiracy theories, Deplorables, and Defectability. In *Taking conspiracy theories seriously*, ed. M.R.X. Dentith, 203–215. Lanham: Rowman and Littlefield International.

Poth, Nina, and Krzysztof Dolega. 2023. Bayesian belief protection: A study of belief in conspiracy theories. *Philosophical Psychology*: 1–26.

Radden, Jennifer. 2011. *On delusion*. London and New York: Routledge.

Rahman, Tahir, et al. 2013. Infanticide in a case of Folie à Deux. *American Journal of Psychiatry* 170 (10): 1110–1112.

Ramachandran, V. S. 1996. The Evolutionary Biology of Self-Deception, laughter, dreaming and Depression: Some cues from Anosognosia. *Medical Hypotheses* 47 (5): 347–362.

Ross, W. D. 1930. *The right and the good*. Oxford: Oxford University Press.

Sober, Elliott. 2002. Intelligent design and probability reasoning. *International Journal for the Philosophy of Religion* 52:65–80.

Vitriol, Joseph A., and Jessecae K. Marsh. 2018. The illusion of explanatory depth and endorsement of conspiracy beliefs. *European Journal of Social Psychology* 48:955–969.

Williams, Daniel. 2022. Signalling, commitment and strategic absurdities. *Mind and Language* 37:1011–1029.

Williams, Daniel. 2023. Bad beliefs: Why they happen to highly intelligent, vigilant, devious, self-deceiving, coalitional apes. *Philosophical Psychology* 36 (4): 819–833.