

# Against radical epistemic environmentalism (or, why uncritically deferring to authority is still irrational)

Robert Mark Simpson and Toby Handfield

Forthcoming in *Episteme*

*Abstract.* Neil Levy's book *Bad Beliefs* defends a prima facie attractive approach to social epistemic policy – namely, an environmental approach, which prioritises the curation of a truth-conducive information environment above the inculcation of individual critical thinking abilities and epistemic virtues. However, Levy's defence of this approach is grounded in a surprising and provocative claim about the rationality of deference. His claim is that it's rational for people to unquestioningly defer to putative authorities, because these authorities hold expert status. As friends of the environmental approach, we try to show why it will be better for that approach to not be argumentatively grounded in this revisionist claim about when and why deference is rational. We identify both theoretical and practical problems that this claim gives rise to.

## 1. Introduction

Why do some groups hold beliefs that mostly align with scientific evidence and expert opinion, while others don't? Why do whole communities buy into young earth creationism, or absurd conspiracies, while living alongside others who reject such patent falsehoods?

One type of answer here adverts to differences among the people in different groups. It says young-earthers are stupid, or dogmatic, or crazy. This answer favours an *individualistic* approach to social epistemic policy. It favours policies that try to make people more rationally responsive to evidence – and less stupid, dogmatic, etc. But this approach downplays how similar people's belief-forming processes are across different commun-

ities. As much as some of us try hard to think for ourselves, still, a lot of us, a lot of the time, form our beliefs in a more socially-scaffolded way. We align our beliefs with what's accepted by respected sources in our own communities. Granted, if the respected sources in your community form their beliefs in a way that's well-tuned to the evidence, it seems better to defer to them, than if you're in a community whose respected sources form their beliefs in a more erratic fashion. The point, though, is that most of us form our beliefs via social deference, in a way that's inattentive to these differences. If we are assessing how rational individuals are, in how they go about forming their beliefs, most scientifically literate people are going to receive a similar assessment to most young-earththers, because most of them are, like the young-earththers, uncritically accepting the views of the people that their in-group recognises as authorities.<sup>1</sup>

An alternative approach to social epistemic policy focuses on building a truth-conducive informational environment, e.g. by ensuring that school syllabi, media, and other key nodes in our epistemic networks are disseminating material that's credible and well-evidenced. The basic idea here is that because there's little individual variation in epistemic abilities between people in different groups – and little to be gained by trying to improve people's epistemic abilities – there is a greater potential for veritistic gains by focusing on environmental variables.<sup>2</sup>

This *moderate environmental approach* to social epistemic policy has a lot going for it. The approach is environmental in that it tells us to prioritise setting up a truth-conducive environment. But it's moderate, in that it doesn't totally rule out the possibility of beneficial interventions targeted at individuals. This approach concentrates on environmental interventions, simply because, in a veritistically bad environment, promoting critical thinking etc. tends to be less efficient in realising our epistemic goals. Environmental 'clean-up' has more impact. At any rate, there are very good reasons to think so, given what we know about how powerful people's cognitive biases are, and how difficult it is for people to evaluate the credibility of different sources that they're exposed to.

In this paper we criticise a more radical environmental approach to social epistemic policy, espoused by Neil Levy in *Bad Beliefs* (2022). Levy's approach is radical in that it is

---

<sup>1</sup> Our point is similar to a key idea in Nguyen's (2020) influential account of echo chambers. An Echo-chambered person, A, treats the fact that B disagrees with his views as grounds for discounting B's credibility, and this can bootstrap A into unshakeable dogmatism. But it's not that different to how ordinary, seemingly rational people behave. For example, if B believes in lizard people conspiracies, I may (plausibly, rationally) take that as a reason to discount B's credibility. There's also a link here to Kripke's (2011) paradox of dogmatism. Suppose I accept p's truth; p's truth entails that evidence against p is misleading, which seems (paradoxically) to give me a reason to ignore such evidence. In short, a tendency to rely upon sources that affirm one's current views is both commonplace and – if one's sources are reliable – rationally permissible, even though this tendency is hard to distinguish, internally, from forms of dogmatism that seem obviously rationally impermissible.

<sup>2</sup> *Veritistic gains* is just another way of saying "gains with respect to the acquisition of true beliefs". An emphasis on building a truth-conducive environment is central to defences of epistemic paternalism, e.g. from Goldman (1991) and Ahlstrom-Vij (2013). These authors argue for policies that e.g. withhold evidence that is likely to be widely misinterpreted, and they cite rules of evidence in legal trials to illustrate (i) why such information control is likely to reduce false belief, and (ii) why its paternalistic character doesn't make it illegitimate.

partly based on a radical view about the rationality of uncritical deference. On Levy's view, people who defer to in-group 'authorities' are in general believing rationally. And this is still the case, he argues, if this uncritical deference leads people to hold beliefs that flout scientific evidence. Contrary to what we ordinarily suppose, it isn't irrational to uncritically assent to what your community's favourite preacher or podcaster says. Indeed, for Levy, attempting to critically assess their views is actually an irrational response.

This claim about rational deference can be used to help defend a particularly radical environmentalist approach to social epistemic policy. If young-earthers and their ilk are in fact being rational, then there is nothing to be gained, vis-à-vis the promotion of our epistemic goals, by urging them towards more rational doxastic practices. They are *already* believing rationally. If we want people to eschew beliefs that flout the evidence, then the only kind of social epistemic policy that it makes sense to pursue is one that seeks to depollute the information environment.

The reason we want to criticise Levy's account isn't because we reject an environmental approach to social epistemic policy, but because we endorse it. Our worry is that Levy's radical ideas about deference end up discrediting an environmental approach. How? First, by saddling it to independently unattractive claims about the nature of rationality. Second, by subverting some of the judgments that are involved in setting up the sort of truth-conducive environment that Levy himself endorses.

So our aim here isn't simply to critique uncritical deference. We're challenging the idea that epistemic environmentalism and radical deference norms come as a package deal. The environmental approach has considerable appeal. It recognises that individual epistemic conduct is shaped by structural factors, and prioritises interventions that target those factors. Nothing about this approach requires us to endorse Levy's radical claims about the rationality of uncritical deference. There are no entailment relations between the claim that "it's rational to uncritically defer to in-group authorities" and the claim that "providing true information is more effective than promoting critical thinking skills for achieving collective epistemic aims." Nor is there any other compelling reason why these commitments have to stand or fall together. If our goal is promoting true beliefs, we should favour both information policies and belief-formation norms that conduce to that aim. Levy's error lies in thinking that his uncritical deference norm better serves this aim than alternative, more nuanced – perhaps more boring – approaches to navigating between deference and discrimination. Disentangling these elements of Levy's view helps clarify the key attractions of epistemic environmentalism.

The plan for the paper is as follows.

- In §2 we continue the stage-setting. We explain Levy's notion of Bad Beliefs, we say a bit more to characterise and motivate the environmental approach, and we explain Levy's argument about why Bad Beliefs formed via social deference are rational.
- In §3 we present our first objection – that Levy's account saddles the environmental approach with a view of rational deference that's independently unattractive. We start by explaining just how easy Levy makes it to qualify as rational when you defer to a putative in-group authority. We then consider how Levy distinguishes rational

and irrational bases of deference, and argue that his way of drawing this distinction doesn't rectify the initial worries about 'easy rationality' for in-group deference.

- In §4 we present our second objection – that Levy's radical account of rational deference discredits some of the critical thinking necessary to create the truth-conducive environment that the environmental approach calls for. To create that environment we need a significant number of people to adopt precisely the critical, questioning, *un*-deferential attitudes towards authority – the very attitudes whose rationality Levy is denying.

## 2. Stage-setting

### 2.1 *Bad Beliefs and the Structural Perspective*

*Bad Beliefs*, as Levy defines them (2022: xi), are beliefs which (i) conflict with what domain-relevant authorities believe, and (ii) flout the publicly-available evidence.<sup>3</sup> Paradigmatic examples include the belief that climate change isn't a real phenomenon, the belief that Barack Obama isn't a US citizen, or the belief that the earth is only a few thousand years old. There's plenty of evidence for the falsity of these views, and relevant experts agree they are false. Yet many people hold these beliefs nevertheless. Those are Bad Beliefs.

It seems natural to view Bad Beliefs as straightforwardly irrational. If there's a weight of credible evidence that's been made available via the media, via online sources like Wikipedia, or in academia, then we presumably have some rational duty to align our beliefs with what that evidence shows, especially when relevant experts digest and explain that evidence for us, offering clear advice about what to believe.<sup>4</sup> Believing Obama was born in Africa, or that climate change is a conspiratorial hoax, in the face of all that, seems like a paradigm case of irrationality.

Levy challenges this way of thinking about Bad Beliefs by challenging the individualistic framing that it presupposes. For Levy, Bad Beliefs are a structural problem – they can't be effectively understood or addressed by focusing on individuals' doxastic practices.

To analogise: the reason why more people are obese nowadays, in many countries, isn't that individuals have become less 'dietarily virtuous'. It's that economic, technological, and political factors are producing an obesogenic environment. If we want to reduce obesity, we have to rebuild that environment, e.g. by limiting the easy availability of cheap,

---

<sup>3</sup> From this point references to Levy (2022) will be cited in the main text via page numbers in square brackets.

<sup>4</sup> This sort of thinking is famously encapsulated in W. K. Clifford's (1877) precept, that it's always wrong to believe anything on the basis of insufficient evidence. It's also evident in more recent defences of the idea that there's a general epistemic duty to seek evidence relevant to one's beliefs (e.g. Hall and Johnson 1998).

addictive, unhealthy foods, or increasing the availability (and/or affordability, and/or appeal) of healthier foods. It's not illuminating to fixate on individual choices downstream from more potent structural factors.<sup>5</sup>

Similarly, to reduce the frequency of beliefs that flout evidence and expert opinion, we need to rebuild our information environments, so that people aren't inundated with misinformation and have reliable default information sources. Just as it's largely futile to urge people in an obesogenic environment to be more dietarily virtuous, it's also largely futile, in a veritistically bad environment, to urge people to be more epistemically virtuous.

The crux of the environmental approach isn't to deny the value of basic education. It's to resist the idea that non-truth-conducive environments can be effectively counteracted by promoting individual reasoning skills like critical thinking or digital literacy. Levy associates such individualistic approaches with virtue epistemology, and he treats Quassim Cassam (2018) as his foil and exemplar of this approach. But the approach Levy is criticising isn't just an epistemologist's contrivance. It shows up regularly in political theory and public policy work on content moderation and online governance. It's commonplace, in that literature, to see people talking about the new norms of online discourse that people must adopt, in order to navigate around echo chambers, misinformation, and related epistemic hazards that seem to become heightened online (e.g. Chambers 2021; Cohen and Fung 2021). The environmental approach regards such interventions as mere band-aids that treat surface-level symptoms while ignoring root causes.

## 2.2 *The Rationality of Bad Beliefs*

Bad Beliefs may be a structural issue, calling for a structural solution. But isn't uncritically believing a podcaster or preacher, against expert opinion and credible evidence, still irrational? Even if obesity is produced by obesogenic environments, eating lots of sugary, ultra-processed food is still unhealthy. And even if Bad Beliefs are produced by polluted environments, accepting misinformation still seems irrational. After all, isn't it just paradigmatically irrational to believe things that the relevant evidence disconfirms?

Yes and no. Bad Beliefs come from trusting putative authorities who misjudge the evidence due to corruption, bias, or delusion. But even so, Levy argues, deferring to them is

---

<sup>5</sup> Blake-Turner (2020) develops a related but distinct environmental approach to epistemic pollution, arguing for a kind of 'strict liability' regarding epistemic blame. On their view, we should hold people accountable for spreading false information even when they do so faultlessly – that is, even when their belief-forming practices weren't obviously irrational or negligent. While this differs from Levy's view (which sees many bad beliefs as rational), both approaches highlight how individual-level evaluations of belief formation might need to be subordinated to broader concerns about maintaining a healthy epistemic environment.

part of a practice that, when iterated at scale, conduces to collective epistemic success. We most effectively acquire epistemic goods via simple deference, or “outsourcing”.<sup>6</sup>

Levy’s argument for this draws on evidence of how humanity’s epistemic success is a consequence of *cumulative culture*. In contrast to the cultures of other intelligent animals that can solve complex problems and engage in sophisticated communication,

Only [human beings’] cumulative culture builds on these innovations, enabling cognitive achievements that go beyond what any individual or any generation can achieve... Cumulative culture opens up horizons for knowledge that are closed to individuals, no matter how individually gifted they are. [40]

The most striking empirical findings, in this connection, and those most congenial to Levy’s account, find that humans are more disposed than other primates to engage in apparently pointless imitative behaviour.<sup>7</sup> This supports our intellectual success by making it easier for us to intergenerationally transmit knowledge, even in cases where the recipients can’t yet grasp why the transmitters believe what they’re transmitting. If humans are so smart, Levy asks, then how can other primates “outperform us in identifying efficiencies and more successful routes to a goal?” [44]. The answer is: because our culture is *extra*-imitative.

Imitation is an adaptation for culture. It allows us to acquire knowledge and practices developed by... individuals dispersed across space and time. It allows us to acquire, and then to build on, deeply social knowledge: adaptive behavior that could not have been developed by any individual *de novo*. [44]<sup>8</sup>

---

<sup>6</sup> While Levy’s argument is grounded in findings from cognitive science, his point is one that’s been a familiar touchstone in contemporary social epistemology, in the wake of Hardwig’s (1985) critique of epistemic autonomy. For Hardwig, any conception of rationality that makes strong self-reliance a condition of rationality is self-defeating, because it makes it rational for people to radically deplete the body of evidence on which they’re basing their beliefs, and opting for this depletion itself seems intuitively irrational.

<sup>7</sup> Here’s Levy’s description of a key finding in this regard: “human beings are disposed to copy even those components of behavior that don’t appear to be required for goal pursuit. Nagell, Olguin, & Tomasello [1993]... demonstrated a novel technique to human children and chimps. They used a rake, *time side down*, to draw sweets that were otherwise out of reach toward themselves. Using a rake that way is very inefficient: many sweets slip through the gaps in the tines. Given the opportunity to perform the task themselves, chimps flipped the rake so that the flat side acted as a more efficient tool, with fewer sweets escaping. But human children tended to imitate the action just as demonstrated” [43].

<sup>8</sup> Among the sources cited by Levy in support of this kind of picture, Joseph Henrich’s book *The Secret of Our Success* (2015) is one particularly prominent and influential account. Some more recent research moderates the apparent significance of overimitation in the acquisition of practically useful knowledge. In a series of studies, Harvey Whitehouse and colleagues find that overimitation is much more pronounced for activities that have no discernible practical goal. Whitehouse suggests the phenomenon of overimitation is a feature of a “ritual stance” in which we try to learn behavior patterns for purposes of conformity to local norms and identities, and that it is less prominent when we adopt an “instrumental stance”, where we are more open to experimentation and omitting apparently irrelevant elements (Whitehouse 2021, chapter 1).

Solo inquiry yields less than when we instead defer to the prevalent (often inherited) judgements of others around us. On Levy's view, this kind of outsourcing is entirely rational: "being open to cues for what others believe is being open to *reasons*" [77].<sup>9</sup>

So why isn't believing against evidence paradigmatically irrational? Because generally, relatively uncritical deference conduces to collective epistemic successes by enabling cumulative culture. We might prefer to call this *rational irrationality* [153], or *ecological rationality* [142–43] – locally irrational deference that yields success. But Levy provocatively suggests that uncritical deference is rational all the way down, denying any sharp distinction between ecological rationality (what conduces to collective success) and individual rationality (what's locally responsive to evidence) [73]. A climate change denier, who uncritically defers to a podcaster or politician who is viewed as an authority within that person (the denier's) community, is rational. Their belief is bad, in flouting evidence, but rational, in being duly responsive to the social evidence, enabling cumulative culture's long run epistemic success.<sup>10</sup>

### 3. Against Uncritical Deference

Levy argues – forcefully, creatively – for an environmental approach to social epistemic policy. By our lights, though, the environmental approach is undermined by being saddled to Levy's account of rational deference. The *prima facie* worrying issue with his account is that it makes it too easy to count as rational in deferring to others. The deeper issue is that Levy's revisionist conception of rationality – which collapses the conceptual gap between what's rational for the individual to believe, and what conduces to a group's collective epistemic success – doesn't in fact vindicate his permissive view of rational deference. It doesn't support the controversial claims that are meant to differentiate Levy's social epistemic policy prescriptions, from the prescriptions that are favoured by more individualistically-minded epistemologists.

---

<sup>9</sup> As this remark indicates, there's an internal tension in the way that Levy characterises his central theoretical claim. Alex Worsnip (2022) highlights this in his review of *Bad Beliefs*. Levy defines Bad Beliefs as beliefs held despite the widespread public availability of... evidence that supports more accurate beliefs [xi]. But he goes on to insist that Bad Beliefs are (often) rational, since their beliefs are based on the endorsement of putative authorities, which is *itself* evidence. The tension is that Levy seems to want to say that Bad Beliefs both *are* and *are not* supported by the evidence that's available to the believer. Like Worsnip, though, we don't want to make this tension a key plank in our critique. There's presumably some way for Levy to finesse his claims to allay this worry, without any real loss to the account's plausibility or significance.

<sup>10</sup> Woodard (2024) highlights one counterintuitive implication of this sort of picture. Suppose A is motivated to defer to B because A wants B to love them (and they think their deference will help), and suppose B is extremely reliable, so that A's deference to B results in A forming lots of true beliefs. Intuitively, A's beliefs aren't rational, because A's motivation in adopting them is in some important sense alethically indifferent. Plausibly, a condition of a belief-forming process's being rational, is that its enactment is alethically oriented: the agent forming their beliefs via this method does so because they think it will lead to true beliefs. This seems like a problem for Levy, but our critique in what follows will be granting that part of Levy's account for the sake of argument.

Levy thinks it's rational for Christians to believe in a young earth, for instance, because such belief results from relatively uncritical deference to local authorities. He sees this as rational because it conduces to our species' collective epistemic success. But in fact uncritical deference ought to be adjudged irrational by Levy's own standards. Why? Because *indiscriminate* deference tends to set back a group's collective epistemic success, and there's no way to differentiate *relatively uncritical* deference from indiscriminate deference – not without appealing to the kind of individualistic norms of critical thinking whose rationality Levy wants to deny.

### 3.1 *Uncritical Deference*

So, the basic worry: Levy's account makes it too easy to count as rational in deferring to in-group authorities. He thinks the young-earthier isn't holding this belief in spite of evidence against it, but that she's holding it *on the basis of* supporting evidence – the evidence constituted by the fact that it's affirmed by her in-group's local authorities.

As a starting point, notice how this places a lot of justificatory weight on one source of evidence, and very little weight on other sources. To paraphrase Alex Worsnip (2022),

To get the result that young-earthiers are believing rationally, Levy needs to explain why young-earthiers are rational in giving so much weight to the testimony of other young-earthiers. Young-earthiers are aware that lots of people aren't young-earthiers – including many people who are part of their “community”, in some sense. So Levy's claim has to be not just that we're rational in deferring to the community, but that we're rational in deferring to our co-partisans, even when we're well-aware that many people outside this immediate community disagree.

The worry here should be familiar. Groupthink seems clearly epistemically bad, and Levy's account of rational deference seems to license it. Even if Levy is correct that outsourcing is beneficial (when iterated at scale), it doesn't follow that *chauvinistic* or *parochial* outsourcing is beneficial. Indeed, some of his own examples, whose point is to show how outsourcing is rational, also serve as illustrations of why chauvinistic outsourcing isn't rational. Consider his example of the 19<sup>th</sup> century colonial explorers who dismiss life-saving dietary knowledge offered by indigenous people [37]. The lesson Levy draws is that we do better, epistemically, by outsourcing, in the way these explorers didn't. But the explorers *were* outsourcers, in general – they wouldn't have been able to embark on their endeavour in the first place, if they hadn't relied on all manner of testimonial knowledge about, for example, navigation. They missed out on vital dietary knowledge not because they failed to outsource generally, but because they refused to outsource to out-group members – hence their plight does not seem very different from that of young earthiers.

Outsourcing is rational, for Levy, because it enables the accumulation of knowledge via cumulative culture. Constantly trying to critically audit the reports or opinions of apparently authoritative people in your in-group isn't a fruitful approach to belief-formation, when it's iterated at scale. We should endorse the rationality of *relatively* uncritical



deference, Levy argues, *insofar as* this belief-forming heuristic conduces to the epistemic benefits of cumulative culture.

But the devil is in the detail of the italicised caveats – *relatively*, and *insofar as*. Outsourcing as practiced by the chauvinistic colonial explorers cut them off from knowledge. In order for Levy’s account of rational deference to work, it needs a criterion that can be used to identify *overly* uncritical, chauvinistic or parochial, forms of deference – the forms that are tantamount to pernicious groupthink, and which inhibit collective epistemic success. The criterion has to differentiate these counterproductive forms of deference from the preferable, only *relatively* uncritical forms of deference – the ones that conduce to collective epistemic success. And this criterion can’t just recapitulate the ideals of critical thinking and epistemic virtue in opposition to which Levy is pitching his account.

### 3.2 Modesty and Discrimination

Levy has a story to tell here, but we don’t think this story can do the work that’s needed. (We will focus on exposition in this sub-section, and then return to criticism in §3.3.)

Solo inquiry isn’t an effective practice – long-term, at a community level – for realising epistemic goods. So we should outsource. But there’s still good and bad outsourcing, by Levy’s own lights. Rational deference requires us to engage in processes of *discrimination*: seeking to differentiate the deference-worthy authorities from non-authorities. Thus Levy’s account has to answer the question of how novices are to assess expertise – building on recent epistemological work on this issue, whether of a bent which is more optimistic (e.g. Goldman 2001, Anderson 2011), or more pessimistic (e.g. Millgram 2015, Nguyen 2020).

Levy’s proposal here calls upon a familiar, attractive norm: we should respect our intellectual limitations. We aren’t being rational if we’re trying to carry out intellectual tasks which outstrip those limitations. Let’s call this norm, as loosely-defined, *modesty*.<sup>11</sup>

So, which grounds for adjudicating expertise are and aren’t rational? The modesty norm says we have to respect our intellectual limits. Some discriminations are easy to get right; others are far harder. Schematically, modesty tells us it’s rational to make the easy discriminations, but, generally, irrational to attempt the harder ones. The easy discriminations are the ones where we’re just looking for *markers* of expertise. The idea isn’t that it’s easy to judge who the real experts are in a given domain, i.e. who has domain-relevant knowledge and complementary intellectual abilities. Indeed, those are the immodest

---

<sup>11</sup> Something like this ideal is central to Fantl’s (2018) critique of open-mindedness. If you meet someone making apparently cogent arguments for an outré idea, you may think the epistemically virtuous response is to consider their arguments open-mindedly. But why presume that you would be able to identify errors in their arguments, if errors were there? For Fantl, what seems like a virtuously open-minded approach to the social aspects of belief-formation is actually a risky, hubristic approach, where you make yourself more susceptible to being argued into believing falsehoods, because you overestimate your ability to identify unsound or badly-evidenced arguments when you encounter them.

discriminations that we're meant to avoid, which Levy's account sees as irrational. The idea, rather, very simply, is that it's usually easy to judge who the *putative* experts are, in a group or institution that you're a part of. Consider university students. They don't find it hard to work out which people are the professors, and which people are administrators or security staff. The university's institutional format makes these discriminations easy. And so, apart from rare mistakes – which are easily correctible – students reliably identify whom the experts are that they're meant to be learning from.

The hard discriminations, by contrast, are those where one is trying to judge *manifestations* of expertise, as per the conundrum Goldman (2001) sets out in his Novice/2-Expert problem. Suppose Lia takes a course in evolutionary biology, where she's told by Professor Dawkins that the central tenets of evolutionary biology are basically incorrigible, and that this is why they're accepted by every expert who understands the relevant data. The evidence vindicates those tenets. Suppose Lia then takes a course in philosophy of science, where Professor Feyerabend tells her that biologists accept evolutionary theory not (or not only) because the evidence supports it, but as a result of social processes that are strikingly similar to processes of cultural habituation or indoctrination.

How should Lia discriminate? It's easy to identify markers of expertise, but both Professors have those markers. What she needs to do is to assess their manifestations of expertise. And that seems nearly impossible. Unless Lia has a way to review their track records of predictive success, or identify biasing factors that undermine the reliability of one of them – unless she can do these things armed only with her novice's understanding of the data and methods that define their respective fields – she seems to have no good way to discriminate.<sup>12</sup> Maybe if a lot goes well, over many years, Lia can eventually acquire enough expertise to rationally judge whether Dawkins or Feyerabend is doing better at understanding and interpreting the evidence that's pertinent to their dispute. But for most of us, most of the time, the attempt to make such discriminations will, in terms of its rational pedigree, be tantamount to making a random stab in the dark.

This part of Levy's thinking is summed up in his commentary on Cassam's views about how epistemic virtues can help people resist misinformation and conspiracies. Levy agrees with Cassam that we may “deploy the virtues to choose between competing experts” [103]. But Levy puts a qualifying asterisk next to this acknowledgement. Insofar as Cassam's prescription calls for individuals to “directly and virtuously adjudicat[e] the second order evidence”, i.e. the evidence about which expert is actually more reliable, “it's still too demanding and too individualistic.” Thus, Levy thinks, “we face the same risks of losing knowledge by engaging at this level,” and hence, “dogmatism remains a better strategy” [103].

---

<sup>12</sup> Goldman (2001), Anderson (2011), and Nguyen (2020) are all cautiously optimistic about the novice's ability to make such judgements. Looking to track records or indirect evidence of bias are two of Goldman's suggested strategies.

But surely this is incredibly risky, isn't it? To say that *dogmatism* is the better strategy – better in the sense of its mitigating our risk of losing knowledge? Levy holds firm. Such outsourcing “remains a better strategy”, where the strategy at issue, specifically, involves “relatively unquestioning deference to authoritative sources *because they're authoritative*, and not because we've assessed their degree of expertise ourselves” [103, our emphasis].

### 3.3 *The Unquestioning Stance*

The novice can judge markers of expertise, but they don't have the expertise to directly judge manifestations of expertise. This is why relatively unquestioning deference to authoritative sources is the rational approach to outsourcing, on Levy's account. To attempt to defer only in a critical, discriminating fashion, is to infringe against the norm of modesty.

The problem we were pressing in §3.1 remains, though. If the putative authorities disagree, and if we defer to our preferred authorities uncritically, we are undermining the epistemic benefits of collective culture, just as surely as if we immodestly attempted to judge the experts' manifestations of expertise. Under both approaches, we are deciding whom to defer to in a way that's tantamount to reckless guessing. If what Levy's approach to discrimination involves, in the end, is that you can ascribe authority to whichever view or idea happens to chime for you, then discriminating deference ends up being no better – with respect to our collective epistemic goals – than hubristically trying to figure everything out yourself.<sup>13</sup> And Levy's account seemingly isn't giving us the resources to do better than this.

To illustrate the worry, consider a follower of a doomsday cult, who has been told by their leader that the world will end in some sort of spiritual-celestial cataclysm, on a specific date. Suppose the leader is clever, knowledgeable, and charismatic – they have a mixture of social traits that mark them out as an authority within their community. And suppose that lots of smart people (teachers, lawyers, financial advisors) are deferring to the leader.

Now suppose the foretold world-ending date goes by, and that the cult leader has concocted an after-the-fact rationalization that explains why the doomsday event didn't happen on that date, and why it's actually going to occur in twelve months' time. In order to give the example a 'two rival experts' format, let's suppose that our follower – who is now vaguely worried that they're being deceived – goes to ask another apparently credible person whether *they* think the world is going to end on this newly appointed date, in a

---

<sup>13</sup> Granted, A's uncritical deference to his co-partisan, A, may not be such a bad strategy if A is deferring to B on some moral question, and if a shared moral outlook is what underpins A's and B's co-partisan status. Something like this thought is evident in defences of co-partisan deference by Rini (2017) and Lepoutre (2020). But this doesn't really help Levy's account, given that he sees uncritical deference to in-group authorities as rational on all sorts of topics.

year's time. And suppose this other putative expert says precisely what we would expect: "no, the world isn't going to end then, you've fallen into some sort of cult."

Recall that Levy's account of what it's rational to do, in such situations, is to evince relatively unquestioning deference to authorities, because they're authoritative – not because you've assessed their intellectual credibility directly. But of course there are two putative authorities in the frame of reference here, so the fact that one is an authority cannot settle the question of which one should be (relatively unquestioningly) deferred to.

Intuitively, it seems clear what this person ought rationally to do. They should see that someone who rearranges their worldview impromptu, and produces *post hoc* rationalizations for why they're doing so, is probably a liar or bullshitter – at any rate, not the reliable expert they're making themselves out to be – and accordingly, stop deferring to that person.

If the follower reasons like this they will be adjudicating expertise in a manner that doesn't neatly fit into the markers vs. manifestations schema. The cult leader may present a bunch of complicated math or astronomy, to try to explain why their prediction of the apocalyptic date shifted. The follower may have no ability to assess that putative evidence. The follower's growing mistrust of the leader isn't because they have spotted granular faults in the leader's reasoning. But equally, the leader may have all of the social markers of expertise that they had previously. So the follower's growing mistrust isn't stemming from a shift in the leader's social markers of expertise. The follower is just perceiving that things don't add up. The best explanation of the pattern of reasoning and testimony that they're observing from the leader isn't that the leader is a real expert who's telling the truth. The best explanation is that they're a smart person who has, for whatever reason, become a zealot on behalf of a belief-system that doesn't line up with reality.<sup>14</sup>

People make these kinds of discriminations fairly often. They obviously aren't foolproof, and it's possible that some more individualistically-minded virtue epistemologists overestimate the median person's ability to step into this kind of gestalt critical vantage point, and to make sound judgements upon doing so. The modesty norm might be a sensible corrective to that overestimation. Nevertheless, people seem to make these discriminations, successfully, with some regularity. More to the point, for Levy's purposes, the execution of these kinds of discriminating judgements seems to be conducive to the realization of cumulative culture's epistemic benefits. True, we reap some collective epistemic rewards from a tendency towards unquestioning deference. But those rewards would be jeopardised if this imitative tendency were not partly counterbalanced by some preparedness to question and demur.

---

<sup>14</sup> Nguyen (2020) defends something like this idea – that the individual's epistemic autonomy might be expressed not so much in trying to assess people's direct manifestations of expertise, or their indirect markers of expertise, but that it might instead involve some kind of broader gestalt interpretation of how various bodies of knowledge, rival perspectives, and putative sources of expertise all hang together.

To be fair, Levy never endorsed a stance of purely unquestioning deference to authority. He endorsed *relatively* unquestioning deference. So then the issue is: what does that caveat permit? How close does it bring Levy's account to the individualistically-minded epistemologists that he's criticising – the ones who think people should use critical thinking to adjudicate between experts? At a more concessive moment in his argument, Levy says

If virtue epistemology can help... it's not by substituting for apt deference to others and socially distributed cognition; instead, it's by playing a (small) role in helping us to do these things better. Virtue epistemologists... appear to aim to bring us each to inculcate the virtues in ourselves and then... to tackle hard problems largely on our own. [91]

Insofar as virtue epistemologists are telling us to eschew collaborative learning, and to figure out everything ourselves, they are liable to criticism. But as Levy's discussion shows, the real locus of disagreement between him and his opponents isn't whether to outsource, but when and how to outsource. The clear point of difference between them, besides subtleties of emphasis (Levy believes the role of individual virtue is small; his opponents presumably see its role as larger), is that Levy wants to say – and his opponents generally want to deny – that it's rational to defer to putative authorities *because* they are putative authorities.

Levy's stance on this looks bad in the cult case. The fact that the leader is a putative authority isn't a good reason to defer to him – not when there are other experts who demur, when the evidence doesn't support his claims, and when there are grounds for questioning his judgement. This is what individualistically-minded epistemologists are urging us to do: to question authority when there are grounds for doing so, rather than taking the fact that someone is an authority among our in-group as a sufficient reason to defer to them. The person who believes climate change is a hoax just because their favourite podcaster said so is being too credulous. If that extreme level of unquestioning deference counts as rational, then rational deference counts for little. It loses the theoretical link to cumulative culture and epistemic success that was supposed to be theoretically underpinning Levy's whole proposal.

#### 4. Structural Solutions

Our second criticism is that Levy's view about the rationality of uncritical deference undermines his own environmentalist approach to social epistemic policy. The environment we want is one where it's easy for people to believe true, well-evidenced things, largely by going along with what they are told. Concretely, Levy argues for improving our norms of scientific publishing and media coverage to discourage disseminating biased or spurious findings [125–31]. But the policies he's endorsing to this end require us to

discriminate between credible and non-credible expertise.<sup>15</sup> This discrimination, when done successfully, seems like a paradigm example of rational judgment. It involves carefully evaluating the track records, methodologies, and incentives of putative experts to determine their reliability. Levy's account, which treats deference as rational merely because it accedes to the views of in-group authorities, seemingly cannot credit the rationality of this type of discrimination.

One might try to defend Levy's position by carefully distinguishing several questions about epistemic norms and epistemic environments (thanks to a reviewer for suggesting this framing):

- (i) How should individuals form beliefs when aiming at knowledge?
- (ii) How should communities collectively form beliefs when aiming at knowledge?
- (iii) How should individuals work to improve the epistemic environment?
- (iv) How should communities work to improve the epistemic environment?

The thought would be that Levy's radical deference principle applies only to questions (i) and perhaps (ii) – it tells us how beliefs should be formed within a given environment. When we turn to questions (iii) and (iv) about improving the environment itself, different norms might apply. On this view, there's no contradiction in saying both that individuals should generally defer to authority when forming ordinary beliefs, and that some individuals need to exercise critical discrimination when working to establish which authorities are credible.

But this attempted defense actually just highlights the core problem with Levy's account. The problem isn't that Levy fails to distinguish these different questions – it's that his defense of uncritical deference makes it difficult to maintain these distinctions in a principled way. If uncritical deference to perceived authorities is rational simply because they are perceived as authorities, then this norm would seem to apply equally to all contexts of belief formation. There would be no principled basis for saying “defer uncritically here, but think critically there.” The very attempt to carve out special contexts where more discriminating norms apply would require the kind of critical thinking whose rationality Levy's account is purporting to deny.

What's needed instead is precisely what this objection gestures at: an approach that recognises different epistemic norms operating in different contexts, in an ecosystemic way. But embracing such an approach means abandoning Levy's far-reaching defense of the rationality of uncritical deference. The lesson here isn't that Levy should have more

---

<sup>15</sup> The same is going to apply to basically any epistemic environmentalist norm or policy, not just the particular norms around scientific publishing that Levy is focusing on, e.g. it will apply similarly to a policy that calls for experts and institutions who engage in dishonest public speech to be placed on a list epistemic polluters (see e.g. Ryan 2018). What makes a social epistemic policy environmentalist, for present purposes, is that its proximate aim is to increase the preponderance of true information in public discourse, rather than trying to improve the rationality or critical thinking of information consumers. In order to achieve its aim, any epistemic environmentalist policy has to have effective ways of distinguishing true information (and/or credible information sources).

carefully distinguished different epistemic questions – it's that making such distinctions requires rejecting his central claim about the rationality of deferring to in-group authorities simply because they're authorities.

Restructuring the epistemic environment to make default sources of information accurate necessitates a massive program of institutional discrimination. Countless judgments need to be made, and continually re-made, about which authorities and sources are credible. These judgments have to be responsive to evidence about the reliability of those sources. The people making the judgments can't do this while uncritically deferring to authority. Instead, we need some reasoners capable of reliably discriminating the comparative expertise of putative authorities, even if they are not domain experts themselves. They need the critical thinking skills to spot potential biases, conflicts of interest, or methodological flaws that could compromise an authority's credibility.<sup>16</sup>

Part of what is going awry in Levy's picture is that, having equated rationality (in matters of deference, discrimination, and critical thinking) with what conduces to collective epistemic success, he doesn't put sufficient emphasis on the ways in which collective success can benefit from heterogeneity in people's belief-forming practices.<sup>17</sup> We have significant evidence that a diversity of approaches and perspectives is required to optimise epistemic performance across a group.

To illustrate, consider a well-studied tradeoff in designing intelligent learning algorithms: the "explore-exploit" tradeoff (Cohen et al 2007, Hills et al 2015). In a noisy environment where we can't be sure that we have learned the entire truth, an epistemic agent must decide how to allocate their time between using the evidence they have already acquired (exploiting) versus engaging in experimentation to devise better beliefs (exploring). The optimal balance will depend on the agent's uncertainty, and on the risks and rewards associated with new discoveries. Too much exploitation risks stagnation; too much exploration wastes resources on fruitless investigations. In a social context, the tradeoff becomes more complex, as an individual's best strategy depends on what others in the community are doing. If everyone else is conformist exploratory critical thinking will be more valuable. But if others are already pushing the boundaries then exploiting existing knowledge may be better.

In a social environment where agents can learn from either social sources or from their own observation, the problem becomes even more complex again. In addition to intra-

---

<sup>16</sup> The points we're making here are broadly similar to the claims Hazlett (2016) defends, concerning the social value of non-deferential belief.

<sup>17</sup> Nguyen (2023) develops a similar point in his discussion of "hostile epistemology". He argues that while individuals must rely on cognitive shortcuts and heuristics to cope with an overwhelming world, these very shortcuts create vulnerabilities that can be exploited. The solution isn't to eliminate these shortcuts (which would be impossible for finite beings), but to cultivate a diversity of strategic responses within the epistemic community. As he puts it, "Limited beings in a hostile epistemic environment are locked in an unending epistemic arms race" (21) – an arms race that requires different individuals and groups to develop different cognitive strategies rather than converging on a single approach to belief formation.

personal tradeoffs, there are now interpersonal tradeoffs. If the community already has many agents engaged in exploration, it may be best to simply watch and learn from them; and conversely if the community is highly conformist and only exploiting past discoveries, it may be best to undertake some experimentation.<sup>18</sup>

The contrast between uncritical deference and independent critical thinking mirrors this explore–exploit tradeoff. The optimal balance, for the collective, between these two modes is not something that can be determined *a priori*, independent of actual context. Surely not everybody needs to critically examine every issue for themselves – that way lies a cacophony of idiosyncratic views. But nor is uncritical deference the answer – that runs a high risk of stagnation and dogmatism. A healthy epistemic community needs a mix of strategies, with some deference to keep the group anchored, and some independent questioning to keep it responsive to new evidence.

Levy might argue that even if some critical thinking is good for the collective, it could still be irrational for the individual. Some epistemic models show that individual and collective incentives can diverge, leading to epistemic tragedies of the commons (Mayo-Wilson et al 2011, Zollman 2020). In these cases, individuals motivated purely by self-interest adopt strategies that are collectively suboptimal. Perhaps critical thinkers are like this: irrational martyrs sacrificing their own epistemic welfare for the greater good. But this argument requires abandoning Levy’s central claim that rationality is tied to collective success. The point of his account was to shrink the conceptual distance between individual rationality and collective optimality. Conceding that the two can come apart, far from salvaging his view of rationality, puts pressure on its core thesis.

## 5. Conclusion

The attractive features of epistemic environmentalism don’t require us to posit a tight link between individual epistemic rationality and group-level success. The core of the case for the environmental approach is that deferring to local authorities will always be an efficient means of transmitting knowledge, so the best strategy to promote true beliefs is to ensure those authorities are espousing true and well-evidenced beliefs.

Levy gives a supercharged justification for this environmental approach: if uncritical deference to local authorities is rational, then individualistic interventions seem totally futile. There’s little to be gained trying to help people be rational social epistemic agents, if people’s natural impulse to defer to in-group authorities is already rational. But the

---

<sup>18</sup> Related ideas in the literature include the idea that “diversity trumps ability”: a team of agents with suboptimal, but diverse, learning heuristics can often outperform a team of agents who all share the same, optimal heuristic. This has been argued for on the basis of a celebrated modelling result due to Hong and Page (2004), though see Grim et al. (2019) for some limitations. There is also a good deal of empirical evidence on the benefit of cultural and other forms of diversity in the workplace, though a comprehensive recent meta-analysis suggests that the benefit, though positive, is of very small magnitude (Wallrich et al. 2024).



environmental approach doesn't need such a supercharged justification. Individualistic approaches can still be wrongheaded, even if they aren't utterly futile. As we see it, there is *something* to be gained in trying to help people be more rational, it's just that there's generally more to be gained – vis-à-vis the realisation of our collective epistemic aims – in prioritising environmental interventions.

We have argued that Levy's radicalism is not just unnecessary, but counterproductive. In §3 we argued that it ties the environmental approach to an unattractive view of rational deference – endorsing uncritical deference that actually hurts individual and group epistemic aims. In §4 we argued that it undermines some of the judgements that are involved in the construction of a favourable information environment. Some non-experts need to critically assess expertise, and doing this should, contra Levy's view, be credited as rational, provided that it's done in a way that's critically-minded and consistently responsive to evidence.

*Acknowledgements.* Thanks to the editors and a referee for helpful feedback on an earlier version of this paper. Thanks also to Michael Hannon and Elise Woodard for discussion on this topic. The authors work on this paper was supported by an Australian Research Council Discovery Project (DP190100041) on 'Governing the Knowledge Commons'.

## References

- Ahlstrom-Vij K. (2013). *Epistemic Paternalism: A Defence*. London: Palgrave Macmillan.
- Anderson E. (2011). 'Democracy, public policy, and lay assessments of scientific testimony.' *Episteme* 8(2), 144-164.
- Blake-Turner C. (2020). 'Fake news, relevant alternatives, and the degradation of our epistemic environment.' *Inquiry*, 1-21.
- Cassam Q. (2018). *Vices of the Mind: From the Intellectual to the Political*. Oxford: Oxford University Press.
- Chambers S. (2021). 'Truth, deliberative democracy, and the virtues of accuracy: is fake news destroying the public sphere?' *Political Studies* 69(1), 147-163.
- Clifford W. K. (1999). 'The ethics of belief.' In T. J. Madigan (ed.), *The Ethics of Belief and Other Essays*. Amherst: Prometheus. [Originally published 1877]
- Cohen J. D., McClure S. M., & Yu A. J. (2007). 'Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration.' *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1481), 933-942.
- Cohen J., & Fung A. (2021). 'Democracy and the digital public sphere.' In L. Bernholz, H. Landemore, and R. Reich (eds), *Digital Technology and Democratic Theory*. Chicago: University of Chicago Press.
- Fantl J. (2018). *The Limitations of the Open Mind*. Oxford: Oxford University Press.
- Goldman A. I. (1991). 'Epistemic paternalism: communication control in law and society.' *The Journal of Philosophy* 88(3): 113-131.
- Goldman A. I. (2001). 'Experts: which ones should you trust?' *Philosophy and Phenomenological Research* 63(1), 85-110.

- Grim P., Singer D. J., Bramson A., Holman B., McGeehan S., & Berger W. J. (2019). 'Diversity, ability, and expertise in epistemic communities.' *Philosophy of Science* 86(1), 98-123.
- Hall R. J., & Johnson C. R. (1998). 'The epistemic duty to seek more evidence.' *American Philosophical Quarterly* 35(2), 129-139.
- Hardwig J. (1985). 'Epistemic dependence.' *The Journal of Philosophy* 82(7), 335-349.
- Hazlett A. (2016). 'The social value of non-deferential belief.' *Australasian Journal of Philosophy* 94(1), 131-151.
- Henrich J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton: Princeton University Press.
- Hills T. T., Todd P. M., Lazer D., Redish A. D., & Couzin I. D. (2015). 'Exploration versus exploitation in space, mind, and society.' *Trends in Cognitive Sciences* 19(1), 46-54.
- Hong L., & Page S. E. (2004). 'Groups of diverse problem solvers can outperform groups of high-ability problem solvers.' *Proceedings of the National Academy of Sciences of the USA* 101(46), 16385-16389.
- Kripke S. A. (2011). 'On two paradoxes of knowledge.' In *Philosophical Troubles: Collected Papers Volume I*. Oxford: Oxford University Press.
- Lepoutre M. (2020). Democratic group cognition. *Philosophy & Public Affairs* 48(1), 40-78.
- Levy N. (2022). *Bad Beliefs: Why They Happen to Good People*. Oxford: Oxford University Press.
- Mayo-Wilson C., Zollman K. J. S., & Danks D. (2011). 'The independence thesis: when individual and social epistemology diverge.' *Philosophy of Science* 78(4), 653-677.
- Millgram E. (2015). *The Great Endarkenment: Philosophy for an Age of Hyperspecialization*. Oxford: Oxford University Press.
- Nagell K., Olguin R. S., & Tomasello M. (1993). 'Processes of social learning in the tool use of chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*).' *Journal of Comparative Psychology* 107(2), 174-186.
- Nguyen C. T. (2018). 'Expertise and the fragmentation of intellectual autonomy.' *Philosophical Inquiries* 6(2), 107-124.
- Nguyen C. T. (2020). 'Echo chambers and epistemic bubbles.' *Episteme* 17(2), 141-161.
- Nguyen C. T. (2023). 'Hostile epistemology.' *Social Philosophy Today* 39(1), 9-32.
- Rini R. (2017). 'Fake news and partisan epistemology.' *Kennedy Institute of Ethics Journal* 27(2), 43-64.
- Ryan S. (2018). 'Epistemic environmentalism.' *Journal of Philosophical Research* 43, 97-112.
- Wallrich L., Opara V., Wesolowska M., Barnoth D., & Yousefi S. (2024). 'The relationship between team diversity and team performance: reconciling promise and reality through a comprehensive meta-analysis registered report.' *OSF*, <https://doi.org/10.31234/osf.io/nscd4>.
- Whitehouse H. (2021). *The Ritual Animal: Imitation and Cohesion in the Evolution of Social Complexity*. Oxford: Oxford University Press.
- Woodard E. (2024). 'What's wrong with partisan deference?' In Alex Worsnip (ed), *Oxford Studies in Epistemology, Volume 8*. Oxford: Oxford University Press.
- Worsnip A. (2022). Review of Neil Levy, *Bad Beliefs: Why They Happen to Good People*. *Notre Dame Philosophical Reviews* 2<sup>nd</sup> November 2022.
- Zollman K. J. S. (2020). 'The theory of games as a tool for the social epistemologist.' *Philosophical Studies* 178(6), 1381-1401.