

Australasian Journal of Philosophy

ISSN: 0004-8402 (Print) 1471-6828 (Online) Journal homepage: www.tandfonline.com/journals/rajp20

Capturing our Attention

Neil Levy

To cite this article: Neil Levy (04 Aug 2026): Capturing our Attention, Australasian Journal of Philosophy, DOI: 10.1080/00048402.2026.2707327

To link to this article: https://doi.org/10.1080/00048402.2026.2707327

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 04 Aug 2026.

Submit your article to this journal

Article views: 128

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rajp20

AUSTRALASIAN JOURNAL OF PHILOSOPHY
https://doi.org/10.1080/00048402.2026.2707327

OPEN ACCESS Check for updates

Capturing our Attention

Neil Levy

Macquarie University and University of Oxford

ABSTRACT

What are the harms of misinformation? It's widely believed that it causes false belief, including false beliefs in quite bizarre claims. I argue that this threat is greatly exaggerated: while plausible misinformation is often believed, the implausible misinformation that is often the focus of discussion is genuinely believed relatively rarely. It doesn't follow that implausible misinformation is harmless. I argue that it has the capacity to capture attention, and in doing so it thwarts or hijacks debates. I argue that we possess sophisticated capacities of epistemic vigilance, which work reasonably well to distinguish reliable from unreliable information, but for a variety of reasons we do not have and are unlikely to be able to develop parallel defences against attentional capture. I conclude by arguing that we need social and institutional responses to attentional capture.

ARTICLE HISTORY Received 16 April 2024; Revised 14 May 2025

KEYWORDS Misinformation; belief; epistemic vigilance; attention

Until recently, attention has not had the attention it deserves in philosophy. Philosophy of mind was first to begin making up for this neglect (Mole 2011; Watzl 2011; Wu 2014; Jennings 2020). More recently, the importance of attention (and salience, to which it is closely related) for normative theory has increasingly been recognized. One strand of research extends the Lewisian 'scorekeeping' approach (Lewis 1979) to understand how hate speech makes prejudiced assumptions salient in a conversational context (Mcgowan 2003; Simpson 2013; Lepoutre 2021; McGowan 2022; see Fraser 2023 for criticism). A different (though compatible) strand grounds prejudice (Munton 2023), virtues and vices (Gardiner 2022; Watzl 2022), and certain sorts of harms (Whiteley 2022) in patterns of attention, while yet a third strand focuses on how emotions capture our attention, prompting the search to seek out facts or considerations that might justify the emotion (Brady 2016). Epistemology's traditional focus has been on knowledge and belief. If Sebastian Watzl is right that attention and salience are explanatorily central in an account of the mind and in the explanation of experience and behavior to a degree that is, as least, roughly commensurate to the explanatory role of belief' (Watzl 2022: 92), perhaps epistemology should follow other areas in making attention a focus.[1]

[1] It might be argued that a domain of inquiry that is not concerned with belief or knowledge does not qualify as epistemology. Even if we restrict epistemology in this sort of way, we can't neglect attention: attention plays a pervasive role in belief formation and in knowledge (Saint-Croix forthcoming).

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

In this paper, I will focus on one particular harm that falls within the broad purview of epistemology, and which can only be understood by focusing on attention. I will argue that the harms of implausible misinformation arise very significantly from its effects on attention. I will also argue that this is bad news: we are less well-equipped to protect ourselves against inappropriate attentional capture than against attempts to induce false beliefs in us. We possess a suite of capacities for epistemic vigilance, but lack equivalent capacities to protect us against pernicious influences on our attention.

In section one, I will describe the contemporary debate over the harms of misinformation. Misinformation, including wildly implausible misinformation, is widely held to be harmful because it results in the formation of false beliefs. As I will argue, however, there is evidence that false beliefs based on implausible misinformation are very much less common than is commonly thought. Despite the fact that it is not widely believed, implausible misinformation may nevertheless be very harmful. In section two, I argue that implausible misinformation is harmful, at least in part, due to the way in which it captures and skews our attention. In section three, I argue that these harms are especially difficult to address. We have mechanisms that provide us with very significant protections when it comes to belief, but we lack such mechanisms with regard to attentional capture. We are therefore especially vulnerable with regard to the hijacking of attention. The final section ventures some tentative responses to the problem, but the primary aim is to argue for its significance and urgency.

1. Misinformation and Belief

Over the past decade, epistemologists and social scientists alike have worried about the pervasiveness of misinformation across all forms of media.[2] Some have gone so far as to speak of a misinformation *crisis* (Houlden, Hodson, *et al.* 2021; Allchin 2022; Germani, Pattison, and Reinfelde 2022). On this view, many people's media diet consists largely of unreliable information, and as a consequence they do not agree on basic facts, not even when these facts are scientifically uncontroversial (that vaccines do not cause autism, for example). Misinformation, allegedly, causes us to live in different worlds (Overgaard and Collier 2023).

Underlying these concerns is a particular conception of the harms of misinformation. Misinformation is harmful when and because it causes *false beliefs*.[3] For Flynn, Nyhan and Reifler (2017), for example, misinformation is harmful because it leads to *misperceptions*: 'factual beliefs that are false or contradict the best available evidence in the public domain' (128). Such misperceptions are harmful, in turn, because they distort debate: we cannot agree on a solution to a problem if we cannot agree on the nature of the problem (whether, for example, global temperatures are rising).

[2] I will understand 'misinformation' broadly, encompassing deliberate attempts to deceive and also information that is false because it is produced by methods that are unreliable (de Ridder 2024). As far as I can tell, nothing I say will depend on the details of any specific account of misinformation.

[3] Simion (2023) and Ichikawa (2024) are recent dissenters from this consensus, but they still see the harm as doxastic. For them, the harm of misinformation consists in its propensity to *undermine* belief. For Simion, 'disinformation' *is* 'ignorance generating content' (Simion 2023: 6); for Ichikawa, the suspension of judgment is a major harm of misinformation.

Misperceptions also cause harms more directly—for example, by leading people to avoid vaccines for themselves or for their children.

There's little doubt that misinformation can and does mislead. We are heavily reliant on testimony for our knowledge of the world; given this reliance, false or misleading testimony should and does lead to false belief under a variety of circumstances. To that extent, the harms on which the debate, within and beyond epistemology, has focused are real. But the evidence strongly suggests very many of those who report believing misinformation or who share it, apparently seriously, don't really believe it.

There are multiple lines of evidence for this conclusion, only some of which I'll briefly review. First, there's evidence that partisan belief report—that is, report of beliefs that are congenial to one side in politics or one side of the culture wars—is sometimes *expressive responding* (Berinsky 2018; Hannon 2021). Expressive responding consists in reporting a belief one does not hold in order to express a conative attitude. For example, one might report believing *that the economy is improving* not because one believes that it really is, but to express support for the side of politics that is currently in government. Expressive responding might explain why though Democrats and Republicans report divergent beliefs in the state of their economy, their economic behavior seems better predicted by the actual state of the economy than their reported beliefs (Bullock and Lenz 2019).

The most compelling evidence of expressive responding is provided by Schaffner and Luks (2018; see Ross and Levy 2023 for a replication). In the immediate aftermath of the controversy surrounding the (false) claim by Trump's then press secretary that his inauguration had been the largest on record, Schaffner and Luks gave their respondents unlabelled photos of the Trump and Obama inaugurations, and simply asked them 'which photo has more people'? They expected that some respondents would identify the photographs and take the opportunity to express their support for Trump by expressive responding. Indeed, a significant minority of Trump voters (but not Clinton voters or independents) identified the photo of the Trump event as depicting the larger crowd. It's not really plausible that these participants were reporting a genuine belief: the disparity between the crowds depicted was too stark. Expressive responding is a far more plausible explanation.[4]

Expressive responding has received a significant amount of attention. Another motivation for insincere belief report has received much less: sheer trolling. Some people have openly reported that they create or share fake news 'for the lulz' (Kunzru 2020), but there's been little systematic research on this sort of behavior (see Lopez and Hillygus 2018 for a rare exception). On some topics, though, there's evidence that trolling is a powerful force. Consider another first Trump-era topic:

[4] This example is one of several I will discuss featuring Donald Trump. In the context of this journal in particular, a broader range of examples, from several countries, would be welcome. Trump is, nevertheless, an irresistible example: he makes free and unapologetic use of misinformation (sometimes disavowed—'a lot of people are saying'—and sometimes not) in a way that no other high-profile politician in a democratic country has been able to emulate. As a referee for this journal points out, however, capturing attention is a goal of most politicians, and there are plenty of examples of politicians using inflammatory claims to this end. In the Australian context, Pauline Hanson has been most successful with this ploy, with inflammatory claims like 'Australia is being swamped by Muslims' (BBC 2016), and with misinformation, such as the claim that the voice to parliament proposal would require white Australians to cede land to indigenous people (Knowles 2023), and the claim that Labor was engaged in systematic voter fraud (Worthington and Workman 2022). Nigel Farage has been very successful in raising his profile in the UK through similar tactics.

the use of household bleach to fight Covid infections. After Trump publicly speculated on its value, there were numerous media reports that credulous Trump supporters were ingesting or injecting bleach (Joseph 2020; Reimann 2020; Reuters staff 2020). Going beyond anecdotal reports, the CDC conducted an internet panel survey of just over 500 US adults, and found that 4% reported ingesting bleach (Gharpure, Hunter, *et al.* 2020). If the population surveyed was representative, that's millions of people consuming bleach.

But are the reports reliable? Litman *et al.* (2023) conducted two follow-up surveys, each with a larger participant pool than the CDC study. They replicated the result with regard to bleach, finding in addition that other participants reported drinking or gargling other household cleaners. But results were very different when they removed problematic responders, such as the 5.8% of respondents to survey 1 who reported having suffered a fatal heart attack, who failed attention checks, or who reported knowing what the made-up word 'wuttlet' means. This procedure reduced the pool of those reporting treating Covid with household cleaners to just 12 individuals. Nine of these 12 reported accidentally choosing 'yes' in response to the question whether they'd ingested cleaning products. A further two reported they'd drunk these products by accident. That left a single individual who claimed intentionally to have ingested cleaning products to fight Covid. This same individual reported their weight as 1900 pounds. In the light of this evidence, we have every reason to think that approximately zero people deliberately drank bleach to combat Covid, media reports notwithstanding.

Trolling and expressive responding are both motivations for people to report their beliefs insincerely. In addition, there may be sincere, or at any rate not insincere, false belief reports. One reason people may not report their genuine belief is that a very large proportion of the time, people simply do not have a belief to report, but report one nevertheless, either because they don't like reporting ignorance or because they lack awareness of their true belief states. It's long been recognized that a great many survey responses are constructed on the spot. Asked a question that one has never considered before in any depth, one *constructs* an attitude, by sampling the considerations that come to mind (Zaller and Feldman 1992). The attitude then reported might subsequently come to be the person's belief: perhaps they recall it on later occasions, search for evidence supporting it, and so on. But it may reflect something far more ephemeral: the attitude constructed may fail to be stable over time.

Graham (2023) found there was an asymmetry between reports of true and of false beliefs: the former exhibited far more stability over time than the latter. He argues that this asymmetry arises from the way in which false beliefs are more likely to be constructed than are true beliefs. When people report knowledge, the report is more often grounded in exposure to reliable information and the resulting state is stable. When people report misperceptions, they typically construct the attitude and—presumably because they sample different considerations at different times—the resulting attitude is not stable. These beliefs are, in Lackey's (2021) terminology, *judgmentally fragile*: were the person to deliberate on a different occasion, they might easily report a different belief. There's a strong case for holding that the reported attitude is not a belief: not only is it fragile, it's unlikely to drive consequential behavior (Mercier 2020). When stakes rise, people consider a broader range of considerations, weigh them more carefully, and are more willing to admit uncertainty.

These different mechanisms likely explain a large proportion of the reports of belief in misinformation on some topics (see Levy, 2022, for yet another mechanism).

The evidence surveyed is, in many ways, good news. It is good news, because it suggests that people are very much less gullible, and much more rational, than is often suggested (see Altay, Berriche and Acerbi 2023 and Pfänder and Altay 2025, for further supportive evidence). While there can be no doubt that misinformation causes false beliefs in large numbers of people, because people are much less gullible than is often thought, implausible—not to say bizarre—misinformation is unlikely to result in widespread belief. People may, for example, genuinely believe *that climate change is not a serious problem*, because *that* claim is in line with common sense and common experience. But they are much less likely to believe (with Donald Trump) *that climate change is a Chinese hoax*, or (with QAnon) *that senior Democrats are sacrificing children to Satan*.

While there can be no doubt that people do sometimes form false beliefs on the basis of misinformation, the circumstances in which they do so lends little support to the notion that they have come to be irrational, or that we live in a 'post-truth' world or that a large proportion of the citizens of many countries have (in the words of one pollster, commenting on a report that 50% of Americans believed in or were unsure about the truth of the QAnon conspiracy), gone 'bonkers' (Rose 2020). Most people are sufficiently rational that they are unwilling to accept flagrant inconsistencies and the sort of implausibilities that plague the more bizarre conspiracy theories (postulating that most of the media and almost the entire scientific establishment are engaged in systematic deception, for example). When misinformation reports bizarre claims, it's probably not widely believed, though it may be widely endorsed nevertheless.[5]

2. Misinformation and Attention

Should we conclude, then, that no matter how harmful plausible misinformation may sometimes be, *implausible* misinformation is harmless? I suspect not: though it is rarely believed, implausible misinformation likely gives rise to a number of harms.[6] Chief among these, I will argue, is how such misinformation occupies and skews our attention.

[5] It should be acknowledged that what strikes one person as bizarre might be quite plausible to another, depending on their background beliefs. These beliefs, in turn, depend heavily on culture and subculture. Thus plausibility differs across societies and across religious and political divides. In this light, it may not be possible to develop a general account of the difference between plausible and implausible claims. However, since a large proportion of those who claim to accept implausible misinformation share much the same cultural and subcultural background as those who explicitly reject it, we can be confident that it does not genuinely strike them as plausible.

[6] Maarten van Doorn (2023) has identified a number of harms that misinformation may give rise to, independently of whether it is believed. Its prevalence may induce general scepticism about the veracity of the information one encounters, it may (thereby) make it difficult to develop the virtue of intellectual autonomy, and it may increase support for unacceptably paternalistic policies. While these concerns are serious, I think they are tractable when they arise from implausible misinformation. Most people give little credence to such misinformation; it should therefore have little power to induce general scepticism or to interfere with the development of epistemic autonomy. More sceptical reporting of belief claims in the media, so that people are no longer tempted to believe in the gullibility of *others*, should significantly reduce the temptation of unacceptably paternalistic policies.

Implausible misinformation is especially likely to be attention-grabbing. What makes a report attention-grabbing is not well understood, but there's no doubt that being *surprising* is centrally important. On the influential predictive coding account of the mind, attention just is a mechanism for making surprising input more precise, in order better to be able to assess it (Feldman and Friston 2010; Hohwy 2013). The properties that render misinformation implausible strongly conduce to making it surprising: it's the very fact that it is out of the run of common or expected occurrences that makes it implausible. Implausible misinformation is often attention-grabbing for other reasons. In particular, it often arouses strong emotion and triggers coalitional feelings (Stubbersfield 2023).

Implausible misinformation is far from the only sort of claim that can grab our attention (indeed, I will illustrate some of the harms of attentional capture below with examples that do not all involve misinformation). True reports can be surprising, after all. True reports and plausible misinformation may also trigger coalitional feelings and arouse strong emotions. Reports of behavior that arouses indignation is attention-grabbing, whether the behavior is plausible or implausible, true or false.[7] But bizarre misinformation is especially likely to have these properties. It is produced, after all, to be memorable; its production is shaped not by the demands of truth (which is often boring) but the demands of storytelling. Bizarre misinformation that is widely repeated is most especially likely to be attention-grabbing, for obvious reasons: its success is more or less decisive evidence for its being attention-grabbing (Acerbi 2019). Plausible misinformation has harms that are well understood: we all know the dangers of false belief. Implausible misinformation may not give rise to *that* harm, but it has harms specific to it.

There is, of course, nothing wrong with being attention-grabbing *per se*. Some events and some reports should be attended to. Attention allows us to plan for the future and cope with the unexpected. When something surprising really happens, we attend because its occurrence suggests we need to update our causal model of the world. When something that arouses our indignation occurs, we attend because justice matters to us. When something happens that arouses our coalitional feelings, we attend because we may be threatened. But attention may be shaped perniciously, too. We may be brought to attend to events (real or otherwise) that do not deserve our attention (because they are relatively trivial, for example, or because they are unrepresentative). Perhaps we could hope to formulate a general account of what makes an event or a report worth attending to; perhaps, on the other hand, such an account would be irreducibly normative, such that it would be little more help than 'attend to all and only those events that are worth attending to'. Fortunately, we don't need such an account here: we can work with clear examples in which attention is captured by events or reports that we would be better not attending to (or which should go

[7] I am grateful to a reviewer for this journal for bringing me to recognize how true reports (and plausible misinformation) may have some of the pernicious effects that are characteristic of implausible misinformation. The reviewer provides the useful example of Obama's tan suit. Reports that Obama wore a tan suit during a press briefing were true. They were also attention-grabbing and skewed debates in pernicious ways. I suspect that this 'controversy' lacked the power to lastingly skew attention, though, because the event itself is not sufficiently surprising or sufficiently evocative of emotion to occupy a large place in the public consciousness for long. Implausible misinformation may be finely honed to hold attention more strongly and for longer.

unattended).[8] Bizarre misinformation is not only especially likely to be attention-grabbing; it is also especially likely to attract *unwarranted* attention, since its claims are neither true nor up for serious consideration as candidates for truth.

Let's turn, now, to the harms of this sort of attentional capture. We've already mentioned some of the ways in which attention is normatively central: to the development of virtues and vices, to the constitution of prejudices, to certain sorts of epistemic injustices, and so on. I will introduce the kind of harm I'm concerned with here by comparison to another we've also already mentioned: how making something salient[9] can introduce pernicious presuppositions into conversations (McGowan 2003, 2022; Simpson 2013).

In ground-breaking work, Mary McGowan (2003, 2004) has argued that speech can be oppressive when it introduces presuppositions into conversations that—*because* they are not asserted—are difficult to challenge, but alter what is conversationally acceptable. Such oppressive speech may work by making harmful stereotypes salient. Indeed, even making race or gender inappropriately salient may change the conversational score, and constitute a kind of sneaky norm enactment (McGowan 2022): for example, when a physician's race is made salient, her credibility might be undermined.

It's important to emphasise that this kind of harm is independent of belief. In McGowan's example of the well-meaning man who makes his colleague's race salient, no one need believe that the colleague is any less of a medical expert after her race is made salient. His assertion may nevertheless alter what is conversationally appropriate. Perhaps misinformation may have similar effects on the public (and private) conversations that constitute political exchange; I leave it to those with expertise in speech acts to assess whether the framework can be scaled up in this manner. I want to point to a related but distinct phenomenon: how making a topic salient changes the *shape* of a debate. It changes what is up for discussion and how it should be discussed.

The properties that make implausible misinformation memorable make it salient, and thereby change the debate; these effects may, once again, occur independently of whether it is seen as credible.[10] For example, when misinformation is introduced, it may become incumbent on some debate participants to address it, even though

[8] Sometimes, we may be called upon to attend to things that it would be better were they unattended to: the fact that many people are attending to something may confer genuine significance on it, and yet it remains true that it would be better were it paid little attention (see below for further discussion). Some culture wars issues are like this.

[9] What's the relationship between salience and attention? Roughly, I think of salience as an extra-agential determinant of attention. Attention is goal-oriented and often under personal-level control; salience is a fact about the extra-agential environment (though of course it may well be *indirectly* under agential control) which strongly influences what we attend to, often automatically (see Dicey Jennings 2022, for discussion).

[10] In recent papers focused on counterspeech, Cepollaro, Lepoutre and Simpson (2023: 4) and Lepoutre (2024: 573-4) have noted that harmful speech may sometimes be harmful due to its capacity to command attention. The harms both papers focus on are doxastic harms: salience may increase the fluency with which we process information, which is alleged to make it more plausible and may serve to give the impression that it has more evidential support than is really the case. While these claims may be true, when misinformation is truly bizarre any such effects do not bring the claims close to the threshold for being taken seriously as candidates for belief.

Both papers also recognize, however, that there may also be a non-doxastic harm brought about when speech becomes salient: unbelieved misinformation may distract us from more pressing concerns. I am grateful to a reviewer for this journal of bringing these papers to my attention, and therefore alerting me to the fact that one of the non-doxastic harms I discuss here has previously been identified.

they don't take it seriously—and even if *no one* really takes its seriously.[11] How many people genuinely believed Donald Trump's claim that Haitian immigrants were kidnapping and eating dogs and geese (Liddell 2024)? While some likely did go on to accord it a significant degree of credibility, I suspect most did not (perhaps including Trump himself). JD Vance, who seems to have been instrumental in bringing the claim to Trump's attention, distanced himself from it at the same time as he introduced it (Chait 2024). A lack of belief did not prevent it becoming a central topic of discussion, including—or especially—among those who explicitly dismissed it as nonsense. At minimum, it reorients the conversation away from productive terrain (what facts and what measures can we agree on?) and down unproductive alleyways.

Making misinformation salient may therefore be an effective tactic for disrupting debates. This makes misinformation a handy tool for those who don't want their side to reach any sort of compromise with the other side. For example, misinformation portraying Democrats as Satanists (or Republicans as flat-Earthers) might be introduced to ensure discussion breaks down. If the accusation could simply be ignored, since the claims are so outlandish, it might do no more harm, but its attention-grabbing nature makes it hard to simply dismiss (as the success of QAnon, to which the 'Satanism' claim is central, at capturing the attention of many newspaper columnists and pollsters illustrates). If we are to continue to engage with those who pretend to take such claims seriously, or even those who are allied with those pretenders (including Trump himself; Ibrahim 2020), we might be forced to confront the claims in one way or another. Indeed, many outlets have devoted articles and think pieces to these claims (for example, Donegan 2020; Caldwell 2021; Venkataramakrishnan and Murphy 2021). Such claims continue to obtrude into consciousness, and make productive conversation difficult.

Finally (but very importantly), attention is a limited resource. We can only attend to a certain number of things at once. Attending therefore has opportunity costs. The conflict in Ukraine slipped off our radar while the war in Gaza raged; at the time of writing, Trump has riveted the world once more, and neither conflict is paid much attention. What we attend to determines what our priorities are. The fact that attention is a limited resource ensures that those who can capture our attention can exert a powerful influence over our priorities. That's a truly awesome power.

It's an open question, to my mind, whether these effects are well characterized as changing the conversational score; at least, their effects do not turn on the introduction of any presuppositions.[12] They change the shape of public debate independent of any such effect.

In the next section, I will argue that the effects misinformation has on our attention are especially difficult to address. For good reasons, it is very difficult to avoid attending to what is made salient; attention is far more vulnerable than belief.

[11] Or rather, even if no one takes it seriously as a candidate for belief. The harms I'm concerned with largely arise only if a significant proportion of people *pretend* to take it seriously. As a referee for this journal points out, if a claim is widely enough seen as too implausible to be given credence and almost no one is willing to defend it, it might become a running joke. In that sort of case, its being attended to might be harmless.

[12] A referee for this journal notes that McGowan has an expansive account of the conversational score, on which it 'tracks those things that matter with respect to what counts as appropriate for the ongoing conversation' (McGowan 2019: 28). On this account, making something salient for the conversation ipso facto changes the score. Perhaps we might see the view here as pointing not to a different kind of mechanism but rather to how a different sort of content changes the conversational score.

3. The Attention/Belief Asymmetry

It may be, as some researchers have suggested, that anxiety about misinformation arises in the main from the concern that others are gullible (Altay and Acerbi 2024), and if it were more widely known that bizarre misinformation is rarely believed the anxiety would diminish. Misinformation might therefore become less salient and recede to some degree from public view. In this section, however, I will argue that token instances of bizarre misinformation are very effective at capturing our attention, no matter what we believe, and that we lack good defences against inappropriate attentional capture.

Our vulnerability to attentional capture contrasts with our relative security against false beliefs. There are good reasons why we are well protected in the latter case: we are deeply and pervasively epistemic dependent and always have been: we've *always* been dependent on the division of epistemic labour, and on tools and techniques that few (or no) individuals could develop for themselves from scratch (Boyd, Richerson, and Henrich 2011; Henrich 2015). This deep dependence renders us vulnerable, and therefore necessitates epistemic *vigilance*.

Accordingly, we have a suite of dispositions to filter information for credibility. We prefer the testimony of those who share our values and those who give evidence of competence. The testimony of the competent is more likely to be reliable than the testimony of the less competent (Mascaro and Sperber 2009; Sperber, Clément, *et al.* 2010; Harris 2012). The testimony of those who share our values is less likely to aim at exploiting us (Levy 2019). We thereby protect our beliefs against fools and knaves.

The mechanisms of epistemic vigilance might be seen as constituting our epistemic immune system. When misinformation goes viral, we rely on these mechanisms to filter it out as a candidate for belief. Some testimony is *rationally* pre-empted for us (Begby 2021; Levy 2021): we place (almost) no credence in it, either because we regard the source as having no credibility or because it conflicts strongly with our settled beliefs. Some testimony is seen as requiring corroboration before it is accepted; again source or content (or, of course, both) determines the weight we place on it. Some testimony might actually be assigned a *negative* weight: there are circumstances in which receipt of the testimony that p might increase my credence that not-p (Suppose you're inspecting a house with a view to possible purchase. The realtor tells you that there's a lot of interest in the house and urges you to move quickly with an inflated offer, or risk missing out. Upon receipt of their testimony, your credence *that there's a lot of serious interest in this property* might reasonably fall: if there really was such interest, the realtor wouldn't be so keen on impressing the fact on you, because they'd be confident of a sale in any case).

It may be that our mechanisms of epistemic vigilance have yet to catch up with the contemporary informational environment. These mechanisms evolved in small-scale societies where the opportunities for deception were fewer than today (Sterelny 2012). Perhaps mechanisms of cognition that were reliable in such environments now misfire, because the world is far more hostile epistemically (Stanovich 2018). Nevertheless, that we possess such mechanisms seems beyond dispute, and the evidence that few people are taken in by bizarre misinformation suggests that they continue to protect us reasonably well against the formation of false beliefs.[13] But we are much more vulnerable to pernicious attentional capture.

[13] In a series of papers, Eric Mandelbaum and colleagues have argued that we reflexively believe every proposition we entertain, even when we already have conflicting beliefs (Mandelbaum 2014; Bendaña and

Our epistemic immune system consists in a suite of mechanisms that are (plausibly) adaptations to protect us *against* false beliefs. Our attentional system is designed to opposite ends: it consists in mechanisms that are designed automatically to bring us to attend to information that is important to our goals. This renders us vulnerable to attentional capture. Attentional capture occurs, centrally, because others attending to something provides us with a genuine good reason to attend to it, either because it's important for us (for a variety of reasons) to know what others attend to, or because their attending to a topic confers significance on it, even when it would not deserve attention were it not attended to. I expand on how this occurs below.

3.1 Cognitive Architecture

Part of the reason we lack attentional vigilance mechanisms is that our deployment of epistemic vigilance *requires* that we attend to what our conversational partners attend to. We engage in mindreading to deploy epistemic vigilance: assessing others' reliability as testifiers and their motivations to aid or to deceive us requires that we attempt to detect their beliefs, desires and goals (Mascaro and Sperber 2009). But mindreading requires attending to what other agents are attending to.

Consider the false-belief task, the single most studied test of mindreading capacity. In the false-belief task, participants watch one agent move a desired object from where it was originally placed to an unlikely location (for example, children might see the agent remove the chocolate from the candy box and place it in a pencil box). They are then asked to predict where an agent who saw the object being placed at the first location, but not its subsequent relocation, would look for the chocolate. Children begin reliably to pass this test around the age of 4; prior to that age, they predict that the agent will look where the chocolate actually is (Mitchell and Lacohée 1991; Wellman, Cross, and Watson 2001). While it is now controversial whether the test reliably tracks developmental trajectory (Heyes 2014), it's the logic of the task, not its reliability, that matters to us here. What the task nicely brings out is how the attribution of false beliefs to an agent requires tracking what they attend to: the agent falsely believes the object is in place *a* because she did not see that the object was moved to *b*. Accordingly, mindreading depends on attending to what others are attending to, including what we attend to jointly, using cues like eye gaze (Baron-Cohen 1997). Being epistemically vigilant *requires* that we attend to what others attend to.

3.2 The Significance of the Attended

Animals pay attention to what other animals, including those outside their own species, attend to because they have a great many threats and opportunities in common. For instance, members of one species may be attentive to the alarm calls

Mandelbaum 2021; Porot and Mandelbaum 2021). I don't (take myself to) believe this claim: I think the empirical foundations it's built on are shaky. Even if they're right, however, the 'beliefs' they identify have a different functional role from the beliefs I have in mind: they guide only a narrow slice of automatic behavior. I'm using the term 'belief' to pick out what Van Leeuwen (2023) calls *factual belief*, where a factual belief governs inference and behavior whenever the agent is concerned with getting things right. We can use terminology however we like, of course, but factual beliefs play the central role in most behavior. Mandelbaum's belief fragments might underlie implicit biases, but implicit biases are poor predictors of behavior (Oswald, Mitchell, *et al.* 2013).

of other species, because they have predators in common (Magrath, Haff, *et al.* 2015). What is true across species holds more strongly *within* species, since the overlap in threats and opportunities is much greater within than across species. So we can be expected to orient toward stimuli that others attend to. Members of the outgroup may be mistaken about many things, but in so far as we share an environment (where an environment might be physical, cultural, social, or virtual), *their* attending to something is a reason for *me* to attend to it.

3.3 Attention Confers Significance

Perhaps counterintuitively, the fact that something is widely attended to just *is* a reason to attend to it, even if it wouldn't be worthy of attention otherwise. *That something is the focus of attention* makes it important (Gardiner 2022). There are at least two mechanisms that make the attended ipso facto important. First, we're cultural animals; culture isn't optional for us but genuinely formative. When something is very widely attended to in our culture, it becomes a reference point for us: something in relation to which we're required to orient ourselves. While there are subcultural groupings that are genuinely oblivious to what the mainstream attends to, this obliviousness is both a symptom and a cause of alienation: such groupings are marginalized. They are unable to participate fully in the conversations whereby we make and remake ourselves. When obliviousness is broad enough, such subcultures suffer not so much from testimonial injustice—where their voices go unheeded or are discounted—but from a lack of a voice at all (in fact, this sort of genuine obliviousness is usually a symptom of a deeper alienation and deprivation).[14]

A second mechanism whereby attention produces something as worth attending to might be illuminated by comparison to what Quine and Ullian (1978) call the web of belief. Some beliefs are more central to the web and some more peripheral. Those on the periphery have fewer linkages to other beliefs, and may alter or atrophy without ramifications for the web as a whole. Those that are more central are more densely connected, and alterations to them have broader effects. Attending to a claim tends to make it more central to the cognitive web (which extends beyond our beliefs) and tends to bring it about that it is more densely connected to other attitudes and views—including many with no obvious logical connection to it.[15]

3.4 The Dynamics of Debate

In so far as we engage with the outgroup at all, it is difficult to avoid a focus on what they focus on. This is both a condition of engagement and a dialectical necessity. It's a

[14] It's important to distinguish between genuine obliviousness and the performative obliviousness of those who decide some cultural phenomenon is beneath them (as well as those who don't pretend to obliviousness but choose not to participate in a discussion). The distinctions mark the difference between choosing and being unable to engage in conversations.

[15] People sometimes argue that patterns of attention reflect prejudices. For example, it is sometimes said that the attention Americans pay to human rights abuses in Israel—compared to the relative neglect of human rights abuses in Saudi Arabia or Kazakhstan (say) reflects anti-Semitism. Patterns of attention may indeed reflect (Siegel 2017), or constitute (Munton 2023), prejudices. But attention may be self-sustaining. We attend to Israel, in important part, *because we have attended to it*. Having attended to it made it central to our web of political attitudes, and gave us a context in which to understand it. Its having been attended to makes it salient for us in a way that tends to assure continued attention.

condition of engagement: we can't be said to be engaging with the other side at all if they keep insisting on some point we refuse to address. It's a dialectical necessity: bystanders might reasonably conclude that the issue is an important one and that our silence on it indicates that we have no good response to it.

Consider an example from outside the domain of misinformation (strictly speaking). It has become a sport of right-wing media, at least in the UK, to ask prominent politicians to define 'woman'. The aim, of course, is to put politicians on the left in an impossible position: either they refuse to answer the question, or they give an answer that will be fodder for headlines ('Keir Starmer Can't Define a Woman' Bet 2024). Whether or not it's important to be able to answer this question is debatable. But the very fact that the right in the UK has made it central ensures that it's hard to set it aside: conversations on feminism or trans inclusion are easily derailed by the insistence we settle whether transwomen are women. A refusal to debate the question might be seen by bystanders (and those who insist on the debate) as an admission that we have no good response (as indeed the press gleefully claims; Penna 2022). In so far as the other side insists on the question, failing to respond ensures the breakdown of conversation.

Of course, we might be quite happy to break off the conversation with those who cynically weaponize this sort of question. We might think they're unlikely to be productive conversational partners in any case. But the success of a group at making a question salient infects conversations that do not involve them. The issue is obtrusive whenever issues of the obstacles confronting women or of trans inclusion are raised. Bystanders and (naïve or ill-motivated) interlocutors can exercise a kind of heckler's veto: the salience of the question means that it's easily raised, and that once raised it tends to crowd out more productive topics. [16]

I've identified four different kinds of reasons why we're vulnerable to attentional hijacking. When individuals with a prominent platform or significant sections of the population publicly attend to a person, an event, or a claim, potential conversational partners are also forced to attend to it. We must attend in order to engage in mindreading (and thereby filter testimony reliably), because attention is evidence that what is attended to is important, because attention *confers* significance and because we cannot engage in debate without attending to what our interlocutors insist on attending to. Attention is mandatory, or near enough. In this respect, it contrasts strongly with belief. Our epistemic immune system protects us from believing what others believe when we assess it as unlikely to be true; *that they believe that p* does not bring *us* to believe that *p* when *they* are unreliable, by our lights. But *that they attend to x* does bring us to attend to *x*, even when we assess *x* as not worth attending to in virtue of its antecedent place in human life or intrinsic interest.

We are therefore uniquely vulnerable to attentional capture. It is a powerful force: Donald Trump owes his presidency to his capacity to attract attention, both from those who support him (in large part *because* he's outrageous) and those who oppose him (as well as the dwindling number of those who neither support nor oppose him). His ability to attract attention is reflected, for example, in the enormous amount of

[16] Further, it is easy to increase the salience of this sort of topic, but hard to decrease it, for reasons that are not well understood. Simpson (2013) suggests this sort of asymmetry might arise because identity-prejudicial remarks or insinuations introduce networks of associations, and that attempts to cancel or undercut associations tend to reinforce them.

media coverage he receives. In the 2016 election, for example, his primary campaign received more coverage than those of all the Democrat candidates combined (Stelter and Olshansky 2015) and more than twice the coverage of Hillary Clinton once she was the candidate (Confessore and Yourish 2016). I don't want to be talking about Donald Trump, but his capacity to dominate attention makes him inevitable. We lack effective defences against this sort of attentional capture. But it's very harmful: it might even threaten democracy.

4. Conclusion and The Way Forward

Bizarre misinformation is believed much less often than it is asserted or shared. Because it's not often believed, it doesn't give rise to some of the harms we fear from it. Those who share it and endorse it haven't really checked out from our common reality; they aren't likely to engage in the harmful actions that would be licensed by beliefs like *those*. But it's far from harmless.

Misinformation has properties that give it an edge in cultural epidemiology: successful misinformation goes viral because it's memorable and attention-grabbing. It's attention-grabbing for all of us; those who find it congenial in some manner (even if not plausible) and those who find it repulsive alike. By capturing our attention, it changes the shape of debates and debate dynamics, and—perhaps most importantly—shapes our priorities. The power to capture our attention is an awesome one.

I've suggested that there's a deep asymmetry between belief and attention when it comes to our capacity to filter out what we assess as unjustified or otherwise bad. If I'm right, it would be foolish to rely on our cognitive capacities to counter the capture of attention or the way in which it changes the shape of discussion. Our best bet, I suspect, is to look beyond such capacities, to social and institutional responses to unwarranted attentional capture.

Of course, individuals are not entirely powerless to counter attentional capture. Since attention is a limited resource, we can crowd out unwanted foci of attention simply by attending to the issues or questions we take to be worth attending to. But the limitations of this approach are obvious: many of the harms of attentional capture stem from the way it shapes debate and public inquiry, and the lone individual's resistance doesn't begin to address these harms. We need a critical mass of people attending together, to force the topics that really matter to continue to attract attention. That requires joint action.

It also requires that we attend to institutional and social mechanisms that govern who has more or less power to demand attention (to themselves, or to what they attend to). The landscape here is complex, with subtle interactions between content and identity. For example, those with historical advantages, stemming from race, gender, and class, have greater attentional power in general, but other groups may have an attentional power with regard to certain topics (Black linguistic innovations are especially likely to shape the future development of slang in the United States, for example; Adams 2012). Mechanisms that ensure fairer allocations of attention and attentional power are urgently needed. When individual capacities fall short, we must compensate with social mechanisms that ensure that such capacities are not needed.

We might, for example, develop better mechanisms for measuring how representative an opinion is. Consider Twitter 'trending topics'. *That a topic is trending* is a

reason for other users to attend to it. But that a topic is trending might represent far fewer people attending to it in a first-order fashion than we tend to think. Further, and probably more significantly, much of the engagement is likely bot-driven. Bots may be programmed to amplify certain topics, or they may simply respond to what others attend to, causing a snowball effect. Better measures of how much attention is genuinely first-order—reflecting people's assessment *that the topic is worth attending to*—and how much is bot-driven might reduce the power of attentional hijacking. Social media is often a hijacking amplifier: we need to develop algorithms that dampen such hijacking.

It may be (as a reviewer for this journal suggests) that attentional hijacking is a strategy that tends to exhaust itself. But the reason it exhausts itself is not comforting. Trump and Farage and other early adopters of this strategy were hugely successful in capturing our attention by violating norms on what could be said. Their success, however, has shattered those norms. Trump has lost a great deal of his power to shock (even as he consolidates power in many other ways). This loss is not comforting, because speech *should* be constrained by norms. It *should* be shocking to say certain things. Right now, there may be fewer opportunities to capture attention by shocking speech, but that's because we've lost the background against which that speech is shocking.

Restoring speech norms in the light of which speech like Trump's is shocking is an enormously difficult task. It requires coordinated action, and those who have benefitted from capturing our attention in the past are unlikely to cooperate with us in the task. They can easily play spoiler. To whatever degree we succeed in reestablishing these norms, we also reestablish an environment in which there are once more opportunities for the hijacking of attention. We face an uneven battle in developing social and institutional mechanisms for fighting back.

Acknowledgements

I am grateful to two anonymous reviewers for the *Australasian Journal of Philosophy* and to Adam Carter for exceptionally helpful comments.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

Work on this paper was funded by the John Templeton Foundation (grant #62631) and the Arts and Humanities Research Council (AH/W005077/1).

ORCID

Neil Levy http://orcid.org/0000-0002-5679-1986

References

Acerbi, Alberto (2019) *Cultural Evolution in the Digital Age*. Oxford University Press.
Adams, Michael (2012) *Slang: The People's Poetry*. Oxford University Press.

Allchin, Douglas (2022) 'Ten Competencies for the Science Misinformation Crisis', *Science Education* **107**: 261–74. doi:doi.org/10.1002sce.21746.

Altay, Sacha and Alberto Acerbi (2024) 'People Believe Misinformation Is a Threat Because They Assume Others Are Gullible', *New Media & Society* **26**: 6440–61. doi:10.1177/14614448231153379.

Altay, Sacha, Manon Berriche, and Alberto Acerbi (2023) 'Misinformation on Misinformation: Conceptual and Methodological Challenges', *Social Media + Society* **9**: 1–13. doi:10.1177/20563051221150412.

Baron-Cohen, Simon (1997) *Mindblindness: An Essay on Autism and Theory of Mind*. MIT Press.

BBC (2016) 'Pauline Hanson Defiant as Speech Criticised in Australia', *BBC News* 15 September.

Begby, Endre (2021) 'Evidential Preemption', *Philosophy and Phenomenological Research* **102**: 515–30. doi:https://doi.org/10.1111phpr.12654.

Bendaña, Joseph and Eric Mandelbaum (2021) 'The Fragmentation of Belief', in Cristina Borgoni, Dirk Kindermann and Andrea Onofri, eds., *The Fragmented Mind*: 78–107. Oxford University Press.

Berinsky, Adam J (2018) 'Telling the Truth about Believing the Lies? Evidence for the Limited Prevalence of Expressive Survey Responding', *The Journal of Politics* **80**: 211–24. doi:10.1086/694258.

Bet, Marina (2024) 'Keir Starmer Unable to Define a Woman AGAIN as He Insists Trans Debate Must Be "Treated with Respect"', *The Sun* 21 March.

Boyd, Robert, Peter J Richerson, and Joseph Henrich (2011) 'The Cultural Niche: Why Social Learning Is Essential for Human Adaptation', *Proceedings of the National Academy of Sciences* **108**: 10918–25. doi:10.1073/pnas.1100290108.

Brady, Michael S (2016) *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford University Press.

Bullock, John G and Gabriel Lenz (2019) 'Partisan Bias in Surveys', *Annual Review of Political Science* **22**: 325–42. doi:10.1146/annurev-polisci-051117-050904.

Caldwell, Noah (2021) 'America's Satanic Panic Returns—This Time Through QAnon', *NPR* 18 May.

Cepollaro, Bianca, Maxime Lepoutre, and Robert Mark Simpson (2023) 'Counterspeech', *Philosophy Compass* **18**: e12890. doi:10.1111/phc3.12890.

Chait, Jonathan (2024) 'How J.D. Vance and the Online Right Sabotaged Trump at the Debate', *New York Magazine*. https://nymag.com/intelligencer/article/trump-debate-eating-pets-j-d-vance-cats-dogs-springfield.html.

Confessore, Nicholas and Karen Yourish (2016) '$2 Billion Worth of Free Media for Donald Trump', *The New York Times* 15 March.

Dicey Jennings, Carolyn (2022) *Attention and Mental Control*. Cambridge University Press. doi.org/10.10179781108982269.

Donegan, Moira (2020) 'QAnon Conspiracists Believe in a Vast Pedophile Ring. The Truth Is Sadder', *The Guardian* 20 September.

van Doorn, Maarten (2023) 'Advancing the Debate on the Consequences of Misinformation: Clarifying Why It's Not (Just) about False Beliefs', *Inquiry* 1–27. doi:10.1080/0020174X.2023.2289137.

Feldman, Harriet and Karl Friston (2010) 'Attention, Uncertainty, and Free-Energy', *Frontiers in Human Neuroscience* **4**. doi:10.3389/fnhum.2010.00215.

Flynn, D J, Brendan Nyhan, and Jason Reifler (2017) 'The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics', *Political Psychology* **38**: 127–50. doi:10.1111/pops.12394.

Fraser, Rachel (2023) 'How to Talk Back: Hate Speech, Misinformation, and the Limits of Salience', *Politics, Philosophy & Economics* **22**: 315–35. doi:10.1177/1470594X231167593.

Gardiner, Georgi (2022) 'Attunement: On the Cognitive Virtues of Attention', in Mark Alfano, Colin Klein and Jeroen de Ridder, eds., *Social Virtue Epistemology*: 48–72. Routledge.

Germani, Federico, Andrew B Pattison, and Monta Reinfelde (2022) 'WHO and Digital Agencies: How to Effectively Tackle COVID-19 Misinformation Online', *BMJ Global Health* **7**: e009483. doi:10.1136/bmjgh-2022-009483.

Gharpure, Radhika, Candis M Hunter, Amy H Schnall, Catherine E Barrett, Amy E Kirby, *et al.* (2020) 'Knowledge and Practices Regarding Safe Household Cleaning and Disinfection for COVID-19 Prevention-United States, May 2020', *MMWR. Morbidity and Mortality Weekly Report* **69**: 705–9. doi:10.15585/mmwr.mm6923e2.

Graham, Matthew H (2023) 'Measuring Misperceptions?', *American Political Science Review* **117**: 80–102. doi:10.1017/S0003055422000387.
Hannon, Michael (2021) 'Disagreement or Badmouthing? The Role of Expressive Discourse in Politics', in Elizabeth Edenberg and Michael Hannon, eds., *Political Epistemology*: 297–318. Oxford University Press. https://doi.org/10.1093/oso/9780192893338.003.0017.
Harris, Paul L (2012) *Trusting What You're Told*. Harvard University Press.
Henrich, Joseph (2015) *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
Heyes, Cecilia (2014) 'False Belief in Infancy: A Fresh Look', *Developmental Science* **17**: 647–59. doi:10.1111/desc.12148.
Hohwy, Jakob (2013) *The Predictive Mind*. OUP Oxford.
Houlden, Shandell, Jaigris Hodson, George Veletsianos, Darren Reid, and Chris Thompson-Wagner (2021) 'The Health Belief Model: How Public Health Can Address the Misinformation Crisis beyond COVID-19', *Public Health in Practice* **2**: 100151. doi:10.1016/j.puhip.2021.100151.
Ibrahim, Nur (2020) 'Did Trump Fail To Denounce QAnon Conspiracy Theories?', *Snopes*. https://www.snopes.com//fact-check/trump-fail-to-denounce-qanon/.
Ichikawa, Jonathan (2024) *Epistemic Courage*. Oxford University Press.
Jennings, Carolyn Dicey (2020) *The Attending Mind*. Cambridge University Press.
Joseph, Andrew (2020) 'CDC: Some Americans Are Misusing Cleaning Products — Including Drinking Them — in Effort to Kill Coronavirus', *STAT*.
Knowles, Rachael (2023) 'One Nation Accused of Pushing Misinformation in Edited Viral Video', *NITV*.
Kunzru, Hari (2020) 'For the Lulz', *New York Review of Books* **67**: 4–8.
Lackey, Jennifer (2021) *The Epistemology of Groups*. Oxford University Press.
Lepoutre, Maxime (2021) *Democratic Speech in Divided Times*. Oxford University Press.
Lepoutre, Maxime (2024) 'Narrative Counterspeech', *Political Studies* **72**: 570–89. doi:10.1177/00323217221129253.
Levy, Neil (2019) 'Due Deference to Denialism: Explaining Ordinary People's Rejection of Established Scientific Findings', *Synthese* **196**: 313–27. doi:doi.org/10.1007s11229-017-1477-x.
Levy, Neil (2021) 'Echoes of Covid Misinformation', *Philosophical Psychology* **0**: 1–18. doi:10.1080/09515089.2021.2009452.
Levy, Neil (2022) 'Conspiracy Theories as Serious Play', *Philosophical Topics* **50**: 1–19.
Lewis, David (1979) 'Scorekeeping in a Language Game', *Journal of Philosophical Logic* **8**: 339.
Liddell, James (2024) 'Trump Drags New Animal into His Debunked Claims Haitian Migrants Are Eating Pets in Ohio', *The Independent*, 13 September.
Litman, Leib, Zohn Rosen, Rachel Hartman, Cheskie Rosenzweig, Sarah L Weinberger-Litman, *et al.* (2023) 'Did People Really Drink Bleach to Prevent COVID-19? A Guide for Protecting Survey Data against Problematic Respondents', *PLOS ONE* **18**: e0287837. doi:10.1371/journal.pone.0287837.
Lopez, Jesse and D Sunshine Hillygus (2018) *Why So Serious?: Survey Trolls and Misinformation*. SSRN Scholarly Paper. Social Science Research Network.
Magrath, Robert D, Tonya M Haff, Pamela M Fallow, and Andrew N Radford (2015) 'Eavesdropping on Heterospecific Alarm Calls: From Mechanisms to Consequences', *Biological Reviews* **90**: 560–86. doi:10.1111/brv.12122.
Mandelbaum, Eric (2014) 'Thinking Is Believing', *Inquiry* **57**: 55–96. doi:10.1080/0020174X.2014.858417.
Mascaro, Olivier and Dan Sperber (2009) 'The Moral, Epistemic, and Mindreading Components of Children's Vigilance towards Deception', *Cognition* **112**: 367–80. doi:10.1016/j.cognition.2009.05.012.
McGowan, Mary Kate (2003) 'Conversational Exercitives and the Force of Pornography', *Philosophy and Public Affairs* **31**: 155–89. doi:10.1111/j.1088-4963.2003.00155.x.
McGowan, Mary Kate (2004) 'Conversational Exercitives: Something Else We Do with Our Words', *Linguistics and Philosophy* **27**: 93–111. doi:10.1023/b:ling.0000010803.47264.f0.
McGowan, Mary Kate (2019) *Just Words: On Speech and Hidden Harm*. Oxford University Press.
McGowan, Mary Kate (2022) 'On Salience and Sneakiness', in Sophie Archer, ed., *Salience: A Philosophical Inquiry*: 177–92. Routledge.
Mercier, Hugo (2020) *Not Born Yesterday: The Science of Who We Trust and What We Believe*. Princeton University Press.

Mitchell, Peter and Hazel Lacohée (1991) 'Children's Early Understanding of False Belief', *Cognition* **39**: 107–27. doi:10.1016/0010-0277(91)90040-B.
Mole, Christopher (2011) *Attention Is Cognitive Unison: An Essay in Philosophical Psychology*. Oxford University Press.
Munton, Jessie (2023) 'Prejudice as the Misattribution of Salience', *Analytic Philosophy* **64**: 1–19. doi:10.1111/phib.12250.
Oswald, Frederick L, Gregory Mitchell, Hart Blanton, James Jaccard, and Philip E Tetlock (2013) 'Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies', *Journal of Personality and Social Psychology* **105**: 171.
Overgaard, Christian Staal Bruun and Jessica R Collier (2023) 'In Different Worlds: The Contributions of Polarization and Platforms to Partisan (Mis)Perceptions', *New Media and Society* **27**: 319–37. doi:doi.org/10.11771461444823117655l.
Penna, Dominic (2022) 'Sir Keir Starmer Refuses to Answer When Asked Whether a Woman Can Have a Penis', *The Telegraph* 28 March.
Pfänder, Jan and Sacha Altay (2025) 'Spotting False News and Doubting True News: A Systematic Review and Meta-Analysis of News Judgements', *Nature Human Behaviour* 1–12. doi:10.1038/s41562-024-02086-1.
Porot, Nicolas and Eric Mandelbaum (2021) 'The Science of Belief: A Progress Report', *Wiley Interdisciplinary Reviews. Cognitive Science* **12**: e1539. doi:10.1002/wcs.1539.
Quine, W V and J S Ullian (1978) *The Web of Belief*. McGraw-Hill Education.
Reimann, Nicholas (2020) 'Some Americans Are Tragically Still Drinking Bleach As A Coronavirus "Cure"', *Forbes* 24 April.
Reuters staff (2020) 'Gargling with Bleach? Americans Misusing Disinfectants to Prevent Coronavirus, Survey Finds', *Reuters* 6 June.
de Ridder, Jeroen (2024) 'What's so Bad about Misinformation?', *Inquiry* **67**: 2956-78. doi:10.1080/0020174X.2021.2002187.
Rose, Joel (2020) 'Even If It's "Bonkers," Poll Finds Many Believe QAnon And Other Conspiracy Theories', *NPR* 30 December.
Ross, Robert M and Neil Levy (2023) 'Expressive Responding in Support of Donald Trump: An Extended Replication of Schaffner and Luks (2018)', *Collabra: Psychology* **9**: 68054. doi:10.1525/collabra.68054.
Saint-Croix, Catharine (forthcoming) 'The Epistemology of Attention', in Kurt Sylvan, Ernest Sosa, Jonathan Dancy and Matthias Steup, eds., *The Blackwell Companion to Epistemology, 3rd Edition*. Wiley Blackwell.
Schaffner, Brian F and Samantha Luks (2018) 'Misinformation or Expressive Responding? What an Inauguration Crowd Can Tell Us about the Source of Political Misinformation in Surveys', *Political Opinion Quarterly* **82**: 135–47.
Siegel, Susanna (2017) *The Rationality of Perception*. Oxford University Press.
Simion, Mona (2023) 'Knowledge and Disinformation', *Episteme* 1–12. doi:10.1017/epi.2023.25.
Simpson, Robert Mark (2013) 'Un-Ringing the Bell: Mcgowan on Oppressive Speech and the Asymmetric Pliability of Conversations', *Australasian Journal of Philosophy* **91**: 555–75. doi:10.1080/00048402.2012.704053.
Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, *et al.* (2010) 'Epistemic Vigilance', *Mind & Language* **25**: 359–93. doi:10.1111/j.1468-0017.2010.01394.x.
Stanovich, Keith E (2018) 'Miserliness in Human Cognition: The Interaction of Detection, Override and Mindware', *Thinking & Reasoning* **24**: 423–44.
Stelter, Brian and Ken Olshansky (2015) 'How Much Does Donald Trump Dominate TV News Coverage? This Much', *CNN* 6 December.
Sterelny, Kim (2012) *The Evolved Apprentice: How Evolution Made Humans Unique*. MIT Press.
Stubbersfield, Joe (2023) 'Conspiracy Theories: A Cultural Evolution Theory Approach', in Jamshid J Tehrani, Jeremy Kendal and Rachel Kendal, eds., *The Oxford Handbook of Cultural Evolution*: C72S1-C72P225. Oxford University Press.
Van Leeuwen, Neil (2023) *Religion as Make-Believe: A Theory of Belief, Imagination, and Group Identity*. Harvard University Press.
Venkataramakrishnan, Siddharth and Hannah Murphy (2021) 'Quitting QAnon: Why It Is so Difficult to Abandon a Conspiracy Theory', *Financial Times* 14 April.

Watzl, Sebastian (2011) 'The Nature of Attention', *Philosophy Compass* **6**: 842–53. doi:10.1111/j.1747-9991.2011.00433.x.
Watzl, Sebastian (2022) 'The Ethics of Attention: An Argument and a Framework', in Sophie Archer, ed., *Salience: A Philosophical Inquiry*: 89–113. Routledge.
Wellman, Henry M, David Cross, and Julanne Watson (2001) 'Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief', *Child Development* **72**: 655–84. doi:10.1111/1467-8624.00304.
Whiteley, Ella (2022) 'Harmful Salience Perspectives', in Sophie Archer, ed., *Salience: A Philosophical Inquiry*: 193–212. Routledge.
Worthington, Elise and Michael Workman (2022) 'AEC Warns Pauline Hanson's One Nation over False Voter Fraud Claims in Cartoon Attacking Labor', *ABC News* 29 April.
Wu, Wayne (2014) *Attention*. Routledge.
Zaller, John and Stanley Feldman (1992) 'A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences', *American Journal of Political Science* **36**: 579. doi:10.2307/2111583.