Disinformation, Politically Motivated Reasoning, and Knowledge Resistance

MONA SIMION

Cogito Epistemology Research Centre, University of Glasgow, UK
Email: mona.simion@glasgow.ac.uk

We have increasingly sophisticated ways of acquiring and communicating knowledge, yet, paradoxically, we are currently facing an unprecedented global ignorance crisis that affects our personal and societal well-being, as well as the stability of our democracies. There are two key triggers to this crisis, i.e. two crucial obstacles to learning: first, the widespread sharing of disinformation, which, in conjunction with an overly trusting audience, contributes to widely spread false beliefs, and correspondingly reckless political and social behaviour. At the same time, though, and at least as critical, is the prevalence of knowledge resistance and distrust in expertise. What we need to solve this high-stakes puzzle is a social epistemological framework that is able to explain the complex mechanisms underlying these surprising and unprecedented epistemic phenomena. This article will aim to sketch the contours of such a framework.

Introduction

We have increasingly sophisticated ways of acquiring and communicating knowledge, yet, paradoxically, we are currently facing an unprecedented global ignorance crisis that the World Health Organization (WHO) has declared an 'infodemic' (https://www.who.int/health-topics/infodemic#tab=tab_1). This crisis affects our personal and societal well-being, as well as the stability of our democracies. There are two key triggers to this problem: first, the widespread sharing of disinformation, which, in conjunction with an overly trusting audience, contributes to widely spread false beliefs, and correspondingly reckless political

and social behaviour. At the same time, though, and at least as critical, is the prevalence of knowledge resistance and distrust in expertise.

These data are puzzling: they seem to suggest that, with the unprecedented advance in information sharing technologies, audiences have become, at the same time, too gullible and too sceptical. Let's call this the Gullible Sceptic Puzzle. What we need to solve this high-stakes puzzle is a social epistemological framework that is able to explain the complex mechanisms underlying these surprising and unprecedented epistemic phenomena. This article will aim to sketch the contours of such a framework.

The Social Psychology of Knowledge Resistance

This section will offer a critical analysis of the main resources offered by recent results in social psychology for explaining the puzzling data.

The Politically Motivated Reasoning Hypothesis

A predominant hypothesis in social psychology (e.g., Kahan 2013; Kahan *et al.* 2016; Lord *et al.* 1979; Molden and Higgins 2012; Taber and Lodge 2006) seeks to explain the puzzling epistemic behaviour registered in consumers of information in recent years with reference to politically motivated reasoning. Under the banner of this wider hypothesis, we find various research results that have been taken, in various ways, to support the view that a thinker's prior political convictions (including politically directed desires and attitudes about political group-membership) best explain why they are inclined to reject expert consensus when they do (Kahan 2013, Kahan *et al.* 2011).

Early studies in the psychological literature that set the groundwork for this explanatory thesis initially focused on how political ideology influences the evaluation of evidence. For example, Lord *et al.* (1979) report a study in which subjects were provided with the same set of arguments for and against capital punishment and were asked to assess the strength of these arguments. Subjects' assessment of the strength of the arguments then strongly correlated with their existing views about the rights and wrongs of capital punishment. In short, subjects *already disposed* to object to capital punishment were more persuaded by the arguments against it, and the opposite was the case for those initially predisposed to favour capital punishment. (See also Kunda 1987 for a discussion of how political ideology seems to have a bearing on causal inference patterns.)

A second-wave of research in this area, led largely by Dan Kahan and his colleagues, has suggested that political ideology not only influences how we think about the persuasiveness of arguments for and against those ideologies themselves, but that our inclination to accept (or reject) scientific consensus across a range of areas is highly sensitive to what political ideology we already accept. For example, Kahan and his collaborators present studies aimed at demonstrating that

background political ideology impacts whether we align with, or go against, expert consensus on topics ranging from global warming to the safety of nuclear power (Kahan *et al.* 2011; Kahan 2014; Kahan *et al.* 2016; cf., Carter and McKenna 2020). In light of this second wave of research, the received thinking about the data underlying the Gullible Sceptic Puzzle takes it to be principally a manifestation of politically motivated reasoning (Kahan 2013). This position, while widely discussed in social psychology, has received comparably less attention in philosophy. Furthermore, typically, philosophers who have discussed it have explored the consequences of this empirical hypothesis, while taking its merits at face value (e.g., Carter and McKenna 2020).

However, on closer and recent inspection, the hypothesis is theoretically, empirically, and epistemically problematic. Theoretically, the worry is that the view scores very low on prior plausibility, due to implying widely spread irrationality in an otherwise highly cognitively successful population (e.g., Sperber *et al.* 2010). Empirically, there are worries that, in extant studies, political group identity is often confounded with prior beliefs about the issue in question; and, crucially, reasoning can be affected by such beliefs in the absence of any political group motivation. This renders much existing evidence for the hypothesis ambiguous (Tappin *et al.* 2021). Epistemologically, the worry is that the hypothesis is ineffective in making crucial distinctions among a number of phenomena, such as: (1) Concerning epistemic status: between epistemically impermissible resistance to evidence on one hand, and justified evidence rejection on the other. After all, if the extant priors that are correlated with political group identity are justified priors, and if evidence resistance is sourced in these justified priors rather than in motivated reasoning, we will have failed to distinguish justified evidence rejection from unjustified evidence resistance. (2) Concerning triggers: between instances of motivated reasoning on one hand, and epistemically deficient reasoning featuring cognitive ('cold') biases and unjustified premise beliefs on the other (Simion 2024).

Furthermore, these worries come with associated high practical stakes for policy and practice: difficulties in answering the question as to what triggers resistance to evidence have very significant negative impact on our prospects of identifying the best ways to address this phenomenon and to avoid its unfortunate practical consequences. If resistance to evidence has one main source – for instance, a particular type of mistake in reasoning, such as motivated reasoning – the strategy to address this problem will be unidirectional and targeted mostly at the individual level. In contrast, should we discover that a pluralistic picture is more plausible when it comes to what triggers resistance to evidence – whereby this phenomenon is, for example, the result of a complex interaction of social, emotive, and cognitive phenomena – we would have to develop much more complex interventions, at both individual and societal level.

Knowledge Resistance and Epistemic Vigilance

One noteworthy way that knowledge resistance manifests is in the context of a hearer's receipt of testimony from a speaker; two kinds of examples which have received particular attention include cases of (i) resistance to expert testimony (e.g., widespread resistance to scientific evidence about climate change, as well as during the onset of the COVID-19 pandemic; Kearney *et al.* 2020); and (ii) resistance to testimony from marginalized groups, which provide the central point of reference in the literature on testimonial injustice (Fricker 2007)). In both kinds of cases, the hearer's response to testimony is epistemically defective.

An important strand in the social psychology of testimonial knowledge transmission suggests the above phenomenon could be explained via the misfiring of an otherwise beneficial epistemic vigilance mechanism. Research due to Dan Sperber and colleagues (2010) and related work by Hugo Mercier (2020) suggests that the risks that we as testimonial recipients face in being accidentally or intentionally misinformed are risks that we are well positioned to navigate via a suite of cognitive mechanisms for epistemic vigilance to sort, sift, and discern information coming from other human beings (whether immediately or mediately). It is this suite of mechanisms that is postulated – on the epistemic vigilance programme – as important in explaining both the honesty of speakers and the reliability of their testimony.

If Sperber *et al.* (2010) and Mercier (2020) are right, and we do benefit from a suite of mechanisms that make us epistemically vigilant, the data underlying the Gullible Sceptic Puzzle may be easily explained as a misfiring of our epistemic vigilance mechanisms – maybe due to the cognitive overload that recent technological advances have exposed us to. If these vigilance mechanisms are misfiring, they will lead us to respond with distrust and disbelief when trust and belief are the appropriate response, and the other way around: to be gullible when we shouldn't be.

Yet, a wave of research on deception recognition paints a mostly pessimistic picture about the plausibility of the very existence of vigilance mechanisms in us. A wide range of studies testing our capacities for deception recognition show that we are very bad at it: our prospects of getting it right barely surpass chance (e.g., Kraut 1980; Vrij 2000; Bond and DePaulo 2006). To see just how well-established this result is in the relevant psychological literature, consider the following telling passage from Levine *et al.* (1999: 126): 'the belief that deception detection accuracy rates are only slightly better than fifty-fifty is among the most well documented and commonly held conclusions in deception research'. Crucially, it is not hard to see that, if these studies are right, and we detect deception with an accuracy rate that is barely above chance, both the hypothesis that we have evolved cognitive mechanisms for epistemic vigilance to help us secure the reliability of testimonial exchanges, and the idea that resistance to evidence is the result of our vigilance mechanisms misfiring, become rather implausible.

More recently though, some voices in the deception detection literature have grown disenchanted with the received view on the issue. In particular, Blair *et al.* (2010) argue that the past 40 years of research in deception detection have neglected the role of contextual cues. According to Blair *et al.*, accuracies significantly higher than chance can be consistently achieved when hearers are given access to meaningful contextual information. On the face of it, this seems like it might be the sort of result vigilance champions need. The vigilance mechanisms, the thought would go, have evolved to work in conjunction with the contextual information Blair *et al.* discuss.

Unfortunately, though, upon closer examination, these results will not do the trick for the epistemic vigilance champion. To see why, it is important to look more closely at the type of contextual information that has been given to the subjects for the purposes of this study, and ask the question: 'How plausible is it that this kind of information – i.e., information that is shown to increase reliability in deception detection – is the kind of information that would still require extra input from vigilance mechanisms?' After all, if the study gives information such as, 'This is a reliable testifier', this is the kind of information that seems to justify testimonial belief on its own: it's simply evidence that the testifier is telling the truth. Conversely, if the study provides the subject with evidence that the testifier in question is unreliable, again, one need not host epistemic vigilance mechanisms in order to justifiably withhold belief (Simion 2020).

The Blair *et al.* study identifies three types of what they dub 'contextual content' that raises the success rates for deception detection (Blair *et al.* 2010: 424–425): (1) Contradictory content: e.g., if a testifier claims to have been at home on a given night, but the hearer was told by a trusted source that she saw the testifier out at a restaurant on the night in question, it is likely that the testifier's statements will be flagged as deceptive. (2) Statistically normal content: e.g., knowledge about the testifier's normal activities; if the testifier's statements or performance are implausible given this statistically normal information, the statements are more likely to be flagged as potentially deceptive. (3) Information that increases the perceived probability of deceit, e.g., a situation in which a number of shortages have occurred at a bank. The shortages stop when one of the employees goes on vacation and begin again when the employee returns. This information may cause the interviewer to believe that the employee's statements are deceptive.

These results are, of course, hardly surprising, either empirically or epistemologically (Simion 2020): it seems trivially true that, if given the right kind and amount of contextual information in advance, most of us should be and are able to go so far as to be impeccable deception detectors, on mere garden-variety epistemic grounds – no extra-mechanisms needed: as a limit case, if I know in advance that everybody is lying, for instance, I will likely be very good – indeed, infallible – at detecting deceit. What matters for us here, however, is whether the kind of information that does the trick in the study at hand is the kind of information that would plausibly increase the general reliability of our alleged vigilance mechanisms – rather than deliver sufficient evidence for/against a particular piece of testimony on its own. I contend, however,

that the plausible answer is clearly the latter: no special vigilance-like psychological skills are required in these cases, the evidence is enough to justify the response. Furthermore, interestingly, one out of three Blair *et al.* experiments failed to confirm their hypothesis (Blair *et al.* 2010: 427): this was the experiment that gave participants the most limited and subtle contextual information. Thus, the experiment that most closely resembled a garden-variety testimonial exchange, where the hearer does not have a whole lot of antecedent knowledge about the speaker, failed to deliver high rates of successful deceit detection. This, again, does not look very promising for the vigilance hypothesis.

If this is right – i.e., if the hypothesis that we host special epistemic vigilance mechanisms is implausible to begin with – the hypothesis that the data underlying the Gullible Sceptic Puzzle are instances of our vigilance mechanisms misfiring remains unvindicated as well.

Cognitive Proper Function and the Gullible Sceptic Puzzle

What we have seen so far is that extant research in social psychology suffers from both empirical and epistemological shortcomings in identifying the triggers behind the phenomenon we are interested in: on the one hand, epistemologically, we need to distinguish between unjustified evidence resistance – sourced in all kinds of epistemically impermissible belief/suspension formation, such as motivated reasonings, biases, etc – and epistemically justified evidence rejection – sourced in justified prior beliefs. On the other hand, even zooming in on epistemically problematic instances of the phenomenon, it is not clear how much evidence resistance is sourced in cold rather than hot biases, or in updating on unjustified priors rather than biases.

These difficulties in answering the question as to what triggers resistance to evidence have, in turn, a very significant negative impact on our prospects of identifying the best ways to address resistance to evidence. If resistance to evidence has one main source – for instance, a particular type of mistake in reasoning, such as motivated reasoning – the strategy to address this problem will be targeted at the individual level. In contrast, should we discover that a pluralistic picture is more plausible when it comes to what triggers resistance to evidence, we would have to develop much more complex interventions, at both individual and societal levels. Finally, if it turns out that the vast majority of instances of alleged evidence resistance is actually explained by epistemically justified evidence rejection – say, because cognizers find themselves in environments polluted with misleading defeaters for the evidence at stake – our interventions should only target the relevant epistemic environment, rather than any particular cognizer or belief-formation mechanisms.

This section will offer a theory that explains the data underlying the Gullible Sceptic Puzzle without appealing to widely spread irrational epistemic behaviour in the population. The theory appeals, instead, to the proper function of our cognitive

systems and the normal epistemic conditions in which they evolved. In my view, the widespread irrationality hypothesis assumed by the politically motivated reasoning account of evidence resistance is incorrect: humans are very reliable cognitive machines, in spite of relatively isolated instances of biased cognitive processing or heuristics-based reasoning. Irrational resistance to evidence is rare, and is an instance of input-level epistemic malfunctioning, often encountered in biological traits, the proper function of which is input-dependent when located in abnormal environments.

Our cognitive systems are systems whose proper function is input-dependent, just like the proper function of our respiratory system, for instance (Simion 2023a, 2024). The proper function of our respiratory system implies that it takes up easily available oxygen from the environment, with the aim of fulfilling its function of sending oxygen into the bloodstream. The proper function of our cognitive system implies that it updates beliefs and degrees of confidence in light of easily available evidence, with the aim of function fulfilment: generating knowledge (Simion 2016, 2019, 2021; Kelp & Simion 2021; Milikan 1984). In turn, proper function is environment-dependent: traits ought to function in a way in which they fulfil their etiological functions reliably enough in normal environmental conditions, where traits' etiological functions are the functions that they evolved to serve in the organism. Thus, our cognitive systems are properly functioning when they work in ways that reliably generate knowledge in normal environmental conditions – i.e., the conditions in which they evolved to generate knowledge.

Note, however, that we now inhabit a very different epistemic environment from the environment that our cognitive mechanisms evolved in: recent technological advances have not only placed us in an epistemic heaven of easy access information: they have also placed us in the midst of an information and disinformation overload. Since our cognitive mechanisms have not evolved in such a heavyweight informational environment, their proper function and function fulfilment are under threat.

Evidence Resistance

I have extensively argued in the past that evidence resistance is an instance of epistemic malfunction of our cognitive system – similar to other input-level malfunctions occurring in other biological traits. It is a type of malfunction that is to be expected in environments with information overload – where the cognizer's cognitive capacities are by far overwhelmed by the quantity of available information. It can occur either due to doxastic defeat, or independently of it. Doxastic defeat (also sometimes referred to as psychological defeat in the literature) is defeat that lacks epistemic normative power, but induces belief loss or downwards confidence adjustment nevertheless. The paradigmatic case of this has to do with proper updating on unjustified priors: I unjustifiably believe that all vaccines are unsafe, and update accordingly to 'The Covid vaccine is not safe'.

Some equate proper updating with rationality, due to the epistemic value of coherence, and distinguish it from epistemic justification; most, however, shy away from offering such epistemic praise to cognizers that are fully coherent but completely disconnected from reality: take the perfectly coherent Nazi, for instance. Are we comfortable to call them perfectly rational? It would seem theoretically more suited to assign positive evaluative properties to a slightly incoherent version thereof – on both epistemic and moral grounds. As the reader might have already guessed, then, my preference lies squarely with the second camp – i.e., the camp that doesn't attribute much epistemic value to coherence alone, and thus is sceptical about taking proper updating to be the mark of rationality. Importantly, doxastic defeat need not occur via proper updating: improper updating is also an option – i.e., giving extant priors more evidential weight than they would deserve (even were they to be justified). Anchoring bias in all of its incarnations is a paradigmatic case.

Finally, evidence resistance need not be the result of updating at all – be it proper or improper. One such non-doxastically sourced, less common, and most simple variety can be an unexplained one-off instance of evidence resistance: maybe I'm looking straight at the table in front of me and, owing to extreme tiredness or lack of focus, I fail to notice the cup lying on it in plain view. Or say that I am very depressed, and thus find it impossible to update on all the evidence that my life is going really well.

Most commonly, though, non-doxastically sourced evidence resistance will be sourced in some variety of bias. Biases come in various shapes, and can present as cognitive ('cold') biases (such as, for example, mental noise, heuristics) or motivational ('hot') biases, such as wishful thinking. To be clear, in many instances, this variety of evidence resistance will be biologically beneficial, evolved due to its biological benefits, and thus arguably practically rational. Compatibly, though, biased reasoning is epistemically deficient reasoning. Testimonial injustice is a paradigmatic case of evidence resistance due to bias: the hearer fails to give the testifier the level of credibility that she deserves due to sexist bias that leads them to downgrade them as testifier.

Justified Evidence Rejection

Cognitive malfunction instantiated as evidence resistance is odd in our species' cognitive life: for the most part, we are highly reliable cognizers, which is what largely explains why we are such a successful species (Mercier 2020; Sperber *et al.* 2010). What is, however, often encountered in the population are cases of epistemic proper function without function fulfilment. These are cases of rationally justified evidence rejection, owing to overwhelming (misleading) evidence present in the (epistemically polluted) environment of the agent.

In the new epistemic environment we inhabit, a lot of misleading evidence and defeat will come our way from sources of systematic, well-designed disinformation. Disinformation is widespread and harmful, epistemically and practically. It is most harmful when it affords justified uptake – i.e., when it targets rational

cognizers – *because* we are such reliable cognitive machines. Designing disinformation campaigns that target the irrational is not a very ambitious endeavour. My results (Simion 2023b) show that disinformation rarely comes in the form of straightforwardly false content, which would be easier to spot by an epistemically well-functioning cognizer. Rather, disinformation consists of content with a disposition to generate ignorance in normal conditions in the context at stake. Clever disinformation campaigns employ true assertions to generate ignorance, in subtle ways, including the following.

(1) *Disinforming via exploiting pragmatic phenomena*. True assertions carrying false implicatures will display a high capacity to generate false beliefs in the audience. I come on the news and assert: 'There is disagreement in science about climate change'. Strictly speaking, my assertion is true: there is a very small, insignificant number of climate change deniers in the scientific community. My asserting this on the news, however, triggers a Gricean relevance implication that the content is newsworthy – i.e., that the amount and kind of disagreement is significant enough to make the subject of the news. In this way, the communicated content is not just the asserted content, but rather that there are significant, newsworthy levels of disagreement in science about climate change. Since audiences are justified to believe communicated content based on sources that they have good reason to trust, the assertion has a high disposition to generate ignorance in its audience.

   Another way in which disinformation can be spread via making use of pragmatic phenomena is by introducing false presuppositions. I tell you 'The disagreement between scientists about the safety of vaccines convinced me not to get vaccinated'. Once more, my assertion does not merely communicate what it strictly speaking means, but it also presupposes that there is disagreement in science about the safety of vaccines – indeed, disagreement that is significant enough to warrant not getting vaccinated.

(2) *Disinforming via misleading defeat*. This category of disinformation has the capacity of stripping the audience of held knowledge via defeating justification.

(3) *Disinforming via content that has the capacity of inducing epistemic anxiety*. This category of disinformation has the capacity of stripping the audience of knowledge via belief defeat. The paradigmatic way to do this is via artificially raising the stakes about the context or introducing irrelevant alternatives as being relevant: 'Are you really sure climate change is happening? After all, sometimes scientists are wrong…' The way this variety of disinforming works is via falsely implying that these error possibilities are relevant to the context, when in fact they are not: one does not require Cartesian certainty in order to separate cardboard from plastic bottles, the costs are minimal, and the expected disutility of accelerated climate change very high.

These are a few examples in which, in a polluted epistemic environment, a properly functioning cognitive system can fail to fulfil its knowledge generating

function. What all of these ways of disinforming have in common is that they generate ignorance – either by generating false beliefs, knowledge loss, or a decrease in warranted confidence. When agents rationally reject reliable scientific testimony, they often do so in virtue of two types of epistemic phenomena: rebutting epistemic defeat (evidence against the proposition asserted by the expert), and undercutting epistemic defeat (evidence that the expert testifier is unreliable). Rebutting epistemic defeat consists, often, in testimony from sources one is rational to trust (Kelp & Simion 2023) that contradicts scientific testimony on the issue. These sources will be rationally trusted by the agent because of an excellent track record of testimony: they are overall reliable testifiers in the cognitive agent's community (which is why it is rational for the agent to trust them), but who are mistaken about the matter at hand: reliability is not infallibility, it allows for failure.

Thus, not all science sceptics need be exhibiting cognitive malfunction: they need not be unjustifiably, nor irrationally rejecting scientific evidence. A science sceptic could be rejecting scientific testimony about, for example, the safety of vaccines because her environment is polluted with misleading defeaters: say that she lives in a community where an overwhelming majority of testimony that she gets suggests that vaccines are not safe. Say, also, that these testifiers are otherwise reliable testifiers, with an impeccable track record (who just get things wrong on this particular occasion – after all, reliability does not imply infallibility). By any account of testimonial justification in the epistemological literature, updating beliefs and credence on misleading evidence and defeat from testifiers one knows to be reliable is justified: according to anti-reductionism, that is because the hearer has no defeaters to this testimony; according to reductionists, because the hearer has inductive evidence of the reliability of these testifiers. Science sceptics can be justified to believe vaccines are not safe when they have (in this case, misleading) rebutting defeaters for the scientific testimony that vaccines are safe. The defeater need not be a full defeater: laypeople testimony might not be heavy enough – epistemically – to outweigh expert testimony. But the sceptic will have reason to lower their confidence in the safety of vaccines: their (partial) rejection of scientific evidence will be epistemically justified.

This is a case of misleading defeat. Of course, defeat to scientific testimony, generating epistemically permissible evidence rejection, can also be non-misleading: consider a case in which vaccinating toddlers is recommended by the experts to the sole benefit of the population at large (for generating herd immunity) – since toddlers are not vulnerable to the virus that the vaccine targets. At the same time, say that the vaccine is shown to have some side effects – albeit in very rare cases – the cause of which remains under-researched due to lack of funding: since these cases are rare, there is little incentive to invest in identifying the root of the problem. Furthermore, say that our science sceptic is well aware of all of these facts, and thus rejects scientific testimony that the vaccine is safe for their toddler, and decides not to vaccinate them. This is a standard case of non-misleading rebutting defeat: the sceptic is not only justified to reject the expert testimony that the vaccine is safe for their toddler, but they are also, arguably, morally right to do so.

Justified evidence rejection need not only come through evidence against the proposition at stake – i.e. rebutting defeat. It can come about – and most often, I believe, it does come about – from undercutting defeat: reason to believe the expert source is not trustworthy. Consider again vaccine scepticism: sociological studies investigating vaccine hesitancy in black and Caribbean communities in the UK, for instance, suggest distrust in the safety of vaccines ultimately boils down in distrust of the NHS and medical science (Adekola *et al.* 2022). The thought is, in a nutshell, that a solid inductive basis – a history of discrimination – suggests that the interests of these communities are not at the forefront of concern of these actors. If so, this inductive evidence constitutes itself in undercutting defeat to the expert testimony in question. And, again, undercutting defeat, while often misleading when it comes to scientific expert testimony, it need not be such.

These above are ways in which one can epistemically justifiably (partially or fully) reject evidence from highly reliable sources. Likely, these will be the most ubiquitous instances on the ground, underlying the Gullible Sceptic Puzzle: again, we are highly reliable cognitive machines. Bracketing very isolated cases of biased and heuristics-based cognition (which are often biological adaptations themselves), we are very good at responding to our epistemic environment: one can see this from the fantastic practical successes we enjoy as a species – which would not be possible without the associated epistemic high performance.

Conclusion

I have argued against positing widely spread irrationality in order to explain the puzzling data behind the current ignorance crisis. Rather, I have argued, we should look at the epistemic environment for the salient causal factor: our cognitive capacities have evolved to fulfil their function of generating knowledge in epistemic environments very different from the one we are faced with today, where information and disinformation are just clicks away. Unsurprisingly, information overload will lead to either malfunction – in cases of evidence resistance – or proper function without function fulfilment – in cases in which we update our beliefs on misleading evidence and defeat from our polluted epistemic environment.

These results, in turn, illuminate the best strategies to address the phenomenon of evidence resistance. Two major types of interventions are required.

(1) For combatting rational evidence rejection: engineering enhanced social epistemic environments. This requires combatting rebutting defeaters via evidence flooding: evidence-resistant communities, inhabiting polluted epistemic environments, cannot be reached via the average communication strategies designed to reach the mainstream population, inhabiting a friendly epistemic environment (with little to no misleading evidence). What is required is (i) a quantitatively enhanced reliable evidence flow: this is a purely quantitative measure, aimed to outweigh rebutting defeaters in the agent's

environment. More evidence in favour of the scientifically well supported facts will, in rational agents, work to outweigh the misleading evidence they have against the facts; (ii) a qualitatively enhanced reliable evidence flow: this is a qualitative measure that aims to outweigh misleading evidence via evidence from sources that the agent trusts – that are trustworthy vis-à-vis the agent's environment (see below on context-variant trustworthiness); (iii) quantitatively and qualitatively enhanced evidence aimed at combatting undercutting defeat (misleading evidence against the trustworthiness of reliable sources): flooding evidence-resistant communities with evidence from sources they trust in favour of the trustworthiness of sources they fail to trust due to misleading undercutting defeaters; (iv) building enhanced disinformation-detection tools to capture disinformation in all of its facets, rather than mere paradigmatic instances thereof, which involve false assertions. At a minimum, we need to build Fact Checkers that track pragmatic deception mechanisms, as well as evidential probability lowering potentials against an assumed (common) evidential background of the audience.

(2) For combatting (relatively isolated) cases of irrational evidence resistance due to uptake cognitive malfunction: increasing availability of cognitive flexibility training (e.g., in workplaces, schools, alongside anti-bias training) (Chaby *et al*. 2019; Sassenberg *et al*. 2022). Cognitive flexibility training helps with enhancing open-mindedness to evidence that runs against one's held beliefs, and to alternative decision pathways.

References

Adekola J, Fischbacher-Smith D, Okey-Adibe T and Audu J (2022) Strategies to build trust and COVID-19 vaccine confidence and engagement in Scotland. *International Journal of Disaster Risk Science* **13**, 890–902.

Blair JP, Levine T and Shaw A (2010) Content in context improves deception detection accuracy. *Human Communication Research* **36**, 423–442.

Bond CF and DePaulo BM (2006) Accuracy of deception judgments. *Personality and Social Psychology Review* **10**, 214–234.

Carter JA and McKenna R (2020) Skepticism motivated: on the skeptical import of motivated reasoning. *Canadian Journal of Philosophy* **50**(6), 702–718.

Chaby LE, Karavidha K, Lisieski MJ, Perrine SA and Liberzon I (2019) Cognitive flexibility training improves extinction retention memory and enhances cortical dopamine with and without traumatic stress exposure. *Frontiers in Behavioral Neuroscience* **13**, 24.

Fricker M (2007) *Epistemic Injustice: Power & the Ethics of Knowing*. Oxford: Oxford University Press.

Kahan D (2013) Ideology, motivated reasoning, and cognitive reflection. *Judgement and Decision Making* **8**, 407–424.

Kahan D (2014) Making climate-science communication evidence-based – all the way down. In Boykoff M and Crow D (eds), *Culture, Politics and Climate Change*. New York: Routledge, pp. 203–220.

Kahan D, Jenkins-Smith H and Braman D (2011) Cultural cognition of scientific consensus. *Journal of Risk Research* **14**(2), 147–174.

Kahan D, Hoffman D, Evans D, Devins N, Lucci E and Cheng K (2016) 'Ideology' or 'situation sense'? An experimental investigation of motivated reasoning and professional judgment', *University of Pennsylvania Law Review* **164**(349), 349–438.

Kearney MD, Chiang SC and Massey PM (2020) The Twitter origins and evolution of the COVID-19 'plandemic' conspiracy theory. *Harvard Kennedy School Misinformation Review* **1**(3).

Kelp C and Simion M (2021) *Sharing Knowledge: A Functionalist Account of Assertion*. Cambridge: Cambridge University Press.

Kelp C and Simion M (2023) What is trustworthiness? *Nous* **57**(3), 667–683.

Kraut R (1980) Humans as lie detectors: some second thoughts. *Journal of Communication* **30**, 209–216.

Kunda Z (1987) Motivated inference: self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology* **53**(4), 635–647.

Levine TR, Park HS and McCornack SA (1999) Accuracy in detecting truths and lies: documenting the 'veracity effect'. *Communication Monographs* **66**(2), 125–144. https://doi.org/10.1080/03637759909376468

Lord CG, Ross L and Lepper MR (1979) Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, **37**(11), 2098–2109.

Mercier H (2020) *Not Born Yesterday. The Science of Who We Trust and What We Believe*. Princeton, NJ: Princeton University Press.

Millikan RG (1984) *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Molden DC and Higgins ET (2012) Motivated thinking. In Holyoak KJ and Morrison RG (eds), *The Oxford Handbook of Thinking and Reasoning*. Oxford: Oxford University Press, pp. 390–409.

Sassenberg K, Winter K, Becker D, Ditrich L, Scholl A and Moskowitz G (2022) Flexibility mindsets: reducing biases that result from spontaneous processing. *European Review of Social Psychology* **33**(1), 171–213.

Simion M (2016) Perception, history and benefit. *Episteme* **13**(1), 61–76.

Simion M (2019) Knowledge-first functionalism. *Philosophical Issues*. Online First.

Simion M (2020) Testimonial contractarianism: a knowledge-first social epistemology. *Nous*. Online First.

Simion M (2021) *Shifty Speech and Independent Thought: Epistemic Normativity in Context*. Oxford: Oxford University Press.

Simion M (2023a) Resistance to evidence and the duty to believe. *Philosophy and Phenomenological Research.* Online First.

Simion M (2023b) Knowledge and disinformation. *Episteme*. Online First.

Simion M (2024) *Resistance to Evidence*. Cambridge: Cambridge University Press.

Sperber D, Clement F, Heintz C, Mascaro O, Mercier H, Origgi G and Wilson D (2010) Epistemic vigilance, *Mind & Language* **25**(4), 359–393.

Taber CS and Lodge M (2006) Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* **50**(3), 755–769.

Tappin BM, Pennycook G and Rand DG (2021) Rethinking the link between cognitive sophistication and politically motivated reasoning. *Journal of Experimental Psychology: General* **150**(6), 1095–1114.

Vrij A (2000) *Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice*. New York: Wiley.

About the Author

Mona Simion is Professor of Philosophy and Deputy Director of the Cogito Epistemology Research Centre at the University of Glasgow. She is a member of the Executive Committee of the Aristotelian Society, the Management Committee of the British Society for Theory of Knowledge, the Steering Committee of the Social Epistemology Network, and on the Editorial Board of the *Philosophical Quarterly*. She is the winner of the Young Epistemologist Prize 2021. Her research is in epistemology, philosophy of language, moral and political philosophy and feminist philosophy. She is the author of three monographs: *Shifty Speech and Independent Thought* (Oxford University Press 2021), *Sharing Knowledge* (Cambridge University Press, 2021, with Christoph Kelp) and *Reasons, Justification, and Defeat* (Oxford University Press 2021, with Jessica Brown).