How Good Are We At Evaluating Communicated Information?

HUGO MERCIER

Abstract

Are we gullible? Can we be easily influenced by what others tell us, even if they do not deserve our trust? Many strands of research, from social psychology to cultural evolution suggest that humans are by nature conformist and eager to follow prestigious leaders. By contrast, an evolutionary perspective suggests that humans should be vigilant towards communicated information, so as not to be misled too often. Work in experimental psychology shows that humans are equipped with sophisticated mechanisms that allow them to carefully evaluate communicated information. These open vigilance mechanisms lead us to reject messages that clash with our prior beliefs, unless the source of the message has earned our trust, or provides good arguments, in which case we can adaptively change our minds. These mechanisms make us largely immune to mass persuasion, explaining why propaganda, political campaigns, advertising, and other attempts at persuading large groups nearly always fall in deaf ears. However, some false beliefs manage to spread through communication. I argue that most popular false beliefs are held reflectively, which means that they have little effect on our thoughts and behaviors, and that many false beliefs can be socially beneficial. Accepting such beliefs thus reflects a much weaker failure in our evaluation of communicated information than might at first appear.

History abounds with poor decisions that have been chalked up to gullibility – when people accept whatever they're told, even if they have no good reason for doing so. Demagogues have been accused of manipulating crowds into making risky, costly, or evil decisions. Religious leaders are supposed to convert whole crowds, and to turn peaceful flocks into crusading armies. Conspiracy theories appear to drive people to the craziest actions, whether it is storming a pizzeria in search of a paedophile ring, or burning 5 G towers to stave off cancer risk. Social psychology, in particular the Asch conformity experiments (Asch, 1956) and the Milgram obedience experiments (Milgram, 1963), is also often used to suggest that people are 'shackled by social pressure [and] overly deferential to authority' (Brennan, 2012, p. 8).

Could it possibly make sense to let others influence us so much, and so easily? The best explanation for the apparent ease with which humans are influenced can be found in the gene-culture coevolution theory (Henrich, 2015; Richerson & Boyd, 2005). This

doi:10.1017/S1358246121000096 © The Royal Institute of Philosophy and the contributors 2021

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

theory stresses the importance of cultural learning for human survival, but also the difficulty of knowing whom to learn from. To solve this problem, gene-culture coevolution suggests that humans have evolved a set of heuristics such as 'copy what the majority does', or 'copy what prestigious people do'. These heuristics would so greatly facilitate learning that people would follow them even if they lead to absurd conclusions, such as accepting blatantly false beliefs or engaging in costly behaviours.

Here, I want to challenge the idea that humans are easy to influence. First, I suggest that the costs of being easy to influence would be much too high: evolutionary theory suggests on the contrary that we should be quite difficult to influence, and in particular very resistant to harmful messages. Second, I review work in experimental psychology exploring the psychological mechanisms we use to decide when we should listen to others and believe them. Third, I review work in political science, advertising, history, showing that most attempts at mass persuasion fail completely, or have very small effects. Fourth, I attempt to explain why, if we are so hard to influence, and so good at evaluating what others tell us, some false beliefs still manage to spread widely.

1. The evolution of open vigilance

Within evolutionary biology, the evolution of communication is a well-studied field. Its main result is that, for communication to be evolutionarily stable, it has to benefit, on average, senders (those who emit the communicative signals), and receivers (those who receive them) (Maynard Smith & Harper, 2003; Scott-Phillips, 2008). If this isn't the case, senders evolve to stop sending signals, or receivers evolve to stop receiving signals. However, except for individuals whose interests are perfectly aligned (e.g. cells in a body), conflicts of interests between senders and receivers mean that senders often have an incentive to send signals that benefit them, but that do not benefit receivers – typically, by sending unreliable signals. As a result, for communication to remain stable, there must exist some mechanism that stops senders from taking advantage of receivers. The presence of such mechanisms can make communication possible in the most unlikely places, and their absence can spell doom for communication even when we expect it to work best.

The relationship between preys and predators appears purely adversarial, with no overlapping interests. In fact they do have one interest in common: if the prey is going to escape the predator

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

anyway, then they are both better off if the predator doesn't chase the prey, since they both just waste much energy for nothing. However, if preys could send a signal conveying to predators 'I'm so fast you'll never catch me', all preys, regardless of their actual speed and stamina, would send that signal, and the predators would ignore it. Amazingly, several mechanisms – known as predator-deterrent signals – have evolved to allow preys to convince predators to not chase them. These mechanisms ensure that only preys that would, in fact, be likely to escape the predators can send the signals. For example, when Thomson's gazelles spot a pack of wild dogs, they often don't just run away; instead, they stot, jumping as high as possible on the same spot. This stotting often discourages the dogs from chasing the gazelles. Why? Because only a very fit gazelle, one that would likely escape the dogs, can stot well enough to convince the dogs to give up (Caro, 1986).

By contrast with the predator-prey relationship, the relationship between a mother and the foetus she carries appears entirely mutualistic. And yet they are waging a war over resources. The foetus is hungry for resources, in particular the sugar it gets from the mothers' blood, and tries to grab as much as possible. The mother's body happily provides resources to the foetus, but also wants to keep some for other pursuits. A tug of war ensues, in which the foetus, through the placenta, floods the mothers' blood with hormones to raiser her blood sugar level, while the mother produces high quantities of insulin to bring that blood sugar level back down. Because there's no mechanism to ensure the foetus only requests an amount that would be optimal for the mother, both sides waste resources to send signals that have no or very little net effect (Haig, 1993).

Humans rely hugely on communication for most of their decisions – what food to eat, whom to befriend, where to go, and so forth. This has likely been true for hundreds of thousands of years, as our ancestors already hunted, waged war, and raised children in collaboration with one another, and thus must have relied on communication. Moreover, we often cooperate and communicate with people who aren't our kin, and who have very different interests from ours. In these conditions, how is it possible that human communication has remained stable for so long? There must be mechanisms that ensure receivers of information benefit, on the whole, from communication, in particular by warding off unreliable and costly signals. We cannot afford to be gullible, overly easy to influence. Instead, we must be equipped with cognitive mechanisms that carefully evaluate what we're told, so we can decide what to accept and what to reject. I've

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

called these mechanisms *open vigilance*, but they have been originally called *epistemic vigilance* (Sperber et al., 2010).

2. Cognitive underpinnings of open vigilance

Before delving into the details of how these cognitive mechanisms of open vigilance function, it's worth looking at their overall organization, contrasting two ways in which they could be organized. In the arms race view, senders and receivers become increasingly sophisticated in their attempts to, respectively, convey, and ward off unreliable information. For the receivers, that means starting from a state of relative gullibility and, as they become more sophisticated, learning to reject more information. This view equates lack of sophistication and gullibility; historically, it has been applied to animals, children, women, slaves, workers, the masses, which were all seen as both unsophisticated and easy to influence (i.e. they copy everything, they are easily manipulated by demagogues, they are subject to silly fads, etc.). Nowadays, this view is mostly present in dual process models which see the mind as divided between more basic System 1 processes, our intuitions and emotions, which are supposed to make us gullible, and System 2 processes, our reason, which supposedly allows us to critically evaluate what we're told, and reject what we would otherwise have naively accepted (Gilbert, 1991; Kahneman, 2011).

Another organization, in many ways opposite to the one suggested by the arms race view, draws on a different analogy: the evolution of omnivorous diets. Receivers would start from a state in which they only consume very specific forms of communication and, as they evolve to better evaluate what they're told, they become more open to a much wider range of communication – much like omnivores must be able to finely discriminate between edible and toxic foods in order to have a diverse but safe diet. Both perspectives make opposite predictions about how open vigilance mechanisms function, in particular what happens when the more sophisticated of these mechanisms cannot function properly. The first view, which draws on the arms race analogy, suggests that in the absence of more sophisticated mechanisms, we revert to a state of gullibility. The second view, which draws on the omnivorous diet analogy, suggests the exact opposite: in the absence of more sophisticated mechanisms, we revert to a state of conservatism in which we are very difficult to influence.

Unfortunately, there has been repeated attempts at forced persuasion grounded in the idea that stopping people from thinking

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

carefully makes them gullible. The most infamous were subliminal advertising, in which some stimuli are flashed so quickly that we are not aware of them, and brainwashing, in which the victims are typically deprived of sleep and food, and subjected to indoctrination sessions, again with the hope that their higher thinking abilities would be impaired, making them more open to persuasion. The evidence clearly shows that neither of these techniques are efficient *at all*. Subliminal messages have no significant effect on us (Trappey, 1996), and brainwashing never worked on anyone (Carruthers, 2009). These (and other) results suggest that our open vigilance mechanisms are organized in such a way that the most basic of these mechanisms, those that are always active, are the most conservative, rejecting much of the information they encounter, and that more sophisticated mechanisms, when they are triggered, can overcome this initial negative reaction and make us accept more information.

The conservative mechanism that is at the heart of open vigilance is plausibility checking. It consists in comparing what we're told (or what we read) to what we already believe. Plausibility checking then drives us to reject information that clashes with our prior beliefs. Taking our prior beliefs into account is eminently reasonable – if you're told someone you respect has done something awful, say, you shouldn't just that this new information for granted. Moreover, we cannot help but relying on our priors: to understand a statement, you must use your beliefs and knowledge, and any clash is automatically recognized. As a result, plausibility checking provides a sound protection against unreliable messages. The issue, however, is that it is too conservative, as it makes changing our minds very difficult.

Several cognitive mechanisms can overcome our intuitive rejection of information that clashes with our prior beliefs. Reasoning, like plausibility checking, deals with the content of information, more specifically, with the quality of the arguments provided in support of a conclusion. Numerous experiments have shown that people are not only able to reject poor arguments (at least when the arguments challenge their beliefs), but also to recognize good arguments and change their minds accordingly (Mercier & Sperber, 2017). This is best demonstrated by small discussion groups, in which the exchange of arguments allows group members to adopt the best answer available, the answer that is supported by the best arguments, even if that means abandoning a confidently held belief (Laughlin, 2011; Mercier, 2016).

Some arguments rely on beliefs the audience already has, showing how the audience's beliefs are more consistent with accepting rather than rejecting the arguments' conclusion. This is for instance the case with mathematical or logical problems. Most arguments,

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

however, require some trust for their premises to be accepted. Consider the argument 'vaccines do not cause autism because many studies have shown no such link': for this argument to be effective, the audience has to (i) trust the speaker that such studies exist and, (ii) trust that the studies were well conducted. This brings us to the second family of open vigilance mechanism, which deals not with content, but with the source of the messages.

When it comes to evaluating the source of a message, we seek to answer two main questions: Is the source right? Does the source have our best interests at heart? In briefly reviewing work on how people evaluate sources, I will draw on the developmental psychology literature, to illustrate how early-emerging these open vigilance mechanisms are (Clément, 2010; Harris et al., 2018).

How can we know whether a source is more or less likely to be right than us? An important cue is informational access. From at least the age of three, children are more likely to believe someone who has had visual access to relevant information (e.g. the content of a box) over someone who hasn't (Robinson, Champion, & Mitchell, 1999). In that situation, the children could see who had had informational access. In most situations, we rely instead on what people tell us: 'I saw Helen this morning, she told me she was pregnant' is a good reason to believe the speaker is well informed about Helen's pregnancy (assuming they aren't lying, see below). Three-year-olds are also able to take such reported informational access into account (Mercier, Bernard, & Clément, 2014).

Besides informational access, the main cue we use to decide whether someone knows best is their competence, or expertise, which in turn can be assessed in different ways. The most reliable manner of assessing someone's competence is to look at their track record: have they been consistently right in the past, in the relevant domain? Again, three-year-olds already believe an informant who has been right in the past over one who has been wrong (Clément et al., 2004). Small children are also able to use less direct cues of competence, in particular those based on domain-specific expertise. For instance, preschoolers know that they should direct questions about toys to another child rather than an adult, and questions about food to an adult rather than a child (VanderBorght & Jaswal, 2009). They also know that they should ask questions about living things to a doctor rather than a mechanic, and questions about objects to the mechanic rather than the doctor (Keil, Stein, Webb, Billings, & Rozenblit, 2008; Lutz & Keil, 2002.)

Finally, the last type of cue we use to ascertain who is right is consensus: if more people – who are themselves at least somewhat likely

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

to be right – share an opinion, then we are more likely to believe them. More specifically, people appear quite apt at adjusting the weight they grant majority opinion as a function of the relative and absolute size of the majority (Morgan et al., 2012), skills that start to develop in the preschool years (Morgan et al., 2015). However, our ability to accurately assess majority opinion is dependent on how the majority opinion is presented: if we have fine-grained information about specific individuals and their opinions, we can weight them appropriately; by contrast, if all we have is statistical information (e.g. '97% of climate scientists say climate change is real'), then we tend to underweight majority opinion – even if it can still have an impact (Mercier & Miton, 2019; Mercier & Morin, 2019).

Figuring out whether someone holds accurate beliefs, however, is only the first step: if they have no intention of communicating these beliefs to us, we shouldn't let ourselves be influenced by what they say. Much work has been dedicated to the question of lie detection: are people able to spot whether someone is lying to them using behavioural cues such as gaze aversion? Decades of research have shown that people, including experts, are no better (or barely better) than chance at spotting liars using such cues (DePaulo et al., 2003; Hartwig & Bond, 2011; Vrij, 2000). Don't these results clash with the current framework, which argues that people are well able to evaluate communicated information? In fact, there are no reasons why liars would behave in any recognizable manner, and thus why people should be able to spot them using behavioural cues. This doesn't mean that receivers are helpless; on the contrary, they can rely on a variety of cues to decide who they can trust. Some cues pertain to the sender's past behaviour: do they have a track record of being helpful? For instance, three-year-olds are more likely to believe an informant who has been nice to one who has been mean (Mascaro & Sperber, 2009). However, even people who've been good to us can't be trusted when our incentives diverge – you wouldn't trust your best friend if they tell you, playing poker, that they have a great hand and you should fold. Like adults, children spontaneously discount self-interested statements (Mills & Keil, 2005; 2008.).

3. Mass persuasion

I have argued on evolutionary grounds that humans should be endowed with cognitive mechanisms allowing them both to accept

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

communicated information that is beneficial to them, and to reject that which is harmful. The experimental psychology literature confirms that humans, both adults and children, pay attention to a variety of sound cues pertaining to the content and source of messages to decide what to accept. Moreover, these mechanisms are organized in such a way that in the absence of cues suggesting we should accept a message – good arguments, a history of trust in the source – we revert to the rejection of messages that clash with our prior beliefs. This means that most attempts at mass persuasion should be largely ineffective: when you are harangued by a demagogue, see an ad, or listen to propaganda on the radio, you don't know the source personally, and you can't exchange arguments with them. You have, as a rule, not much ground for changing your mind. Yet it is commonly believed that humans are easily manipulated by mass persuasion, whether it comes from demagogues, preachers, advertising, or propaganda. What gives? These beliefs in mass persuasion's power of influence are simply false. A mountain of evidence shows that mass persuasion attempts fail to persuade.

Since ancient Greece, scholars have been concerned with the possibility that a demagogue could pervert democracy by suborning the masses to their will. There is no evidence that this is the case. Instead, demagogues gain in popularity if they are skilled at figuring out what people want to hear, at giving people a voice. For instance, historian Ian Kershaw shows Hitler was elected in 1933 because he 'embodied an already well-established, extensive, ideological consensus' (Kershaw, 1987, p. 46). For example, during his electoral campaign, Hitler had to strongly tune down his anti-Semitism, stressing instead his more popular anti-communism. Moreover, quantitative studies of his campaign speeches suggests they had little or no impact on the Nazi vote share (Selb & Munzert, 2018).

Maybe preachers have more success, converting whole crowds, imposing a strict ideology, convincing people to engage in a variety of costly behaviours, from fasting to the tithe? The rise of Christianity, going from one individual to most of the population of the Roman Empire, and then its domination over Western Europe for centuries, appears to illustrate the power of religious mass persuasion. In fact, like all religions, Christianity spread slowly – a growth rate below 5% a year –, and through personal networks, as people recruited family members or close associates (Stark, 1996; Stark, 1984). Even when it had become the dominant, official religion, Christianity – like all dogmatic religions – found it very difficult to end pagan practices, which it often ended up recycling instead (Abercrombie et al., 1980; Delumeau, 1977; Le Bras,

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

1955; Thomas, 1971). Finally, the population complained about any imposition by the Church, from paying the tithe to going to confession, and they used many Church mandated activity as excuses – going to church to barter or gossip, using religious holidays to drink and party, etc. (see, e.g., Murray, 1974).

What happens, though, when an authoritarian regime controls all the media and can drown the population in propaganda? Even then, mass persuasion doesn't seem to have much persuasive effect. The example of Nazi propaganda is the best studied. Using a variety of materials – from archival documents to surveys – researchers have found that Nazi propaganda often failed completely, and that when it did appear to have some effect, it was unrelated to the intensity of propaganda, being modulated instead by prior opinions (Kershaw, 1983; Adena, Enikolopov, Petrova, Santarosa, & Zhuravskaya, 2015; Voigtländer & Voth, 2015). For example, Nazi propaganda might have allowed people who were already anti-Semitic to act on their prejudices – by signalling that the government would back them up – but it had the opposite effect on areas with lower anti-Semitism (Adena *et al*., 2015). The same patterns are found in other authoritarian regimes, from Stalinist Russia to Maoist China – two countries which now largely rely on propaganda techniques that do not require straightforward persuasion, but attempt instead to sow confusion and distrust, or simply to distract people from issues the regime finds problematic (Roberts, 2018).

Even if they are, in many ways, vastly more sophisticated than authoritarian propaganda, political campaigns in democracies also have surprisingly limited effects. As in other fields, the gold standard for evidence is the randomized control trial: for example, deploying a given campaign act – distributing fliers, robocalling, etc. – in only a random half of electoral districts, and measuring whether voting intention or voting behaviours changed. Studies using these methods suggest that the effects of campaign acts are often null, and at best small (Kalla & Broockman, 2018). This is particularly true in widely publicized elections – presidential elections, say – and the effects of political campaigns increase when voters cannot rely on simple heuristics such as voting for their preferred party to make a decision (e.g. for primaries or ballot measures). On the whole, the media (and, to some extent, political campaigns) do have an effect on voters, but 'by conveying candidates' positions on important issues' (Gelman & King, 1993, p. 409): during a campaign, most people will learn at least some basic facts about the candidates, such as what party they belong to, and this will inform their decisions – a phenomenon quite distinct from mass persuasion.

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

In terms of amounts invested, advertising dwarves political campaigns – yet it does not seem to have stronger persuasive power. A recent study of the effects of TV advertising, which was comprehensive and methodologically superior to its predecessors, revealed that the effects of the vast majority of TV ads were small at best, with 'a sizable percentage of statistically insignificant or negative estimates' (Shapiro et al., 2019). Another recent and large scale study – with billions of observations – of online display ad effectiveness also points to very small or null effects for the vast majority of ad campaigns, with a few positive outliers, and half of the (limited) effect of advertising being to merely make people buy earlier a product they would have bought anyway (Johnson et al., 2017).

There are different ways for mass persuasion attempts to succeed. Propaganda demonstrates the strength of the regime. Spending on political ads displays that a politician has a well-funded campaign. But mass persuasion fails to persuade the audience of its outward message. This makes sense, given that mass persuasion attempts, as a rule, do not provide us with any good reason to change our mind and accept the messages.

4. Explaining the cultural success of (some) false beliefs

Even if they do not spread though mass persuasion, it is undeniable that some false beliefs have proven culturally successful, from flat earth theory to rumours of paedophiles infiltrating the government, or a belief that bleeding someone will cure just about every ailment they might suffer from. Doesn't the success of these apparently absurd ideas belie the claimed effectiveness of our open vigilance mechanisms? How can people accept such hogwash and not be gullible?

The last six chapters of *Not Born Yesterday* are dedicated to answering these questions, leading up to the conclusion that not only is the success of some false beliefs consistent with the existence of efficient open vigilance mechanisms, but that in fact these mechanisms are crucial to make sense of why some false beliefs spread (while most don't). Not every one of these arguments can be summarized here, and I will focus on two points: 'belief' takes different forms, and holding false beliefs can be socially beneficial.

When we talk about 'beliefs,' we put in the same category very different things, such as 'Olivia thinks relativity theory is correct' and 'Olivia thinks her keys are in the drawer.' However, cognitively, beliefs come in two different varieties: intuitive and reflective beliefs (Sperber, 1997). Intuitive beliefs are the type of beliefs we

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

share with other animals, that are formed (or could have been formed) through perception and basic inference, and that we use to draw further inferences and guide our behaviour. For instance, Olivia might have seen her keys in the drawer recently, and use that belief to guide her inferences and behaviour (e.g. where to look for her keys if she needs them). By contrast, reflective beliefs take a different form, they are metarepresentations: instead of simple believing '*p*' one can believe '*p* is true.' That humans can hold such metarepresentations is obvious in the case of views we do not actually entertain, such as '*p* is false.' In that case, there is a clear distinction between '*p*,' which we do not believe, and '*p* is false,' which we believe, and it is clear that we will not act on the basis of '*p*,' as it is insulated from other beliefs, inferences, and planning mechanisms. What is less obvious is that the same can be true for things we do believe. For example, there are contents that our minds are not well-equipped to process (reflective concepts, see, Sperber, 1997), and which are bound to remain within a metarepresentational embedding. If I tried to add 'relativity theory' to my intuitive beliefs, I would have no idea what to do: what inferences to draw from it, what behaviour should follow. The best I can do is hold the reflective belief 'relativity theory is correct.'

Within this framework it is thus perfectly possible for people to genuinely hold beliefs that have very little impact on their other beliefs or their behaviours, because they are only held reflectively. As is argued at length in the book, many popular false beliefs are held in this reflective manner – and we can tell because people do not behave as one would expect if they held the same beliefs intuitively. Consider the example of pizzagate, a widespread conspiracy theory according to which high ranking Democrats were abusing children in the basement of a Washington D.C. pizzeria. One man, Edgar Maddison Welch, stormed the pizzeria with his assault weapon, demanding that the children be freed. It is plausible that he had formed an intuitive belief that children were being abused, the police was too corrupt to act, etc. – in which case, his behaviour was, arguably, moral and rational. By contrast, the millions of Americans who, according to polls, endorsed pizzagate mostly failed to act on that belief. At best, they wrote one star reviews on the pizzeria's page, noting that their that they harbour paedophiles, *and that their pizza isn't great*. Such behaviour would be completely incongruous for someone who intuitively believed that children were being abused. We can thus safely conclude that the vast majority of people who believed in pizzagate only did so reflectively.

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

Holding a false belief reflectively, instead of intuitively, is much less of a failure of open vigilance, given that there are, as a rule, few practical consequences for doing so. Indeed, in quite a few cases there might actually be benefits for (reflectively) holding false beliefs, benefits that further explain why these beliefs might pass the filters of open vigilance. Some false beliefs are worth entertaining because they would be very relevant if they were true (Altay et al., 2020; Blaine & Boyer, 2018). In particular, information about threats – such as potential paedophiles – is worth considering, even if we're not sure it's true. Moreover, since people are keen on being aware of potential threats, they tend to reward, reputationally, those who share such beliefs, deeming them more competent (Boyer & Parren, 2015). As a result, false beliefs about threats can spread, even if they do not have much effect on people's behaviour.

Another way through which false beliefs can spread while having a limited direct causal impact is as justifications. People like to be able to justify their beliefs and behaviours. As a result, when there is disagreement over a belief or a behaviour in a community, this creates a market for beliefs that justify the belief or behaviour. People who suggest such justifications can be rewarded reputationally or professionally, and their audience is not particularly critical of the justifications, since these justifications have no effect on the audience's beliefs or behaviours. A striking example is provided by Galen's humoral theory of disease, which was prominent in the Western medical canon for centuries, and associated with the practice of bloodletting. At first, it might seem that this theory has done much damage, by prompting people to rely on a harmful therapy. In fact, bloodletting is one of the most common forms of medical therapies worldwide, being present in dozens of cultures that have never heard of Galen (Miton et al., 2015). It appears that bloodletting is an intuitive form of therapy, and that Galen's theories mostly justified the practice, rather than cause its success.

There are many other reasons why professing false beliefs might be beneficial: that they help us look nicer (Altay & Mercier, 2020) or more dominant (de Araujo et al., 2020), that they allow us to credibly signal our commitment to a group (Mercier, 2020), and many others (see, e.g., McKay & Dennett, 2009). Given that there is, as a rule, little personal cost to holding these beliefs – as they are held reflectively – and that there can be reputational benefits, it is not so surprising that they are allowed to pass through by our epistemic vigilance mechanisms.

The belief that large swathes of humanity are gullible, in spite of its historical pedigree and theoretical rationale, is simply false.

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

Convergent evidence from many fields – psychology, political science, marketing, anthropology, history, economics – shows that people are well able to evaluate communicated information. In particular, they are largely inured to mass persuasion attempts, in which no good reason for changing their minds is provided. Instead of focusing on why people sometimes accept false information, it appears much more productive to understand how we can help good information spread by tapping into our open vigilance mechanisms.[1]

Institut Jean Nicod
Ecole Normale Supérieure
hugo.mercier@gmail.com

References

Abercrombie, N., Hill, S., & Turner, B. S. (1980). *The dominant ideology thesis*. Allen & Unwin.

Altay, S., de Araujo, E., & Mercier, H. (2020). *'If this account is true, it is most enormously wonderful': Interestingness-if-true and the sharing of true and false news*.

S. Altay, & H. Mercier, *Happy Thoughts: The Role of Communion in Accepting and Sharing Epistemically Suspect Beliefs* (2020).

S. E. Asch, 'Studies of independence and conformity: A minority of one against a unanimous majority', *Psychological Monographs*, 70 (1956), 1–70.

T. Blaine & P. Boyer, 'Origins of sinister rumors: A preference for threat-related material in the supply and demand of information', *Evolution and Human Behavior*, 39 (2018), 67–75.

P. Boyer & N. Parren, 'Threat-related information suggests competence: A possible factor in the spread of rumors', *PloS One*, 10 (2015), e0128421.

J. Brennan, *The ethics of voting*, (Princeton University Press, 2012).

T. M. Caro, 'The functions of stotting: A review of the hypotheses', *Animal Behaviour*, 34 (1986), 649–62.

S. L. Carruthers, *Cold War Captives: Imprisonment, Escape, and Brainwashing*, (University of California Press, 2009).

F. Clément, 'To Trust or not to trust? Children's social epistemology', *Review of Philosophy and Psychology*, 1 (2010), 1–19.

[1] *Acknowledgements*. This work was funded by two ANR grants, ANR-17-EURE-0017 and ANR-10-IDEX-0001-02.

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

F. Clément, M.A. Koenig, & P. Harris, 'The ontogeny of trust', *Mind and Language*, 19 (2004), 360–79.
E. de Araujo, S. Altay, A. Bor, & H. Mercier, *Dominant Jerks: Sharing Offensive Statements can be Used to Demonstrate Dominance* (2020).
J. Delumeau, *Catholicism Between Luther and Voltaire* (Westminster Press, 1977).
B. M. DePaulo, J. J. Lindsay, B. E. Malone, L. Muhlenbruck, K. Charlton & H. Cooper 'Cues to deception', *Psychological Bulletin*, 129 (2003), 74–118.
A. Gelman & G. King, 'Why are American presidential election campaign polls so variable when votes are so predictable?', *British Journal of Political Science*, 23 (1993), 409–451.
D. T. Gilbert, 'How mental systems believe', *American Psychologist*, 46 (1991), 107–119.
D. Haig, 'Genetic conflicts in human pregnancy', *Quarterly Review of Biology*, (1993) 495–532.
P. L. Harris, M. A. Koenig, K. H. Corriveau, V.K. Jaswal, 'Cognitive foundations of learning from testimony', *Annual Review of Psychology*, 69 (2018).
M. Hartwig, & C.H. Bond, 'Why do lie-catchers fail? A lens model meta-analysis of human lie judgments'. *Psychological Bulletin*, 137 (2011), 643.
J. Henrich, *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*, (Princeton University Press, 2015).
G.A. Johnson, R.A. Lewis, E.I. Nubbemeyer, 'Ghost ads: Improving the economics of measuring online ad effectiveness', *Journal of Marketing Research*, 54 (2017), 867–84.
D. Kahneman, *Thinking, fast and slow*, (Farrar Straus & Giroux, 2011).
J. L. Kalla & D.E. Broockman, 'The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments', *American Political Science Review*, 112 (2018), 148–66.
F. C. Keil, C. Stein, L. Webb, V.D. Billings, & L. Rozenblit, 'Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds', *Cognitive Science*, 32 (2008), 259–300.
I. Kershaw, 'How effective was Nazi propaganda'. In D. Welch (Ed.), *Nazi propaganda: The power and the limitations* (Croom Helm, 1983) 180–205.

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

I. Kershaw, *The Hitler Myth: Image and Reality in the Third Reich*, (Oxford University Press, 1987).
M. A. Koenig, 'Beyond semantic accuracy: Preschoolers evaluate a speaker's reasons', *Child Development*, 83 (2012), 1051–1063.
P. R. Laughlin, *Group problem solving*, (Princeton University Press, 2011).
G. Le Bras, *Etudes de sociologie religieuse* (Presses Universitaires de France, 1955).
D. J. Lutz & F.C. Keil, 'Early understanding of the division of cognitive labor', *Child Development* (2002), 1073–1084.
O. Mascaro & D. Sperber, 'The moral, epistemic, and mindreading components of children's vigilance towards deception', *Cognition*, 112 (2009), 367–380.
J. Maynard Smith & D. Harper, *Animal signals*, (Oxford University Press, 2003).
R.T. McKay, & D.C. Dennett, 'The evolution of misbelief', *Behavioral and Brain Sciences*, 32 (2009), 493–510.
H. Mercier, 'The argumentative theory: Predictions and empirical evidence', *Trends in Cognitive Sciences*, 20 (2016), 689–700.
H. Mercier, *Not Born Yesterday: The Science of Who we Trust and What we Believe*, (Princeton University Press, 2020).
H. Mercier, S. Bernard, & F. Clément, 'Early sensitivity to arguments: How preschoolers weight circular arguments', *Journal of Experimental Child Psychology*, 125 (2014), 102–109.
H. Mercier & H. Miton, 'Utilizing simple cues to informational dependency', *Evolution and Human Behavior*, 40 (2019), 301–314.
H. Mercier & O. Morin, 'Majority rules: How good are we at aggregating convergent opinions?' *Evolutionary Human Sciences*, 1 (2019), e6.
H. Mercier & D. Sperber *The Enigma of Reason*, (Harvard University Press, 2017).
S. Milgram, 'Behavioral study of obedience', *Journal of Abnormal and Social Psychology*, 67 (1963), 371–78.
C. M. Mills, F.C. Keil, 'The Development of Cynicism', *Psychological Science*, 16 (2005), 385–90.
C.M. Mills, & F.C. Keil, 'Children's developing notions of (im)partiality', *Cognition*, 107 (2008), 528.
H. Miton, N. Claidière, & H. Mercier, 'Universal cognitive mechanisms explain the cultural success of bloodletting', *Evolution and Human Behavior*, 36 (2015), 303–312.
T. J. H. Morgan, K.N. Laland, & P.L. Harris, 'The development of adaptive conformity in young children: Effects of uncertainty and consensus', *Developmental Science*, 18 (2015), 511–24.

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press

T. J. H. Morgan, L.E. Rendell, M. Ehn, W. Hoppitt, & K.N. Laland, 'The evolutionary basis of human social learning'. *Proceedings of the Royal Society of London B: Biological Sciences*, 279 (2012), 653–62.
A. Murray, 'Religion among the poor in thirteenth-century France: The testimony of Humbert de Romans'. *Traditio*, (1974) 285–324.
P.J. Richerson & R. Boyd, *Not by genes alone*, (University of Chicago Press, 2005).
M. E. Roberts, *Censored: Distraction and Diversion Inside Chinas Great Firewall*, (Princeton University Press, 2018).
E.J. Robinson, H. Champion, & P. Mitchell, 'Children's ability to infer uterrance veracity from speaker informedness', *Developmental Psychology*, 35 (1999), 535–46.
T.C. Scott-Phillips, 'Defining biological communication', *Journal of Evolutionary Biology*, 21 (2008), 387–95.
P. Selb & S. Munzert, 'Examining a Most Likely Case for Strong Campaign Effects: Hitler's Speeches and the Rise of the Nazi Party, 1927–1933', *American Political Science Review*, 112 (2018), 1050–1066.
B. Shapiro, G.J. Hitsch, & A. Tuchman, 'Generalizable and robust TV advertising effects' (2019), *Available at SSRN 3273476*.
D. Sperber, 'Intuitive and reflective beliefs', *Mind and Language*, 12 (1997), 67–83.
R. Stark, 'The rise of a new world faith', *Review of Religious Research*, (1984) 18–27.
R. Stark, *The Rise of Christianity: A Sociologist Reconsiders History*, (Princeton University Press, 1996).
K. Thomas, *Religion and the Decline of Magic*, (Weidenfeld and Nicolson, 1971).
C. Trappey, 'A meta-analysis of consumer choice and subliminal advertising', *Psychology & Marketing*, 13 (1996), 517–30.
M. VanderBorght & V.K. Jaswal, 'Who knows best? Preschoolers sometimes prefer child informants over adult informants', *Infant and Child Development: An International Journal of Research and Practice*, 18 (2009), 61–71.
N. Voigtländer & H.J. Voth, 'Nazi indoctrination and anti-Semitic beliefs in Germany', *Proceedings of the National Academy of Sciences*, 112 (2015), 7931–36.
A. Vrij, *Detecting lies and deceit: The psychology of lying and the implications for professional practice*, (Wiley, 2000).

https://doi.org/10.1017/S1358246121000096 Published online by Cambridge University Press