

Regulating the Spread of Online Misinformation

Étienne Brown

Pre-print. Please cite the published version in the *Routledge Handbook of Political Epistemology*.

Abstract: Attempts to influence people's beliefs through misinformation have a long history. In the age of social media, however, there is a growing fear that the circulation of false or misleading claims will be more impactful than ever now that sophisticated technological means are available to those who desire to spread them. Should democratic societies worry about misinformation? If so, is it possible and desirable for them to control its spread by regulating it? This chapter offers an answer to these questions. First, I propose a definition of misinformation and explain how it proliferates in online contexts. Second, I consider four reasons to worry about misinformation by discussing its likely impact on people's political opinions, emotions, physical safety and personal autonomy. Third, I assess three strategies for regulating misinformation—individual self-regulation, platform-based innovations and governmental action—and suggest that the most effective ones are those that spare human agents from having to successively review individual pieces of content.

Brief description of the author: Étienne Brown is Assistant Professor in the Department of Philosophy at San José State University. His current research focuses on misinformation and fake news. He holds a Ph.D. in philosophy from the Sorbonne.

Since the Russian attempt to influence the outcome of the 2016 U.S. Election, there is a growing fear that the circulation of made-up news and misleading information on social media disrupts the democratic process by inciting citizens to make political judgments based on false beliefs. According to a 2019 survey by the Pew Research Center (Mitchell et al., 2019), 50% of U.S. adults believe that misinformation is a critical problem that needs to be fixed, and 79% percent of them consider that steps should be taken to restrict it. In Western Europe, countries such as Germany and France have enacted legal measures which authorize public officials to order the removal of pieces of misinformation from social media. While a growing number of legal measures against misinformation are currently being implemented, some researchers remain unconvinced that misinformation fundamentally threatens democracy. In their recent study of exposure to untrustworthy websites in the 2016 Election, Andrew Guess, Brendan Nyhan

and Jason Reifler (2020) suggest that “the widespread speculation about the prevalence of exposure to untrustworthy websites has been overstated.”

These two perspectives on misinformation—the alarmist and the deflationary—raises philosophical questions that are worth considering. Deflationary perspectives invite us to consider whether liberal democracies truly have reasons to worry about misinformation and, if they do, what these reasons are. By way of contrast, concerned citizens and public officials who conceive of misinformation as a problem that ought to be fixed invite us to reflect upon the regulatory options that are available to us. This entry takes up these two tasks. First, I define key terms such as misinformation, disinformation and fake news, and offer a brief explanation of how misinformation circulates in online contexts. I then assess the reasons we have to limit its spread. Lastly, I consider three possible ways to combat misinformation in liberal democracies—individual self-regulation, platform-based innovations and governmental action—and suggest that the most effective ones are those that spare human agents from having to successively review individual pieces of content.

1. What is misinformation and how does it spread?

Let us begin by broadly defining misinformation as the communication of false or misleading information. How does it compare to related terms such as disinformation and fake news? The simplest way to contrast *misinformation* with *disinformation* is to use the former as a broad category under which we can subsume the latter¹. In this light, disinformation amounts to *intentional* misinformation, i.e., the communication of false or misleading claims by actors who believe that such claims are false or misleading. As for fake news, I suggest that a news story qualifies as fake when (i) it contains false or misleading claims, (ii) it imitates the format of

traditional news media and (iii) its originators do not intend to communicate true claims². As false information can have a negative impact on the belief and behaviour of individuals whether or not it counts as *disinformation* or *fake news*, this entry focuses on misinformation *per se* as opposed to its subsets.

Different actors are involved in the spread of misinformation. To see this, consider how misinformation spreads online in a chronological manner. First, individuals or automated accounts (i.e. “bots”) create the content and originally publish it on social media. In general, the motive behind intentional misinformation is either political or financial. Foreign actors sometimes attempt to influence rival countries’ politics by diffusing false information which supports a particular cause or creates antagonism between citizens. In 2016, for instance, Russia’s Internet Research Agency created thousands of “sock puppet” social media accounts (that is, accounts based on false identities) which posed as radical American political groups and published fabricated articles in an attempt to destabilize the country. In other cases, originators of misinformation create and diffuse it to make financial gains. Indeed, misinformation often functions as clickbait, that is, content which is designed to trigger internet users’ interest and make them click on a thumbnail link. This generates traffic to the linked website, which itself translates into profit through advertising revenue.

Once misinformation has been created, posted and diffused, it is often relayed by inadvertent users who do not perceive it as misleading or, alternatively, know that it is misleading but nonetheless have reasons to share it. For instance, social media users might write posts about false rumours they have encountered online or simply “reshare” misleading posts and fake news articles. As Starbird (2019) underlines, misinformation campaigns often include “a majority of

‘unwitting agents’ who are unaware of their role” in amplifying content. That said, the role of bots in the spread of online misinformation should not go unnoticed. According to a recent study published in *Nature* (Shao et al., 2019), bots play a “disproportionate role in spreading articles from low-credibility sources,” especially in the early spreading moments before an article goes viral. One strategy bots are programmed to use amounts to targeting “users with many followers through replies and mentions” in order to maximize exposure. In a nutshell, the spread of misinformation on social media is usually a joint endeavour between humans who desire to misinform others, careless social media users and automated accounts.

2. Why worry about misinformation?

Although there is evidence that falsehood diffuses “significantly farther, faster, deeper, and more broadly than the truth” on social media (Vosoughi et al., 2017), it may still be the case that misinformation has little influence on the beliefs and reasoning process of its consumers³. Yet, if this influence turns out to be substantial, there are several reasons that count in favour of limiting its spread. This section briefly considers four of them.

Misinformation Prevents Democratic Citizens to Make Good Political Judgments

Scholars interested in freedom of expression often argue that the free flow of quality information is essential to democratic self-government as citizens need to form true beliefs in order to make good political judgments (Barendt, 2007, pp. 18–21). This is the rationale behind Alexander Meiklejohn’s (1961, p. 255) famous quote according to which the First Amendment protects “the freedom of those activities of thought and communication by which we govern.” From this perspective, it is quite easy to see why we might want to regulate misinformation. Here, the worry is that the worry is that forming false beliefs can impair people’s capacity to

make good political decisions. The assumption which undergirds such reasoning is that political judgments which exclusively rely on true beliefs are more likely to produce good normative outcomes than those that rely on false beliefs. A straightforward example of this is defamatory misinformation that besmirches electoral candidates. For instance, one frequently cited fake news story—commonly known as “Pizzagate”—connected Hilary Clinton and other well-known members of the Democratic Party to a fictitious child sex ring hidden in the basement of a Pizza restaurant in Washington, D.C. If one believes that a presidential candidate is a sex trafficker, one is significantly less likely to vote for her (and this independently of the quality of her policy proposals).

As some political theorists have suggested that (at least part of) the legitimacy of democracy derives from its ability to promote good normative outcomes (Estlund, 2007; Landemore, 2013), the worry that misinformation will prevent citizens from making good political judgments can also foster uncertainty about the value of democratic self-government (Brown, 2018). In 20th century political epistemology, the thought that misinformation undercuts the people’s ability to govern itself has been influential. The scholar who best expressed this worry remains Walter Lippman (1922, p. 249), who wrote that “under the impact of propaganda [...] it is no longer possible, for example, to believe in the original dogma of democracy”⁴.

Misinformation Makes Politics More Antagonistic

Misinformation also risks eroding civic friendship. To see this, consider echo chambers, that is, online spaces in which an inside group denies the status of potential truth tellers to outsiders through epistemic discrediting⁵. Simply put, epistemic discrediting amounts to assigning “some epistemic demerit, such as unreliability, epistemic maliciousness, or dishonesty” to others

(Nguyen, 2018, p. 6), and misinformation often plays a central role in it. If one's goal is to convince an audience to reject all claims made by outsiders, an available strategy is to denigrate the latter by falsely claiming that they have malicious intentions or committed evil deeds. This represents a second manner in which misinformation can hurt democracy: even if it does not directly lead citizens to make bad political judgments, it can strengthen epistemic structures which themselves fosters distrust between political, religious or racial groups⁶.

Interestingly, the spread of misinformation online might also be a *consequence* of political antagonism. In a study of partisan polarization on Twitter, Osmundsen et. al found that people who report hating their political opponents are the most likely to share fake news⁷. Their interpretation of this finding is that “when strong partisans decide what to share on social media, they place less value on the veracity of information as long as it satisfies their partisan motivations: derogating the out-party.” In other words, the spread of online misinformation might be partisan business as usual, a fact which itself suggests that polarization, echo chambers and the spread of low-quality information should be studied jointly⁸.

Misinformation Leads to Physical Harm

A third reason to fight online misinformation is that people can get hurt as a result of its spread. Consider, for instance, false rumours that circulate on private encrypted messenger services. On July 1, 2018, five men belonging to a nomadic group in India were beaten to death by a mob in Rainpada, India, after having handed a biscuit to a young girl (McLaughlin, 2018). Before the killings, rumours warning of kidnappers and organ-harvesters roving in the area had circulated on WhatsApp. As Mark Zuckerberg—the CEO of Facebook—himself recognized, fake news stories diffused by state-sponsored sock-puppet accounts have also played an important

role in the genocidal killings and displacement of the Rohingya people in Myanmar (Mozur, 2018). In such cases, photographs taken out of context are often mixed with hate speech that target minority groups. For instance, the Rainpada rumours included photos of lifeless children laid out in rows who were not the victims of a “murderous kidnapping ring” but rather of a “chemical attack on the town of Ghouta, Syria, five years earlier and thousands of miles away.” This demonstrates that misinformation sometimes works through what I propose to call *pictorial implicature*, that is, the act of showing a photograph that is a representation of something other than what is implied by the showing.⁹

Misinformation can also lead to harm when it fools people into taking decisions that compromises their health or that of others. Misleading information that fosters vaccine hesitancy by suggesting that vaccines cause autism is a clear example of this. Lastly, misinformation about climate change may turn out to be a deadly killer if we take the lives of future generations into accounts, for it risks slowing down attempts at mitigation and adaptation that could protect densely populated areas from catastrophic effects.

Misinformation Hinders Personal Autonomy

Lastly, misinformation sometimes hinders our capacity to pursue the goals we freely set for ourselves by impairing our capacity to appreciate reasons. Consider once again the case of parents who have been misinformed about vaccines. Presumably, such parents have a desire to protect their children’s health and, because of this, have a reason to have their children vaccinated (assuming that the vaccine went through normal phases of testing). Yet, misinformation may prevent them from appreciating this reason by suggesting that vaccines do not protect their children’s health, but in fact endanger it. This particular example illustrates two

general points about practical rationality and testimony. First, we need to appreciate reasons to effectively pursue our goals. Second, we continuously rely on information provided to us by others in order to appreciate reasons. When our informational landscape is polluted by claims that are false or misleading, however, our capacity to appreciate reasons is impaired and our desires are more likely to be frustrated as result.

A different way of making this claim is to point out that misinformation undermines personal autonomy. On a reasons-responsive account of autonomy, the level of personal autonomy a person enjoys depends on her capacity to appreciate reasons. As Buss and Westlund (2018) underline:

If doing *Y* is constitutive of doing *Z*, then if I authorize myself to be moved by the desire to do *Y* because I mistakenly believe that doing *Y* is a way of *not* doing *Z*, then there is an obvious sense in which I have not authorized myself to do what I am now doing when I am moved by the desire to do *Y*.

The vaccine sceptic example—in which a person considers that having their children vaccinated is a way of *not* protecting their children’s health—is a simple illustration of this idea. Arguments in favour of regulating misinformation that focus on autonomy will be especially appealing to Kantian deontologists. In Kant’s (2012) view, lying is morally wrong precisely because it robs others of their freedom to choose rationally. It is also quite easy to see how intentionally deceiving others often amounts to treating them as mere means rather than as ends. If I trick you into doing something that furthers my interest through deliberate misinformation, then I use you as a tool that can help me better fulfil my desires while disregarding your need to rely on quality information in order to satisfy your own¹⁰.

3. How should misinformation be regulated?

Assuming that we have reasons to limit the spread of online misinformation, what regulatory options are available to us? In the remainder of this chapter, I assess three ways of doing so. A first option amounts to allowing or helping internet users modify their online habits to minimize the chance that they will form false beliefs on the basis of misleading claims. A second option is to let social media companies modify their platform so that fewer people are exposed to it. A third option is to demand that governments enact laws that punish individuals who wittingly engage in misinformation or disincentivize them from doing so.

Individual Self-Regulation

Do internet users have a duty to inoculate themselves against online misinformation? Given that misinformation and epistemic discrediting often go together, should they not exclude all individuals who defame others from their epistemic network? Moreover, do they have the epistemic responsibility to ensure that information to which they are exposed is reported by multiple sources individuals before allowing it to influence their beliefs? Individual self-regulation is an appealing option to philosophers who consider that individuals are primarily responsible for their poor epistemic habits, and that it is unfair to impose the costs of misinformation on others. In *Know-it-all Society* (2019, p. 4), for instance, Michael Lynch criticizes our tendency to “dismiss evidence for victory and truth for power” in online contexts. In his view, our arrogance and carelessness with regards to truth is not a technical problem, but a human one: “if we want to solve it [...] we have to change our attitudes” (2019, p. 5).

Self-regulation is arguably the less costly option in terms of both financial resources and liberties compared with platform-based innovations (which entail financial costs for technology

companies) and government regulation (which carries the risk of censorship). Yet, some deep-seated psychological tendencies appear to stand in the way of cognitive self-improvement in the fight against misinformation¹¹. For instance, fake news is often published on relatively unknown websites and psychological studies suggest that “when people know little about a source, they treat information from that source as credible” (Rapp, 2016). A second relevant psychological tendency that can limit our attempt to immunize ourselves against misinformation is misattribution. Put simply, people sometimes initially categorize some claims as false because they encountered them in (what they know to be) a fake news article, but eventually misattribute this claim to a reliable source and re-categorize such claims as true (Marsh, Cantor & Brashier, 2016).

The very architecture of social media platforms also presents some obstacles for users who desire to improve their epistemic habits. Indeed, even internet users who make the conscious choice to become intellectually virtuous and reduce their consumption of misinformation might struggle to do so as a result of algorithmic filtering. If such users have a history of consuming conspiracy theories stemming from sources of dubious quality, for instance, such content will continue to appear in their feed for the near future. Even if they stop engaging with such theories or report them as misinformation, they may still be exposed to misleading stories if their epistemic network is mainly composed of users who like sharing them.

Of course, exposure is not all there is to the story. When individuals have formed false beliefs as a result being exposed to misinformation, can they not correct them at a later point in time? Empirical evidence about corrections is mixed. While meta-analyses of fact checks suggest that they improve the accuracy of belief (Chan et al., 2017), a central worry is that internet users will

not be exposed to them. Consider first what psychologists have dubbed the confirmation (or myside) bias. People have a tendency to seek “evidence in ways that are partial to existing beliefs” and they are therefore not likely to attempt to disconfirm their false beliefs by seeking corrections (Nickerson, 1998). This hypothesis is supported by the finding that corrections almost never reach their targeted audience. For instance, Guess, Nyhan and Reifler (2020, p. 476) found that less than “half of the 44.3% of Americans who visited an untrustworthy website” during the length of their recent research “also saw any fact check from one of the dedicated fact-checking websites.”

Moreover, people are generally overconfident and have a “bias blind spot,” that is, tend to rate themselves as less susceptible to cognitive biases compared to the average individual and their peers (Pronin, Lin and Ross, 2002). The worry here is that this will prevent them from seeking new information that corroborates or contradicts their beliefs. Nathan Ballantyne has also argued that we often engage in *debunking reasoning*. When others present us with evidence that contradicts our opinions, when tend to attribute biases to them “in order to prevent their dissent from lowering our confidence in our view” (2015, p. 142). This has led some epistemologists to argue that the prospects for epistemic self-improvement are dire (Ahlstrom-Vij, 2013).

If self-motivated cognitive improvement is trickier than it appears at first sight, can we not foster such self-improvement through external constraints such as media literacy education programs? Some philosophers have recently suggested that we teach intellectual virtues such as scepticism and intellectual humility to middle school and high-school students to decrease the likelihood that misinformation will affect their beliefs (Brown, 2019b). For instance, they could

be tasked with identifying instances of reliable journalism and misinformation as well as situations in which epistemic discrediting impairs the rational evaluation of claims and arguments. Admittedly, the empirical evidence which supports the claims that critical thinking or intellectual virtue education is an effective solution to misinformation is not robust. One encouraging finding comes from Rozenbeek and van der Linden (2019), who designed a fake news game that “confers psychological resistance against misinformation.” In the game, players take on the role of a fake news producer who employs different strategies to deceive its audience: “polarisation, invoking emotions, spreading conspiracy theories, trolling people online, deflecting blame, and impersonating fake accounts”¹². According to the researchers, playing the game “significantly reduced the perceived reliability of tweets that embedded several common online misinformation strategies.” In general, it would be relatively simple for instructors to integrate modules on misinformation developed by cognitive psychologists to their teaching (provided that they have access to the right technology). Even if such modules do not fulfil their promises in the end, the costs of *attempting* to inoculate students against misinformation remain low.

Platform-Based Innovations

A second regulatory option is to let social media companies modify their platform so that fewer people are exposed to misinformation or, at least, exposed to it without knowing that they are. Such an option will be appealing to those who consider that technology companies are partly responsible for the spread of online misinformation and should take on the burden of eliminating it. Consider two different ways to conceive of social media platforms. According to the first conception, such platforms are akin to public squares. If someone spreads false information while

standing on a public square, it would be strange to hold urban planners who designed the square responsible for this person's behaviour. After all, the very function of public squares is to allow people to publicly express their views, and we normally judge that people are responsible for what they say. Yet, social media platforms do not truly function like public squares. One important difference is that unlike neutral spaces in which people can express their views, they promote and demote content through algorithmic filtering. This gives us reason to consider that social media companies are more akin to publishers. According to this second conception, social media companies design algorithms which determine what content individual users encounter just as an editor decides what content will be published on the front page of a newspaper. If a fake news story finds itself at the top of millions of users' feed, it therefore seems reasonable to claim that social media companies share at least part of the responsibility for this.

What can technology companies concretely do to limit the spread of online misinformation? We can situate the options available to them on a paternalism continuum. At the least paternalistic end of the continuum are transparency requirements. Indeed, platforms can provide users with information about the content to which they are exposed. When users encounter a news story on Facebook, for instance, they have the opportunity to click on a context button and learn more about the source which published the article. The social media giant has also announced that it will start labelling content "that has been rated false or partly false by a third-party fact-checker." (Rosen et al., 2019). For its part, Twitter has introduced labels to help users identify misleading information and disputed claims¹³. A more radical proposal comes from philosopher Regina Rini (2017), who suggests that social media platforms give reputation scores to their users. Such scores "could be displayed in a subtle way, perhaps with a colored icon beside

user photos.” Users could then choose to have content posted from users with low reputations scores downranked or filtered out, but they would not be obligated to do so. This innovative proposal is worth considering, but it could be exploited by sophisticated deceivers. For instance, an organization could gradually build an excellent reputation score and then use it as leverage to launch a misinformation campaign that may become even more effective than if epistemic trustworthiness scores had not been made available in the first place.

Moving up the paternalism spectrum, social media companies could remove or demote content reported as misinformation by their community of users or by third-party fact-checkers. For instance, Facebook both employs fact-checkers and allows its users to report news stories they encounter as “false news.” While removing and demoting misleading content is more paternalistic than simply increasing transparency, it is arguably not more paternalistic than how social media platforms function generally given that *all* content to which social media users are exposed is ranked by algorithms. There is, however, an important risk tied to this measure. When users expect that false information will be flagged, false information that *fails* to be flagged can be seen as more accurate than it would otherwise have been. This is what Pennycook et al. (2020) call the “implied truth effect.” Another moderately paternalistic measure amounts to redirecting users to quality information when they are about to engage with news stories. During the coronavirus outbreak, for instance, Twitter has combated rumours that oregano oil proves effective against coronavirus by steering U.S. users searching for coronavirus-related hashtags to the website of Center for Disease Control and Prevention.

As for pragmatic challenges that social media companies face in their attempts to fight misinformation, one worry is that the amount of third-party fact checkers they currently employ

pales in comparison of the vast amount of content that needs to be evaluated. In January 2020, Rodrigo (2020) estimated that Facebook's six fact-checking partners "have 26 full-time staff and fact-checked roughly 200 pieces of content per month," an amount which many experts find insufficient. To solve this problem, social media giants can rely on user reports, but these risk not being accurate. For instance, nothing prevents a user to report a perfectly accurate news story which contradicts its ideological beliefs as "false news." What is more, letting social media companies self-regulate is unlikely to lead them to enact measures that would significantly hurt their business model, which consists of promoting content in exchange for advertising revenues. In 2020, Facebook has faced severe criticisms for refusing to remove political advertisements that have been found to contain false claims by its third-party fact-checking partners¹⁴. This explains why some governments have recently decided to step in the fight against misinformation.

Third Option: Governmental Action

By far, the most controversial way to regulate the spread of online misinformation is through law and policy, for speech restrictions always come with the reasonable worry that public officials will abuse them or that they will amount to undue censorship. Nevertheless, some Western European countries have chosen this route. In Germany, the Network Enforcement Act allows the government to impose fines of up to 50 million euros on social media companies if they fail to remove illegal content including fake news. In late 2018, the Macron government in France also passed the "law against the manipulation of information," which authorizes judges to order the removal of pieces of misleading content from social media platforms in a period of three months leading up to a national election. Such a law was heavily criticized by the French Senate on the grounds that it amounts to a violation of citizens' right to freedom of expression.

Let us consider this objection in more detail. To refute it, one might point out that influential theories of freedom of expression rest on the idea that (at least part of) the value of speech derives from its ability to promote truth. As misinformation and fake news are unlikely to do so, this provides us with reason to believe that they do not amount to the kind of speech that we ought to protect in liberal democracies. Yet, these arguments are unlikely to convince Millians who consider that truth is more likely to emerge when we refrain from regulating speech and let it collide with error. This is an empirical claim that remains difficult to assess, but disagreement over it may yield radically different views about the limits of freedom of expression. Proponents of the so-called autonomy defense of free speech are also likely to oppose the French law¹⁵. Even if the legal prohibition of misinformation proved effective at hindering the formation of false beliefs, they would condemn it on the grounds that governmental attempts to preselect the evidence on which citizens base their judgments are unduly paternalistic and incompatible with our right to autonomy (Scanlon, 1972).

Beyond free speech worries, it is worth asking whether the French law is likely to prove effective at reducing the amount of misinformation which circulates on social media. Let us imagine a situation in which judges mostly defer to judgments made by third-party fact checkers when deciding which pieces of content should be removed. As we have seen, fact checkers may not be available in sufficient numbers to review the vast amount of content that might amount to misinformation. Of course, the situation would only worsen if judges themselves took up the task of evaluating content. What is more, even if fact checkers and judges identified most misleading content, there is no guarantee that they would be able to do so *before* pieces of misinformation have spread and influenced users' beliefs.

Laws against misinformation have also been proposed in the American context. For instance, Senator Mark Warner (2018) has proposed to pass laws that require social media companies to “clearly and conspicuously identify and label bots” as well as accounts based on false identities. An alternative proposal is to make social media platforms liable for state-law torts (defamation, false light, public disclosure of private facts, etc.)¹⁶. Such a law would provide technology companies with a strong incentive to remove deep fakes (the content of which is generally defamatory), but once again presupposes that such companies have the resources to review an enormous amount of content.

Before concluding, let us therefore consider a proposal that avoids this pitfall. Unfortunately, social media provide organizations that willfully engage in misinformation with highly effective tools. Perhaps the most effective is targeted advertising, which allows organizations to ensure that their posts will be seen by individuals who share specific characteristics with each other (age, gender, geographical location, political orientation, etc.). Here, the underlying assumption is that specific instances of misinformation will be of interest to a particular demographic. Indeed, it seems reasonable to suppose that a fake news story which defames high-ranking Democrats is more likely to be believed by Republicans who have shown interest for conspiracy theories than by Democrats who have not. For this reason, one way to decrease the efficacy of misinformation without having to review specific instances of it would be to prohibit targeted advertising altogether, that is, to make it impossible for originators of misinformation to direct it at the individuals they think are the most likely to believe it (or be interested by it). While a prohibition against targeted advertising might be a step in the right direction, it would not be a panacea, for there are myriad other means by which misinformation can spread and influence people’s beliefs.

Perhaps the most daunting way in which it does is by permeating closed and encrypted communicative spaces such as WhatsApp groups, which remain extremely difficult to monitor and, as we have seen, can spread rumours which lead to physical harm. Such a prohibition would also significantly hurt the business model of social media platforms.

Is misinformation a serious enough public issue for the government to interfere with the freedom of private businesses to organize and operate for profit in a competitive market? My final suggestion is that we will not be able to answer such a question before answering some of those considered in this entry. First, how likely is it that misinformation will influence our beliefs? Second, assuming that it does, to what extent do false beliefs threaten our capacity to be safe, to pursue our personal goals and to fulfil our democratic duties? Third, what other regulatory options are available to us? Regarding the latter, it remains possible that rapid advances in machine learning will make regulations against misinformation superfluous. If we develop artificial agents that are capable of swiftly and accurately identify pieces of misinformation, it will become possible for technology companies to use them and better review the vast amount of information (and misinformation) which currently circulates on their platform¹⁷. Until they do, however, it is likely that worries about misinformation will incite political epistemologists to doubt the people's capacity to self-govern. As we have seen, rapid advancements in communications technology usually give rise to Lippmannian misgivings about democracy. This is why philosophical reflection upon misinformation ultimately matters. In the end, assessing the political dangers posed by false beliefs can lead us to regain confidence in the rule of the many or—if we fail to do so—to seek viable alternatives to it.

NOTES

¹ This broad definition of misinformation entails that honest mistakes, factual errors in reporting and even some instances of political satire count as instances of it. I do not see this as a problem if we use narrower concepts (disinformation, fake news, etc.) to refer to different kinds of misinformation. If we define misinformation broadly, not all forms of misinformation will be legitimate targets of regulation. For instance, some might consider that it would be a mistake to attempt to regulate satirical news website like *The Onion* on the grounds that contributors to the website have no wrongful intention. That said, communication scholars have recently argued that we should at least label satirical misinformation as such on the grounds that “too many people think satirical news is real.” See Garrett, Bond and Poulsen (2019).

² Some philosophers have offered definitions of fake news which focus on the reliability of the production process instead of on the intention of publishers. See, for instance, Michaelson, Pepp and Sterken (2019). For an extensive discussion of the definition of fake news, see Axel Gelfert’s chapter in this handbook.

³ For an extensive discussion of psychological studies which suggest that misinformation can be effective, see Levy (2017) as well as Levy and Ross’s chapter in this handbook (“The cognitive science of fake news”).

⁴ On the relationship between propaganda and misinformation, see Megan Hyska’s chapter in this handbook.

⁵ On epistemic bubbles and echo chambers, see Hana Kiri Gunn’s chapter in this handbook.

⁶ Highly selective one-sided information that does not contain false claims can also make politics more antagonistic, and my claim here is not that we should worry about misinformation *more than we do* about other types of political communication that risk pitting citizens against each other like spin, dog-whistles or straightforward insults (“crooked Hilary,” etc.).

⁷ I thank Michael Hannon and Jeroen de Ridder for drawing my attention to this study and, more generally, for their insightful remarks on this chapter.

⁸ It is worth noting that “making politics more antagonistic” is a contingent feature of misinformation, not a necessary one. In fact, misinformation could even be used to *lessen* partisan antipathy. For instance, political actors who seek to create alliances between groups that have traditionally shown hostility toward each other could falsely portray the members of one group as having committed heroic deeds that benefit the other group.

⁹ This “something else” can be a real event that happened in a different context (as in the Rainpada case) or a fictional event.

¹⁰ See *On a supposed right to lie from philanthropy* in the Cambridge edition of Kant’s *Practical Philosophy* for his influential discussion of lying.

¹¹ Levy and Ross’s contribution to this handbook offers a more detailed discussion of this point.

¹² Some limitations of this study should be pointed out. First, it has not been replicated. Second, it is hard to tell how long the “inoculation effect” will last. Third, the sample of experimental subjects was self-selected.

¹³ According to Twitter’s policy, misleading information (but not disputed claims) can be removed from the platform if its propensity for harm is considered severe rather than moderate. In May 2020, Twitter applied this “propensity for harm” criterion to a tweet by Donald Trump (in which he remarked that “when the looting starts, the shooting starts”) by labeling it as “glorifying violence.”

¹⁴ This policy stands in stark contrast with Twitter’s 2019 to ban all political advertisements from its platform. That said, both Facebook and Twitter have banned content featuring politicians making false claims about COVID-19 that were deemed harmful. In August 2020, for instance, Facebook removed a video interview with Donald Trump in which he claimed that children are “almost immune” to the virus.

¹⁵ See (Brison, 1998) for a discussion of the autonomy defense of free speech.

¹⁶ Section 230 of the Communications Decency Act currently immunizes social media platforms from state tort and criminal liability in the United States.

¹⁷ Andersen and S e (2019) have argued that we should refrain from using automated fact-checking on the grounds that it would be difficult (if not impossible) for artificial agents to understand the intentions behind posts, and the precise meaning of such posts.

REFERENCES

- Ahlstrom-Vij, K. (2013) 'Why We Cannot Rely On Ourselves For Epistemic Improvement', *Philosophical Issues*, 23(1), pp. 276–296.
- Andersen, J. and Sjøe, S. O. (2019) 'Communicative Actions We Live By: The Problem with Fact-Checking, Tagging or Flagging Fake News – The Case of Facebook', *European Journal of Communication*, 35(2), pp. 146–139.
- Barendt, E. (2007) *Freedom of Speech*. 2nd Edition. New York: Oxford University Press.
- Brisson, S. J. (1998) 'The Autonomy Defense of Free Speech', *Ethics*, 108(2), pp. 312–339.
- Ballantyne, N. (2015) 'Debunking Biased Thinkers (Including Ourselves)', *Journal of the American Philosophical Association* 1(1), pp. 141–162.
- Brown, É. (2018) 'Propaganda, Misinformation, and the Epistemic Value of Democracy', *Critical Review* 30(3–4), pp. 194–218.
- Brown, É. (2019a) "'Fake News" and Conceptual Ethics', *Journal of Ethics and Social Philosophy*, 16(2), pp. 144–154.
- Brown, É. (2019) 'Civic Education in the Post-Truth Era: Intellectual Virtues and the Epistemic Threats of Social Media', in Macloed, C. and Tappolet, C. (eds) *Philosophical Perspectives on Education: Shaping Citizens*. New York: Routledge, pp. 45–67.
- Buss, S. and Westlund, A. (2018) 'Personal Autonomy', in Zalta, E. N. (ed) *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/personal-autonomy>.
- Carrieri, V., Madio, L. and Principe, F. (2019) 'Vaccine Hesitancy and (Fake) News: Quasi Experimental Evidence From Italy', *Health Economics*, 28(11), pp. 1377–1382.
- Chan, M. P. S., Jones, C. R., Jamieson, K. H. and Albarracín, D. (2017) 'Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation', *Psychological Science* 28(11), pp. 1531–1546.
- Estlund, D. M. (2008) *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press.
- Garrett, R. K., Bond, R. and Poulsen, S. (2019) 'Too Many People Think Satirical News Is Real', *The Conversation*. Available at: <https://theconversation.com/too-many-people-think-satirical-news-is-real-121666>.
- Guess, A. M., Nyhan, B. and Reifler, J. (2020) 'Exposure to Untrustworthy Websites in the 2016 Election', *Nature Human Behaviour* 4(5), pp. 472–480.

-
- Kant, I. (2012) *Practical Philosophy*, Gregor, M. J. (ed), Cambridge: Cambridge University Press.
- Landemore, H. (2013) *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton: Princeton University Press.
- Levy, N. (2017) 'The Bad News About Fake News', *Social Epistemology Review & Reply Collective* 6(8), pp. 20–36.
- Lippman, W. (1922) *Public Opinion*. New York: Harcourt, Brace and Company.
- Lynch, M. P. (2019) *Know-It-All Society: Truth and Arrogance in Political Culture*. New York: Liveright Publishing Corporation.
- Marsh, E. J., Cantor, A. D., and Brashier, N. M. (2016) 'Believing that Humans Swallow Spiders in Their Sleep: False Beliefs as Side-Effects of the Processes that Support Accurate Knowledge', *Psychology of Learning and Motivation* 64, pp. 93–132.
- Meiklejohn, A. (1961) 'The First Amendment Is an Absolute', *The Supreme Court Review* 1961, pp. 245–266.
- Mclaughlin, T. (2018) 'How WhatsApp Fuels Fake News and Violence in India', *Wired*. Available at: www.wired.com/story/how-whatsapp-fuels-fake-news-and-violence-in-india.
- Michaelson, E., Pepp, J. and Sterken, R. K. (2019) 'What Is New About Fake News?', *Journal of Ethics and Social Philosophy* 16(2), pp. 67–94.
- Mitchell, A., Gottfried, J., Stocking, G., Walker, M. and Fedeli, S. (2019) 'Many Americans Say Made-Up News Is a Critical Problem That Needs To Be Fixed', *Pew Research Center*. Available at: www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed.
- Mozur, P. (2018) 'A Genocide Incited on Facebook With Posts From Myanmar's Military', *The New York Times*. Available at: www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.
- Nguyen, C. T. (2018) 'Echo Chambers and Epistemic Bubbles', *Episteme* 17(2), pp. 1–21.
- Nickerson, R. S. (1998) 'Confirmation Bias: A Ubiquitous Phenomenon in Many Guises', *Review of General Psychology* 2(2), pp. 175–220.
- Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A. and Petersen, M. B. (2020) 'Partisan Polarization Is the Primary Psychological Motivation Behind "Fake News" Sharing on Twitter', *PsyArXiv*. Available at: <https://psyarxiv.com/v45bk/>.
- Pennycook, G., Bear, A., Collins, E. T., and Rand., D. G. (2020) 'The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings', *Management Science*. Available at: <https://pubsonline.informs.org/doi/pdf/10.1287/mnsc.2019.3478>.

-
- Pronin, E., Lin, D. Y. and Ross, L. (2002) 'The Bias Blind Spot: Perceptions of Bias in Self Versus Others', *Personality and Social Psychology Bulletin* 28(3), pp. 369–381.
- Rapp, D. N. (2016) 'The Consequences of Reading Inaccurate Information', *Current Directions in Psychological Science* 25(4), pp. 281–285.
- Rini, R. (2017) 'Fake news and Partisan Epistemology', *Kennedy Institute of Ethics Journal* 27(2), pp. 43–64.
- Rodrigo, C. M. (2020) 'Critics fear Facebook fact-checkers losing misinformation fight', *The Hill*. Available at: <https://thehill.com/policy/technology/478896-critics-fear-facebook-fact-checkers-losing-misinformation-fight>.
- Rosen, G., Harbath, K., Gleicher, N. and Leathern, R. (2019) 'Helping to Protect the 2020 US Election', *Facebook*. Available at: <https://about.fb.com/news/2019/10/update-on-election-integrity-efforts>.
- Rozenbeek, J., and van der Linden S. (2019) 'Fake news game confers psychological resistance against online misinformation', *Palgrave Communications* 5(65). Available at: www.nature.com/articles/s41599-019-0279-9.
- Scanlon, T. (1972) 'A Theory of Freedom of Expression', *Philosophy & Public Affairs* 1(2), pp. 204–226.
- Shao, C., Giovanni L. C., Varol, P., Yang, K-C., Flammini, A. and Menczer, F. (2018) 'The Spread of Low-Credibility Content By Social Bots', *Nature Communications* 9(4787). Available at: <https://www.nature.com/articles/s41467-018-06930-7.pdf>.
- Starbird, K. (2019) 'Disinformation's Spread: Bots, Trolls and All of Us', *Nature* 571 (7766), p. 449. Available at: <https://media.nature.com/original/magazine-assets/d41586-019-02235-x/d41586-019-02235-x.pdf>.
- Vosoughi, S., Roy, D. and Aral, S. (2018) 'The Spread of True and Fake news Online', *Science* 359(6380), pp. 1146–1151.
- Warner, M. R. (2018) 'Potential Policy Proposals for Regulation of Social Media and Technology Firms'. Available at: www.warner.senate.gov/public/_cache/files/d/3/d32c2f17-cc76-4e11-8aa9-897eb3c90d16/65A7C5D983F899DAAE5AA21F57BAD944.social-media-regulation-proposals.pdf.