



<http://social-epistemology.com>
ISSN: 2471-9560

Misinformation and the Limits of Individual Responsibility

Boyd Millar, Washington University in St. Louis. millar.boyd@gmail.com

Millar, Boyd. 2021. "Misinformation and the Limits of Individual Responsibility." *Social Epistemology Review and Reply Collective* 10 (12): 8-21. <https://wp.me/p1Bfg0-6kD>.

Abstract

The issue of how best to combat the negative impacts of misinformation distributed via social media hangs on the following question: are there methods that most individuals can reasonably be expected to employ that would largely protect them from the negative impact that encountering misinformation on social media would otherwise have on their beliefs? If the answer is “yes,” then presumably individuals bear significant responsibility for those negative impacts; and, further, presumably there are feasible educational remedies for the problem of misinformation. However, I argue that the answer is “no.” Accordingly, I maintain that individuals do not bear significant responsibility for the negative impacts at issue; and, further, I maintain that the only effective remedies for the problem of misinformation involve changing the information environment itself.

1. Introduction

Misinformation is ubiquitous on social media;¹ and this ubiquity of misinformation on social media has significant negative consequences. For instance, a recent study found that in the lead-up to the 2020 U.S. Presidential election, media pages trafficking in misinformation were averaging hundreds of millions of views a month on Facebook; in fact, this study found that in July and August of 2020, such pages outperformed the most popular traditional news media pages in terms of engagement.² And, plausibly, the sheer quantity of false claims about election fraud circulating on social media has had a significant impact on Americans’ false beliefs regarding election integrity: almost a year later, 35% of registered voters said that the results of the 2020 presidential election should probably or definitely be overturned.³

Misinformation regarding COVID-19 provides perhaps an even more distressing example. Claims about miracle cures, and imagined vaccine risks, have spread incredibly widely on social media.⁴ And research has shown that being exposed to COVID-19 misinformation has significant impacts on what individuals believe and how they behave—for instance, by increasing vaccine hesitancy.⁵

Very generally, there are two approaches to the enormously important question of how best to combat the negative impacts of misinformation distributed via social media. According to what Michel Croce and Tommaso Piazza (2021, 2) label *educational approaches*, we should focus on individual information consumers: we should design interventions—public information campaigns, changes to school curriculum, etc.—that will bring about relatively long-term changes to individuals’ behavior or cognitive traits so that they are better equipped to face the misinformation they are guaranteed to encounter on social media. Alternatively,

¹ See Vosoughi, Roy, and Aral (2018), and Bradshaw et al. (2020).

² Avaaz (2021).

³ Morning Consult/Politico (2021).

⁴ Avaaz (2020).

⁵ Loomba et al. (2021).

according to what Croce and Piazza (2021, 2) label *structural approaches*, we should focus on the information environment itself: social media platforms should be modified so as to limit the distribution of misinformation, or to minimize the impact it has on ordinary users.

These two approaches are obviously compatible—there is no particular reason we could not employ *both* educational and structural remedies. However, whether educational remedies are worth pursuing depends on the extent to which individual information consumers are responsible for the problem. For present purposes, then, the crucial question is:

Are there methods that most individuals can reasonably be expected to employ that would largely protect them from the negative impact that encountering misinformation on social media would otherwise have on their doxastic attitudes?

If the answer is “yes,” then presumably individuals bear significant responsibility for the negative impacts that exposure to misinformation via social media has on the accuracy of their doxastic attitudes (since there are reasonable steps they could have taken that would have avoided those impacts); and, further, it’s natural to assume that educational approaches will help significantly (so long as there are feasible ways to train or encourage people to take the requisite reasonable steps). But if the answer is “no,” then presumably individuals do not bear significant responsibility for the negative impacts that exposure to misinformation via social media has on the accuracy of their doxastic attitudes; and, further, it’s natural to reject educational approaches as unhelpful. If individuals are already doing everything they can reasonably be expected to do, attempting to make them better information consumers serves little purpose.

I will defend a negative answer to the question at issue—that is, I will argue that it’s not the case that there are methods that most individuals can reasonably be expected to employ that would largely protect them from the negative impact that encountering misinformation on social media would otherwise have on their doxastic attitudes. I defend this claim by considering and rejecting a number of proposed methods of the relevant sort.

In §2, I consider a recent proposal from Croce and Piazza (2021): social media users can reverse the negative impacts of encountering misinformation by consuming more traditional, high-quality news media. I argue that there is overwhelming evidence from psychology that encountering additional accurate information will not largely undo the negative impact of encountering inaccurate information.

In §3, I consider alternative proposals that focus on changing the way that individuals respond to the social media content they consume. I argue that by no means would the proposed strategies largely protect individuals from the negative impacts of encountering misinformation. Ultimately, while we should grant that some educational interventions could help lessen the negative impact of misinformation distributed via social media, we should insist that any relevant improvements would be distinctly marginal—so marginal, in fact, as to make most any educational intervention hardly worth the expense and effort.

There are a few points I should clarify before proceeding. First, as in Millar (2019), I will not focus exclusively on stereotypical *fake news*—my topic is the much broader category of

misinformation. By *misinformation* I mean false claims that can be demonstrated to be false on the basis of publicly available evidence. (Misinformation can also be recognized by means of a reference-fixing description: false claims of the sort that elicit widespread agreement amongst professional fact-checking organizations.) In addition, as is Millar (2019), I will not focus exclusively on outright belief of misinformation—my topic is the negative impact that encountering misinformation has on the accuracy of one’s doxastic attitudes more generally. Accordingly, I should note in advance that Croce and Piazza are explicitly concerned with a narrower topic: “sharing and believing fake news” (2021, 1). Finally, I should emphasize that while the present paper is exclusively concerned with social media and similar internet technologies, I am not assuming that misinformation distributed via social media does more harm than misinformation distributed via traditional media.⁶

2. Croce and Piazza’s Proposal

One might think that there is a simple method guaranteed to protect individuals from the negative impact of encountering misinformation on social media that anyone can employ: just avoid using social media altogether. Such a method would be sure to succeed: if you never use social media you will never encounter any of the misinformation that is ubiquitous on social media. However, the question at issue is whether there is some method of avoiding the negative impact of misinformation on social media that most individuals can reasonably be expected to employ; and most people can’t reasonably be expected to eliminate their social media use.⁷ First, Facebook, Twitter, and other similar internet technologies, are presently the cheapest and most convenient sources of information; and most people can’t reasonably be expected to expend the time and energy that would be required to obtain information from traditional media exclusively. Second, and more importantly, such platforms have become one of the principal means by which individuals stay connected to friends and loved ones; and most people can’t reasonably be expected to significantly impair their relationships with all other human beings in order to avoid being exposed to misinformation.

Croce and Piazza’s (2021, 6-8) proposal avoids these difficulties by demanding much less of individual social media users. According to them, individuals do not need to avoid social media altogether in order to counteract the negative impacts of encountering misinformation; instead, they need only supplement their social media use by adopting a more “varied information diet.” Specifically, Croce and Piazza claim that individuals can protect themselves from the misinformation they encounter on social media by widening “their sources of information, at least by consuming news from traditional media such as newspapers and magazines (in their print or online versions), books, TV and radio newscasts” (2021, 6).⁸ Later, they add that the relevant traditional sources must also conform

⁶ For discussion, see Benkler, Faris, and Roberts (2018).

⁷ See Millar (2019, 532).

⁸ Sindermann, Cooper, and Montag (2020, 46) briefly discuss a similar proposal. Coady defends a related conclusion: he maintains that one reason not to worry about the amount of fake news online is that “people now have more resources available to them to evaluate the veracity of information they come across” (2019, 51).

to well-established journalistic norms—so, for instance, Fox News is disqualified by its blatant partisanship (2021, 7).

But why should regularly consuming at least some high-quality news content protect individuals from the misinformation they encounter on social media? Croce and Piazza suggest two mechanisms. First, the more traditional news content you consume, the more likely you will be to obtain counterevidence that debunks the misinformation you have encountered on social media. And, second, in cases where you do not find some previously-encountered piece of misinformation being debunked in the mainstream news media, the mere fact that it is not being reported by traditional news outlets provides you with evidence that the relevant piece of misinformation is indeed false (Croce and Piazza 2021, 6).⁹ Accordingly, if Croce and Piazza are correct, then there may be promising educational remedies for the problem of misinformation: perhaps interventions can be designed to educate “social media users about the epistemic benefits of a varied information diet” and thereby encourage them to “manage their online epistemic conduct more responsibly” (2021, 8).

However, there are compelling empirical reasons to reject Croce and Piazza’s proposal: the psychological evidence suggests that the negative impacts that encountering falsehoods has on the accuracy of our doxastic attitudes cannot be largely corrected by exposure to accurate information. First, consider cases where you encounter misinformation on social media that you end up believing as a result. Given the prevalence of misinformation on social media, you might end up believing a particular demonstrable falsehood for any number of different reasons. Because the information you encounter on social media is filtered by personalization algorithms, you will often encounter false claims that are strongly supported by your existing beliefs; and you are also likely to encounter the same false claims repeatedly. In addition, many of the demonstrable falsehoods you encounter will receive some persuasive social endorsement: the information might come from some trusted individual or organization, and it might be endorsed by very many individuals belonging to groups with which you identify. Suppose, then, that due to a combination of such factors, while using social media you come to believe that voter fraud was widespread in the 2020 Presidential election, or that COVID-19 vaccines are dangerous and ineffective. And suppose that, next, you decide to follow Croce and Piazza’s advice—you spend some time reading *The New York Times*, or watching CNN.

With respect to the accuracy of your doxastic attitudes, the best-case scenario is one in which you happen to encounter a story that thoroughly debunks the misinformation you’ve come to believe. However, even in this best-case scenario, the misinformation’s negative impact isn’t likely to be largely reversed. First, under certain conditions, your false belief will simply persist in the face of the correction you’ve encountered: for instance, if you trust the

⁹ An additional mechanism that Croce and Piazza don’t mention: the more high-quality news content you consume, the more true beliefs you will form about current events; and the more true beliefs you form about current events, the better able you will be to prevent misinformation on social media from influencing our doxastic attitudes by comparing purported news stories to your store of existing beliefs. For a review of the evidence concerning *plausibility checking*, see Mercier (2017, 104-105).

source of the original false claims more than you trust the source of the correction; or, if the falsehood at issue is closely connected to your deeply held political or social beliefs.¹⁰

Second, while the most likely outcome of encountering the correction is that your doxastic attitudes will become more accurate, this improvement is likely to be only *partial* and *temporary*. Research has shown that individuals who read a correction of some falsehood tend to have more accurate doxastic attitudes than they would have if they hadn't read the correction—but not so accurate as they would have if they had never encountered the falsehood in the first place.¹¹ Moreover, research has also shown that even this partial improvement tends to diminish rather quickly.¹² Consequently, even in this best-case scenario, the false beliefs regarding voter fraud or COVID-19 vaccines that you acquired using social media are not likely to be largely eliminated.

Relatedly, encountering and accepting a correction of previously-encountered misinformation does not somehow erase or expunge the relevant false information from memory. Instead, even when you accept a correction, the false information that you previously encountered but now reject continues to influence your doxastic attitudes—a phenomenon known as the *continued influence effect*.¹³ Recently, Gordon et al. (2019) attempted to determine the mechanism behind this phenomenon using neuroimaging: they found that when you encounter a correction of false information, rather than simply being deleted from the brain, the false information remains stored in memory and continues to be activated when you deliberate on relevant topics. Such research suggests that in cases where you happen to encounter a correction of a piece of misinformation, and where you find that correction entirely convincing, the relevant misinformation will continue to have a systematic negative impact on the accuracy of your doxastic attitudes.

Neither can we assume that the best-case scenario will be the typical scenario. At least very often, when you come to believe some piece of misinformation you've encountered on social media, you won't find the relevant claims covered by the mainstream news outlets you consult. According to Croce and Piazza, in such cases, you have thereby acquired compelling evidence that the relevant claims are false: “the fact that the mainstream media didn't report a piece of alleged news that, if true, they would have published, provides indirect evidence that the piece is fake news” (2021, 6). However, even if we grant to Croce and Piazza that such lack of coverage constitutes strong evidence that the misinformation you've encountered is indeed false, we shouldn't expect your possession of this evidence to largely eliminate your false belief: acquiring such evidence will have less influence than encountering an actual correction, and we've just seen that encountering a correction is not likely to largely reverse the impact of originally accepting the relevant misinformation.

¹⁰ For a review of the evidence concerning the conditions under which beliefs persist in the face of corrections, see Jerit and Zhao (2020, 81-85).

¹¹ See Walter et al. (2020).

¹² See, Swire et al. (2017) and Porter and Wood (2019, 42-43).

¹³ See Lewandowsky et al. (2012) and Thorson (2016).

And, moreover, it's not at all clear how strong the evidence in question really is. At least with respect to quite a lot of misinformation, you will have plausible alternative explanations for why you didn't find the relevant topics covered by mainstream outlets when you consulted them. Perhaps the story, while interesting or entertaining, didn't strike the journalists and editors of mainstream news organizations as sufficiently important to cover; or, perhaps some of these organizations did cover the story, but not on the days or at the times that you happened to read, watch, or listen; or, perhaps the original story was obtained using methods that journalists and editors of mainstream news organizations disapprove of; or, perhaps mainstream news organizations don't want to call attention to the story thanks to their political biases or economic incentives.¹⁴ Accordingly, the mere fact that you haven't come across any coverage of a given noteworthy story on mainstream news outlets does not provide you with strong reasons to conclude that that story is false.

In addition, not only will the best-case scenario not be the typical scenario, the worst-case scenario will be extremely common: at least very often, the false beliefs that you acquired using social media will be *reinforced* when you consume high-quality traditional news content. In the United States, most every reputable mainstream news organization is committed to the principle that news coverage should be *balanced*—and balance is understood to mean that journalists must present at least two “sides” of any disputed issue, devote roughly equal time to each side, and avoid weighing in regarding which side is more likely to be correct.¹⁵ Even on seemingly non-political, scientific questions, mainstream news organizations very often achieve balance by contrasting the views of genuine experts with the views of political figures, lobbyists, or concerned citizens.¹⁶

Moreover, such news organizations typically regard statements made by politicians and other public figures as newsworthy, regardless of whether those statements are true or false—in fact, outrageous falsehoods are often regarded as particularly newsworthy. Due to this combination of factors, when you consume news from traditional sources, you will often be repeatedly exposed to many of the same falsehoods that you have encountered on social media. For instance, in a recent study, Benkler et al. (2020) found that Donald Trump was able to reach enormous numbers of Americans with misinformation regarding voter fraud in the 2020 Presidential election largely thanks to the traditional news media's persistent coverage of his tweets and other relevant public statements. Now, of course, journalists themselves do not endorse the relevant misinformation in such cases; but from the perspective of the impact that news coverage has on the accuracy of the audience's doxastic attitudes, this fact is irrelevant. If you happen to be someone who trusts conservative political figures more than you trust journalists and official experts, a news story that relates Trump's

¹⁴ Croce and Piazza claim that “ordinary people” do not assume that mainstream media “routinely hide the truth [from] the public” on the basis of such political or economic motives—only individuals trapped in a “pathological” informational situation would make such an assumption (2021, 7). But this claim isn't plausible. In the United States, mistrust of the mainstream media is enormously widespread. For instance, a recent comprehensive study found that Americans trust the news media less than any other nation in the world; and less than half of Americans trust such mainstream outlets as ABC News, CNN, and *The New York Times* (Newman et al., 2021, 112-113). In fact, most Americans think that mainstream news outlets don't just withhold information: a 2018 poll found that 92% of Republicans, 72% of independents, and 53% of Democrats claim that “traditional news outlets knowingly report false or misleading stories at least sometimes” (Fischer, 2018).

¹⁵ See, for example, Boykoff and Boykoff (2004).

¹⁶ See Merkley (2020).

claim that voter fraud was widespread, and then provides balance by explaining that many experts maintain that the election was secure, has little chance of improving the accuracy of your beliefs on the topic.¹⁷

Finally, thus far we've focused on cases in which you encounter misinformation on social media that you end up believing as a result; but it's important to emphasize that the falsehoods you encounter have systematic negative impacts on the accuracy of your doxastic attitudes even when you don't believe them. While consuming high-quality news content will increase your stock of true beliefs concerning current events, research has shown that each time you encounter some falsehood inconsistent with what you know to be true, your grip on the truth is weakened.¹⁸ For instance, Pennycook, Cannon, and Rand (2018) demonstrated that an individual who recognizes that a particular implausible fake news story is inaccurate will rate that story to be less inaccurate each time she encounters it.

Worse still, being exposed to false information undermines the positive effect that being exposed to accurate information would otherwise have. So, if you have encountered large quantities of misinformation while using social media, even if you have managed to avoid believing any of it, the mere fact that you have encountered these falsehoods will negatively impact how you respond to the accurate news you encounter while consuming traditional media—the benefit you receive from consuming accurate information via traditional media won't be as great as it would have been had you never consumed all that misinformation via social media.

Even a single encounter with false information can have significant effects. For instance, van der Linden et al. (2017) showed that, typically, reading the statement that “97% of climate scientists have concluded that human-caused climate change is happening” significantly improved the accuracy of an individual's estimate of the current level of scientific consensus regarding climate change. But when such an individual subsequently read the false statement that “there is no consensus on human-caused climate change,” encountering this falsehood *completely eliminated* the accuracy-improving impact that reading the true statement would otherwise have had.

Ultimately, then, while most social media users who employ Croce and Piazza's proposed method will thereby improve the accuracy of their doxastic attitudes, they will not be largely protected from the misinformation they encounter on social media. That is, by expanding your “information diet” to include more high-quality traditional news content, you will likely acquire more accurate doxastic attitudes than if you obtained all of your information via social media; but such an improvement would be marginal at best. Regardless of what the rest of your information diet looks like, if you consume significant amounts of social media content, you are likely to end up believing at least some of the significant amounts of misinformation you will thereby be exposed to; and the misinformation that you encounter on social media and avoid believing will still have a systematic negative impact on the

¹⁷ For evidence concerning the role that trust of elite co-partisans plays in acceptance of misinformation, see Swire et al. (2017), and Calvillo et al. (2020).

¹⁸ See Fazio et al. (2015).

accuracy of your doxastic attitudes. So, while individuals can reasonably be expected to consume news content from traditional media, this method will not largely protect them from the negative impact that encountering misinformation on social media will have on their doxastic attitudes.

(A significantly less important worry that's worth mentioning: even if the proposed information diet protected individuals from the impact of exposure to misinformation, it's not obvious that an educational intervention could be developed that would successfully encourage large numbers of people to increase their consumption of high-quality news content. Croce and Piazza characterize their proposal as a "feasible educational remedy" (2021, 8); but they don't actually specify the type of intervention they have in mind. Given the widespread distrust of mainstream news organizations in the United States, it's difficult to imagine the sort of educational campaign that would effectively train large numbers of Americans to distinguish high-quality from low-quality mainstream outlets—i.e., to recognize that *The New York Times* and CNN conform to well-established journalistic norms, but Fox News does not—and also convince them to spend significantly more time consuming high-quality mainstream content. On the other hand, this difficulty won't be as pronounced in other countries.)

3. Some Alternative Proposals

If it isn't reasonable to expect individuals to avoid social media altogether, and if merely supplementing your social media use with high-quality news consumption doesn't largely protect you from the negative impacts of misinformation, then perhaps individuals ought to take certain preventative measures before using social media. For instance, perhaps individuals ought to cultivate the sort of cognitive traits that would enable them to remain largely unaffected by the misinformation they encounter on social media; or perhaps individuals ought to employ certain strategies or techniques while consuming social media to avoid being taken in by the falsehoods they come across.

Some recent research on epistemic vices suggests a proposal of the former sort. Meyer, Alfano, and de Bruin recently found that certain *epistemic vices*—in particular, *indifference*, or "lack of motivation to find the truth," and *rigidity*, or "insensitivity to evidence"—are strongly correlated with acceptance of COVID-19 misinformation (2021, 5-6). The epistemic vices at issue were measured using the authors' Epistemic Vice Scale: for instance, indifference is measured by the extent to which subjects agree with statements such as "I am not very interested in understanding things," and rigidity is measured by the extent to which subjects agree with statements such as "I often have strong opinions about issues I don't know much about," and "I tend to feel sure about my views even if I don't have much evidence" (2021, 5). Meyer, Alfano, and de Bruin discovered that individuals who received the highest scores on the Epistemic Vice Scale were far more likely to believe COVID-19 misinformation than individuals who received lower scores (2021, 14). In fact, they found that epistemic vice was more strongly correlated with susceptibility to COVID-19 misinformation than were any other demographic or psychological traits, such as age, political affiliation, need for cognition, performance on the Cognitive Reflection Test, and so on (2021, 14-16).

Given this strong correlation, one might conclude that individuals' doxastic attitudes are negatively impacted by the misinformation they consume on social media because they exhibit the sorts of epistemic vices at issue. And if so, this fact would suggest a method that social media users might reasonably be expected to employ in order to protect themselves from the negative impacts of misinformation: individuals should take steps to eliminate or mitigate the relevant epistemic vices. In other words, if individuals believe the misinformation they encounter on social media because they exhibit indifference and rigidity, then a simple method they could use to protect themselves from misinformation would be to become less indifferent and less rigid. So, if there are strategies one can use to reduce one's epistemic vices, then the present research suggests promising educational remedies: perhaps individuals can be protected from misinformation by reducing or eliminating their epistemic vices with the right kind of training.¹⁹

However, there are at least three significant difficulties with this proposal. First and foremost, Meyer, Alfano, and de Bruin's research only establishes a correlation between epistemic vice and accepting *particularly implausible* misinformation. Their study utilized falsehoods concerning COVID-19 such as "adding pepper to your meals prevents COVID-19," "5G mobile networks spread COVID-19," and "exposing yourself to the sun or to temperatures higher than 77°F prevents the coronavirus disease" (2021, 9). These claims were taken from a World Health Organization "myth-busting" webpage, so presumably they had spread widely online; but by no means are these claims representative of the kind of misinformation that reaches the most people on social media and has the biggest impact on what people believe.

For instance, the present research provides no reason to think that becoming less indifferent and rigid would prevent someone from believing that the COVID-19 vaccines are dangerous and ineffective, or that voter fraud was rampant in the 2020 Presidential election. Consequently, even if there were some method that individuals could reasonably be expected to employ that would eliminate their epistemic vices, doing so would not largely protect them from much of the most consequential misinformation that they are likely to encounter while using social media.

A second difficulty is that Meyer, Alfano, and de Bruin's research focuses exclusively on belief in misinformation. But, as we've already seen (§2), the falsehoods you are exposed to have a systematic negative impact on the accuracy of your doxastic attitudes even when you don't believe them. So, even if you happen to be extremely low with respect to indifference and rigidity, and even if you recognize that a particular piece of misinformation is indeed false, your grip on the truth will still be undermined each time you are exposed to it. And, moreover, your lack of epistemic vices will not prevent the misinformation you encounter on social media from counteracting the benefits that exposure to accurate information would otherwise have.

¹⁹ Meyer, Alfano, and de Bruin endorse only a more qualified conclusion along these lines: they suggest that if epistemic vice plays a role in the acceptance of misinformation, and "if epistemic vice can be countered using educational or other interventions, then the public health response to COVID-19 may be bolstered by this line of research" (2021, 2-3).

A third difficulty is that it's not clear whether there are any effective methods via which individuals could eliminate the relevant epistemic vices. Individuals with the vices at issue are the kind of people who freely admit "I often have strong opinions about issues I don't know much about," and "I tend to feel sure about my views even if I don't have much evidence." What kind of training, or therapy, could convince such a person that, in fact, one should only believe what one's evidence supports? Presumably individuals could be prevented from becoming indifferent and rigid with the right sort of guidance; but it's not obvious that someone who already possesses these vices will be able (and motivated) to change such perverse views of permissible epistemic conduct. And so, if these vices can't be eliminated, then—even if we were to grant that eliminating the relevant epistemic vices would protect individuals from most misinformation—individuals don't have a method they can reasonably be expected to employ that would largely protect them from the negative impact of misinformation.

Alternatively, one might think that, in order to protect themselves from misinformation, individuals needn't dramatically overhaul their firmly-established character traits—perhaps they simply need to exercise more care or vigilance when consuming information via social media. Recent research by Gordon Pennycook and colleagues suggests a proposal of this sort. For instance, Pennycook and Rand (2019) found that individuals who are more prone to engage in the right sort of reflection—as measured by performance on the Cognitive Reflection Test—are better able to distinguish accurate news stories from stereotypical fake news stories. And Bago, Rand, and Pennycook (2020) found that when individuals take the time to engage in deliberation, they are better able to identify fake news headlines as inaccurate. Summarizing a significant quantity of related research, Pennycook and Rand maintain that one of the principal reasons that individuals fail to recognize that fake news stories are false is that "they do not stop to reflect sufficiently on their prior knowledge (or have insufficient or inaccurate prior knowledge)" (2021, 393).

Such research, then, suggests a method that social media users might reasonably be expected to employ in order to protect themselves from the negative impacts of misinformation: when consuming information via social media, individuals should engage in sufficient reflection and deliberation—they should be careful to bring their stored knowledge to bear as a kind of misinformation filter. If such a method were able to protect individuals from misinformation, then promising educational remedies might exist: for instance, Pennycook and Rand suggest that educational "interventions that are directed at making the public more thoughtful consumers of news media may have promise" (2019, 48). More generally, one might think that critical thinking training might successfully encourage large numbers of people to engage in the requisite sort of reflection while using social media.²⁰

However, there are at least three significant difficulties with this proposal. The first is that, once again, the research at issue is focused exclusively on belief in misinformation. Even if we grant that there are methods by which individuals can increase their tendency to engage in the relevant sort of reflection while using social media, the best outcome we could expect is that they would end up believing fewer false stories than they would have otherwise. But,

²⁰ See Heersmink (2018, §5).

again, the falsehoods we encounter have systematic negative impacts on the accuracy of our doxastic attitudes even when we avoid believing them.

A second difficulty is that it's not clear whether there are any effective methods via which individuals can increase their tendency to engage in the relevant sort of reflection. It seems plausible that educational interventions could be designed that train individuals to take the time to reflect and deliberate while consuming information. But whether any such intervention will ultimately be effective is an open empirical question—at present there simply isn't good evidence that educational interventions improve critical thinking abilities over the long term.²¹ (As Pennycook and Rand (2021, 397) explain, there is considerable evidence that exposing individuals to “accuracy prompts”—nudges that remind social media users to focus their attention on accuracy—significantly improves the rate at which they identify fake news stories. But such an intervention is a *structural* rather than an *educational* remedy—the purpose is to make the information environment safer to navigate, not to bring about relatively long-term changes to individuals' behavior or cognitive traits.)

Finally, the most fundamental difficulty with the present proposal is that reflection and deliberation can only hope to protect you from a small fraction of the misinformation you are likely to encounter on social media. Taking the time to reflect on your existing knowledge while consuming social media will only protect you from misinformation so long as the relevant false claims are inconsistent with what you already know. But, first, in many cases, you simply won't possess any relevant beliefs. Some of the most consequential misinformation that circulates widely on social media concerns issues that most people don't know much about; for instance, most people don't know much about the science of COVID-19 vaccines, or the details of election security, and so they can't filter out many false claims concerning these topics. And, second, in many cases, the misinformation you encounter will be consistent with your existing beliefs (sometimes because your relevant existing beliefs will be false). For instance, research has shown that individuals are much more likely to judge that a stereotypical fake news story is accurate when it conforms to their political commitments: such falsehoods seem more plausible to someone with the relevant background beliefs.²²

This point is particularly important in the present context: thanks to personalization algorithms, when you use social media, you will regularly encounter misinformation that is consistent with your existing beliefs. Accordingly, the proposed method—stopping to reflect and deliberate while consuming social media—can only prevent individuals from believing misinformation that happens to be incompatible with true beliefs they already possess; and so, this method offers individuals no protection against a significant proportion of the misinformation they will encounter on social media.

4. Conclusion

The question we started with was: are there methods that most individuals can reasonably be expected to employ that would largely protect them from the negative impact that

²¹ See El Soufi and See (2019).

²² See Pennycook, Cannon, and Rand (2018, 1875), and Pennycook and Rand (2019, 47).

encountering misinformation on social media would otherwise have on their doxastic attitudes? We've now seen that there are compelling reasons to conclude that the answer is "no." There are methods available to social media users that would likely yield improvements: if you consume more high-quality traditional news content, suppress certain epistemic vices, or take the time to reflect and deliberate while consuming social media content, you will likely end up with fewer false beliefs than you would have otherwise. But such improvements would be marginal at best. Whatever else you do, so long as you regularly utilize social media, you will be repeatedly exposed to misinformation that is strongly consistent with what you already believe, and that is endorsed by individuals and groups you admire and trust. And, as such, you are likely to end up believing at least some portion of this misinformation; and the misinformation that you avoid believing will still have a systematic negative impact on the accuracy of your doxastic attitudes. No ordinary human being can inhabit such an information environment without suffering significant harm.

Accordingly, any educational intervention designed to encourage large numbers of people to employ one of the aforementioned methods would likely have only a minor impact—and so would not be worth the expense and effort. So, while educational approaches are compatible with structural approaches, and are likely to do some good, we have independent reasons for concluding that they are not worth pursuing. Instead, in order to effectively combat the negative impacts of misinformation distributed via social media, rather than focus on individuals, we ought to focus on the information environment itself. In particular, we ought to focus on whether we can identify (morally and politically acceptable) modifications to social media platforms that would significantly limit the distribution of misinformation, or that would minimize the impact misinformation has on ordinary users.

References

- Avaaz. 2020. "How Facebook Can Flatten the Curve of the Coronavirus Infodemic." *Avaaz Report*, April 15, 2020. https://avaazimages.avaaz.org/facebook_coronavirus_misinformation.pdf.
- Avaaz. 2021. "Facebook: From Election to Insurrection." *Avaaz Report*, March 18, 2021. https://avaazimages.avaaz.org/facebook_election_insurrection.pdf.
- Bago, Bence, David Rand, and Gordon Pennycook. 2020. "Fake News, Fast and Slow: Deliberation Reduces Belief in False (but Not True) News Headlines." *Journal of Experimental Psychology: General* 149 (8): 1608–1613.
- Benkler, Yochai, Robert Faris, and Hal Roberts. 2018. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. New York: Oxford University Press.
- Benkler, Yochai, Casey Tilton, Bruce Etling, Hal Roberts, Justin Clark, Robert Faris, Jonas Kaiser, and Carolyn Schmitt. 2020. "Mail-In Voter Fraud: Anatomy of a Disinformation Campaign." *Berkman Center Research Publication No. 2020-6*. <https://cyber.harvard.edu/publication/2020/Mail-in-Voter-Fraud-Disinformation-2020>.
- Boykoff, Maxwell and Jules Boykoff. 2004. "Balance as Bias: Global Warming and the US Prestige Press." *Global Environmental Change* 14 (2): 125–136.
- Bradshaw, Samantha, Philip Howard, Bence Kollanyi, and Lisa-Maria Neudert. 2020. "Sourcing and Automation of Political News and Information over Social Media in the United States, 2016-2018." *Political Communication* 37 (2): 173–193.

- Calvillo, Dustin, Bryan Ross, Ryan Garcia, Thomas Smelter, and Abraham Rutchick. 2020. "Political Ideology Predicts Perceptions of the Threat of COVID-19 (and Susceptibility to Fake News About It)." *Social Psychological and Personality Science* 11 (8): 1119–1128.
- Coady, David. 2019. "The Trouble With 'Fake News'." *Social Epistemology Review and Reply Collective* 8 (10): 40–52.
- Croce, Michel and Tommaso Piazza. 2021. "Consuming Fake News: Can We Do Any Better?" *Social Epistemology*. 1–10. <https://doi.org/10.1080/02691728.2021.1949643>.
- El Soufi, Nada and Beng Huat See. 2019. "Does Explicit Teaching of Critical Thinking Improve Critical Thinking Skills of English Language Learners in Higher Education? A Critical Review of Causal Evidence." *Studies in Educational Evaluation* 60: 140–162.
- Fazio, Lisa, Nadia Brashier, B. Keith Payne, and Elizabeth Marsh. 2015. "Knowledge Does Not Protect Against Illusory Truth." *Journal of Experimental Psychology: General* 144 (5): 993–1002.
- Fischer, Sara. 2018. "92% of Republicans Think Media Intentionally Reports Fake News." *Axios*, June 27, 2018. <https://www.axios.com/trump-effect-92-percent-republicans-media-fake-news-9c1bbf70-0054-41dd-b506-0869bb10f08c.html>.
- Gordon, Andrew, Susanne Quadflieg, Jonathan Brooks, Ullrich Ecker, and Stephan Lewandowsky. 2019. "Keeping Track of 'Alternative Facts': The Neural Correlates of Processing Misinformation Corrections." *NeuroImage* 193: 46–56.
- Heersmink, Richard. 2018. "A Virtue Epistemology of the Internet: Search Engines, Intellectual Virtues and Education." *Social Epistemology* 32 (1): 1–12.
- Jerit, Jennifer and Yangzi Zhao. 2020. "Political Misinformation." *Annual Review of Political Science* 23: 77–94.
- Lewandowsky, Stephan, Ullrich Ecker, Colleen Seifert, Norbert Schwarz, and John Cook. 2012. "Misinformation and Its Correction: Continued Influence and Successful Debiasing." *Psychological Science in the Public Interest* 13 (3): 106–131.
- Loomba, Sahil, Alexandre de Figueiredo, Simon Piatek, Kristen de Graaf, and Heidi Larson. 2021. "Measuring the Impact of COVID-19 Vaccine Misinformation on Vaccination Intent in the UK and USA." *Nature Human Behaviour* 5: 337–348.
- Mercier, Hugo. 2017. "How Gullible Are We? A Review of the Evidence from Psychology and Social Science." *Review of General Psychology* 21 (2): 103–122.
- Merkley, Eric. 2020. "Are Experts (News)Worthy? Balance, Conflict, and Mass Media Coverage of Expert Consensus." *Political Communication* 37 (4): 530–549.
- Meyer, Marco, Mark Alfano, and Boudewijn de Bruin. 2021. "Epistemic Vice Predicts Acceptance of Covid-19 Misinformation." *Episteme*. 1–22.
- Millar, Boyd. 2019. "The Information Environment and Blameworthy Belief." *Social Epistemology* 33 (6): 525–537.
- Morning Consult/Politico. 2021. "National Tracking Poll #2110134." <https://www.politico.com/f/?id=0000017c-c00e-d3c9-a77c-cb9eab6a0000&nname=playbook&nid=0000014f-1646-d88f-a1cf-5f46b7bd0000&nrid=0000014e-f115-dd93-ad7f-f91513e50001&nlid=630318>.
- Newman, Nic, Richard Fletcher, Anne Schulz, Simge Andı, Craig T. Robertson, and Rasmus Kleis Nielsen. 2021. "Reuters Institute Digital News Report 2021." https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf.

- Pennycook, Gordon and David Rand. 2021. "The Psychology of Fake News." *Trends in Cognitive Sciences* 25 (5): 388–402.
- Pennycook, Gordon and David Rand. 2019. "Lazy, Not Biased: Susceptibility to Partisan Fake News is Better Explained by Lack of Reasoning than by Motivated Reasoning." *Cognition* 188: 39–50.
- Pennycook, Gordon, Tyrone Cannon, and David Rand. 2018. "Prior Exposure Increases Perceived Accuracy of Fake News." *Journal of Experimental Psychology: General* 147 (12): 1865–1880.
- Porter, Ethan and Thomas Wood. 2019. *False Alarm: The Truth about Political Mistruths in the Trump Era*. Cambridge: Cambridge UP.
- Sindermann, Cornelia, Andrew Cooper, and Christian Montag. 2020. "A Short Review on Susceptibility to Falling for Fake Political News." *Current Opinion in Psychology* 36: 44–48.
- Swire, Briony, Adam Berinsky, Stephan Lewandowsky, and Ullrich Ecker. 2017. "Processing Political Misinformation: Comprehending the Trump Phenomenon." *Royal Society Open Science* 4 (3): 160802.
- Thorson, Emily. 2016. "Belief Echoes: The Persistent Effects of Corrected Misinformation." *Political Communication* 33: 460–480.
- van der Linden, Sander, Anthony Leiserowitz, Seth Rosenthal, and Edward Maibach. 2017. "Inoculating the Public against Misinformation about Climate Change." *Global Challenges* 1: 1600008.
- Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The Spread of True and False News Online." *Science* 359: 1146–1151.
- Walter, Nathan, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag. 2020. "Fact-Checking: A Meta-Analysis of What Works and for Whom." *Political Communication* 37 (3): 350–375.