

# Misinformation: susceptibility, spread, and interventions to immunize the public

Sander van der Linden  

The spread of misinformation poses a considerable threat to public health and the successful management of a global pandemic. For example, studies find that exposure to misinformation can undermine vaccination uptake and compliance with public-health guidelines. As research on the science of misinformation is rapidly emerging, this conceptual Review summarizes what we know along three key dimensions of the infodemic: susceptibility, spread, and immunization. Extant research is evaluated on the questions of why (some) people are (more) susceptible to misinformation, how misinformation spreads in online social networks, and which interventions can help to boost psychological immunity to misinformation. Implications for managing the infodemic are discussed.

In early 2020, the World Health Organization (WHO) declared a worldwide ‘infodemic’. An infodemic is characterized by an overabundance of information, particularly false and misleading information<sup>1</sup>. Although researchers have debated the effect of fake news on the outcomes of major societal events, such as political elections<sup>2,3</sup>, the spread of misinformation has much clearer potential to cause direct and notable harm to public health, especially during a pandemic. For example, research across different countries has shown that the endorsement of COVID-19 misinformation is robustly associated with people being less likely to follow public-health guidance<sup>4–7</sup> and having reduced intentions to get vaccinated<sup>4,5</sup> and to recommend the vaccine to others<sup>4</sup>. Experimental evidence has found that exposure to misinformation about vaccination resulted in about a 6-percentage-point decrease in the intention to get vaccinated among those who said that they would otherwise “definitely accept a vaccine”, undermining the potential for herd immunity<sup>8</sup>. Analyses of social-network data estimate that, without intervention, anti-vaccination content on social platforms such as Facebook will dominate discourse in the next decade<sup>9</sup>. Other research finds that exposure to misinformation about COVID-19 has been linked to the ingestion of harmful substances<sup>10</sup> and an increased propensity to engage in violent behaviors<sup>11</sup>. Of course, misinformation was a threat to public health long before the pandemic. The debunked link between the MMR vaccine and autism was associated with a significant drop in vaccination coverage in the United Kingdom<sup>12</sup>, Listerine manufacturers falsely claimed that their mouthwash cured the common cold for many decades<sup>13</sup>, misinformation about tobacco products has influenced attitudes toward smoking<sup>14</sup> and, in 2014, Ebola clinics were attacked in Liberia because of the false belief that the virus was part of a government conspiracy<sup>15</sup>.

Given the unprecedented scale and pace at which misinformation can now travel online, research has increasingly relied on models from epidemiology to understand the spread of fake news<sup>16–18</sup>. In these models, the key focus is on the reproduction number ( $R_0$ )—in other words, the number of individuals who will start posting fake news (that is, secondary cases) following contact with someone who is already posting misinformation (the infectious individual). It is therefore helpful to think of misinformation as a viral pathogen that can infect its host, spreading rapidly from one individual to another

within a given network, without the need for physical contact. One benefit of this epidemiological approach lies in the fact that early detection systems could be designed to identify, for example, superspreaders, which would allow for the timely deployment of interventions to curb the spread of viral misinformation<sup>18</sup>.

This Review will provide readers with a conceptual overview of recent literature on misinformation, along with a research agenda (Box 1) that covers three major theoretical dimensions aligned with the viral analogy: susceptibility, spread, and immunization. What makes individuals susceptible to misinformation in the first place? Why and how does it spread? And what can we do to boost public immunity?

Before reviewing the extant literature to help answer these questions, it is worth briefly discussing what the term ‘misinformation’ means, because inconsistent definitions affect not only the conceptualization of research designs but also the nature and validity of key outcome measures<sup>19</sup>. Indeed, misinformation has been referred to as an ‘umbrella category of symptoms’<sup>20</sup> not only because definitions vary, but also because the behavioral consequences for public health might differ depending on the type of misinformation. The term ‘fake news’ is often especially regarded as problematic because it insufficiently describes the full spectrum of misinformation<sup>21</sup> and has become a politicized rhetorical device in itself<sup>22</sup>. Box 2 provides a more detailed discussion of the problems associated with different scholarly definitions of misinformation<sup>23</sup> but for the purpose of this Review, I will simply define misinformation in its broadest possible sense: ‘false or misleading information masquerading as legitimate news, regardless of intent’<sup>24</sup>. Although disinformation is often differentiated from misinformation insofar as it involves a clear intention to deceive or harm other people, intent can be difficult to establish, so in this Review my treatment of misinformation will cover both intentional and unintentional forms of misinformation.

## Susceptibility

Although people use many cognitive heuristics to make judgments about the veracity of a claim (for example, perceived source credibility)<sup>25</sup>, one particularly prominent finding that helps explain why people are susceptible to misinformation is known as the ‘illusory truth’ effect: repeated claims are more likely to be judged as true than non-repeated (or novel) claims<sup>26</sup>. Given the fact that many

## Box 1 | Agenda and recommendations for future research

Research question 1: What factors make people susceptible to misinformation?

- Better integrate accuracy-driven with social, political, and cultural motivations to explain people's susceptibility to misinformation.
- Define, develop, and validate standardized instruments for assessing general and domain-specific susceptibility to misinformation.

Research question 2: How does misinformation spread in social networks?

- Outline with greater clarity the conditions under which 'exposure' is more or less likely to lead to 'infection', including the impact of repeated exposure, the micro-targeting of fake news on social media, contact with superspreaders, the role of echo chambers, and the structure of the social network itself.
- Provide more accurate population-level estimates of exposure to misinformation by (1) capturing more diverse types of misinformation and (2) linking exposure to fake news across different kinds of traditional and social-media platforms.

Research question 3: Can we inoculate or immunize people against misinformation?

- Focus on evaluating the relative efficacy of different debunking methods in the field, as well as how debunking (therapeutic) and prebunking (prophylactic) interventions could be combined to maximize their protective properties.
- Model and evaluate how psychological inoculation methods can spread online and influence real-world sharing behavior on social media.

falsehoods are often repeated by the popular media, politicians, and social-media influencers, the relevance of illusory truth has increased substantially. For example, the conspiracy theory that the coronavirus was bio-engineered in a military laboratory in Wuhan, China, and the false claim that "COVID-19 is no worse than the flu" have been repeated many times in the media<sup>27</sup>. The primary cognitive mechanism responsible for the fact that people are more likely to think that repeated claims are true is known as processing fluency: the more a claim is repeated, the more familiar it becomes and the easier it is to process<sup>28</sup>. In other words, the brain uses fluency as a signal for truth. Importantly, research shows that (1) prior exposure to fake news increases its perceived accuracy<sup>29</sup>; (2) illusory truth can occur for both plausible and implausible claims<sup>30</sup>; (3) prior knowledge does not necessarily protect people against illusory truth<sup>31</sup>; and (4) illusory truth does not appear to be moderated by thinking styles such as analytical versus intuitive reasoning<sup>32</sup>.

Although illusory truth can affect everyone, research has noted that some people are still more susceptible to misinformation than others. For example, some common findings include the observation that older individuals are more susceptible to fake news<sup>33,34</sup>, potentially owing to factors such as cognitive decline and greater digital illiteracy<sup>35</sup>, although there are exceptions: in the context of COVID-19, older individuals appear less likely to endorse misinformation<sup>4</sup>. Those with a more extreme and right-wing political orientation have also consistently shown to be more susceptible to misinformation<sup>3,4,33,36,37</sup>, even when the misinformation in question is non-political<sup>38,39</sup>. Yet, the link between ideology and misinformation susceptibility is not always consistent across different cultures<sup>4,37</sup>. Other factors such as greater numeracy skills<sup>4</sup> and cognitive

## Box 2 | The challenges with defining and operationalizing misinformation

One of the most frequently cited definitions of fake news is "fabricated information that mimics news media content in form but not in organizational process or intent"<sup>119</sup>. This definition implies that what matters in determining whether a story is misinformation or not is the journalistic or editorial process. Other definitions echo similar sentiments insofar the view is taken that misinformation producers do not adhere to editorial norms<sup>120</sup> and that the defining attribution of 'fake-ness' happens at the level of the publisher and not at the level of the story<sup>3</sup>. However, others have taken a completely different view by defining misinformation either in terms of the veracity of its content or the presence or absence of common techniques used to produce it<sup>109</sup>.

It could be argued that some definitions are overly narrow because news stories do not need to be completely false in order to be misleading. A highly salient example comes from the *Chicago Tribune*, a generally credible outlet, which re-published a story in January 2021 with the headline "A healthy doctor died two weeks after getting a COVID-19 vaccine". This story would not be classified as false on the basis of the source or even the content, as the events were true when considered in isolation. However, it is highly misleading—even considered unethical—to suggest that the doctor died specifically because of the COVID-19 vaccine when there was no evidence to make such a causal connection at the time of publication. This is not an obscure example: it was viewed over 50 million times on Facebook in early 2021 (ref. <sup>121</sup>).

Another potential challenge with purely content-based definitions is that when expert consensus on a public-health topic is rapidly emerging and subject to uncertainty and change, the definition of what is likely to be true and false can shift over time, making overly simplistic 'real' versus 'fake' categorizations a potentially unstable property. For example, although news media initially claimed that ibuprofen could worsen coronavirus symptoms, this claim was later retracted as more evidence became available<sup>122</sup>. The problem is that researchers often ask people how accurate or reliable they find a selective series of true and fake headlines that were either created or selected by the researchers on the basis of different definitions of what constitutes misinformation.

There is also variation in outcome measures; sometimes the relevant outcome measure is misinformation susceptibility, and sometimes it is the difference between fake and real news detection, or so-called 'truth discernment'. The only existing instrument that uses a psychometrically validated set of headlines is the recent Misinformation Susceptibility Test, a standardized measure of news veracity discernment that is normed to the test population<sup>123</sup>. Overall, this means that hundreds of emerging studies on the topic of misinformation have outcome measures that are non-standardized and not always easily comparable.

and analytic thinking styles<sup>36,40,41</sup> have consistently been revealed to have a negative correlation with misinformation susceptibility—although other scholars have identified partisanship as a potential moderating factor<sup>42–44</sup>. In fact, these individual differences have given rise to two competing overarching theoretical explanations<sup>45,46</sup> for why people are susceptible to misinformation. The first theory is often referred to as the classical 'inattention' account; the second is often dubbed the 'identity-protective' or 'motivated cognition' account. I will discuss emerging evidence for both theories in turn.

**The inattention account.** The inattention or 'classical reasoning' account argues that people are committed to sharing accurate

content but the context of social media simply distracts people from making news-sharing decisions that are based on a preference for accuracy<sup>45</sup>. For example, consider that people are often bombarded with news content online, much of which is emotionally charged and political, which, coupled with the fact that people have limited time and resources to think about the veracity of a piece of news, might significantly interfere with their ability to accurately reflect on such content. The inattention account is based on a ‘classical reasoning perspective insofar as it draws on dual-process theories of human cognition, which suggest that people rely on two qualitatively different processes of reasoning<sup>47</sup>. These processes are often referred to as System 1, which is predominantly automatic, associative, and intuitive, and System 2, which is more reflective, analytical, and deliberate. A canonical example is the Cognitive Reflection Test (CRT), which administers a series of puzzles in which the intuitive or first answer that comes to mind is often wrong and thus a correct answer requires people to pause and reflect more carefully. The basic point is that activating more analytical System 2-type reasoning can override erroneous System 1-type intuitions. Evidence for the inattention account comes from the fact that people who score higher on the CRT<sup>36,41</sup>, who deliberate more<sup>48</sup>, who have greater numeracy skills<sup>4</sup>, and who have higher knowledge and education<sup>37,49</sup> are consistently better able to discern between true and false news—regardless of whether the content is politically congruent<sup>36</sup>. In addition, experimental interventions that ‘prime’ people to think more analytically or consider the accuracy of news content<sup>50,51</sup> have been shown to improve the quality of people’s news-sharing decisions and decrease acceptance of conspiracy theories<sup>52</sup>.

**The motivated reasoning account.** In stark contrast to the inattention account stands the theory of (politically) motivated reasoning<sup>53–55</sup>, which posits that information deficits or lack of reflective reasoning are not the primary driver of susceptibility to misinformation. Motivated reasoning occurs when someone starts out their reasoning process with a pre-determined goal (for example, someone might want to believe that vaccines are unsafe because that belief is shared by their family members), so individuals interpret new (mis)information in service of reaching that goal<sup>53</sup>. The motivated account therefore argues that the types of commitments that people have to their affinity groups is what leads them to selectively endorse media content that reinforces deeply held political, religious, or social identities<sup>56,57</sup>. There are several variants of the politically motivated reasoning account, but the basic premise is that people pay attention to not just the accuracy of a piece of news content, but also the goals that such information may serve. For example, a fake news story could be viewed as much more plausible when it happens to offer positive information about someone’s political group, or equally when it offers negative information about a political opponent<sup>42,57,58</sup>. A more extreme and scientifically contentious version of this model, also known as the ‘motivated numeracy’<sup>59</sup> account, suggests that more reflective and analytical System 2 reasoning abilities do not help people make more accurate assessments but in fact are frequently hijacked in service of identity-based reasoning. Evidence for this claim comes from the fact that partisans with the highest numeracy and education levels tend to be the most polarized on contested scientific issues, such as climate change<sup>60</sup> or stem-cell research<sup>61</sup>. Experimental work has also shown that when people are asked to make causal inferences about a data problem, such as the benefits of a new skin rash treatment, people with greater numeracy skills performed better when the problem was non-political. By contrast, people became more polarized and less accurate when the same data were presented as results from a new study on gun control<sup>59</sup>. These patterns were more pronounced among those with higher numeracy skills. Other research has found that politically conservative individuals are much more likely to (mistakenly) judge misinformation as true when the information

is presented as coming from a conservative source than when that same information is presented as coming from a liberal source, and vice versa for politically liberal individuals—highlighting the key role of politics in truth discernment<sup>62</sup>.

**Susceptibility: limitations and future research.** It is worth mentioning that both accounts face significant critiques and limitations. For example, independent replications of interventions designed to nudge accuracy have revealed mixed findings<sup>63</sup>, and questions have been raised about the conceptualization of partisan bias in these studies<sup>43</sup>, including the possibility that the intervention effects are moderated by people’s political identities<sup>44</sup>. In turn, the motivated numeracy account has faced several failed and mixed replications<sup>64–66</sup>. For example, one large nationally representative study in the United States showed that, although polarization on global warming was indeed greatest among the highest educated partisans at baseline, this effect was neutralized and even reversed by an experimental intervention that induced accuracy motivations by highlighting the scientific consensus on global warming<sup>66</sup>. These findings have led to the discovery of a much larger confound in the motivated-reasoning literature, in that partisan bias could simply be due to selective exposure rather than motivated reasoning<sup>66–68</sup>. This is so because the role of politics is confounded with people’s prior beliefs<sup>66</sup>. Although people are polarized on many issues, this does not mean that they are unwilling to update their (misinformed) beliefs in line with the evidence. Moreover, people might refuse to update their beliefs not because of a motivation to reject the information (because it is incongruent with their political worldview) but simply because they find the information not credible, either because they discount the source or the veracity of the content itself for what appear to be legitimate reasons to those individuals. This ‘equivalence paradox’<sup>69</sup> makes it difficult to causally disentangle accuracy from motivation-based preferences.

Future research should therefore not only carefully manipulate people’s motivations in the processing of (mis)information that is politically (dis)concordant, but also offer a more integrated theoretical account of susceptibility to misinformation. For example, it is likely that for political fake news, identity-motivations are going to be more salient; however, for misinformation that tackles non-politicized issues (such as falsehoods about cures for the common cold), knowledge deficits, inattention, or confusion might be more likely to play a role. Of course, it is possible for public-health issues—such as COVID-19—to become politicized relatively quickly, in which case the prominence of motivational goals in driving susceptibility to misinformation might increase. Accuracy and motivational goals are also frequently in conflict. For example, people might understand that a news story is unlikely to be true, but if the misinformation promotes the goals of their social group, they might be more inclined to forgo their desire for accuracy in favor of a motivation to conform with the norms of their community<sup>56,57</sup>. In other words, in any given context, the importance people assign to accuracy versus social goals is going to determine how and when they are going to update their beliefs in light of misinformation. There is much to be gained by advancing more contextual theories that focus on the interplay between accuracy and socio-political goals in explaining why people are susceptible to misinformation.

## Spread

**Measuring the infodemic.** To return to the viral analogy, researchers have adopted models from epidemiology, such as the susceptible–Infected–Recovered (SIR) model, to measure and quantify the spread of misinformation in online social networks<sup>17,70</sup>. In this context,  $R_0$  often represents individuals who will start posting fake news following contact with someone who is already ‘infected’. Evidence for the potential of an infodemic is taken when  $R_0$  exceeds 1, which signals the potential for exponential growth (when  $R_0$  is lower than 1,

the infodemic will eventually sizzle out) and is evidence pointing to a possible infodemic. Analyses of social-media platforms have shown that all have the potential to drive infodemic-like spread, but some are more capable than others<sup>17</sup>. For example, research on Twitter has found that false news is about 70% more likely to be shared than true news, and it takes true news 6 times longer than false stories to reach 1,500 people<sup>71</sup>. Although fake news can thus spread faster and deeper than true news, it is important to emphasize that these findings are based on a relatively narrow definition of fact-checked news (see Box 2 and ref. <sup>70</sup>), and more recent research has pointed out that these estimates are likely platform-dependent<sup>72</sup>. Importantly, several studies have now shown that fake news typically represents a small part of people's overall media diet and that the spread of misinformation on social media is highly skewed so that a small number of accounts are responsible for the majority of the content that is shared and consumed, also known as 'super-sharers' and 'superconsumers'<sup>3,24,73</sup>. Although much of this work has come from the political domain, very similar findings have been found in the context of the COVID-19 pandemic, during which 'superspreaders' on Twitter and Facebook were exerting a majority of the influence on the platform<sup>74</sup>. A major issue is the existence of echo chambers, in which the flow of information is often systematically biased toward like-minded others<sup>72,75,76</sup>. Although the prevalence of echo chambers is debated<sup>77</sup>, the existence of such polarized clusters has shown to aid the virality of misinformation<sup>75,78,79</sup> and impede the spread of corrections<sup>76</sup>.

**Exposure does not equal infection.** Importantly, exposure estimates based on social-media data often do not seem to line up with people's self-reported experiences. Different polls show that over a third of people self-report frequent, if not daily exposure, to misinformation<sup>80</sup>. Of course, the validity of people's self-reported experiences can be variable, but it raises questions about the accuracy of exposure estimates, which are often based on limited public data and can be sensitive to model assumptions. Moreover, a crucial factor to consider here is that exposure does not equal persuasion (or 'infection'). For example, research in the context of COVID-19 headlines shows that people's judgments of headline veracity had little impact on their sharing intentions<sup>45</sup>. People may thus choose to share misinformation for reasons other than accuracy. For example, one recent study<sup>81</sup> found that people often share content that appears 'interesting if true'. The study indicated that although people rate fake news as less accurate, they also rate it as 'more interesting if true' than real news and are thus willing to share it.

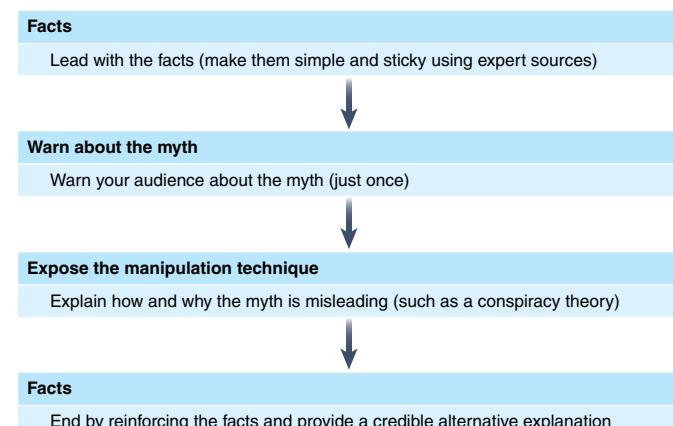
**Spread: limitations and future research.** More generally, the body of research on 'spreading' has faced significant limitations, including critical gaps in knowledge. There is skepticism about the rate at which people exposed to misinformation begin to actually believe it because research on media and persuasion effects has shown that it is difficult to persuade people using traditional advertisements<sup>82</sup>. But existing research has often used contrived laboratory designs that may not sufficiently represent the environment in which people make news-sharing decisions. For example, studies often test one-off exposures to a single message rather than persuasion as a function of repeated exposure to misinformation from diverse social and traditional media sources. Accordingly, we need a better understanding of the frequency and intensity with which exposure to misinformation ultimately leads to persuasion. Most studies also rely on publicly available data that people have shared or clicked on, but people may be exposed and influenced by much more information while scrolling on their social-media feed<sup>45</sup>. Moreover, fake news is often conceptualized as a list of URLs that were fact-checked as true or false, but this type of fake news represents only a small segment of misinformation; people may be much more likely to encounter content that is misleading or manipulative without being overtly

false (see Box 2). Finally, micro-targeting efforts have significantly enhanced the ability for misinformation producers to identify and target subpopulations of individuals who are most susceptible to persuasion<sup>83</sup>. In short, more research is needed before precise and valid conclusions can be made about either population-level exposure or the probability that exposure to misinformation leads to infection (that is, persuasion).

## Immunization

A rapidly emerging body of research has started to evaluate the possibility of 'immunizing' the public against misinformation at a cognitive level. I will categorize these efforts by whether their application is primarily prophylactic (preventative) or therapeutic (post-exposure), also known as 'prebunking' and 'debunking', respectively.

**Therapeutic treatments: fact-checking and debunking.** The traditional, standard approach to countering misinformation generally involves the correction of a myth or falsehood after people have already been exposed or persuaded by a piece of misinformation. For example, debunking misinformation about autism interventions has shown to be effective in reducing support for non-empirically supported treatments, such as dieting<sup>84</sup>. Exposure to court-ordered corrective advertisements from the tobacco industry on the link between smoking and disease can increase knowledge and reduce misperceptions about smoking<sup>85</sup>. In one randomized controlled trial, a video debunking several myths about vaccination effectively reduced influential misperceptions, such as the false belief that vaccines cause autism or that they reduce the strength of the natural immune system<sup>86</sup>. Meta-analyses have consistently found that fact-checking and debunking interventions can be effective<sup>87,88</sup>, including in the context of countering health misinformation on social media<sup>89</sup>. However, not all medical misperceptions are equally amenable to corrections<sup>90</sup>. In fact, these same analyses note that the effectiveness of interventions is significantly attenuated by (1) the quality of the debunk, (2) the passing of time, and (3) prior beliefs and ideologies. For example, the aforementioned studies on autism<sup>84</sup> and corrective smoking advertisements<sup>85</sup> showed no remaining effect after a 1-week and 6-week follow-up, respectively. When designing corrections,



**Fig. 1 | Best-practice recommendations for effectively debunking misinformation**<sup>91,92</sup>. An effective debunking message should open with the facts and present them in a simple and memorable fashion. The audience should then be warned about the myth (do not repeat the myth more than once). The manipulation technique used to mislead people should subsequently be identified and exposed. End by repeating the facts and emphasizing the correct explanation.

simply labeling information as false or incorrect is generally not sufficient because correcting a myth by means of a simple retraction leaves a gap in people's understanding of why the information is false and what is true instead. Accordingly, the recommendation for practitioners is often to craft much more detailed debunking materials<sup>88</sup>. Reviews of the literature<sup>91,92</sup> have indicated that best practice in designing debunking messages involves (1) leading with the truth, (2) appealing to scientific consensus and authoritative expert sources, (3) ensuring that the correction is easily accessible and not more complex than the initial misinformation, (4) a clear explanation of why the misinformation is wrong, and (5) the provision of a coherent alternative causal explanation (Fig. 1). Although there is generally a lack of comparative research, some recent studies have shown that optimizing debunking messages according to these guidelines enhances their efficacy when compared with alternative or business-as-usual debunking methods<sup>84</sup>.

**Debunking: limitations and future research.** Despite these advances, significant concerns have been raised about the application of such post hoc 'therapeutic' corrections, mostly notably the risk of a correction backfiring so that people end up believing more in the myth as a result of the correction. This backfire effect can occur along two potential dimensions<sup>92,93</sup>, one of which concerns psychological reactance against the correction itself (the 'worldview' backfire effect) whereas the other is concerned with the repetition of false information (the 'familiarity' backfire effect). Although early research was supportive of the fact that, for example, corrections about myths surrounding the flu and MMR vaccine can cause already concerned individuals to become even more hesitant about vaccination decisions<sup>94,95</sup>, more recent studies have failed to find evidence for such worldview backfire effects<sup>93,96</sup>. In fact, while evidence of backfire remains widely cited, recent replications have failed to reproduce such effects when correcting misinformation about vaccinations specifically<sup>97</sup>. Thus, although the effect likely exists, its frequency and intensity is less common than previously thought. Worldview backfire concerns can also be minimized by designing debunking messages in a way that coheres rather than conflicts with the recipients' worldviews<sup>92</sup>. Nonetheless, because debunking forces a rhetorical frame in which the misinformation needs to be repeated in order to correct it (that is, rebutting someone else's claim), there is a risk that such repetition enhances familiarity with the myth while people subsequently fail to encode the correction in long-term memory. Although research clearly shows that people are more likely to believe repeated (mis)information than non-repeated (mis)information<sup>26</sup>, recent work has found that the risk of ironically strengthening a myth as part of a debunking effort is relatively low<sup>93</sup>, especially when the debunking messages feature the correction prominently relative to the misinformation. The consensus is therefore that, although practitioners should be aware of these backfire concerns, they should not prevent the issuing of corrections given the infrequent nature of these side effects<sup>91,93</sup>.

Having said this, there are two other notable problems with therapeutic approaches that limit their efficacy. The first is that retrospective corrections do not reach the same amount of people as the original misinformation. For example, estimates reveal that only about 40% of smokers were exposed to the tobacco industry's court-ordered corrections<sup>98</sup>. A related concern is that, after being exposed, people continue to make inferences on the basis of falsehoods, even when they acknowledge a correction. This phenomenon is known as the 'continued influence of misinformation'<sup>92</sup>, and meta-analyses have found robust evidence of continued influence effects in a wide range of contexts<sup>88,99</sup>.

**Prophylactic treatments: inoculation theory.** Accordingly, researchers have recently begun to explore prophylactic or pre-emptive approaches to countering misinformation, that is,

**Fig. 2 | The process of psychological inoculation against misinformation.**

Psychological inoculation consists of two core components: (1) forewarning people that they may be misled by misinformation (to activate the psychological 'immune system'), and (2) prebunking the misinformation (tactic) by exposing people to a severely weakened dose of it coupled with strong counters and refutations (to generate the cognitive 'antibodies'). Once people have gained 'immunity' they can then vicariously spread the inoculation to others via offline and online interactions.

before an individual has been exposed to or has reached 'infectious' status. Although prebunking is a more general term used for interventions that pre-emptively remind people to 'think before they post'<sup>51</sup>, such reminders in and of themselves do not equip people with any new skills to identify and resist misinformation. The most common framework for preventing unwanted persuasion is psychological inoculation theory<sup>100,101</sup> (Fig. 2). The theory of psychological inoculation follows the biomedical analogy and posits that, just as vaccines trigger the production of antibodies to help confer immunity against future infection, the same can be achieved with information. By pre-emptively forewarning and exposing people to severely weakened doses of misinformation (coupled with strong refutations), people can cultivate cognitive resistance against future misinformation<sup>102</sup>. Inoculation theory operates via two mechanisms, namely (1) motivational threat (a desire to defend oneself from manipulation attacks) and (2) refutational pre-emption or prebunking (pre-exposure to a weakened example of the attack). For example, research has found that inoculating people against conspiratorial arguments about vaccination before (but not after) exposure to a conspiracy theory effectively raised vaccination intentions<sup>103</sup>. Several recent reviews<sup>102,104</sup> and meta-analyses<sup>105</sup> have pointed to the efficacy of psychological inoculation as a robust strategy for conferring immunity to persuasion by misinformation, including many applications in the health domain<sup>106</sup>, such as inoculating people against misinformation about the use of mammography in breast-cancer screening<sup>107</sup>.

Several recent advances, in particular, are worth noting. The first is that the field has moved from 'narrow-spectrum' or 'fact-based' inoculation to 'broad-spectrum' or 'technique-based' immunization<sup>102,108</sup>. The reasoning behind this shift is that, although it is possible to synthesize a severely weakened dose from existing misinformation (and to subsequently refute that weakened example with strong counterarguments), it is difficult to scale the vaccine if this process has to be repeated anew for every piece of misinformation. Instead, scholars have started to identify the common building blocks of misinformation more generally<sup>38,109</sup>, including techniques such as impersonating fake experts and doctors, manipulating people's emotions with fear appeals, and the use of conspiracy theories. Research has found that people can be inoculated against these underlying strategies and, as a result, become relatively more immune to a whole range of misinformation that makes use of these tactics<sup>38,102</sup>. This process is sometimes referred to as cross-protection insofar as inoculating people against one strain offers protection against related and different strains of the same misinformation tactic.

A second advance surrounds the application of active versus passive inoculation. Whereas the traditional inoculation process is passive insofar as people pre-emptively receive the specific refutations from the experimenter, the process of active inoculation encourages people to generate their own ‘antibodies’. Perhaps the best-known example of active inoculation are popular gamified inoculation interventions such as *Bad News*<sup>38</sup> and *GoViral!*<sup>10</sup>, where players step into the shoes of a misinformation producer and are exposed—in a simulated social-media environment—to weakened doses of common strategies used to spread misinformation. As part of this process, players actively generate their own media content and unveil the techniques of manipulation. Research has found that resistance to deception occurs when people (1) recognize their own vulnerability to being persuaded and (2) perceive undue intent to manipulate their opinion<sup>11,12</sup>. These games therefore aim to expose people’s vulnerability, motivating an individuals’ desire to protect themselves against misinformation through pre-exposure to weakened doses. Randomized controlled trials have found that active inoculation games help people identify misinformation<sup>38,110,113,114</sup>, boost confidence in people’s truth-discernment abilities<sup>110,113</sup>, and reduce self-reported sharing of misinformation<sup>110,115</sup>. Yet, like many biological vaccines, research has found that psychological immunity also wanes over time but can be maintained for several months with regular ‘booster’ shots that re-engage people with the inoculation process<sup>114</sup>. A benefit of this line of research is that these gamified interventions have been evaluated and scaled across millions of people as part of the WHO’s ‘Stop The Spread’ campaign and the United Nations’ ‘Verified’ initiative in collaboration with the UK government<sup>110,116</sup>.

**Prebunking: limitations and future research.** A potential limitation is that, although misinformation tropes are often repeated throughout history (consider similarities in the myths that the cowpox vaccine would turn people into human–cow hybrids and the conspiracy theory that COVID-19 vaccines alter human DNA), inoculation does require at least some advanced knowledge of what misinformation (tactic) people might be exposed to in the future<sup>91</sup>. In addition, as healthcare workers are being trained to combat misinformation<sup>117</sup>, it is important to avoid confusion in terminology when using psychological inoculation to counter vaccine hesitancy. For example, the approach can be implemented without making explicit reference to the vaccination analogy and instead can focus on the value of ‘prebunking’ and helping people unveil the techniques of manipulation.

Several other important open questions remain. For example, analogous to recent advances in experimental medicine on the application of therapeutic vaccines—which can still boost immune responses after infection—research has found that inoculation can still protect people from misinformation even when they have already been exposed to misinformation<sup>108,112,118</sup>. This makes conceptual sense insofar it may take repeated exposure or a significant amount of time for misinformation to fully persuade people or become integrated with prior attitudes. Yet it remains conceptually unclear at which point therapeutic inoculation transitions into traditional debunking. Moreover, although some comparisons of active versus passive inoculation approaches exist<sup>105,110</sup>, the evidence base for active forms of inoculation remains relatively small. Similarly, whereas head-to-head studies that compared prebunking to debunking suggest that prevention is indeed better than cure<sup>103</sup>, more comparative research is needed. Research also finds that it is possible for people to vicariously pass on the inoculation interpersonally or on social media, a process known as ‘post-inoculation talk’<sup>104</sup>, which alludes to the possibility of herd immunity in online communities<sup>110</sup>, yet no social-network simulations currently exist that evaluate the potential of inoculative approaches. Current studies are also based on self-reported sharing of misinformation. Future

research will need to evaluate the extent to which inoculation can scale across the population and influence objective news-sharing behavior on social media.

## Conclusion

The spread of misinformation has undermined public-health efforts, from vaccination uptake to public compliance with health-protective behaviors. Research finds that although people are sometimes duped by misinformation because they are distracted on social media and are not paying sufficient attention to accuracy cues, the politicized nature of many public-health challenges suggests that people also believe in and share misinformation because doing so reinforces important socio-political beliefs and identity structures. A more integrated framework is needed that is sensitive to context and can account for varying susceptibility to misinformation on the basis of how people prioritize accuracy and social motives when forming judgments about the veracity of news media. Although ‘exposure’ does not equal ‘infection’, misinformation can spread fast in online networks, and its virality is often aided by the existence of political echo chambers. Importantly, however, the bulk of misinformation on social media often originates from influential accounts and superspreaders. Therapeutic and prophylactic approaches to countering misinformation have both demonstrated some success, but given the continued influence of misinformation after exposure, there is much value in preventative approaches, and more research is needed on how to best combine debunking and prebunking efforts. Further research is also encouraged to outline the benefits and potential challenges to applying the epidemiological model to understand the psychology behind the spread of misinformation. A major challenge for the field moving forward will be clearly defining how misinformation is measured and conceptualized, as well as the need for standardized psychometric instruments that allow for better comparisons of outcomes across studies.

Received: 21 November 2021; Accepted: 24 January 2022;

Published online: 10 March 2022

## References

- Zarocostas, J. How to fight an infodemic. *Lancet* **395**, 676 (2020).
- Allcott, H. & Gentzkow, M. Social media and fake news in the 2016 election. *J. Econ. Perspect.* **31**, 211–236 (2020).
- Grinberg, N. et al. Fake news on Twitter during the 2016 US presidential election. *Science* **363**, 374–378 (2019).
- Roozenbeek, J. et al. Susceptibility to misinformation about COVID-19 around the world. *R. Soc. Open Sci.* **7**, 201199 (2020).
- Romer, D. & Jamieson, K. H. Conspiracy theories as barriers to controlling the spread of COVID-19 in the US. *Soc. Sci. Med.* **263**, 113356 (2020).
- Imhoff, R. & Lamberty, P. A bioweapon or a hoax? The link between distinct conspiracy beliefs about the coronavirus disease (COVID-19) outbreak and pandemic behavior. *Soc. Psychol. Personal. Sci.* **11**, 1110–1118 (2020).
- Freeman, D. et al. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychol. Med.* <https://doi.org/10.1017/S0033291720001890> (2020).
- Loomba, S. et al. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat. Hum. Behav.* **5**, 337–348 (2021).
- Johnson, N. et al. The online competition between pro-and anti-vaccination views. *Nature* **58**, 230–233 (2020).
- Aghababaeian, H. et al. Alcohol intake in an attempt to fight COVID-19: a medical myth in Iran. *Alcohol* **88**, 29–32 (2020).
- Jolley, D. & Paterson, J. L. Pylons ablaze: examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *Br. J. Soc. Psychol.* **59**, 628–640 (2020).
- Dubé, É. et al. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. *Expert Rev. Vaccines* **14**, 99–117 (2015).
- Armstrong, G. M. et al. A longitudinal evaluation of the Listerine corrective advertising campaign. *J. Public Policy Mark.* **2**, 16–28 (1983).
- Albarracín, D. et al. Misleading claims about tobacco products in YouTube videos: experimental effects of misinformation on unhealthy attitudes. *J. Medical Internet Res.* **20**, e9959 (2018).

15. Krishna, A. & Thompson, T. L. Misinformation about health: a review of health communication and misinformation scholarship. *Am. Behav. Sci.* **65**, 316–332 (2021).
16. Kucharski, A. Study epidemiology of fake news. *Nature* **540**, 525–525 (2016).
17. Cinelli, M. et al. The COVID-19 social media infodemic. *Sci. Rep.* **10**, 1–10 (2020).
18. Scales, D. et al. The COVID-19 infodemic—applying the epidemiologic model to counter misinformation. *N. Engl. J. Med.* **385**, 678–681 (2021).
19. Vraga, E. K. & Bode, L. Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. *Polit. Commun.* **37**, 136–144 (2020).
20. Southwell et al. Misinformation as a misunderstood challenge to public health. *Am. J. Prev. Med.* **57**, 282–285 (2019).
21. Wardle, C. & Derakhshan, H. *Information Disorder: toward an Interdisciplinary Framework for Research and Policymaking*. Council of Europe report DGI (2017)09 (Council of Europe, 2017).
22. van der Linden, S. et al. You are fake news: political bias in perceptions of fake news. *Media Cult. Soc.* **42**, 460–470 (2020).
23. Tandoc, E. C. Jr et al. Defining ‘fake news’ a typology of scholarly definitions. *Digit. J.* **6**, 137–153 (2018).
24. Allen, J. et al. Evaluating the fake news problem at the scale of the information ecosystem. *Sci. Adv.* **6**, eaay3539 (2020).
25. Marsh, E. J. & Yang, B. W. in *Misinformation and Mass Audiences* (eds Southwell, B. G., Thorson, E. A., & Sheble, L.) 15–34 (University of Texas Press, 2018).
26. Dechêne, A. et al. The truth about the truth: a meta-analytic review of the truth effect. *Pers. Soc. Psychol. Rev.* **14**, 238–257 (2010).
27. Lewis, T. Eight persistent COVID-19 myths and why people believe them. *Scientific American*. <https://www.scientificamerican.com/article/eight-persistent-covid-19-myths-and-why-people-believe-them/> (2020).
28. Wang, W. C. et al. On known unknowns: fluency and the neural mechanisms of illusory truth. *J. Cogn. Neurosci.* **28**, 739–746 (2016).
29. Pennycook, G. et al. Prior exposure increases perceived accuracy of fake news. *J. Exp. Psychol. Gen.* **147**, 1865–1880 (2018).
30. Fazio, L. K. et al. Repetition increases perceived truth equally for plausible and implausible statements. *Psychon. Bull. Rev.* **26**, 1705–1710 (2019).
31. Fazio, L. K. et al. Knowledge does not protect against illusory truth. *J. Exp. Psychol. Gen.* **144**, 993–1002 (2015).
32. De Keersmaecker, J. et al. Investigating the robustness of the illusory truth effect across individual differences in cognitive ability, need for cognitive closure, and cognitive style. *Pers. Soc. Psychol. Bull.* **46**, 204–215 (2020).
33. Guess, A. et al. Less than you think: prevalence and predictors of fake news dissemination on Facebook. *Sci. Adv.* **5**, eau4586 (2019).
34. Saunders, J. & Jess, A. The effects of age on remembering and knowing misinformation. *Memory* **18**, 1–11 (2010).
35. Brashier, N. M. & Schacter, D. L. Aging in an era of fake news. *Curr. Dir. Psychol. Sci.* **29**, 316–323 (2020).
36. Pennycook, G. & Rand, D. G. Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* **188**, 39–50 (2019).
37. Imhoff, R. et al. Conspiracy mentality and political orientation across 26 countries. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-021-01258-7> (2022).
38. Roodenbeek, J. & van der Linden, S. Fake news game confers psychological resistance against online misinformation. *Humanit. Soc. Sci. Commun.* **5**, 1–10 (2019).
39. Van der Linden, S. et al. The paranoid style in American politics revisited: an ideological asymmetry in conspiratorial thinking. *Polit. Psychol.* **42**, 23–51 (2021).
40. De keersmaecker, J. & Roets, A. ‘Fake news’: incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions. *Intelligence* **65**, 107–110 (2017).
41. Bronstein, M. V. et al. Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *J. Appl. Res. Mem.* **8**, 108–117 (2019).
42. Greene, C. M. et al. Misremembering Brexit: partisan bias and individual predictors of false memories for fake news stories among Brexit voters. *Memory* **29**, 587–604 (2021).
43. Gawronski, B. Partisan bias in the identification of fake news. *Trends Cogn. Sci.* **25**, 723–724 (2021).
44. Rathje, S. et al. Meta-analysis reveals that accuracy nudges have little to no effect for US conservatives: Regarding Pennycook et al. (2020). *Psychol. Sci.* <https://doi.org/10.25384/SAGE.12594110.v2> (2021).
45. Pennycook, G. & Rand, D. G. The psychology of fake news. *Trends Cogn. Sci.* **22**, 388–402 (2021).
46. van der Linden, S. et al. How can psychological science help counter the spread of fake news? *Span. J. Psychol.* **24**, e25 (2021).
47. Evans, J. S. B. In two minds: dual-process accounts of reasoning. *Trends Cogn. Sci.* **7**, 454–459 (2003).
48. Bago, B. et al. Fake news, fast and slow: deliberation reduces belief in false (but not true) news headlines. *J. Exp. Psychol. Gen.* **149**, 1608–1613 (2020).
49. Scherer, L. D. et al. Who is susceptible to online health misinformation? A test of four psychosocial hypotheses. *Health Psychol.* **40**, 274–284 (2021).
50. Pennycook, G. et al. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.* **31**, 770–780 (2020).
51. Pennycook, G. et al. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021).
52. Swami, V. et al. Analytic thinking reduces belief in conspiracy theories. *Cognition* **133**, 572–585 (2014).
53. Kunda, Z. The case for motivated reasoning. *Psychol. Bull.* **108**, 480–498 (1990).
54. Kahan, D. M. in *Emerging Trends in the Social and Behavioral sciences* (eds Scott, R. & Kosslyn, S.) 1–16 (John Wiley & Sons, 2016).
55. Bolsen, T. et al. The influence of partisan motivated reasoning on public opinion. *Polit. Behav.* **36**, 235–262 (2014).
56. Osmundsen, M. et al. Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *Am. Polit. Sci. Rev.* **115**, 999–1015 (2021).
57. Van Bavel, J. J. et al. Political psychology in the digital (mis) information age: a model of news belief and sharing. *Soc. Issues Policy Rev.* **15**, 84–113 (2020).
58. Rathje, S. et al. Out-group animosity drives engagement on social media. *Proc. Natl Acad. Sci. USA* **118**, e2024292118 (2021).
59. Kahan, D. M. et al. Motivated numeracy and enlightened self-government. *Behav. Public Policy* **1**, 54–86 (2017).
60. Kahan, D. M. et al. The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nat. Clim. Chang.* **2**, 732–735 (2012).
61. Drummond, C. & Fischhoff, B. Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. *Proc. Natl Acad. Sci. USA* **114**, 9587–9592 (2017).
62. Traberg, C. S. & van der Linden, S. Birds of a feather are persuaded together: perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Pers. Individ. Differ.* **185**, 111269 (2022).
63. Roodenbeek, J. et al. How accurate are accuracy-nudge interventions? A preregistered direct replication of Pennycook et al. (2020). *Psychol. Sci.* **32**, 1169–1178 (2021).
64. Persson, E. et al. A preregistered replication of motivated numeracy. *Cognition* **214**, 104768 (2021).
65. Connor, P. et al. Motivated numeracy and active reasoning in a Western European sample. *Behav. Public Policy* 1–23 (2020).
66. van der Linden, S. et al. Scientific agreement can neutralize politicization of facts. *Nat. Hum. Behav.* **2**, 2–3 (2018).
67. Tappin, B. M. et al. Rethinking the link between cognitive sophistication and politically motivated reasoning. *J. Exp. Psychol. Gen.* **150**, 1095–1114 (2021).
68. Tappin, B. M. et al. Thinking clearly about causal inferences of politically motivated reasoning: why paradigmatic study designs often undermine causal inference. *Curr. Opin. Behav. Sci.* **34**, 81–87 (2020).
69. Druckman, J. N. & McGrath, M. C. The evidence for motivated reasoning in climate change preference formation. *Nat. Clim. Chang.* **9**, 111–119 (2019).
70. Juul, J. L. & Ugander, J. Comparing information diffusion mechanisms by matching on cascade size. *Proc. Natl. Acad. Sci. USA* **118**, e210078611 (2021).
71. Vosoughi, S. et al. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
72. Cinelli, M. et al. The echo chamber effect on social media. *Proc. Natl Acad. Sci. USA* **118**, e2023301118 (2021).
73. Guess, A. et al. Exposure to untrustworthy websites in the 2016 US election. *Nat. Hum. Behav.* **4**, 472–480 (2020).
74. Yang, K. C. et al. The COVID-19 infodemic: Twitter versus Facebook. *Big Data Soc.* **8**, 20539517211013861 (2021).
75. Del Vicario, M. et al. The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559 (2016).
76. Zollo, F. et al. Debunking in a world of tribes. *PLoS ONE* **12**, e0181821 (2017).
77. Guess, A. M. (Almost) everything in moderation: new evidence on Americans’ online media diets. *Am. J. Pol. Sci.* **65**, 1007–1022 (2021).
78. Törnberg, P. Echo chambers and viral misinformation: modeling fake news as complex contagion. *PLoS ONE* **13**, e0203958 (2018).
79. Choi, D. et al. Rumor propagation is amplified by echo chambers in social media. *Sci. Rep.* **10**, 1–10 (2020).
80. Eurobarometer on Fake News and Online Disinformation. *European Commission* <https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation> (2018).
81. Altay, S. et al. If this account is true, it is most enormously wonderful: interestingness-if-true and the sharing of true and false news. *Digit. Journal.* <https://doi.org/10.1080/21670811.2021.1941163> (2021).

82. Kalla, J. L. & Broockman, D. E. The minimal persuasive effects of campaign contact in general elections: evidence from 49 field experiments. *Am. Political Sci. Rev.* **112**, 148–166 (2018).
83. Matz, S. C. et al. Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl Acad. Sci. USA* **114**, 12714–12719 (2017).
84. Paynter, J. et al. Evaluation of a template for countering misinformation—real-world autism treatment myth debunking. *PLoS ONE* **14**, e0210746 (2019).
85. Smith, P. et al. Correcting over 50 years of tobacco industry misinformation. *Am. J. Prev. Med.* **40**, 690–698 (2011).
86. Yousuf, H. et al. A media intervention applying debunking versus non-debunking content to combat vaccine misinformation in elderly in the Netherlands: a digital randomised trial. *EClinicalMedicine* **35**, 100881 (2021).
87. Walter, N. & Murphy, S. T. How to unring the bell: a meta-analytic approach to correction of misinformation. *Commun. Monogr.* **85**, 423–441 (2018).
88. Chan, M. P. S. et al. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. *Psychol. Sci.* **28**, 1531–1546 (2017).
89. Walter, N. et al. Evaluating the impact of attempts to correct health misinformation on social media: a meta-analysis. *Health Commun.* **36**, 1776–1784 (2021).
90. Aikin, K. J. et al. Correction of overstatement and omission in direct-to-consumer prescription drug advertising. *J. Commun.* **65**, 596–618 (2015).
91. Lewandowsky, S. et al. *The Debunking Handbook 2020* <https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf> (2020).
92. Lewandowsky, S. et al. Misinformation and its correction: continued influence and successful debiasing. *Psychol. Sci. Publ. Int.* **13**, 106–131 (2012).
93. Swire-Thompson, B. et al. Searching for the backfire effect: measurement and design considerations. *J. Appl. Res. Mem. Cogn.* **9**, 286–299 (2020).
94. Nyhan, B. et al. Effective messages in vaccine promotion: a randomized trial. *Pediatrics* **133**, e835–e842 (2014).
95. Nyhan, B. & Reifler, J. Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine* **33**, 459–464 (2015).
96. Wood, T. & Porter, E. The elusive backfire effect: mass attitudes' steadfast factual adherence. *Polit. Behav.* **41**, 135–163 (2019).
97. Haglin, K. The limitations of the backfire effect. *Res. Politics* <https://doi.org/10.1177/2053168017716547> (2017).
98. Chido-Amajuoysi et al. Exposure to court-ordered tobacco industry antismoking advertisements among US adults. *JAMA Netw. Open* **2**, e196935 (2019).
99. Walter, N. & Tukachinsky, R. A meta-analytic examination of the continued influence of misinformation in the face of correction: how powerful is it, why does it happen, and how to stop it? *Commun. Res.* **47**, 155–177 (2020).
100. Papageorgis, D. & McGuire, W. J. The generality of immunity to persuasion produced by pre-exposure to weakened counterarguments. *J. Abnorm. Psychol.* **62**, 475–481 (1961).
101. McGuire, W. J. in *Advances in Experimental Social Psychology* (ed Berkowitz, L.) 191–229 (Academic Press, 1964).
102. Lewandowsky, S. & van der Linden, S. Countering misinformation and fake news through inoculation and prebunking. *Eur. Rev. Soc. Psychol.* **32**, 348–384 (2021).
103. Jolley, D. & Douglas, K. M. Prevention is better than cure: addressing anti vaccine conspiracy theories. *J. Appl. Soc. Psychol.* **47**, 459–469 (2017).
104. Compton, J. et al. Inoculation theory in the post-truth era: extant findings and new frontiers for contested science, misinformation, and conspiracy theories. *Soc. Personal. Psychol.* **15**, e12602 (2021).
105. Banas, J. A. & Rains, S. A. A meta-analysis of research on inoculation theory. *Commun. Monogr.* **77**, 281–311 (2010).
106. Compton, J. et al. Persuading others to avoid persuasion: Inoculation theory and resistant health attitudes. *Front. Psychol.* **7**, 122 (2016).
107. Iles, I. A. et al. Investigating the potential of inoculation messages and self-affirmation in reducing the effects of health misinformation. *Sci. Commun.* **43**, 768–804 (2021).
108. Cook et al. Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLoS ONE* **12**, e0175799 (2017).
109. van der Linden, S., & Roozenbeek, J. in *The Psychology of Fake News: Accepting, Sharing, and Correcting Misinformation* (eds Greifeneder, R., Jaffe, M., Newman, R., & Schwarz, N.) 147–169 (Psychology Press, 2020).
110. Basol, M. et al. Towards psychological herd immunity: cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data Soc.* **8**, 20539517211013868 (2021).
111. Sagarin, B. J. et al. Dispelling the illusion of invulnerability: the motivations and mechanisms of resistance to persuasion. *J. Pers. Soc. Psychol.* **83**, 526–541 (2002).
112. van der Linden, S. et al. Inoculating the public against misinformation about climate change. *Glob. Chall.* **1**, 1600008 (2017).
113. Basol, M. et al. Good news about bad news: gamified inoculation boosts confidence and cognitive immunity against fake news. *J. Cogn.* **3**, 2 (2020).
114. Maertens, R. et al. Long-term effectiveness of inoculation against misinformation: three longitudinal experiments. *J. Exp. Psychol. Appl.* **27**, 1–16 (2021).
115. Roozenbeek, J., & van der Linden, S. Breaking Harmony Square: a game that ‘inoculates’ against political misinformation. *The Harvard Kennedy School Misinformation Review* <https://doi.org/10.37016/mr-2020-47> (2020).
116. What is Go Viral? *World Health Organization* <https://www.who.int/news-room/23-09-2021-what-is-go-viral> (WHO, 2021).
117. Abbas, J. COVID-19 conspiracies and beyond: how physicians can deal with patients’ misinformation. *JAMA* **325**, 208–210 (2021).
118. Compton, J. Prophylactic versus therapeutic inoculation treatments for resistance to influence. *Commun. Theory* **30**, 330–343 (2020).
119. Lazer, D. M. et al. The science of fake news. *Science* **359**, 1094–1096 (2018).
120. Pennycook, G. & Rand, D. G. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *J. Pers.* **88**, 185–200 (2020).
121. Benton, J. Facebook sent a ton of traffic to a Chicago Tribune story. So why is everyone mad at them? *NiemanLab* <https://www.niemanlab.org/2021/08/facebook-sent-a-ton-of-traffic-to-a-chicago-tribune-story-so-why-is-everyone-mad-at-them/> (2021).
122. Poutoglou, F. et al. Ibuprofen and COVID-19 disease: separating the myths from facts. *Expert Rev. Respir. Med.* **15**, 979–983 (2021).
123. Maertens, R. et al. The Misinformation Susceptibility Test (MIST): a psychometrically validated measure of news veracity discernment. Preprint at PsyArXiv <https://doi.org/10.31234/osf.io/gk68h> (2021).

## Acknowledgements

I am grateful for support from the IRIS Infodemic Coalition (UK Government, no. SCH-00001-3391) and JITSUVAZ (EU 2020 Horizon no. 964728). I thank the Cambridge Social Decision-Making Lab and credit R. Maertens in particular for his help with designing Fig. 2.

## Competing interests

S.V.D.L. co-designed several interventions in collaboration with the UK Government, DROG, and the WHO, reviewed in this paper, namely *GoViral!* and *Bad News*. He does not receive nor hold any financial interests in these interventions. He has received research funding from or consulted for the UK government, the US Government, the European Commission, Facebook, Google, WhatsApp, Edelman, the United Nations, and the WHO on misinformation and infodemic management.

## Additional information

**Correspondence** should be addressed to Sander van der Linden.

**Peer review information** *Nature Medicine* thanks Brian Southwell and Alison Buttenheim for their contribution to the peer review of this work. Karen O’Leary was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature America, Inc. 2022