



The psychological drivers of misinformation belief and its resistance to correction

Ullrich K. H. Ecker¹✉, Stephan Lewandowsky², John Cook³, Philipp Schmid⁴, Lisa K. Fazio⁵, Nadia Brashier^{6,7}, Panayiota Kendeou⁸, Emily K. Vraga⁹ and Michelle A. Amazeen¹⁰

Abstract | Misinformation has been identified as a major contributor to various contentious contemporary events ranging from elections and referenda to the response to the COVID-19 pandemic. Not only can belief in misinformation lead to poor judgements and decision-making, it also exerts a lingering influence on people's reasoning after it has been corrected — an effect known as the continued influence effect. In this Review, we describe the cognitive, social and affective factors that lead people to form or endorse misinformed views, and the psychological barriers to knowledge revision after misinformation has been corrected, including theories of continued influence. We discuss the effectiveness of both pre-emptive ('prebunking') and reactive ('debunking') interventions to reduce the effects of misinformation, as well as implications for information consumers and practitioners in various areas including journalism, public health, policymaking and education.

Misinformation — which we define as any information that turns out to be false — poses an inevitable challenge for human cognition and social interaction because it is a consequence of the fact that people frequently err and sometimes lie¹. However, this fact is insufficient to explain the rise of misinformation, and its subsequent influence on memory and decision-making, as a major challenge in the twenty-first century^{2–4}. Misinformation has been identified as a contributor to various contentious events, ranging from elections and referenda⁵ to political or religious persecution⁶ and to the global response to the COVID-19 pandemic⁷.

The psychology and history of misinformation cannot be fully grasped without taking into account contemporary technology. Misinformation helped bring Roman emperors to power⁸, who used messages on coins as a form of mass communication⁹, and Nazi propaganda heavily relied on the printed press, radio and cinema¹⁰. Today, misinformation campaigns can leverage digital infrastructure that is unparalleled in its reach. The internet reaches billions of individuals and enables senders to tailor persuasive messages to the specific psychological profiles of individual users^{11,12}. Moreover, social media users' exposure to information that challenges their worldviews can be limited when communication environments foster confirmation of previous beliefs — so-called echo chambers^{13,14}. Although there is some

controversy about echo chambers and their impact on people's beliefs and behaviours^{12,15}, the internet is an ideal medium for the fast spread of falsehoods at the expense of accurate information¹⁶. However, the prevalence of misinformation cannot be attributed only to technology: conventional efforts to combat misinformation have also not been as successful as hoped² — these include educational efforts that focus on merely conveying factual knowledge and corrective efforts that merely retract misinformation.

For decades, science communication has relied on an information deficit model when responding to misinformation, focusing on people's misunderstanding of, or lack of access to, facts¹⁷. Thus, a thorough and accessible explanation of facts should overcome the impact of misinformation. However, the information deficit model ignores the cognitive, social and affective drivers of attitude formation and truth judgements^{18–20}. For example, some individuals deny the existence of climate change or reject vaccinations despite being aware of a scientific consensus to the contrary^{21,22}. This rejection of science is not the result of mere ignorance but is driven by factors such as conspiratorial mentality, fears, identity expression and motivated reasoning — reasoning driven more by personal or moral values than objective evidence^{19,23–26}. Thus, to understand the psychology of misinformation and how it might be countered, it is

✉e-mail: ullrich.ecker@uwa.edu.au
<https://doi.org/10.1038/s44159-021-00006-y>

essential to consider the cognitive architecture and social context of individual decision makers.

In this Review, we describe the cognitive, social and affective processes that make misinformation stick and leave people vulnerable to the formation of false beliefs. We review the theoretical models that have been proposed to explain misinformation's resistance to correction. We provide guidance on countering misinformation, including educational and pre-emptive interventions, refutations and psychologically informed technological solutions. Finally, we return to the broader societal trends that have contributed to the rise of misinformation and discuss its practical implications on journalism, education and policymaking.

Different types of misinformation exist — for example, misinformation that goes against scientific consensus or misinformation that contradicts simple, objectively true facts. Moreover, the term *disinformation* is often specifically used for the subset of misinformation that is spread intentionally²⁷. More research is needed on the extent to which different types of misinformation might be associated with differential psychological impacts and barriers for revision, and to establish the extent to which people infer intentionality and how this might affect their processing of the false information. Thus, in this Review we do not draw a sharp distinction between misinformation and disinformation, or different types of misinformation. We use the term *misinformation* as an umbrella term referring to any information that turns out to be false and reserve the term *disinformation* for misinformation that is spread with intention to harm or deceive.

Drivers of false beliefs

The formation of false beliefs all but requires exposure to false information. However, lack of access to high-quality information is not necessarily the primary precursor to false-belief formation; a range of cognitive, social and affective factors influence the formation of false beliefs (FIG. 1). False beliefs generally arise through the same mechanisms that establish accurate beliefs^{28,29}. When deciding what is true, people are often biased to believe in the validity of information³⁰, and 'go with their gut' and intuitions instead of deliberating^{31,32}. For example, in March 2020, 31% of Americans agreed that COVID-19 was purposefully created and spread³³, despite the

absence of any credible evidence for its intentional development. People are likely to have encountered conspiracy theories about the source of the virus multiple times, which might have contributed to this widespread belief because simply repeating a claim makes it more believable than presenting it only once^{34,35}. This illusory truth effect arises because people use peripheral cues such as familiarity (a signal that a message has been encountered before)³⁶, processing fluency (a signal that a message is either encoded or retrieved effortlessly)^{37,38} and cohesion (a signal that the elements of a message have references in memory that are internally consistent)³⁹ as signals for truth, and the strength of these cues increases with repetition. Thus, repetition increases belief in both misinformation and facts^{40–43}. Illusory truth can persist months after first exposure⁴⁴, regardless of cognitive ability⁴⁵ and despite contradictory advice from an accurate source⁴⁶ or accurate prior knowledge^{18,47}.

Another 'shortcut' for truth might involve defaulting to one's own personal views. Overall belief in news headlines is higher when the news headlines complement the reader's worldview⁴⁸. Political partisanship can also contribute to false memories for made-up scandals⁴⁹. However, difficulties discerning true from false news headlines can also arise from intuitive (or 'lazy') thinking rather than the impact of worldviews⁴⁸. In one study, participants received questions ('If you're running a race and you pass the person in second place, what place are you in?') with intuitive, but incorrect, answers ('first place'). Participants who answered these questions correctly were better able to discern fake from real headlines than participants who answered these questions incorrectly, independently of whether the headlines aligned with their political ideology⁵⁰. A link has also been reported between intuitive thinking and greater belief in COVID-19 being a hoax, and reduced adherence to public health measures⁵¹.

Similarly, allowing people to deliberate can improve their judgements. If quick evaluation of a headline is followed by an opportunity to rethink, belief in fake news — but not factual news — is reduced⁵². Likewise, encouraging people to 'think like fact checkers' leads them to rely more on their own prior knowledge instead of heuristics. For example, prior exposure to statements such as 'Deer meat is called veal' makes these statements seem truer than similar statements encountered for the first time, even when people know the truth (in this case that the correct term is venison⁴⁷). However, asking people to judge whether the statement is true at initial exposure protects them from subsequently accepting contradictions of well-known facts⁵³.

The information source also provides important social cues that influence belief formation. In general, messages are more persuasive and seem more true when they come from sources perceived to be credible rather than non-credible⁴². People trust human information sources more if they perceive the source as attractive, powerful and similar to themselves⁵⁴. These source judgements are naturally imperfect — people believe in-group members more than out-group members⁵⁵, tend to weigh opinions equally regardless of the

Author addresses

¹School of Psychological Science, University of Western Australia, Perth, Western Australia, Australia.

²School of Psychological Science, University of Bristol, Bristol, UK.

³Climate Change Communication Research Hub, Monash University, Melbourne, Victoria, Australia.

⁴Media and Communication Science, University of Erfurt, Erfurt, Germany.

⁵Department of Psychology and Human Development, Vanderbilt University, Nashville, TN, USA.

⁶Department of Psychology, Harvard University, Cambridge, MA, USA.

⁷Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA.

⁸Department of Educational Psychology, University of Minnesota, Minneapolis, MN, USA.

⁹Hubbard School of Journalism and Mass Communication, University of Minnesota, Minneapolis, MN, USA.

¹⁰College of Communication, Boston University, Boston, MA, USA.

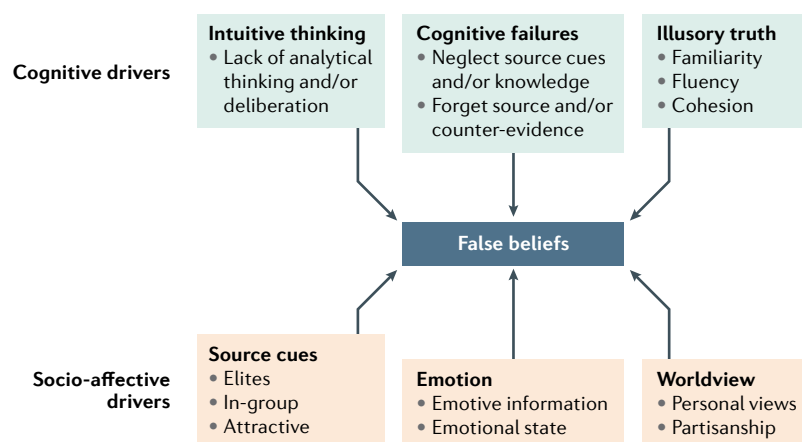


Fig. 1 | Drivers of false beliefs. Some of the main cognitive (green) and socio-affective (orange) factors that can facilitate the formation of false beliefs when individuals are exposed to misinformation. Not all factors will always be relevant, but multiple factors often contribute to false beliefs.

competence of those expressing them⁵⁶ and overestimate how much their beliefs overlap with other people's, which can lead to the perception of a false consensus⁵⁷. Experts and political elites are trusted by many and have the power to shape public perceptions^{58,59}; therefore, it can be especially damaging when leaders make false claims. For example, false claims about public health threats such as COVID-19 made by political leaders can reduce the perceived threat of the virus as well as the perceived efficacy of countermeasures, decreasing adherence to public health measures^{60,61}.

Moreover, people often overlook, ignore, forget or confuse cues about the source of information⁶². For example, for online news items, a logo banner specifying the publisher (for example, a reputable media outlet or a dubious web page) has been found not to decrease belief in fake news or increase belief in factual news⁶³. In the aggregate, groups of laypeople perform as well as professional fact checkers at categorizing news outlets as trustworthy, hyper-partisan or fake⁶⁴. However, when acting alone, individuals — unlike fact checkers — tend to disregard the quality of the news outlet and judge a headline's accuracy based primarily on the plausibility of the content⁶⁵. Similarly, although people are quick to distrust others who share fake news⁶⁵, they frequently forget information sources⁶⁶. This tendency is concerning: even though a small number of social media accounts spread an outsized amount of misleading content^{67–69}, if consumers do not remember the dubious origin, they might not discount the content accordingly.

The emotional content of the information shared also affects false-belief formation. Misleading content that spreads quickly and widely ('virally') on the internet often contains appeals to emotion, which can increase persuasion. For example, messages that aim to generate fear of harm can successfully change attitudes, intentions and behaviours under certain conditions if recipients feel they can act effectively to avoid the harm⁷⁰. Moreover, according to a preprint that has not been peer-reviewed, 'happy thoughts' are more believable than neutral ones⁷¹. People seem to understand the association between

emotion and persuasion, and naturally shift towards more emotional language when attempting to convince others⁷². For example, anti-vaccination activists frequently use emotional language⁷³. Emotion can be persuasive because it distracts readers from potentially more diagnostic cues, such as source credibility. In one study, participants read positive, neutral and negative headlines about the actions of specific people; social judgements about the people featured in the headlines were strongly determined by emotional valence of the headline but unaffected by trustworthiness of the news source⁷⁴.

Inferences about information are also affected by one's own emotional state. People tend to ask themselves 'How do I feel about this claim?', which can lead to influences of a person's mood on claim evaluation⁷⁵. Using feelings as information can leave people susceptible to deception⁷⁶, and encouraging people to 'rely on their emotions' increases their vulnerability to misinformation⁷⁷. Likewise, some specific emotional states such as a happy mood can make people more vulnerable to deception⁷⁸ and illusory truth⁷⁹. Thus, one functional feature of a sad mood might be that it reduces gullibility⁸⁰. Anger has also been shown to promote belief in politically concordant misinformation⁸¹ as well as COVID-19 misinformation⁸². Finally, social exclusion, which is likely to induce a negative mood, can increase susceptibility to conspiratorial content^{83,84}.

In sum, the drivers of false beliefs are multifold and largely overlooked by a simple information deficit model. The drivers include cognitive factors, such as use of intuitive thinking and memory failures; social factors, such as reliance on source cues to determine truth; and affective factors, such as the influence of mood on credibility. Although we have focused on false-belief formation here, the psychology behind sharing misinformation is a related area of active study (BOX 1).

Barriers to belief revision

A tacit assumption of the information deficit model is that false beliefs can easily be corrected by providing relevant facts. However, misinformation can often continue to influence people's thinking even after they receive a correction and accept it as true. This persistence is known as the continued influence effect (CIE)^{85–88}.

In the typical CIE laboratory paradigm, participants are presented with a report of an event (for example, a fire) that contains a critical piece of information related to the event's cause ('the fire was probably caused by arson'). That information might be subsequently challenged by a correction, which can take the form of a retraction (a simple negation, such as 'it is not true that arson caused the fire') or a refutation (a more detailed correction that explains why the misinformation was false). When reasoning about the event later (for example, responding to questions such as 'what should authorities do now?'), individuals often continue to rely on the critical information even after receiving — and being able to recall — a correction⁸⁹. Variants of this paradigm have used false real-world claims or urban myths^{90–92}. Corrected misinformation can also continue to influence the amount a person is willing to pay for a consumer product or their propensity to promote a social media

Box 1 | Why people share misinformation

Online misinformation transmission involves both a receiver (the person encountering the misinformation) and a sender (the person making or sharing the misinformation). Thus, it is crucial to consider why people share misinformation with others. On social media, sharing is often dictated by what captures attention. Moral-emotional words such as ‘fight’, ‘greed’, ‘evil’ and ‘punish’ are prioritized in early visual attention over other arousing words²⁷⁶ and also lead to increased sharing. For example, adding a single moral-emotional word to tweets about contentious political issues such as gun control increases retweets by 20%²⁷⁷. An angry mood can also boost misinformation sharing⁸². Because social media algorithms promote content that is likely to be shared, the interplay of psychological tendencies and technological optimization can thus easily lead to viral spread of misinformation online.

‘Lazy’ or intuitive thinking can also lead people to share content that they might recognize as false if they thought about it more. Accordingly, asking people to explain how they know that news headlines are true or false reduces sharing of false political headlines²⁷⁸, and brief accuracy nudges — simple interventions that prompt people to consider the accuracy of the information they encounter or share — can reduce sharing of false news about politics²⁰⁷ and COVID-19 (REF. 279). These studies suggest that to the extent that people pay attention to accuracy, they are likely to share things they genuinely believe. Most people report that they would need to be paid to share false news; even when stories favour their political views, they worry about possible reputation costs from sharing false news⁶⁵. Those reputation costs are real — over half of social media users report that they have stopped following someone who posted ‘made-up news and information’²⁸⁰.

If a person’s focus is not on information veracity, they might share misinformation for other reasons²⁰¹. Indeed, 14% of respondents in a 2016 US survey admitted to knowingly sharing false news²⁸¹. There are some innocuous reasons to intentionally spread falsehoods; for example, it is tempting to share information that would be ‘interesting (or consequential) if true’²⁸². Likewise, findings from a preprint that has not been peer-reviewed suggest that people might share positive but questionable claims that could make others feel better, such as ‘A cat saved a woman’s life by scaring off a bear trying to attack her’⁷⁷. There are also self-serving motives for sharing, such as to signal group membership²⁸³ or for self-promotion²⁶⁰. Finally, some people share misinformation to fuel moral outrage in others^{277,284}. One non-peer reviewed preprint suggests that some people share hostile political rumours and conspiracy theories to incite chaos; this desire to ‘watch the world burn’ is even stronger following social exclusion²⁸⁵. With these alternative goals in mind, the viral nature of misinformation does not occur despite its low veracity but because of its ability to fulfil other psychological needs¹¹.

post^{93–95}. The CIE might be an influential factor in the persistence of beliefs that there is a link between vaccines and autism despite strong evidence discrediting this link^{96,97} or that weapons of mass destruction were found in Iraq in 2003 despite no supporting evidence⁹⁸. The CIE has primarily been conceptualized as a cognitive effect, with social and affective underpinnings.

Cognitive factors. Theoretical accounts of the CIE draw heavily on models of memory in which information is organized in interconnected networks and the availability of information is determined by its level of activation^{99,100} (FIG. 2). When information is encoded into memory and then new information that discredits it is learned, the original information is not simply erased or replaced¹⁰¹. Instead, misinformation and corrective information coexist and compete for activation. For example, misinformation that a vaccine has caused an unexpectedly large number of deaths might be incorporated with knowledge related to diseases, vaccinations and causes of death. A subsequent correction that the information about vaccine-caused deaths was inaccurate will also be added to memory and is likely to result in some knowledge revision. However, the misinformation

will remain in memory and can potentially be reactivated and retrieved later on.

One school of thought — the integration account — suggests that the CIE arises when a correction is not sufficiently encoded and integrated with the misinformation in the memory network (FIG. 2a). There is robust evidence that integration of the correction and misinformation is a necessary, albeit not sufficient, condition for memory updating and knowledge revision¹⁰⁰. This view implies that a successful revision requires detecting a conflict between the misinformation and the correction, the co-activation of both representations in memory, and their subsequent integration¹⁰². Evidence for this account comes from the success of interventions that bolster conflict detection, co-activation, and integration of misinformation and correction^{103,104}. Assuming that information integration relies on processing in working memory (the short-term store used to briefly hold and manipulate information in the service of thinking and reasoning), the finding that lower working memory capacity predicts greater susceptibility to the CIE is also in line with this account¹⁰⁵ (although it has not been replicated¹⁰⁶). This theory further assumes that as the amount of integrated correct information increases, memory for the correction becomes stronger, at the expense of memory for the misinformation¹⁰². Thus, both the interconnectedness and the amount of correct information can influence the success of memory revision.

An alternative account is based on the premise that the CIE arises from selective retrieval of the misinformation even when corrective information is present in memory (FIG. 2b). For example, it has been proposed that a retraction causes the misinformation representation to be tagged as false¹⁰⁷. The misinformation can be retrieved without the false tag, but the false tag cannot be retrieved without concurrent retrieval of the misinformation. One instantiation of this selective-retrieval view appeals to a dual-process mechanism, which assumes that retrieval can occur based on an automatic, effortless process signalling information familiarity (‘I think I have heard this before’) or a more strategic, effortful process of recollection that includes contextual detail (‘I read about this in yesterday’s newspaper’)¹⁰⁸. According to this account of continued influence, the CIE can arise if there is automatic, familiarity-driven retrieval of the misinformation (for example, in response to a cue), without explicit recollection of the corrective information and associated post-retrieval suppression of the misinformation^{107,109}.

Evidence for this account comes from studies demonstrating that the CIE increases as a function of factors associated with increased familiarity (such as repetition)¹⁰⁷ and reduced recollection (such as advanced participant age and longer study-test delays)⁹². Neuroimaging studies have suggested that activity during retrieval, when participants answer inference questions about an encoded event — but not when the correction is encoded — is associated with continued reliance on corrected misinformation^{110,111}. This preliminary neuroimaging evidence generally supports the selective-retrieval account of the CIE, although it

suggests that the CIE is driven by misinformation recollection rather than misinformation familiarity, which is at odds with the dual-process interpretation.

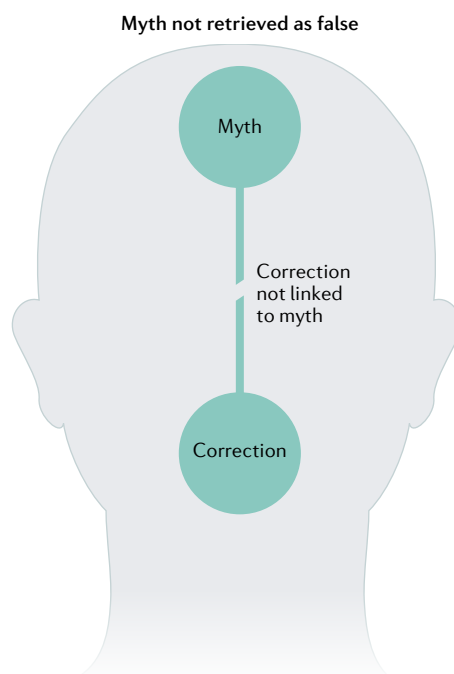
Both of these complementary theoretical accounts of the CIE can explain the superiority of detailed refutations over retractions^{92,112,113}. Provision of additional corrective information can strengthen the activation of correct information in memory or provide more detail to support recollection of the correction^{89,103}, which makes a factual correction more enduring than the misinformation⁹⁰. Because a simple retraction will create a gap in a person's mental model, especially in situations that require a causal explanation (for example, a fire must be caused by something), a refutation that can fill in details of a causal, plausible, simple and memorable alternative explanation will reduce subsequent recall of the retracted misinformation.

Social and affective factors. These cognitive accounts do not explicitly consider the influence of social and affective mechanisms on the CIE. One socio-affective factor is source credibility, the perceived trustworthiness and expertise of the sources providing the misinformation and correction. Although source credibility has been found to exert little influence on acceptance of misinformation if the source is a media outlet^{63,114}, there is generally strong evidence that credibility has significant impact on acceptance of misinformation from non-media sources^{42,88,115}.

The credibility of a correction source also matters for (post-correction) misinformation reliance¹¹⁶, although perhaps less than the credibility of the misinformation source⁸⁸. The effectiveness of factual corrections might depend on perceived trustworthiness rather than perceived expertise of the correction source^{117,118}, although perceived expertise might matter more in science-related contexts, such as health misinformation^{119,120}. It can also be quite rational to discount a correction if the correction source is low in credibility^{121,122}. Further complicating matters, the perceived credibility of a source varies across recipients. In extreme cases, people with strong conspiratorial ideation tendencies might mistrust any official source (for example, health authorities)^{19,26}. More commonly, people tend to trust sources that are perceived to share their values and worldviews^{54,55}.

A second key socio-affective factor is worldview — a person's values and belief system that grounds their personal and sociocultural identity. Corrections attacking a person's worldview can be ineffective¹²³ or backfire^{25,124}. Such corrections can be experienced as attacking one's identity, resulting in a chain reaction of appraisals and emotional responses that hinder information revision^{19,125}. For example, if a message is appraised as an identity threat (for example, a correction that the risks of a vaccine do not outweigh the risks of a disease might be perceived as an identity threat by a person identifying as an anti-vaxxer), this can lead to intense negative emotions that motivate strategies such

a Correction not integrated



b Selective retrieval

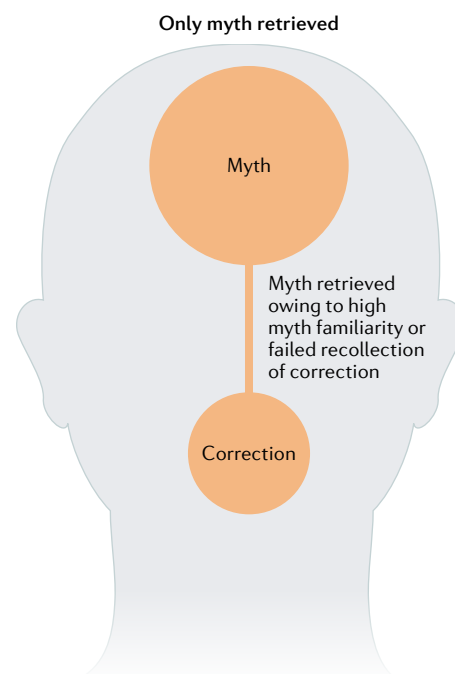


Fig. 2 | Integration and retrieval accounts of continued influence. a | Integration account of continued influence. The correction had the representational strength to compete with or even dominate the misinformation ('myth') but was not integrated into the relevant mental model. Depending on the available retrieval cues, this lack of integration can lead to unchecked misinformation retrieval and reliance. **b** | Retrieval account of continued influence. Integration has taken place but the myth is represented in memory more strongly, and thus dominates the corrective information in the competition for activation and retrieval. Note that the two situations are not mutually exclusive: avoiding continued influence might require both successful integration and retrieval of the corrective information.

as discrediting the source of the correction, ignoring the worldview-inconsistent evidence or selectively focusing on worldview-bolstering evidence^{24,126}. However, how a person's worldview influences misinformation corrections is still hotly debated (BOX 2), and there is a developing consensus that even worldview-inconsistent corrections typically have some beneficial impact^{91,127–131}.

The third socio-affective factor that influences the CIE is emotion. One study found that corrections can produce psychological discomfort that motivates a person to disregard the correction to reduce the feeling of discomfort¹³². Misinformation conveying negative emotions such as fear or anger might be particularly likely to evoke a CIE^{133,134}. This influence might be due to a general negativity bias^{11,135} or more specific emotional influences. For example, misinformation damaging the reputation of a political candidate might spark outrage or contempt, which might promote continued influence of this misinformation (in particular among non-supporters)¹³⁴. However, there seems to be little continued influence of negative misinformation on impression formation when the person subjected to the false allegation is not a disliked politician, perhaps because reliance on corrected misinformation might

be seen as biased or judgemental (that is, it might be frowned upon to judge another person even though allegations have been proven false)¹³⁶.

Other studies have compared emotive and non-emotive events — for example, a plane crash falsely assumed to have been caused by a terror attack, resulting in many fatalities, versus a technical fault, resulting in zero fatalities — and found no impact of misinformation emotiveness on the magnitude of the CIE¹³⁷. Moreover, just as a sad mood can protect against initial misinformation belief⁸⁰, it also seems to facilitate knowledge revision when a correction is encountered¹³⁸. People who exhibit both subclinical depression and rumination tendencies have even been shown to exhibit particularly efficient correction of negative misinformation relative to control individuals, presumably because the salience of negative misinformation to this group facilitates revision¹³⁹.

Finally, there is evidence that corrections can also benefit from emotional recalibration. For example, when misinformation downplays a risk or threat (for example, misinformation that a serious disease is relatively harmless), corrections that provide a more accurate risk evaluation operate partly through their impact on emotions such as hope, anger and fear. This emotional mechanism might help correction recipients realign their understanding of the situation with reality (for example, to realize they have underestimated the real threat)^{113,140}. Likewise, countering disinformation that seeks to fuel fear or anger can benefit from a downward adjustment of emotional arousal; for example, refutations of vaccine misinformation can reduce anti-vaccination attitudes by mitigating misinformation-induced anger¹⁴¹.

Box 2 | The elusive backfire effects

There have been concerns that corrective interventions might cause harm by inadvertently strengthening misconceptions and ironically enhancing reliance on the very misinformation that is being corrected. However, these concerns are largely overstated. Specifically, three types of ostensible 'backfire effects' have been discussed: the overkill backfire effect, the familiarity backfire effect, and the worldview backfire effect⁸⁹.

Only one study has investigated the potential overkill backfire effect, thought to result from a correction using too many counterarguments. This study found that corrections of dubious claims were more (rather than less) potent when more counterarguments were used, so long as those counterarguments were relevant²⁸⁶. Thus, the overkill backfire effect does not have empirical support.

The familiarity backfire effect is thought to result from a correction that unintentionally boosts the familiarity of the misinformation being corrected. This effect is characterized as an increase in misinformation belief following a correction, relative to a pre-correction baseline or no-exposure control condition. There are some findings that repeating corrections might lead to a tendency to recall false claims as true, especially after a 3-day delay or in older adults (age 70+ years)²⁸⁷. Likewise, it has been argued that presenting 'myths versus facts' flyers that repeat to-be-debunked misinformation when correcting it could lead to familiarity backfire effects after a mere 30 min²⁸⁸. However, these findings have not been replicated^{107,289} or remain unpublished. Other putative familiarity backfire effects did not compare the backfire condition with a proper baseline (for reviews see^{92,256}). Strong evidence against familiarity backfire comes from findings that explicit reminders of misinformation enhance the effect of corrections^{104,290}. Although some researchers have argued that familiarity backfire might occur when a correction spreads novel misinformation to new audiences¹⁸⁵, only one study has found support for this claim (and only in one of two experiments)²⁹¹, with other studies finding no evidence^{112,151,222}. Other demonstrations of familiarity backfire effects in the context of vaccine misinformation might be driven by worldview rather than familiarity²⁹². In sum, misinformation familiarity contributes to the CIE but does not typically produce backfire effects.

The backfire effect of greatest concern is arguably the worldview backfire effect, thought to arise when people dismiss and counterargue against corrections of false beliefs that are central to their identity^{126,293}. Early demonstrations of worldview backfire effects^{124,294,295} drew much attention from the academy and beyond, but have proven difficult to replicate^{81,128,130}, partially due to unreliable methods^{256,296}. Although findings of worldview backfire effects continue to be reported occasionally^{25,297}, overall the potential threat of worldview backfire effects seems limited and should not generally discourage debunking.

Interventions to combat misinformation

As discussed in the preceding section, interventions to combat misinformation must overcome various cognitive, social and affective barriers. The most common type of correction is a fact-based correction that directly addresses inaccuracies in the misinformation and provides accurate information^{90,102,112,142} (FIG. 3). A second approach is to address the logical fallacies common in some types of disinformation — for example, corrections that highlight inherently contradictory claims such as 'global temperature cannot be measured accurately' and 'temperature records show it has been cooling' (FIG. 4). Such logic-based corrections might offer broader protection against different types of misinformation that use the same fallacies and misleading tactics^{21,143}. A third approach is to undermine the plausibility of the misinformation or the credibility of its source¹⁴⁴. Multiple approaches can be combined into a single correction — for example, highlighting both the factual and logical inaccuracies in the misinformation or undermining source credibility and underscoring factual errors^{94,95,145}. However, most research to date has considered each approach separately and more research is required to test synergies between these strategies.

More generally, two strategies that can be distinguished are pre-emptive intervention (prebunking) and reactive intervention (debunking). Prebunking seeks to help people recognize and resist subsequently encountered misinformation, even if it is novel. Debunking

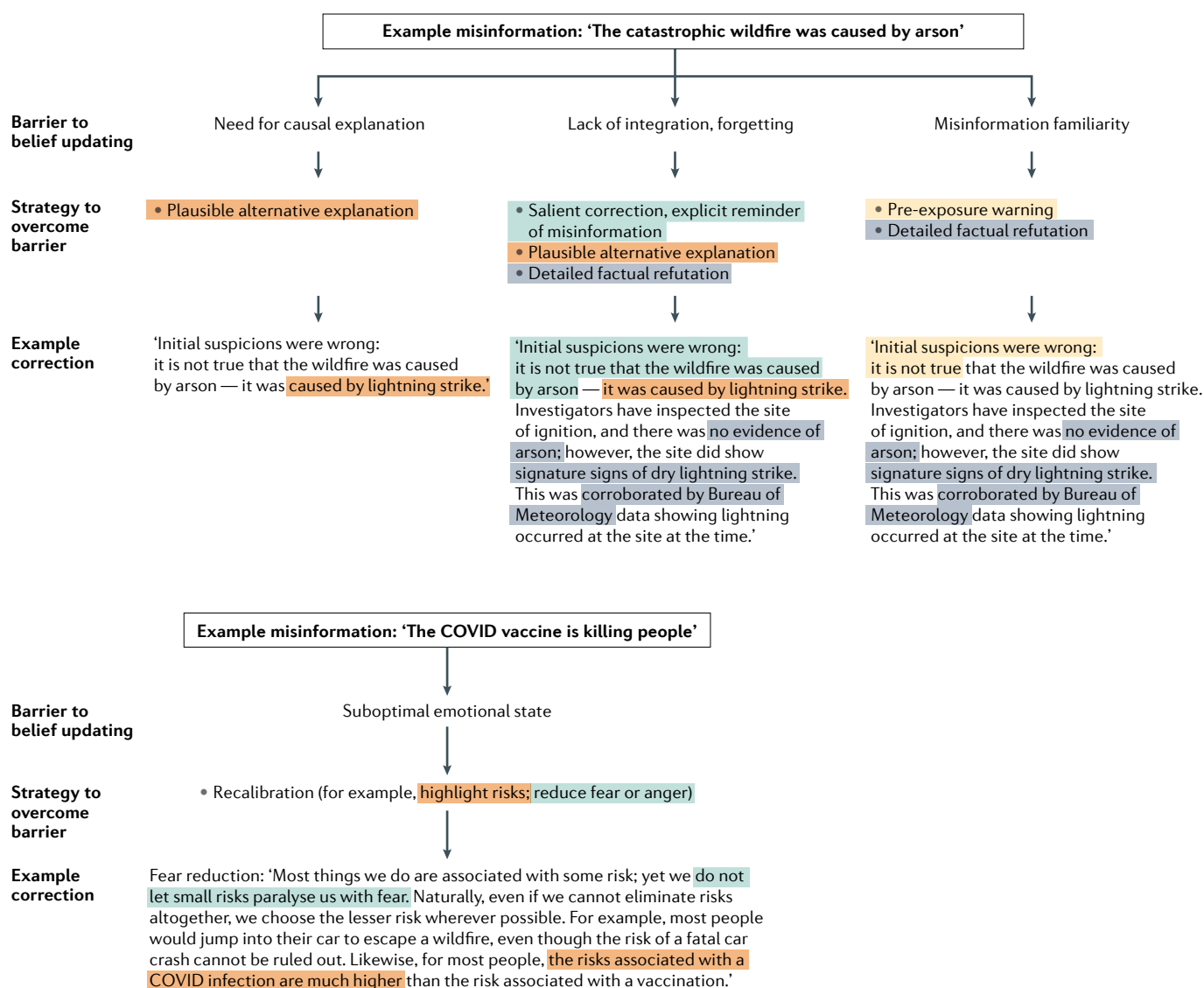


Fig. 3 | **Barriers to belief updating and strategies to overcome them (part 1).** How various barriers to belief updating can be overcome by specific communication strategies applied during correction, using event and health misinformation as examples. Colour shading is used to show how specific strategies are applied in the example corrections.

emphasizes responding to specific misinformation after exposure to demonstrate why it is false. The effectiveness of these corrections is influenced by a range of factors, and there are mixed results regarding their relative efficacy. For example, in the case of anti-vaccination conspiracy theories, prebunking has been found to be more effective than debunking¹⁴⁶. However, other studies have found debunking to outperform prebunking^{87,95,142}. Reconciling these findings might require considering both the specific type of correction and its placement in time. For example, when refuting climate misinformation, one study found that fact-based debunking outperformed fact-based prebunking, whereas logic-based prebunking and debunking were equally effective¹⁴⁷.

Some interventions, particularly those in online contexts, are hybrid or borderline cases. For example, if a misleading social media post is tagged with 'false'¹⁴⁸ and appears alongside a comment with a corrective explanation, this might count as both prebunking (owing to

the tag, which is likely to have been processed before the post) and debunking (owing to the comment, which is likely to have been processed after the post).

Prebunking interventions. The simplest prebunking interventions involve presenting factually correct information^{149,150}, a pre-emptive correction^{142,151} or a generic misinformation warning^{99,148,152,153} before the misinformation. More sophisticated interventions draw on inoculation theory, a framework for pre-emptive interventions^{154–156}. This theory applies the principle of vaccination to knowledge, positing that 'inoculating' people with a weakened form of persuasion can build immunity against subsequent persuasive arguments by engaging people's critical-thinking skills (FIG. 5).

An inoculation intervention combines two elements. The first element is warning recipients of the threat of misleading persuasion. For example, a person could be warned that many claims about climate change are false

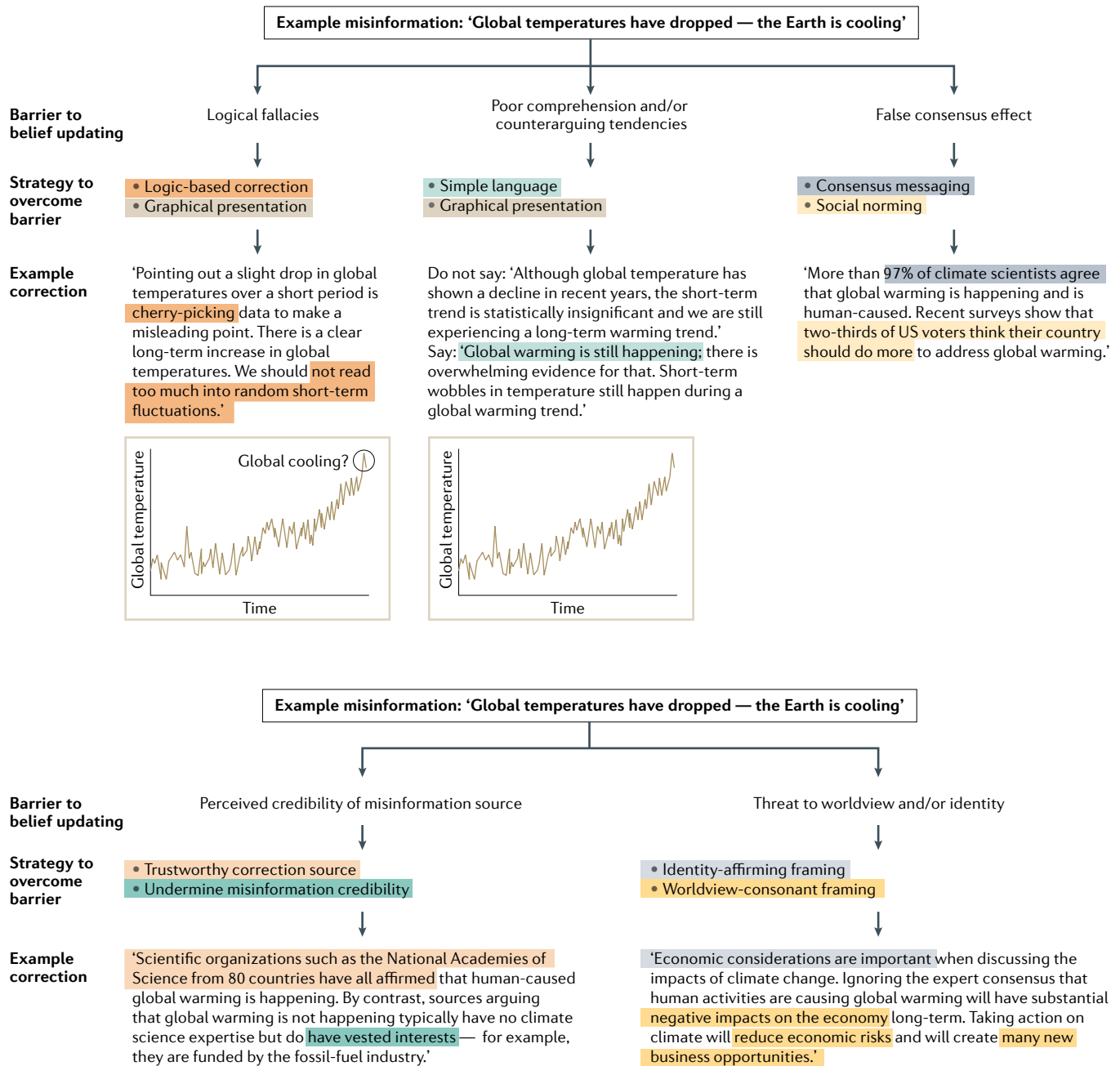


Fig. 4 | **Barriers to belief updating and strategies to overcome them (part 2).** How various barriers to belief updating can be overcome by specific communication strategies applied during correction, using climate change misinformation as an example. Colour shading is used to show how specific strategies are applied in the example corrections.

and intentionally misleading. The second element is identifying the techniques used to mislead or the fallacies that underlie the false arguments to refute forthcoming misinformation^{157,158}. For example, a person might be taught that techniques used to mislead include selective use ('cherry-picking') of data (for example, only showing temperatures from outlier years to create the illusion that global temperatures have dropped) or the use of fake experts (for example, scientists with no expertise in climate science). Understanding how those misleading persuasive techniques are applied equips a person with the cognitive tools to ward off analogous persuasion attempts in the future.

Because one element of inoculation is highlighting misleading argumentation techniques, its effects can generalize across topics, providing an 'umbrella' of protection^{159,160}. For example, an inoculation against a misleading persuasive technique used to cast doubt on science demonstrating harm from tobacco was found to convey resistance against the same technique when used to cast doubt on climate science¹⁴³. Moreover, inoculated people are more likely to talk about the target issue than non-inoculated people, an outcome referred to as post-inoculation talk¹⁶¹. Post-inoculation talk is more likely to be negative than talk among non-inoculated people, which promotes misinformation resistance both

within and between individuals because people's evaluations tend to weight negative information more strongly than positive information¹⁶².

Inoculation theory has also been used to explain how strategies designed to increase information literacy and media literacy could reduce the effects of misinformation. Information literacy — the ability to effectively find, understand, evaluate and use information — has been linked to the ability to detect misleading news¹⁶³ and reduced sharing of misinformation¹⁶⁴. Generally, information literacy and media literacy (which focuses on knowledge and skills for the reception and dissemination of information through the media) interventions are designed to improve critical thinking¹⁶⁵ and the application of such interventions to spaces containing many different types of information might help people identify misinformation¹⁶⁶.

One successful intervention focused on lateral reading — consulting external sources to examine the origins and plausibility of a piece of information, or the credibility of an information source^{115,167,168}. A separate non-peer-reviewed preprint suggests that focusing on telltale signs of online misinformation (including lexical cues, message simplicity and blatant use of emotion) can help people identify fake news¹⁶⁹. However, research to date suggests that literacy interventions do not always mitigate the effects of misinformation^{170–173}. Whereas most work has used relatively passive inoculation and literacy interventions, applications that engage people more actively have shown promise — specifically, app-based or web-based games^{174–177}. More work is needed to consider what types of literacy interventions are most effective for conferring resistance to different types of misinformation in the contemporary media and information landscape¹⁷⁸.

In sum, the prebunking approach provides a great tool to act pre-emptively and help people build resistance to misinformation in a relatively general manner. However, the advantage of generalizability can also be a weakness, because it is often specific pieces of

misinformation that cause concern, which call for more specific responses.

Debunking interventions. Whereas pre-emptive interventions can equip people to recognize and resist misinformation, reactive interventions retrospectively target concrete instances of misinformation. For example, if a novel falsehood that a vaccine can lead to life-threatening side effects in pregnant women begins to spread, then this misinformation must be addressed using specific counter-evidence. Research broadly finds that direct corrections are effective in reducing — although frequently not eliminating — reliance on the misinformation in a person's reasoning^{86,87}. The beneficial effects of debunking can last several weeks^{92,100,179}, although the effects can wear off quicker¹⁴⁵. There is also evidence that corrections that reduce misinformation belief can have downstream effects on behaviours or intentions^{94,95,180,181} — such as a person's inclination to share a social media post or their voting intentions — but not always^{91,96,182}.

Numerous best practices for debunking have emerged^{90,145,183}. First, the most important element of a debunking correction is to provide a factual account that ideally includes an alternative explanation for why something happened^{85,86,99,102,184}. For example, if a fire was thought to have been caused by negligence, then providing a causal alternative ('there is evidence for arson') is more effective than a retraction ('there was no negligence'). In general, more detailed refutations work better than plain retractions that do not provide any detail on why the misinformation is incorrect^{92,100,112,113}. It can be beneficial to lead with the correction rather than repeat the misinformation to prioritize the correct information and set a factual frame for the issue. However, a preprint that has not been peer-reviewed suggests that leading with the misinformation can be just as, or even more, effective if no pithy fact is available¹⁵⁰.

Second, the misinformation should be repeated to demonstrate how it is incorrect and to make the correction salient. However, the misinformation should be prefaced with a warning^{99,148} and repeated only once in order not to boost its familiarity unnecessarily¹⁰⁴. It is also good to conclude by repeating and emphasizing the accurate information to reinforce the correction¹⁸⁵.

Third, even though credibility matters less for correction sources compared with misinformation sources⁸⁸, corrections are ideally delivered by or associated with high-credibility sources^{116–120,186}. There is also emerging evidence that corrections are more impactful when they come from a socially connected source (for example, a connection on social media) rather than a stranger¹⁸⁷.

Fourth, corrections should be paired with relevant social norms, including injunctive norms ('protecting the vulnerable by getting vaccinated is the right thing to do') and descriptive norms ('over 90% of parents are vaccinating their children')¹⁸⁸, as well as expert consensus ('doctors and medical societies around the world agree that vaccinations are important and safe')^{189–192}. One study found a benefit to knowledge revision if corrective evidence was endorsed by many others on social media, thus giving the impression of normative backing¹⁹³.

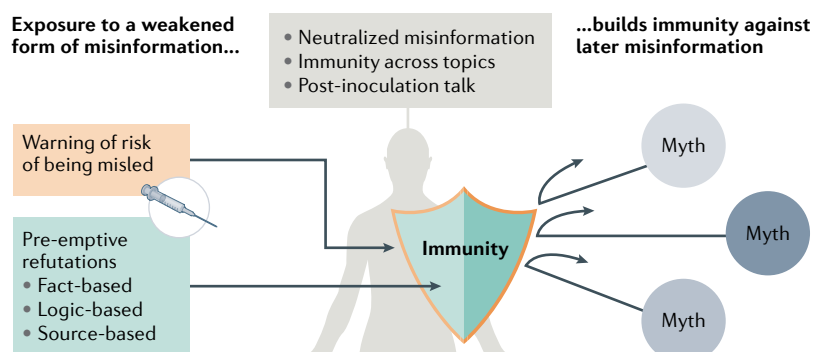


Fig. 5 | Inoculation theory applied to misinformation. 'Inoculation' treatment can help people prepare for subsequent misinformation exposure. Treatment typically highlights the risks of being misled, alongside a pre-emptive refutation. The refutation can be fact-based, logic-based or source-based. Inoculation has been shown to increase misinformation detection and facilitate counterarguing and dismissal of false claims, effectively neutralizing misinformation. Additionally, inoculation can build immunity across topics and increase the likelihood of people talking about the issue targeted by the refutation (post-inoculation talk).

Fifth, the language used in a correction is important. Simple language and informative graphics can facilitate knowledge revision, especially if fact comprehension might be otherwise difficult or if the person receiving the correction has a strong tendency to counterargue^{194–197}. When speaking directly to misinformed individuals, empathic communication should be used rather than wielding expertise to argue directives^{198,199}.

Finally, it has been suggested that worldview-threatening corrections can be made more palatable by concurrently providing an identity affirmation^{145,200,201}. Identity affirmations involve a message or task (for example, writing a brief essay about one's strengths and values) that highlights important sources of self-worth. These exercises are assumed to protect and strengthen the correction recipient's self-esteem and the value of their identity, thereby reducing the threat associated with the correction and associated processing biases. However, evidence for the utility of identity affirmations in the context of misinformation corrections is mixed¹⁹⁴, so firm recommendations cannot yet be made.

In sum, debunking is a valuable tool to address specific pieces of misinformation and largely reduces misinformation belief. However, debunking will not eliminate the influence of misinformation on people's reasoning at a group level. Furthermore, even well-designed debunking interventions might not have long-lasting effects, thus requiring repeated intervention.

Corrections on social media. Misinformation corrections might be especially important in social media contexts because they can reduce false beliefs not just in the target of the correction but among everyone that sees the correction — a process termed observational correction¹¹⁹. Best practices for corrections on social media echo many best practices offline¹¹², but also include linking to expert sources and correcting quickly and early²⁰². There is emerging evidence that online corrections can work both pre-emptively and reactively, although this might depend on the type of correction¹⁴⁷.

Notably, social media corrections are more effective when they are specific to an individual piece of content rather than a generalized warning¹⁴⁸. Social media corrections are effective when they come from algorithmic sources²⁰³, from expert organizations such as a government health agency^{119,204,205} or from multiple other users on social media²⁰⁶. However, particular care must be taken to avoid ostracizing people when correcting them online. To prevent potential adverse effects on people's online behaviour, such as sharing of misleading content, gentle accuracy nudges that prompt people to consider the accuracy of the information they encounter or highlight the importance of sharing only true information might be preferable to public corrections that might be experienced as embarrassing or confrontational^{181,207}.

In sum, social media users should be aware that corrections can be effective in this arena and have the potential to reduce false beliefs in people they are connected with as well as bystanders. By contrast, confronting strangers is less likely to be effective. Given the effectiveness of algorithmic corrections, social media companies and regulators should promote implementation

and evaluation of technical solutions to misinformation on social media.

Practical implications

Even if optimal prebunking or debunking interventions are deployed, no intervention can be fully effective or reach everyone with the false belief. The contemporary information landscape brings particular challenges: the internet and social media have enabled an exponential increase in misinformation spread and targeting to precise audiences^{14,16,208,209}. Against this backdrop, the psychological factors discussed in this Review have implications for practitioners in various fields — journalists, legislators, public health officials and healthcare workers — as well as information consumers.

Implications for practitioners. Combatting misinformation involves a range of decisions regarding the optimal approach (FIG. 6). When preparing to counter misinformation, it is important to identify likely sources. Although social media is an important misinformation vector²¹⁰, traditional news organizations can promote misinformation via opinion pieces²¹¹, sponsored content²¹² or uncritical repetition of politician statements²¹³. Practitioners must anticipate the misinformation themes and ensure suitable fact-based alternative accounts are available for either prebunking or a quick debunking response. Organizations such as the [International Fact-Checking Network](#) or the [World Health Organization](#) often form coalitions in the pursuit of this endeavour²¹⁴.

Practitioners must be aware that simple retractions will be insufficient to mitigate the impact of misinformation, and that the effects of interventions tend to wear off over time^{92,145,152}. If possible, practitioners must therefore be prepared to act repeatedly¹⁷⁹. Creating engaging, fact-based narratives can provide a foundation for effective correction^{215,216}. However, a narrative format is not a necessary ingredient^{140,217}, and anecdotes and stories can also be misleading²¹⁸.

Practitioners can also help audiences discriminate between facts and opinion, which is a teachable skill^{170,219}. Whereas most news consumers do not notice or understand content labels forewarning that an article is news, opinion or advertising^{220,221}, more prominent labelling can nudge readers to adjust their comprehension and interpretation accordingly. For example, labelling can lead readers to be more sceptical of promoted content²²⁰. However, even when forewarnings are understood, they do not reliably eliminate the content's influence^{99,153}.

If pre-emptive correction is not possible or ineffective, practitioners should take a reactive approach. However, not every piece of misinformation needs to be a target for correction. Due to resource limitations and opportunity costs, corrections should focus on misinformation that circulates among a substantive portion of the population and carries potential for harm¹⁸³. Corrections do not generally increase false beliefs among individuals who were previously unfamiliar with the misinformation²²². However, if the risk of harm is minimal, there is no need to debunk misinformation that few people are aware of, which could potentially raise the profile of its source.

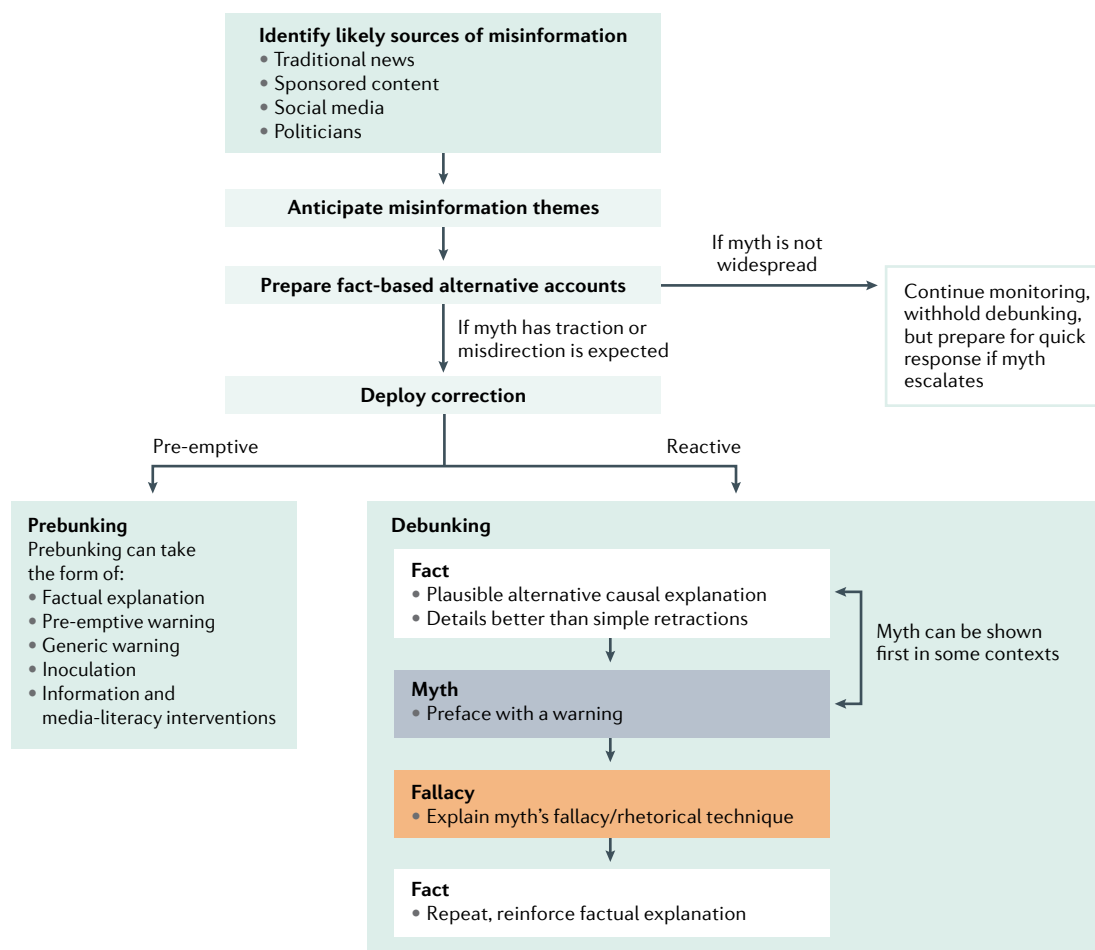


Fig. 6 | Strategies to counter misinformation. Different strategies for countering misinformation are available to practitioners at different time points. If no misinformation is circulating but there is potential for it to emerge in the future, practitioners can consider possible misinformation sources and anticipate misinformation themes. Based on this assessment, practitioners can prepare fact-based alternative accounts, and either continue monitoring the situation while preparing for a quick response, or deploy pre-emptive (prebunking) or reactive (debunking) interventions, depending on the traction of the misinformation. Prebunking can take various forms, from simple warnings to more involved literacy interventions. Debunking can start either with a pithy counterfact that recipients ought to remember or with dismissal of the core ‘myth’. Debunking should provide a plausible alternative cause for an event or factual details, preface the misinformation with a warning and explain any logical fallacies or persuasive techniques used to promote the misinformation. Debunking should end with a factual statement.

Implications for information consumers. Information consumers also have a role to play in combatting misinformation by avoiding contributing to its spread. For instance, people must be aware that they might encounter not only relatively harmless misinformation, such as reporting errors, outdated information and satire, but also disinformation campaigns designed to instil fear or doubt, discredit individuals, and sow division^{2,26,223,224}. People must also recognize that disinformation can be psychologically targeted through profit-driven exploitation of personal data and social media algorithms¹². Thoughtless sharing can amplify misinformation that might confuse and deceive others. Sharing misinformation can also contribute to the financial rewards sought by misinformation producers, and deepen ideological divides that disenfranchise voters, encourage violence and, ultimately, harm democratic processes^{2,170,223,225,226}.

Thus, while engaged with content, individuals should slow down, think about why they are engaging

and interrogate their visceral response. People who thoughtfully seek accurate information are more likely to successfully avoid misinformation compared with people who are motivated to find evidence to confirm their pre-existing beliefs^{50,227,228}. Attending to the source and considering its credibility and motivation, along with lateral reading strategies, also increase the likelihood of identifying misinformation^{115,167,171}. Given the benefits of persuading onlookers through observational correction, everyone should be encouraged to civilly, carefully and thoughtfully correct online misinformation where they encounter it (unless they deem it a harmless fringe view)^{119,206}. All of these recommendations are also fundamental principles of media literacy¹⁶⁶. Indeed, a theoretical underpinning of media literacy is that understanding the aims of media protects individuals from some adverse effects of being exposed to information through the media, including the pressure to adopt particular beliefs or behaviours¹⁷⁰.

Implications for policymakers. Ultimately, even if practitioners and information consumers apply all of these strategies to reduce the impact of misinformation, their efforts will be stymied if media platforms continue to amplify misinformation^{14,16,208–213}. These platforms include social media platforms such as YouTube, which are geared towards maximizing engagement even if this means promoting misinformation²²⁹, and traditional media outlets such as television news channels, where misinformation can negatively impact audiences. For example, two non-peer-reviewed preprints have found that COVID-19 misinformation on Fox News was causally associated with reduced adherence to public health measures and a larger number of COVID-19 cases and deaths^{230,231}. It is, therefore, important to scrutinize whether the practices and algorithms of media platforms are optimized to promote misinformation or truth.

In this space, policymakers should consider enhanced regulation. These regulations might include penalties for creating and disseminating disinformation where intentionality and harm can be established, and mandating platforms to be more proactive, transparent and effective in their dealings with misinformation. With regards to social media specifically, companies should be encouraged to ban repeat offenders from their platforms, and to generally make engagement with and sharing of low-quality content more difficult^{12,232–235}. Regulation must not result in censorship, and proponents of freedom of speech might disagree with attempts to regulate content. However, freedom of speech does not include the right to amplification of that speech. Furthermore, being unknowingly subjected to disinformation can be seen as a manipulative attack on freedom of choice and the right to be well informed²³⁶. These concerns must be balanced. A detailed summary of potential regulatory interventions can be found elsewhere^{237,238}.

Other strategies have the potential to reduce the impact of misinformation without regulation of media content. Undue concentration of ownership and control of both social and traditional media facilitate the dissemination of misinformation²³⁹. Thus, policymakers are advised to support a diverse media landscape and adequately fund independent public broadcasters. Perhaps the most important approach to slowing the spread of misinformation is substantial investment in education, particularly to build information literacy skills in schools and beyond^{240–243}. Another tool in the policymaker's arsenal is interventions targeted more directly at behaviour, such as nudging policies and public pledges to honour the truth (also known as self-nudging) for policymakers and consumers alike^{12,244,245}.

Overall, solutions to misinformation spread must be multipronged and target both the supply (for example, more efficient fact-checking and changes to platform algorithms and policies) and the consumption (for example, accuracy nudges and enhanced media literacy) of misinformation. Individually, each intervention might only incrementally reduce the spread of misinformation, but one preprint that has not been peer-reviewed suggests that combinations of interventions can have a substantial impact²⁴⁶.

More broadly speaking, any intervention to strengthen public trust in science, journalism, and democratic

institutions is an intervention against the impacts of misinformation^{247,248}. Such interventions might include enhancing transparency in science^{249,250} and journalism²⁵¹, more rigorous fact-checking of political advertisements²⁵², and reducing the social inequality that breeds distrust in experts and contributes to vulnerability to misinformation^{253,254}.

Summary and future directions

Psychological research has built solid foundational knowledge of how people decide what is true and false, form beliefs, process corrections, and might continue to be influenced by misinformation even after it has been corrected. However, much work remains to fully understand the psychology of misinformation.

First, in line with general trends in psychology and elsewhere, research methods in the field of misinformation should be improved. Researchers should rely less on small-scale studies conducted in the laboratory or a small number of online platforms, often on non-representative (and primarily US-based) participants²⁵⁵. Researchers should also avoid relying on one-item questions with relatively low reliability²⁵⁶. Given the well-known attitude-behaviour gap — that attitude change does not readily translate into behavioural effects — researchers should also attempt to use more behavioural measures, such as information-sharing measures, rather than relying exclusively on self-report questionnaires^{93–95}. Although existing research has yielded valuable insights into how people generally process misinformation (many of which will translate across different contexts and cultures), an increased focus on diversification of samples and more robust methods is likely to provide a better appreciation of important contextual factors and nuanced cultural differences^{7,82,205,257–263}.

Second, most existing work has focused on explicit misinformation and text-based materials. Thus, the cognitive impacts of other types of misinformation, including subtler types of misdirection such as paltering (misleading while technically saying the truth)^{95,264–266}, doctored images²⁶⁷, deepfake videos²⁶⁸ and extreme patterns of misinformation bombardment²²³, are currently not well understood. Non-text-based corrections, such as videos or cartoons, also deserve more exploration^{269,270}.

Third, additional translational research is needed to explore questions about causality, including the causal impacts of misinformation and corrections on beliefs and behaviours. This research should also employ non-experimental methods^{230,231,271}, such as observational causal inference (research aiming to establish causality in observed real-world data)²⁷², and test the impact of interventions in the real world^{145,174,181,207}. These studies are especially needed over the long term — weeks to months, or even years — and should test a range of outcome measures, for example those that relate to health and political behaviours, in a range of contexts. Ultimately, the success of psychological research into misinformation should be linked not only to theoretical progress but also to societal impact²⁷³.

Finally, even though the field has a reasonable understanding of the cognitive mechanisms and social

determinants of misinformation processing, knowledge of the complex interplay between cognitive and social dynamics is still limited, as is insight into the role of emotion. Future empirical and theoretical work would benefit from development of an overarching theoretical model that aims to integrate cognitive, social and affective factors, for example by utilizing agent-based

modelling approaches. This approach might also offer opportunities for more interdisciplinary work²⁵⁷ at the intersection of psychology, political science²⁷⁴ and social network analysis²⁷⁵, and the development of a more sophisticated psychology of misinformation.

Published online 12 January 2022

1. DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M. & Epstein, J. A. Lying in everyday life. *J. Personal. Soc. Psychol.* **70**, 979–995 (1996).
2. Lewandowsky, S., Ecker, U. K. H. & Cook, J. Beyond misinformation: understanding and coping with the post-truth era. *J. Appl. Res. Mem. Cogn.* **6**, 353–369 (2017).
3. Zarocostas, J. How to fight an infodemic. *Lancet* **395**, 676 (2020).
4. Lazer, D. M. J. et al. The science of fake news. *Science* **359**, 1094–1096 (2018).
5. Bennett, W. L. & Livingston, S. The disinformation order: disruptive communication and the decline of democratic institutions. *Eur. J. Commun.* **33**, 122–139 (2018).
6. Whitten-Woodring, J., Kleinberg, M. S., Thawngmum, A. & Thitsar, M. T. Poison if you don't know how to use it: Facebook, democracy, and human rights in Myanmar. *Int. J. Press Politics* **25**, 407–425 (2020).
7. Roozenbeek, J. et al. Susceptibility to misinformation about COVID-19 around the world. *R. Soc. Open Sci.* **7**, 201199 (2020).
8. Rich, J. in *Private and Public Lies. The Discourse of Despotism and Deceit in the Graeco-Roman World (Impact of Empire 11)* (eds Turner, A. J., Kim On Chong-Cossard, J. H. & Vervaeke, F. J.) Vol. 11 167–191 (Brill Academic, 2010).
9. Hekster, O. in *The Representation and Perception of Roman Imperial Power* (eds de Blois, L., Erdkamp, P., Hekster, O., de Kleijn, G. & Mols, S.) 20–35 (J. C. Gieben, 2013).
10. Herf, J. The Jewish War: Goebbels and the antisemitic campaigns of the Nazi propaganda ministry. *Holocaust Genocide Stud.* **19**, 51–80 (2005).
11. Acerbi, A. Cognitive attraction and online misinformation. *Palgrave Commun.* **5**, 15 (2019).
12. Kozyreva, A., Lewandowsky, S. & Hertwig, R. Citizens versus the internet: confronting digital challenges with cognitive tools. *Psychol. Sci. Public Interest* **21**, 103–156 (2020).
13. Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A. & Bonneau, R. Tweeting from left to right: is online political communication more than an echo chamber? *Psychol. Sci.* **26**, 1531–1542 (2015).
14. Del Vicario, M. et al. The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559 (2016).
15. Garrett, R. K. The echo chamber distraction: disinformation campaigns are the problem not audience fragmentation. *J. Appl. Res. Mem. Cogn.* **6**, 370–376 (2017).
16. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
17. Simis, M. J., Madden, H., Cacciatore, M. A. & Yeo, S. K. The lure of rationality: why does the deficit model persist in science communication? *Public Underst. Sci.* **25**, 400–414 (2016).
18. Fazio, L. K., Brashier, N. M., Payne, B. K. & Marsh, E. J. Knowledge does not protect against illusory truth. *J. Exp. Psychol.* **144**, 993–1002 (2015).
19. Hornsey, M. J. & Fielding, K. S. Attitude roots and jiu jitsu persuasion: understanding and overcoming the motivated rejection of science. *Am. Psychol.* **72**, 459 (2017).
20. Nisbet, E. C., Cooper, K. E. & Garrett, R. K. The partisan brain: how dissonant science messages lead conservatives and liberals to (dis)trust science. *Ann. Am. Acad. Political Soc. Sci.* **658**, 36–66 (2015).
21. Schmid, P. & Betsch, C. Effective strategies for rebutting science denialism in public discussions. *Nat. Hum. Behav.* **3**, 931–939 (2019).
22. Hansson, S. O. Science denial as a form of pseudoscience. *Stud. History Philos. Sci. A* **63**, 39–47 (2017).
23. Amin, A. B. et al. Association of moral values with vaccine hesitancy. *Nat. Hum. Behav.* **1**, 873–880 (2017).
24. Lewandowsky, S. & Oberauer, K. Motivated rejection of science. *Curr. Dir. Psychol. Sci.* **25**, 217–222 (2016).
25. Trevors, G. & Duffy, M. C. Correcting COVID-19 misconceptions requires caution. *Educ. Res.* **49**, 538–542 (2020).
26. Lewandowsky, S. Conspiracist cognition: chaos convenience, and cause for concern. *J. Cult. Res.* **25**, 12–35 (2021).
27. Lewandowsky, S., Stritzke, W. G. K., Freund, A. M., Oberauer, K. & Krueger, J. I. Misinformation, disinformation, and violent conflict: from Iraq and the war on terror to future threats to peace. *Am. Psychol.* **68**, 487–501 (2013).
28. Marsh, E. J., Cantor, A. D. & Brashier, N. M. Believing that humans swallow spiders in their sleep. *Psychol. Learn. Motiv.* **64**, 93–132 (2016).
29. Rapp, D. N. The consequences of reading inaccurate information. *Curr. Dir. Psychol. Sci.* **25**, 281–285 (2016).
30. Pantazi, M., Kissine, M. & Klein, O. The power of the truth bias: false information affects memory and judgment even in the absence of distraction. *Soc. Cogn.* **36**, 167–198 (2018).
31. Brashier, N. M. & Marsh, E. J. Judging truth. *Annu. Rev. Psychol.* **71**, 499–515 (2020).
32. Prike, T., Arnold, M. M. & Williamson, P. The relationship between anomalous belief misperception of chance and the base rate fallacy. *Think. Reason.* **26**, 447–477 (2020).
33. Uscinski, J. E. et al. Why do people believe COVID-19 conspiracy theories? *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-015> (2020).
34. Dechêne, A., Stahl, C., Hansen, J. & Wänke, M. The truth about the truth: a meta-analytic review of the truth effect. *Personal. Soc. Psychol. Rev.* **14**, 238–257 (2010).
35. Unkelbach, C., Koch, A., Silva, R. R. & Garcia-Marques, T. Truth by repetition: explanations and implications. *Curr. Dir. Psychol. Sci.* **28**, 247–253 (2019).
36. Begg, I. M., Anas, A. & Farinacci, S. Dissociation of processes in belief: source recollection, statement familiarity, and the illusion of truth. *J. Exp. Psychol. Gen.* **121**, 446–458 (1992).
37. Unkelbach, C. Reversing the truth effect: learning the interpretation of processing fluency in judgments of truth. *J. Exp. Psychol. Learn. Memory Cogn.* **33**, 219–230 (2007).
38. Wang, W. C., Brashier, N. M., Wing, E. A., Marsh, E. J. & Cabeza, R. On known unknowns: fluency and the neural mechanisms of illusory truth. *J. Cognit. Neurosci.* **28**, 739–746 (2016).
39. Unkelbach, C. & Rom, S. C. A referential theory of the repetition-induced truth effect. *Cognition* **160**, 110–126 (2017).
40. Pennycook, G., Cannon, T. D. & Rand, D. G. Prior exposure increases perceived accuracy of fake news. *J. Exp. Psychol. Gen.* **147**, 1865–1880 (2018).
41. Unkelbach, C. & Speckmann, F. Mere repetition increases belief in factually true COVID-19-related information. *J. Appl. Res. Mem. Cogn.* **10**, 241–247 (2021).
42. Nadarevic, L., Reber, R., Helmecke, A. J. & Köse, D. Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognit. Res. Princ. Implic.* **5**, 56 (2020).
43. Fazio, L. K., Rand, D. G. & Pennycook, G. Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bull. Rev.* **26**, 1705–1710 (2019).
44. Brown, A. S. & Nix, L. A. Turning lies into truths: referential validation of falsehoods. *J. Exp. Psychol. Learn. Memory Cogn.* **22**, 1088–1100 (1996).
45. 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).
46. Unkelbach, C. & Greifeneder, R. Experiential fluency and declarative advice jointly inform judgments of truth. *J. Exp. Soc. Psychol.* **79**, 78–86 (2018).
47. Fazio, L. K. Repetition increases perceived truth even for known falsehoods. *Collabra Psychol.* **6**, 38 (2020).
48. Pennycook, G. & Rand, D. G. The psychology of fake news. *Trends Cognit. Sci.* **25**, 388–402 (2021).
49. Murphy, G., Loftus, E. F., Grady, R. H., Levine, L. J. & Greene, C. M. False memories for fake news during Ireland's abortion referendum. *Psychol. Sci.* **30**, 1449–1459 (2019).
50. 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).
51. Stanley, M. L., Barr, N., Peters, K. & Seli, P. Analytic-thinking predicts hoax beliefs and helping behaviors in response to the COVID-19 pandemic. *Think. Reas.* **27**, 464–477 (2020).
52. Bago, B., Rand, D. G. & Pennycook, G. Fake news, fast and slow: deliberation reduces belief in false (but not true) news headlines. *J. Exp. Psychol. Gen.* **149**, 1608–1613 (2020).
53. Brashier, N. M., Eliseev, E. D. & Marsh, E. J. An initial accuracy focus prevents illusory truth. *Cognition* **194**, 104054 (2020).
54. Briñol, P. & Petty, R. E. Source factors in persuasion: a self-validation approach. *Eur. Rev. Soc. Psychol.* **20**, 49–96 (2009).
55. Mackie, D. M., Worth, L. T. & Asuncion, A. G. Processing of persuasive in-group messages. *J. Pers. Soc. Psychol.* **58**, 812–822 (1990).
56. Mahmoodi, A. et al. Equality bias impairs collective decision-making across cultures. *Proc. Natl Acad. Sci. USA* **112**, 3835–3840 (2015).
57. Marks, G. & Miller, N. Ten years of research on the false-consensus effect: an empirical and theoretical review. *Psychol. Bull.* **102**, 72–90 (1987).
58. Brulle, R. J., Carmichael, J. & Jenkins, J. C. Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the U.S. 2002–2010. *Clim. Change* **114**, 169–188 (2012).
59. Lachapelle, E., Montpetit, É. & Gauvin, J.-P. Public perceptions of expert credibility on policy issues: the role of expert framing and political worldviews. *Policy Stud. J.* **42**, 674–697 (2014).
60. Dada, S., Ashworth, H. C., Bewa, M. J. & Dhatt, R. Words matter: political and gender analysis of speeches made by heads of government during the COVID-19 pandemic. *BMJ Glob. Health* **6**, e003910 (2021).
61. Chung, M. & Jones-Jang, S. M. Red media, blue media, Trump briefings, and COVID-19: examining how information sources predict risk preventive behaviors via threat and efficacy. *Health Commun.* <https://doi.org/10.1080/10410236.2021.1914386> (2021).
62. Mitchell, K. J. & Johnson, M. K. Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory? *Psychol. Bull.* **135**, 638–677 (2009).
63. Dias, N., Pennycook, G. & Rand, D. G. Emphasizing publishers does not effectively reduce susceptibility to misinformation on social media. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-001> (2020).
64. Pennycook, G. & Rand, D. G. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc. Natl Acad. Sci. USA* **116**, 2521–2526 (2019).
65. Altay, S., Hacquin, A.-S. & Mercier, H. Why do so few people share fake news? It hurts their reputation. *N. Media Soc.* <https://doi.org/10.1177/1461444820969893> (2020).

66. Rahhal, T. A., May, C. P. & Hasher, L. Truth and character: sources that older adults can remember. *Psychol. Sci.* **13**, 101–105 (2002).
67. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. & Lazer, D. Fake news on Twitter during the 2016 U.S. presidential election. *Science* **363**, 374–378 (2019).
68. Stanford University Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory. The long fuse: misinformation and the 2020 election. *Stanford Digital Repository* <https://purl.stanford.edu/tr171zs0069> (2021).
69. Jones, M. O. Disinformation superspreaders: the weaponisation of COVID-19 fake news in the Persian Gulf and beyond. *Glob. Discourse* **10**, 431–437 (2020).
70. Tannenbaum, M. B. et al. Appealing to fear: a meta-analysis of fear appeal effectiveness and theories. *Psychol. Bull.* **141**, 1178–1204 (2015).
71. Altay, S. & Mercier, H. Happy thoughts: the role of communion in accepting and sharing epistemically suspect beliefs. *psyarxiv* <https://psyarxiv.com/3s4nr/> (2020).
72. Rocklage, M. D., Rucker, D. D. & Nordgren, L. F. Persuasion, emotion, and language: the intent to persuade transforms language via emotionality. *Psychol. Sci.* **29**, 749–760 (2018).
73. Chou, W.-Y. S. & Budenz, A. Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence. *Health Commun.* **35**, 1718–1722 (2020).
74. Baum, J. & Abdel, R. R. Emotional news affects social judgments independent of perceived media credibility. *Soc. Cognit. Affect. Neurosci.* **16**, 280–291 (2021).
75. Kim, H., Park, K. & Schwarz, N. Will this trip really be exciting? The role of incidental emotions in product evaluation. *J. Consum. Res.* **36**, 983–991 (2010).
76. Forgas, J. P. Happy believers and sad skeptics? Affective influences on gullibility. *Curr. Dir. Psychol. Sci.* **28**, 306–313 (2019).
77. Martel, C., Pennycook, G. & Rand, D. G. Reliance on emotion promotes belief in fake news. *Cognit. Res. Princ. Implic.* **5**, 47 (2020).
78. Forgas, J. P. & East, R. On being happy and gullible: mood effects on skepticism and the detection of deception. *J. Exp. Soc. Psychol.* **44**, 1362–1367 (2008).
79. Koch, A. S. & Forgas, J. P. Feeling good and feeling truth: the interactive effects of mood and processing fluency on truth judgments. *J. Exp. Soc. Psychol.* **48**, 481–485 (2012).
80. Forgas, J. P. Don't worry be sad! On the cognitive, motivational, and interpersonal benefits of negative mood. *Curr. Dir. Psychol. Sci.* **22**, 225–232 (2013).
81. Weeks, B. E. Emotions, partisanship, and misperceptions: how anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *J. Commun.* **65**, 699–719 (2015).
82. Han, J., Cha, M. & Lee, W. Anger contributes to the spread of COVID-19 misinformation. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-39> (2020).
83. Graepner, D. & Coman, A. The dark side of meaning-making: how social exclusion leads to superstitious thinking. *J. Exp. Soc. Psychol.* **69**, 218–222 (2017).
84. Poon, K.-T., Chen, Z. & Wong, W.-Y. Beliefs in conspiracy theories following ostracism. *Pers. Soc. Psychol. Bull.* **46**, 1234–1246 (2020).
85. Johnson, H. M. & Seifert, C. M. Sources of the continued influence effect: when misinformation in memory affects later inferences. *J. Exp. Psychol. Learn. Memory Cogn.* **20**, 1420–1436 (1994).
86. Chan, M.-P. S., Jones, C. R., Jamieson, K. H. & Albarracín, D. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. *Psychol. Sci.* **28**, 1531–1546 (2017).
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. 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).
89. Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N. & Cook, J. Misinformation and its correction: continued influence and successful debiasing. *Psychol. Sci. Public. Interest.* **13**, 106–131 (2012).
90. Barrera, O., Guriev, S., Henry, E. & Zhuravskaya, E. Facts, alternative facts, and fact checking in times of post-truth politics. *J. Public. Econ.* **182**, 104123 (2020).
91. Swire, B., Berinsky, A. J., Lewandowsky, S. & Ecker, U. K. H. Processing political misinformation: comprehending the Trump phenomenon. *R. Soc. Open. Sci.* **4**, 160802 (2017).
92. Swire, B., Ecker, U. K. H. & Lewandowsky, S. The role of familiarity in correcting inaccurate information. *J. Exp. Psychol. Learn. Memory Cogn.* **43**, 1948–1961 (2017).
93. Hamby, A., Ecker, U. K. H. & Brinberg, D. How stories in memory perpetuate the continued influence of false information. *J. Consum. Psychol.* **30**, 240–259 (2019).
94. MacFarlane, D., Tay, L. Q., Hurlstone, M. J. & Ecker, U. K. H. Refuting spurious COVID-19 treatment claims reduces demand and misinformation sharing. *J. Appl. Res. Mem. Cogn.* **10**, 248–258 (2021).
95. Tay, L. Q., Hurlstone, M. J., Kurz, T. & Ecker, U. K. H. A comparison of prebunking and debunking interventions for implied versus explicit misinformation. *Brit. J. Psychol.* (in press).
96. Nyhan, B., Reifler, J., Richey, S. & Freed, G. L. Effective messages in vaccine promotion: a randomized trial. *Pediatrics* **133**, e835–e842 (2014).
97. Poland, G. A. & Spier, R. Fear misinformation, and innuendoes: how the Wakefield paper, the press, and advocacy groups damaged the public health. *Vaccine* **28**, 2361–2362 (2010).
98. Lewandowsky, S., Stritzke, W. G. K., Oberauer, K. & Morales, M. Memory for fact, fiction, and misinformation. *Psychol. Sci.* **16**, 190–195 (2005).
99. Ecker, U. K. H., Lewandowsky, S. & Tang, D. T. W. Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Mem. Cogn.* **38**, 1087–1100 (2010).
100. Kendeou, P., Walsh, E. K., Smith, E. R. & O'Brien, E. J. Knowledge revision processes in refutation texts. *Discourse Process.* **51**, 374–397 (2014).
101. Shtulman, A. & Valcarcel, J. Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition* **124**, 209–215 (2012).
102. Kendeou, P., Butterfuss, R., Kim, J. & Boekel, M. V. Knowledge revision through the lenses of the three-pronged approach. *Mem. Cogn.* **47**, 33–46 (2019).
103. Ithiuphalap, J., Rich, P. R. & Zaragoza, M. S. Does evaluating belief prior to its retraction influence the efficacy of later corrections? *Memory* **28**, 617–631 (2020).
104. Ecker, U. K. H., Hogan, J. L. & Lewandowsky, S. Reminders and repetition of misinformation: helping or hindering its retraction? *J. Appl. Res. Mem. Cogn.* **6**, 185–192 (2017).
105. Brydges, C. R., Gignac, G. E. & Ecker, U. K. H. Working memory capacity, short-term memory capacity, and the continued influence effect: a latent-variable analysis. *Intelligence* **69**, 117–122 (2018).
106. Sanderson, J. A., Gignac, G. E. & Ecker, U. K. H. Working memory capacity, removal efficiency and event specific memory as predictors of misinformation reliance. *J. Cognit. Psychol.* **33**, 518–532 (2021).
107. Ecker, U. K. H., Lewandowsky, S., Swire, B. & Chang, D. Correcting false information in memory: manipulating the strength of misinformation encoding and its retraction. *Psychon. Bull. Rev.* **18**, 570–578 (2011).
108. Yonelinas, A. P. The nature of recollection and familiarity: Aa review of 30 years of research. *J. Mem. Lang.* **46**, 441–517 (2002).
109. Butterfuss, R. & Kendeou, P. Reducing interference from misconceptions: the role of inhibition in knowledge revision. *J. Educ. Psychol.* **112**, 782–794 (2020).
110. Brydges, C. R., Gordon, A. & Ecker, U. K. H. Electrophysiological correlates of the continued influence effect of misinformation: an exploratory study. *J. Cognit. Psychol.* **32**, 771–784 (2020).
111. Gordon, A., Quadflieg, S., Brooks, J. C. W., Ecker, U. K. H. & Lewandowsky, S. Keeping track of 'alternative facts': the neural correlates of processing misinformation corrections. *NeuroImage* **193**, 46–56 (2019).
112. Ecker, U. K. H., O'Reilly, Z., Reid, J. S. & Chang, E. P. The effectiveness of short-format refutational fact-checks. *Br. J. Psychol.* **111**, 36–54 (2020).
113. van der Meer, T. G. L. A. & Jin, Y. Seeking formula for misinformation treatment in public health crises: the effects of corrective information type and source. *Health Commun.* **35**, 560–575 (2020).
114. Wintersieck, A., Fridkin, K. & Kenney, P. The message matters: the influence of fact-checking on evaluations of political messages. *J. Political Mark.* **20**, 93–120 (2021).
115. Amazeen, M. & Krishna, A. Correcting vaccine misinformation: recognition and effects of source type on misinformation via perceived motivations and credibility. *SSRN* <https://doi.org/10.2139/ssrn.3698102> (2020).
116. Vraga, E. K. & Bode, L. I do not believe you: how providing a source corrects health misperceptions across social media platforms. *Inf. Commun. Soc.* **21**, 1337–1353 (2018).
117. Ecker, U. K. H. & Antonio, L. M. Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Mem. Cogn.* **49**, 631–644 (2021).
118. Guillory, J. J. & Geraci, L. Correcting erroneous inferences in memory: the role of source credibility. *J. Appl. Res. Mem. Cogn.* **2**, 201–209 (2013).
119. Vraga, E. K. & Bode, L. Using expert sources to correct health misinformation in social media. *Sci. Commun.* **39**, 621–645 (2017).
120. Zhang, J., Featherstone, J. D., Calabrese, C. & Wojcieszak, M. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Prev. Med.* **145**, 106408 (2021).
121. Connor Desai, S. A., Pilditch, T. D. & Madsen, J. K. The rational continued influence of misinformation. *Cognition* **205**, 104453 (2020).
122. O'Rear, A. E. & Radvansky, G. A. Failure to accept retractions: a contribution to the continued influence effect. *Mem. Cogn.* **48**, 127–144 (2020).
123. Ecker, U. K. H. & Ang, L. C. Political attitudes and the processing of misinformation corrections. *Political Psychol.* **40**, 241–260 (2019).
124. Nyhan, B. & Reifler, J. When corrections fail: the persistence of political misperceptions. *Political Behav.* **32**, 303–330 (2010).
125. Trevors, G. The roles of identity conflict, emotion, and threat in learning from refutation texts on vaccination and immigration. *Discourse Process.* <https://doi.org/10.1080/0163853X.2021.1917950> (2021).
126. Prasad, M. et al. There must be a reason: Osama, Saddam, and inferred justification. *Social. Inq.* **79**, 142–162 (2009).
127. Amazeen, M. A., Thorson, E., Muddiman, A. & Graves, L. Correcting political and consumer misperceptions: the effectiveness and effects of rating scale versus contextual correction formats. *J. Mass. Commun. Q.* **95**, 28–48 (2016).
128. Ecker, U. K. H., Sze, B. K. N. & Andreotta, M. Corrections of political misinformation: no evidence for an effect of partisan worldview in a US convenience sample. *Philos. Trans. R. Soc. B: Biol. Sci.* **376**, 20200145 (2021).
129. Nyhan, B., Porter, E., Reifler, J. & Wood, T. J. Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behav.* **42**, 939–960 (2019).
130. Wood, T. & Porter, E. The elusive backfire effect: mass attitudes' steadfast factual adherence. *Political Behav.* **41**, 135–163 (2018).
131. Yang, Q., Oureshi, K. & Zaman, T. Mitigating the backfire effect using pacing and leading. *arxiv* <https://arxiv.org/abs/2008.00049> (2020).
132. Susmann, M. W. & Wegener, D. T. The role of discomfort in the continued influence effect of misinformation. *Memory Cogn.* <https://doi.org/10.3758/s13421-021-01232-8> (2021).
133. Cobb, M. D., Nyhan, B. & Reifler, J. Beliefs don't always persevere: how political figures are punished when positive information about them is discredited. *Political Psychol.* **34**, 307–326 (2013).
134. Thorson, E. Belief echoes: the persistent effects of corrected misinformation. *Political Commun.* **33**, 460–480 (2016).
135. Jaffé, M. E. & Greifeneder, R. Negative is true here and now but not so much there and then. *Exp. Psychol.* **67**, 314–326 (2020).
136. Ecker, U. K. H. & Rodricks, A. E. Do false allegations persist? Retracted misinformation does not continue to influence explicit person impressions. *J. Appl. Res. Mem. Cogn.* **9**, 587–601 (2020).
137. Ecker, U. K. H., Lewandowsky, S. & Apai, J. Terrorists brought down the plane! No actually it was a technical fault: processing corrections of emotive information. *Q. J. Exp. Psychol.* **64**, 283–310 (2011).

138. Trevors, G., Bohn-Gettler, C. & Kendeou, P. The effects of experimentally induced emotions on revising common vaccine misconceptions. *Q. J. Exp. Psychol.* <https://doi.org/10.1177/17470218211017840> (2021).
139. Chang, E. P., Ecker, U. K. H. & Page, A. C. Not wallowing in misery — retractions of negative misinformation are effective in depressive rumination. *Cogn. Emot.* **33**, 991–1005 (2019).
140. Sangalang, A., Ophir, Y. & Cappella, J. N. The potential for narrative correctives to combat misinformation. *J. Commun.* **69**, 298–319 (2019).
141. Featherstone, J. D. & Zhang, J. Feeling angry: the effects of vaccine misinformation and refutational messages on negative emotions and vaccination attitude. *J. Health Commun.* **25**, 692–702 (2020).
142. Brashier, N. M., Pennycook, G., Berinsky, A. J. & Rand, D. G. Timing matters when correcting fake news. *Proc. Natl Acad. Sci. USA* **118**, e2020043118 (2021).
143. Cook, J., Lewandowsky, S. & Ecker, U. K. H. Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence. *PLoS ONE* **12**, e0175799 (2017).
144. Hughes, M. G. et al. Discrediting in a message board forum: the effects of social support and attacks on expertise and trustworthiness. *J. Comput. Mediat. Commun.* **19**, 325–341 (2014).
145. Paynter, J. et al. Evaluation of a template for countering misinformation — real-world autism treatment myth debunking. *PLoS ONE* **14**, e0210746 (2019).
146. Jolley, D. & Douglas, K. M. Prevention is better than cure: addressing anti-vaccine conspiracy theories. *J. Appl. Soc. Psychol.* **47**, 459–469 (2017).
147. Vraga, E. K., Kim, S. C., Cook, J. & Bode, L. Testing the effectiveness of correction placement and type on Instagram. *Int. J. Press Politics* **25**, 632–652 (2020).
148. Clayton, K. et al. Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behav.* **42**, 1073–1095 (2019).
149. Dai, Y., Yu, W. & Shen, F. The effects of message order and debiasing information in misinformation correction. *Int. J. Commun.* **15**, 21 (2021).
150. Swire-Thompson, B. et al. Evidence for a limited role of correction format when debunking misinformation. *OSF* <https://osf.io/udny9/> (2021).
151. Gordon, A., Ecker, U. K. H. & Lewandowsky, S. Polarity and attitude effects in the continued-influence paradigm. *J. Mem. Lang.* **108**, 104028 (2019).
152. Grady, R. H., Ditto, P. H. & Loftus, E. F. Nevertheless partisanship persisted: fake news warnings help briefly, but bias returns with time. *Cogn. Res. Princ. Implic.* **6**, 52 (2021).
153. Schmid, P., Schwarzer, M. & Betsch, C. Weight-of-evidence strategies to mitigate the influence of messages of science denialism in public discussions. *J. Cogn.* **3**, 36 (2020).
154. Compton, J., van der Linden, S., Cook, J. & Basol, M. Inoculation theory in the post-truth era: extant findings and new frontiers for contested science misinformation, and conspiracy theories. *Soc. Personal. Psychol. Compass* **15**, e12602 (2021).
155. Lewandowsky, S. & van der Linden, S. Countering misinformation and fake news through inoculation and prebunking. *Eur. Rev. Soc. Psychol.* <https://doi.org/10.1080/10463283.2021.1876983> (2021).
156. Roozenbeek, J., van der Linden, S. & Nygren, T. Prebunking interventions based on the psychological theory of inoculation can reduce susceptibility to misinformation across cultures. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-008> (2020).
157. Maertens, R., Roozenbeek, J., Basol, M. & van der Linden, S. Long-term effectiveness of inoculation against misinformation: three longitudinal experiments. *J. Exp. Psychol. Appl.* **27**, 1–16 (2020).
158. van der Linden, S., Leiserowitz, A., Rosenthal, S. & Maibach, E. Inoculating the public against misinformation about climate change. *Glob. Chall.* **1**, 1600008 (2017).
159. Parker, K. A., Ivanov, B. & Compton, J. Inoculation's efficacy with young adults' risky behaviors: can inoculation confer cross-protection over related but untreated issues? *Health Commun.* **27**, 223–233 (2012).
160. Lewandowsky, S. & Yesilada, M. Inoculating against the spread of Islamophobic and radical-Islamist disinformation. *Cognit. Res. Princ. Implic.* **6**, 57 (2021).
161. Ivanov, B. et al. The general content of postinoculation talk: recalled issue-specific conversations following inoculation treatments. *West. J. Commun.* **79**, 218–238 (2015).
162. Amazeen, M. A. & Vargo, C. J. Sharing native advertising on Twitter: content analyses examining disclosure practices and their inoculating influence. *Journal. Stud.* **22**, 916–933 (2021).
163. Jones-Jang, S. M., Mortensen, T. & Liu, J. Does media literacy help identification of fake news? Information literacy helps but other literacies don't. *Am. Behav. Sci.* **65**, 371–388 (2019).
164. Khan, M. L. & Idris, I. K. Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behav. Inf. Technol.* **38**, 1194–1212 (2019).
165. Machete, P. & Turpin, M. The use of critical thinking to identify fake news: a systematic literature review. *Lecture Notes Comput. Sci.* **12067**, 235–246 (2020).
166. Vraga, E. K., Tully, M., Maksil, A., Craft, S. & Ashley, S. Theorizing news literacy behaviors. *Commun. Theory* **31**, 1–21 (2020).
167. Wineburg, S., McGrew, S., Breakstone, J. & Ortega, T. Evaluating information: the cornerstone of civic online reasoning. *SDR* <https://purl.stanford.edu/fv751yt5934> (2016).
168. Breakstone, J. et al. Lateral reading: college students learn to critically evaluate internet sources in an online course. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-56> (2021).
169. Choy, M. & Chong, M. Seeing through misinformation: a framework for identifying fake online news. *arXiv* <https://arxiv.org/abs/1804.03508> (2018).
170. Amazeen, M. A. & Bucy, E. P. Conferring resistance to digital disinformation: the inoculating influence of procedural news knowledge. *J. Broadcasting Electron. Media* **63**, 415–432 (2019).
171. Guess, A. M. et al. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proc. Natl Acad. Sci. USA* **117**, 15536–15545 (2020).
172. Hameleers, M. Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands. *Inf. Commun. Soc.* <https://doi.org/10.1080/1369118x.2020.1764603> (2020).
173. Tully, M., Vraga, E. K. & Bode, L. Designing and testing news literacy messages for social media. *Mass. Commun. Soc.* **23**, 22–46 (2019).
174. Roozenbeek, J. & van der Linden, S. Fake news game confers psychological resistance against online misinformation. *Palgrave Commun.* **5**, 65 (2019).
175. Roozenbeek, J. & van der Linden, S. Breaking Harmony Square: a game that inoculates against political misinformation. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-47> (2020).
176. Micallef, N., Avram, M., Menczer, F. & Patil, S. Fakey. *Proc. ACM Human Comput. Interact.* **5**, 1–27 (2021).
177. Katsaounidou, A., Vrysis, L., Kotsakis, R., Dimoulas, C. & Veglis, A. MATHe the game: a serious game for education and training in news verification. *Educ. Sci.* **9**, 155 (2019).
178. Mihailidis, P. & Viotty, S. Spreadable spectacle in digital culture: civic expression, fake news, and the role of media literacies in post-fact society. *Am. Behav. Sci.* **61**, 441–454 (2017).
179. Carnahan, D., Bergan, D. E. & Lee, S. Do corrective effects last? Results from a longitudinal experiment on beliefs toward immigration in the U.S. *Political Behav.* **43**, 1227–1246 (2021).
180. Wintersieck, A. L. Debating the truth. *Am. Politics Res.* **45**, 304–331 (2017).
181. Moseh, M., Martel, C., Eckles, D. & Rand, D. in *Proc. 2021 CHI Conf. Human Factors Computing Systems* 2688–2700 (ACM, 2021).
182. Swire-Thompson, B., Ecker, U. K. H., Lewandowsky, S. & Berinsky, A. J. They might be a liar but they're my liar: source evaluation and the prevalence of misinformation. *Political Psychol.* **41**, 21–34 (2019).
183. Lewandowsky, S. et al. *The Debunking Handbook 2020* (George Mason Univ., 2020).
184. Kendeou, P., Smith, E. R. & O'Brien, E. J. Updating during reading comprehension: why causality matters. *J. Exp. Psychol. Learn. Memory Cogn.* **39**, 854–865 (2013).
185. Schwarz, N., Newman, E. & Leach, W. Making the truth stick & the myths fade: lessons from cognitive psychology. *Behav. Sci. Policy* **2**, 85–95 (2016).
186. Van Boekel, M., Lassonde, K. A., O'Brien, E. J. & Kendeou, P. Source credibility and the processing of refutation texts. *Mem. Cogn.* **45**, 168–181 (2017).
187. Margolin, D. B., Hannak, A. & Weber, I. Political fact-checking on Twitter: when do corrections have an effect? *Political Commun.* **35**, 196–219 (2017).
188. Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J. & Griskevicius, V. The constructive, destructive, and reconstructive power of social norms. *Psychol. Sci.* **18**, 429–434 (2007).
189. Chinn, S., Lane, D. S. & Hart, P. S. In consensus we trust? Persuasive effects of scientific consensus communication. *Public Underst. Sci.* **27**, 807–823 (2018).
190. Lewandowsky, S., Gignac, G. E. & Vaughan, S. The pivotal role of perceived scientific consensus in acceptance of science. *Nat. Clim. Change* **3**, 399–404 (2013).
191. van der Linden, S. L., Clarke, C. E. & Maibach, E. W. Highlighting consensus among medical scientists increases public support for vaccines: evidence from a randomized experiment. *BMC Public Health* **15**, 1207 (2015).
192. van der Linden, S., Leiserowitz, A. & Maibach, E. Scientific agreement can neutralize politicization of facts. *Nat. Hum. Behav.* **2**, 2–3 (2017).
193. Vlasceanu, M. & Coman, A. The impact of social norms on health-related belief update. *Appl. Psychol. Health Well-Being* <https://doi.org/10.1111/aphw.12313> (2021).
194. Nyhan, B. & Reifler, J. The roles of information deficits and identity threat in the prevalence of misperceptions. *J. Elect. Public Opin. Parties* **29**, 222–244 (2018).
195. Danielson, R. W., Sinatra, G. M. & Kendeou, P. Augmenting the refutation text effect with analogies and graphics. *Discourse Process.* **53**, 392–414 (2016).
196. Dixon, G. N., McKeever, B. W., Holton, A. E., Clarke, C. & Eosco, G. The power of a picture: overcoming scientific misinformation by communicating weight-of-evidence information with visual exemplars. *J. Commun.* **65**, 639–659 (2015).
197. van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D. & Maibach, E. W. How to communicate the scientific consensus on climate change: plain facts, pie charts or metaphors? *Clim. Change* **126**, 255–262 (2014).
198. Steffens, M. S., Dunn, A. G., Wiley, K. E. & Leask, J. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. *BMC Public Health* **19**, 1348 (2019).
199. Hyland-Wood, B., Gardner, J., Leask, J. & Ecker, U. K. H. Toward effective government communication strategies in the era of COVID-19. *Humanit. Soc. Sci. Commun.* **8**, 30 (2021).
200. Sherman, D. K. & Cohen, G. L. Accepting threatening information: self-affirmation and the reduction of defensive biases. *Curr. Direct. Psychol. Sci.* **11**, 119–123 (2002).
201. Carnahan, D., Hao, Q., Jiang, X. & Lee, H. Feeling fine about being wrong: the influence of self-affirmation on the effectiveness of corrective information. *Hum. Commun. Res.* **44**, 274–298 (2018).
202. Vraga, E. K. & Bode, L. Correction as a solution for health misinformation on social media. *Am. J. Public Health* **110**, S278–S280 (2020).
203. Bode, L. & Vraga, E. K. In related news, that was wrong: the correction of misinformation through related stories functionality in social media. *J. Commun.* **65**, 619–638 (2015).
204. Vraga, E. K. & Bode, L. Addressing COVID-19 misinformation on social media preemptively and responsively. *Emerg. Infect. Dis.* **27**, 396–403 (2021).
205. Vijaykumar, S. et al. How shades of truth and age affect responses to COVID-19 (mis)information: randomized survey experiment among WhatsApp users in UK and Brazil. *Humanit. Soc. Sci. Commun.* **8**, 88 (2021).
206. Bode, L. & Vraga, E. K. See something say something: correction of global health misinformation on social media. *Health Commun.* **33**, 1131–1140 (2017).
207. Pennycook, G. et al. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021).
208. Matz, S. C., Kosinski, M., Nave, G. & Stillwell, D. J. Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl Acad. Sci. USA* **114**, 12714–12719 (2017).
209. Vargo, C. J., Guo, L. & Amazeen, M. A. The agenda-setting power of fake news: a big data analysis of the online media landscape from 2014 to 2016. *N. Media Soc.* **20**, 2028–2049 (2018).

210. Allington, D., Duffy, B., Wessely, S., Dhavan, N. & Rubin, J. Health-protective behavior, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychol. Med.* **51**, 1763–1769 (2020).
211. Cook, J., Bedford, D. & Mandia, S. Raising climate literacy through addressing misinformation: case studies in agnotology-based learning. *J. Geosci. Educ.* **62**, 296–306 (2014).
212. Amazeen, M. A. News in an era of content confusion: effects of news use motivations and context on native advertising and digital news perceptions. *Journal. Mass. Commun. Q.* **97**, 161–187 (2020).
213. Lawrence, R. G. & Boydston, A. E. What we should really be asking about media attention to Trump. *Political Commun.* **34**, 150–153 (2016).
214. Schmid, P., MacDonald, N. E., Habersaat, K. & Butler, R. Commentary to: How to respond to vocal vaccine deniers in public. *Vaccine* **36**, 196–198 (2018).
215. Shelby, A. & Ernst, K. Story and science. *Hum. Vaccines Immunother.* **9**, 1795–1801 (2013).
216. Lazić, A. & Željelić, I. A systematic review of narrative interventions: lessons for countering anti-vaccination conspiracy theories and misinformation. *Public Underst. Sci.* **30**, 644–670 (2021).
217. Ecker, U. K. H., Butler, L. H. & Hamby, A. You don't have to tell a story! A registered report testing the effectiveness of narrative versus non-narrative misinformation corrections. *Cognit. Res. Princ. Implic.* **5**, 64 (2020).
218. Van Bavel, J. J., Reinero, D. A., Spring, V., Harris, E. A. & Duke, A. Speaking my truth: why personal experiences can bridge divides but mislead. *Proc. Natl Acad. Sci. USA* **118**, e2100280118 (2021).
219. Merpert, A., Furman, M., Anauati, M. V., Zommer, L. & Taylor, I. Is that even checkable? An experimental study in identifying checkable statements in political discourse. *Commun. Res. Rep.* **35**, 48–57 (2017).
220. Amazeen, M. A. & Wojdyski, B. W. Reducing native advertising deception: revisiting the antecedents and consequences of persuasion knowledge in digital news contexts. *Mass. Commun. Soc.* **22**, 222–247 (2019).
221. Peacock, C., Masullo, G. M. & Stroud, N. J. What's in a label? The effect of news labels on perceived credibility. *Journalism* <https://doi.org/10.1177/1464884920971522> (2020).
222. Ecker, U. K. H., Lewandowsky, S. & Chadwick, M. Can corrections spread misinformation to new audiences? Testing for the elusive familiarity backfire effect. *Cognit. Res. Princ. Implic.* **5**, 41 (2020).
223. McCright, A. M. & Dunlap, R. E. Combatting misinformation requires recognizing its types and the factors that facilitate its spread and resonance. *J. Appl. Res. Mem. Cogn.* **6**, 389–396 (2017).
224. Oreskes, N. & Conway, E. M. Defeating the merchants of doubt. *Nature* **465**, 686–687 (2010).
225. Golovchenko, Y., Hartmann, M. & Adler-Nissen, R. State media and civil society in the information warfare over Ukraine: citizen curators of digital disinformation. *Int. Aff.* **94**, 975–994 (2018).
226. Tandoc, E. C., Lim, Z. W. & Ling, R. Defining fake news. *Digit. Journal.* **6**, 137–153 (2017).
227. Mosleh, M., Pennycook, G., Archar, A. A. & Rand, D. G. Cognitive reflection correlates with behavior on Twitter. *Nat. Commun.* **12**, 921 (2021).
228. Scheufele, D. A. & Krause, N. M. Science audiences misinformation, and fake news. *Proc. Natl Acad. Sci. USA* **116**, 7662–7669 (2019).
229. Yesilada, M. & Lewandowsky, S. A systematic review: the YouTube recommender system and pathways to problematic content. *psyarxiv* <https://psyarxiv.com/6pv5c/> (2021).
230. Burszty, L., Rao, A., Roth, C. & Yanagizawa-Drott, D. Misinformation during a pandemic. *NBER* <https://www.nber.org/papers/w27417> (2020).
231. Simonov, A., Sacher, S., Dubé, J.-P. & Biswas, S. The persuasive effect of Fox News: non-compliance with social distancing during the COVID-19 pandemic. *NBER* <https://www.nber.org/papers/w27237> (2020).
232. Bechmann, A. Tackling disinformation and infodemics demands media policy changes. *Digit. Journal.* **8**, 855–863 (2020).
233. Marsden, C., Meyer, T. & Brown, I. Platform values and democratic elections: how can the law regulate digital disinformation? *Comput. Law Security Rev.* **36**, 105373 (2020).
234. Saurwein, F. & Spencer-Smith, C. Combating disinformation on social media: multilevel governance and distributed accountability in Europe. *Digit. Journal.* **8**, 820–841 (2020).
235. Tenove, C. Protecting democracy from disinformation: normative threats and policy responses. *Int. J. Press Politics* **25**, 517–537 (2020).
236. Reisch, U. The responsibility of social media in times of societal and political manipulation. *Eur. J. Oper. Res.* **291**, 906–917 (2021).
237. Lewandowsky, S. et al. Technology and democracy: understanding the influence of online technologies on political behaviour and decision-making. *Publ. Office Eur. Union* <https://doi.org/10.2760/593478> (2020).
238. Blasio, E. D. & Selva, D. Who is responsible for disinformation? European approaches to social platforms' accountability in the post-truth era. *Am. Behav. Scientist* **65**, 825–846 (2021).
239. Pickard, V. Restructuring democratic infrastructures: a policy approach to the journalism crisis. *Digit. J.* **8**, 704–719 (2020).
240. Barzilai, S. & Chinn, C. A. A review of educational responses to the post-truth condition: four lenses on post-truth problems. *Educ. Psychol.* **55**, 107–119 (2020).
241. Lee, N. M. Fake news, phishing, and fraud: a call for research on digital media literacy education beyond the classroom. *Commun. Educ.* **67**, 460–466 (2018).
242. Sinatra, G. M. & Lombardi, D. Evaluating sources of scientific evidence and claims in the post-truth era may require reappraising plausibility judgments. *Educ. Psychol.* **55**, 120–131 (2020).
243. Vraga, E. K. & Bode, L. Leveraging institutions, educators, and networks to correct misinformation: a commentary on Lewandowsky, Ecker, and Cook. *J. Appl. Res. Mem. Cogn.* **6**, 382–388 (2017).
244. Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R. & Hertwig, R. How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nat. Hum. Behav.* **4**, 1102–1109 (2020).
245. Tsipursky, G., Votta, F. & Mulick, J. A. A psychological approach to promoting truth in politics: the pro-truth pledge. *J. Soc. Political Psychol.* **6**, 271–290 (2018).
246. Bak-Coleman, J. B. et al. Combining interventions to reduce the spread of viral misinformation. *OSF* <https://osf.io/preprints/socarxiv/4jvtm/> (2021).
247. Ognyanova, K., Lazer, D., Robertson, R. E. & Wilson, C. Misinformation in action: fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-024> (2020).
248. Swire-Thompson, B. & Lazer, D. Public health and online misinformation: challenges and recommendations. *Annu. Rev. Public Health* **41**, 433–451 (2020).
249. Boele-Woelki, K., Francisco, J. S., Hahn, U. & Herz, J. How we can rebuild trust in science and why we must. *Angew. Chem. Int. Ed.* **57**, 13696–13697 (2018).
250. Klein, O. et al. A practical guide for transparency in psychological science. *Collabra Psychol.* **4**, 20 (2018).
251. Masullo, G. M., Curry, A. L., Whipple, K. N. & Murray, C. The story behind the story: examining transparency about the journalistic process and news outlet credibility. *Journal. Pract.* <https://doi.org/10.1080/17512786.2020.1870529> (2021).
252. Amazeen, M. A. Checking the fact-checkers in 2008: predicting political ad scrutiny and assessing consistency. *J. Political Mark.* **15**, 433–464 (2014).
253. Hahl, O., Kim, M. & Sivan, E. W. Z. The authentic appeal of the lying demagogue: proclaiming the deeper truth about political illegitimacy. *Am. Sociol. Rev.* **83**, 1–33 (2018).
254. Jaiswal, J., LoSchiavo, C. & Perlman, D. C. Disinformation, misinformation and inequality-driven mistrust in the time of COVID-19: lessons learned from AIDS denialism. *AIDS Behav.* **24**, 2776–2780 (2020).
255. Cheon, B. K., Melani, I. & Hong, Y. How USA-centric is psychology? An archival study of implicit assumptions of generalizability of findings to human nature based on origins of study samples. *Soc. Psychol. Personal. Sci.* **11**, 928–937 (2020).
256. Swire-Thompson, B., DeGutis, J. & Lazer, D. Searching for the backfire effect: measurement and design considerations. *J. Appl. Res. Mem. Cogn.* **9**, 286–299 (2020).
257. Wang, Y., McKee, M., Torbica, A. & Stuckler, D. Systematic literature review on the spread of health-related misinformation on social media. *Soc. Sci. Med.* **240**, 112552 (2019).
258. Bastani, P. & Bahrami, M. A. COVID-19 related misinformation on social media: a qualitative study from Iran. *J. Med. Internet Res.* <https://doi.org/10.2196/18932> (2020).
259. Arata, N. B., Torneo, A. R. & Contreras, A. P. Partisanship, political support, and information processing among President Rodrigo Duterte's supporters and non-supporters. *Philippine Political Sci. J.* **41**, 73–105 (2020).
260. Islam, A. K. M. N., Laato, S., Talukder, S. & Sutinen, E. Misinformation sharing and social media fatigue during COVID-19: an affordance and cognitive load perspective. *Technol. Forecast. Soc. Change* **159**, 120201 (2020).
261. Xu, Y., Wong, R., He, S., Veldre, A. & Andrews, S. Is it smart to read on your phone? The impact of reading format and culture on the continued influence of misinformation. *Mem. Cogn.* **48**, 1112–1127 (2020).
262. Lyons, B., Mèrola, V., Reifler, J. & Stoeckel, F. How politics shape views toward fact-checking: evidence from six European countries. *Int. J. Press Politics* **25**, 469–492 (2020).
263. Porter, E. & Wood, T. J. The global effectiveness of fact-checking: evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *Proc. Natl Acad. Sci. USA* **118**, e2104235118 (2021).
264. Ecker, U. K. H., Lewandowsky, S., Chang, E. P. & Pillai, R. The effects of subtle misinformation in news headlines. *J. Exp. Psychol. Appl.* **20**, 323–335 (2014).
265. Powell, D., Bian, L. & Markman, E. M. When intentions to educate can misinform: inadvertent paltering through violations of communicative norms. *PLoS ONE* **15**, e0230360 (2020).
266. Rich, P. R. & Zaragoza, M. S. The continued influence of implied and explicitly stated misinformation in news reports. *J. Exp. Psychol. Learn. Memory Cogn.* **42**, 62–74 (2016).
267. Shen, C. et al. Fake images: the effects of source intermediary and digital media literacy on contextual assessment of image credibility online. *N. Media Soc.* **21**, 438–463 (2018).
268. Barari, S., Lucas, C. & Munger, K. Political deepfakes are as credible as other fake media and (sometimes) real media. *OSF* <https://osf.io/cdfh3/> (2021).
269. Young, D. G., Jamieson, K. H., Poulsen, S. & Goldring, A. Fact-checking effectiveness as a function of format and tone: evaluating FactCheck.org and FlackCheck.org. *Journal. Mass. Commun. Q.* **95**, 49–75 (2017).
270. Vraga, E. K., Kim, S. C. & Cook, J. Testing logic-based and humor-based corrections for science health, and political misinformation on social media. *J. Broadcasting Electron. Media* **63**, 393–414 (2019).
271. Dunn, A. G. et al. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. *Vaccine* **35**, 3033–3040 (2017).
272. Marinescu, I. E., Lawlor, P. N. & Kording, K. P. Quasi-experimental causality in neuroscience and behavioural research. *Nat. Hum. Behav.* **2**, 891–898 (2018).
273. 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 (2021).
274. Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D. & Rich, R. F. Misinformation and the currency of democratic citizenship. *J. Politics* **62**, 790–816 (2000).
275. Shelke, S. & Attar, V. Source detection of rumor in social network: a review. *Online Soc. Netw. Media* **9**, 30–42 (2019).
276. Brady, W. J., Gantman, A. P. & Van Bavel, J. J. Attentional capture helps explain why moral and emotional content go viral. *J. Exp. Psychol. Gen.* **149**, 746–756 (2020).
277. Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A. & Van Bavel, J. J. Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl Acad. Sci. USA* **114**, 7313–7318 (2017).
278. Fazio, L. Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harv. Kennedy Sch. Misinformation Rev.* <https://doi.org/10.37016/mr-2020-009> (2020).
279. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.* **31**, 770–780 (2020).
280. Pew Research Center. *Many Americans Say Made-up News is a Critical Problem That Needs to be Fixed* https://www.journalism.org/wp-content/uploads/sites/8/2019/06/PJ_2019.06.05_Misinformation_FINAL-1.pdf (2019).
281. Pew Research Center. *Many Americans Believe Fake News is Sowing Confusion* https://www.journalism.org/wp-content/uploads/sites/8/2016/12/PJ_2016.12.15_fake-news_FINAL.pdf (2016).

282. Altay, S., Araujo, Ede & Mercier, H. If this account is true, it is most enormously wonderful: interestingness-if-true and the sharing of true and false news. *Digital Journal*. <https://doi.org/10.1080/21670811.2021.1941163> (2021).
283. Brady, W. J., Crockett, M. J. & Van Bavel, J. J. The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspect. Psychol. Sci.* **15**, 978–1010 (2020).
284. Crockett, M. J. Moral outrage in the digital age. *Nat. Hum. Behav.* **1**, 769–771 (2017).
285. Petersen, M. B., Osmundsen, M. & Arceneaux, K. The “need for chaos” and motivations to share hostile political rumors. *psyarxiv* <https://psyarxiv.com/6m4ts/> (2020).
286. Ecker, U. K. H., Lewandowsky, S., Jayawardana, K. & Mladenovic, A. Refutations of equivocal claims: no evidence for an ironic effect of counterargument number. *J. Appl. Res. Mem. Cogn.* **8**, 98–107 (2019).
287. Skurnik, I., Yoon, C., Park, D. C. & Schwarz, N. How warnings about false claims become recommendations. *J. Consum. Res.* **31**, 713–724 (2005).
288. Schwarz, N., Sanna, L. J., Skurnik, I. & Yoon, C. Metacognitive experiences and the intricacies of setting people straight: implications for debiasing and public information campaigns. *Adv. Exp. Soc. Psychol.* **39**, 127–161 (2007).
289. Cameron, K. A. et al. Patient knowledge and recall of health information following exposure to facts and myths message format variations. *Patient Educ. Counsel.* **92**, 381–387 (2013).
290. Wahlheim, C. N., Alexander, T. R. & Peske, C. D. Reminders of everyday misinformation statements can enhance memory for and belief in corrections of those statements in the short term. *Psychol. Sci.* **31**, 1325–1339 (2020).
291. Autry, K. S. & Duarte, S. E. Correcting the unknown: negated corrections may increase belief in misinformation. *Appl. Cognit. Psychol.* **35**, 960–975 (2021).
292. Pluviano, S., Watt, C. & Della Sala, S. Misinformation lingers in memory: failure of three pro-vaccination strategies. *PLoS ONE* **12**, e0181640 (2017).
293. Taber, C. S. & Lodge, M. Motivated skepticism in the evaluation of political beliefs. *Am. J. Political. Sci.* **50**, 755–769 (2006).
294. Nyhan, B., Reifler, J. & Ubel, P. A. The hazards of correcting myths about health care reform. *Med. Care* **51**, 127–132 (2013).
295. Hart, P. S. & Nisbet, E. C. Boomerang effects in science communication. *Commun. Res.* **39**, 701–723 (2011).
296. Swire-Thompson, B., Miklaucic, N., Wihbey, J., Lazer, D. & DeGutis, J. Backfire effects after correcting misinformation are strongly associated with reliability. *J. Exp. Psychol. Gen.* (in the press).
297. Zhou, J. Boomerangs versus javelins: how polarization constrains communication on climate change. *Environ. Politics* **25**, 788–811 (2016).

Acknowledgements

U.K.H.E. acknowledges support from the Australian Research Council (Future Fellowship FT190100708). S.L. acknowledges support from the Alexander von Humboldt Foundation, the Volkswagen Foundation (large grant ‘Reclaiming individual autonomy and democratic discourse online’) and the

Economic and Social Research Council (ESRC) through a Knowledge Exchange Fellowship. S.L. and P.S. acknowledge support from the European Commission (Horizon 2020 grant agreement No. 964728 JITSUVAX).

Author contributions

U.K.H.E., S.L. and J.C. were lead authors. Section leads worked on individual sections with the lead authors: P.S. on ‘Introduction’; L.K.F. (with N.B.) on ‘Drivers of false beliefs’; P.K. on ‘Barriers to belief revision’; E.K.V. on ‘Interventions to combat misinformation’; M.A.A. on ‘Practical implications’. Authors are ordered in this manner. All authors commented on and revised the entire manuscript before submission. J.C. developed the figures.

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Psychology thanks M. Hornsey, M. Zaragoza and J. Zhang for their contribution to the peer review of this work.

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

RELATED LINKS

International Fact-Checking Network: <https://www.poynter.org/ifcn/>

World Health Organization: <https://www.who.int/news-room/fact-sheets>

© Springer Nature America, Inc. 2022