The Routledge Companion to Media Disinformation and Populism

Howard Tumber, Silvio Waisbord

The effect of corrections and corrected misinformation

Publication details
Emily Thorson, Jianing Li
Published online on: 24 Mar 2021

How to cite: Emily Thorson, Jianing Li. 24 Mar 2021, The effect of corrections and corrected misinformation from: The Routledge Companion to Media Disinformation and Populism

Routledge
Accessed on: 12 Jul 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
51
THE EFFECT OF CORRECTIONS AND CORRECTED MISINFORMATION

Emily Thorson and Jianing Li

Introduction
Misinformation poses a normative problem because it has the potential to distort both beliefs and attitudes, causing people to believe things that are not true as well as to hold attitudes that differ from those that they would have held if they had been correctly informed. Ideally, then, a successful correction will alter both beliefs and attitudes, reverting both to the same state as prior to receiving the misinformation. However, not all corrections achieve both goals. Sometimes, a correction fails to alter beliefs at all. And sometimes, it successfully changes beliefs but has no effect on attitudes.

A large body of work, including some discussed in this volume, has taken on the question of how to ensure that corrections reach the people who need them most. This chapter focuses on different questions: what happens after a person receives a correction to a piece of misinformation? When are corrections successful, and when do they fail? Can misinformation continue to affect attitudes even after it is successfully corrected? This chapter begins by outlining why some corrections are more effective than others. Then, we discuss the ways in which even misinformation that is successfully corrected can shape beliefs and attitudes.

When are corrections successful at debunking misinformation?
This section discusses factors that contribute to a correction’s success at debunking misinformation. On aggregate, people do tend to move their beliefs in the expected direction when given a correction. A much-publicised 2010 study by Brendan Nyhan and Jason Reifler suggested that under certain circumstances, a correction might ‘backfire’, leading people to double down on their incorrect beliefs. However, more recent attempts to replicate and expand on those findings have demonstrated that in practice, the backfire effect is extremely rare. Wood and Porter (2019) conducted five separate experiments in which respondents were exposed to corrections of 52 different misperceptions. In the substantial majority of cases, corrections moved people closer to the truth. A similar study, conducted in the context of the 2016 US presidential election, showed that exposing people to journalistic fact-checks of false claims made by Donald Trump
The effect of corrections

led them to hold more accurate beliefs. This was true for Trump supporters as well as for the sample as a whole (Nyhan et al. 2019).

However, while corrections may on aggregate move people closer to the truth, they are by no means a panacea. First, while it is not the focus of this chapter, the people most likely to need corrections are often the ones least likely to see them (Guess et al. 2020). And second, not everyone is equally likely to accept a correction. The following section details several factors that affect the likelihood of a correction being accepted. These factors include individual-level characteristics (for example, partisanship) as well as aspects of the correction (for example, whether it includes an image).

Individual-level characteristics

Identity and motivated reasoning

When people process a new piece of information (including corrections), they are rarely objective. Rather, their pre-existing beliefs and attitudes shape the extent to which they attend to, process, and believe the new information. This tendency is called motivated reasoning (Kunda 1990; Kraft et al. 2015). The term motivation refers to the human tendency to be motivated by two different end goals. The first goal is accuracy – people generally want to hold accurate beliefs and make ‘correct’ decisions. The second goal is directional: people want to defend pre-existing identities and attitudes. The existence of motivated reasoning means that when a piece of misinformation reflects directly on someone’s identity (for example, if it concerns a controversial political issue), it will be more difficult to correct.

Political misinformation can be especially difficult to correct because it is closely tied to people’s political identities. When a piece of misinformation reinforces a person’s partisanship, they are less likely to accept a correction. For example, Ecker and Ang (2019) found that people were less likely to accept corrections of fictitious misconduct of Australian politicians if the corrections were incongruent with their partisanship. The impact of political identity on the effectiveness of corrections has also been found in comparative political settings. In a study conducted shortly after major combat in the Iraq War ended, Lewandowsky et al. (2005) found that participants in countries that supported the war were less likely to accept corrections of Iraq-related misinformation than were participants in countries that were more sceptical of the war.

Beyond political identity, social categories and cultural identities also contribute to biased processing of corrections. Recent research on ‘cultural cognition’ sheds light on the importance of underlying cultural values in orienting opinion formation and change through social and cognitive processes (Kahan and Braman 2006; Kahan et al. 2007). Motivated reasoning of corrective information not only stems from partisan biases but also results from self-serving biases fuelled by any important identities, core values, or attitudes that people hold strongly. For example, among people with strong anti-vaccine attitudes, debunking false claims on the risks of flu and MMR vaccines can result in decreased intentions to vaccinate themselves or a future child (Nyhan et al. 2014; Nyhan and Reifler 2015). Similarly, when highlighting the Muslim American identity of a religious leader, people are less likely to accept corrections of false news coverage of his claims. This tendency is heightened among those with unfavourable opinions of Islam and high social-dominance orientation (Garrett et al. 2013).
Cognitive reflection and curiosity

Individuals who are better at engaging in analytical thinking may be more capable of accepting corrections, including counter-attitudinal ones. Analytical thinking is often measured by the cognitive reflection test (CRT), on which individuals perform better if they are able to reflect on a question and resist reporting the first response that comes to mind (Frederick 2005). There are two competing arguments on whether cognitive reflection mitigates or exacerbates the propensity to engage in biased processing of corrections (Pennycook and Rand 2019). First, ‘classical reasoning approaches’, including dual-process theory, usually expect that more deliberative or analytical processing can lead to better judgments: for example, studies have shown that people who score high in analytical thinking are less likely to believe in paranormal and conspiracy concepts (Pennycook et al. 2015) and more likely to endorse scientific conclusions on evolution (Gervais 2015). On the other hand, research on motivated reasoning has offered evidence that cognitive reflection actually promotes the formation and maintenance of beliefs congruent with one’s identity and leads to a higher chance of ‘System 2 motivated reasoning’ (Kahan 2013) or opinion polarisation (Kahan et al. 2012).

Research that directly tests these competitive mechanisms in processing corrections is still scarce. Some initial evidence comes from Pennycook and Rand (2019), who found that better performance in analytical thinking was associated with more accurate judgment on whether a news headline was real or fake. Further, such relationship between analytical thinking and misinformation detection was unrelated to how closely the news headline aligned with one’s political ideology – which suggests that the reason for susceptibility to misinformation is ‘lazy, not biased’ (Pennycook and Rand 2019).

Personal interest in or curiosity about a particular topic may also mitigate motivated reasoning. Kahan et al. (2017) developed a scale of science curiosity and found that while more science comprehension led to a partisan-motivated perception of global warming and fracking (in line with ‘System 2 motivated reasoning’ argument), science curiosity counteracted the tendency to engage in motivated reasoning. Further, those who had low to modest science curiosity actively sought out unsurprising, congruent information, and those who had high science curiosity showed a high preference for surprising information, even when it was identity incongruent.

News literacy

People who are less able to navigate the changing media environment may also be more susceptible to misinformation as well as less able to understand and process corrections. In an observational study of sharing behavior during the 2016 election, Guess et al. (2019a) found that people over 65 years old, so-called digital immigrants, shared nearly seven times as much online misinformation as ‘digital natives’ aged 18 to 29. Vraga and Tully (2019) also found that those who with higher ‘news literacy’ (i.e. those with more knowledge about media industries, systems, and effects) shared less news and political content on social media and were more sceptical of information quality on social media.

Recent attempts to improve news literacy among the public have generated both hope and caveats. Simple intervention such as a general warning of online misinformation (Clayton et al. 2019) and news literacy intervention with more detailed tips on spotting misinformation (Guess et al. 2019b) can help people perceive false headlines as less accurate, even when the false headlines are congenial to their political identity. However, the increased scepticism of false headlines also had a spillover effect, leading people to become more sceptical of true news headlines (Clayton et al. 2019; Guess et al. 2019b). While research has not specifically
examined whether this spillover effect extends to corrections, it is plausible that by increasing scepticism of all information, news literacy interventions may also make people less inclined to believe corrections.

Further, the effects of news literacy interventions may not last over time (Guess et al. 2019b). Another concern is the practicality of these interventions in the real-world media environment as, in most studies, participants were exposed to a message provided by researchers in an experimental setting. Tully et al. (2019) found evidence that a ‘promoted tweet’ (Twitter’s paid post) encouraging users to be critical news consumers can mitigate the effect of exposure to misinformation on GMOs and flu vaccines. However, the effect was not consistent across experiments. In addition, most experimental interventions designed to increase news literacy are focused on changing how people process misinformation. More research is needed to understand how news literacy interventions might also affect how they process corrections.

**Characteristics of the correction**

**Source**

A long literature in communication has examined the power of source credibility, including both perceived expertise and trustworthiness, in shaping whether or not a given piece of information is accepted (Page et al. 1987; Eagly and Chaiken 1993; Gilbert et al. 1998). This literature has direct implications for how people process both misinformation and corrections. The source of a piece of information – whether it is a politician, a fact-checking organisation, or a friend – affects whether people believe it. The more credible they deem the source, the more likely they are to accept the information.

Perceptions of source credibility are shaped by a person’s pre-existing attitudes and beliefs, with congenial sources deemed more credible and thus more persuasive. For example, people may find a person who shares more cultural values with them (e.g. hierarchy–egalitarianism, individualism–communitarianism) more credible and more persuasive in their arguments, either pro- or anti-vaccination (Kahan et al. 2010). Guillory and Geraci (2013) offered direct evidence for how source congeniality shaped the effectiveness of corrections. They found that corrections of fictitious misinformation about political bribery were more effective when coming from sources that participants had previously deemed credible. In particular, the perceived trustworthiness of the source sufficiently reduced reliance on misinformation while perceived expertise was less powerful (Guillory and Geraci 2013).

The effect of source congeniality transfers to real-world social contexts too: Twitter users are more likely to accept corrections coming from users with whom they have a mutual follower-following relationship than corrections coming from strangers (Margolin et al. 2018). However, while congeniality matters, people may also take into account the motives of the corrector: for example, a Republican correcting a fellow Republican may be seen as an especially credible source because she appears to be acting in a way that runs counter to her own interests (Berinsky 2017). Finally, while this chapter is focused mainly on corrections, it is worth noting that the original source of the misinformation can be as important, if not more important, than the source of the correction. In an experiment examining false statements made by American presidential candidate Donald Trump, Swire et al. (2017) found that while the source of the correction had little impact on the correction’s effectiveness, the source of the misinformation mattered more.

Since one of the major sources of corrections is fact-checking organisations, it is important to understand both when and why these organisations are deemed credible. Whether or
not the practices of fact-checking organisations are reliable and unbiased has been questioned in academic research and public discourse (Nieminen and Rapeli 2019). Shin and Thorson (2017) used a unique dataset to examine public reactions to fact-checkers: how Twitter users share and comment on fact-checks from PolitiFact. They found that fact-checks of politicians spurred users who shared that politician’s party to criticise fact-checking organisations, including accusing them of bias. These critiques may have downstream effects, reducing the perceived credibility of these organisations (Thorson et al. 2010). Finally, at least in the American context, partisanship also shapes the perceived credibility of fact-checkers. In surveys, Republicans are substantially more likely than Democrats to say that fact-checkers tend to favour one side (Walker and Gottfried 2019). Republicans are also more likely to criticise fact-checking organisations on social media (Shin and Thorson 2017).

Some corrections come not from fact-checks but directly from journalists. In the US, fact-checking practices are adopted by several mainstream news outlets such as AP and USA Today and also occupy a notable, although non-primary, role in journalists’ own use of Twitter (Coddington et al. 2014). The practice of directly correcting a politician’s false statement, often termed ‘journalistic adjudication’, is controversial as it runs counter to the traditional ‘he said, she said’ model of reporting disputed factual claims (Pingree et al. 2014). However, journalistic adjudication can be an effective approach to fact-checking, even when a correction runs counter to pre-existing political attitudes (Lyons 2018). Finally, experts can be an effective source of corrections. In the social media context, corrections of health misinformation that come from expert sources like the Centers for Disease Control (CDC) are more successful than those that come from social media users (Vraga and Bode 2018; van der Meer and Jin 2020).

**Context**

How people encounter a correction – be it on social media, in a news article, or passed on by a friend – can shape its effectiveness. Social media can enable the spread of not only misinformation but also corrections, although people are more likely to share corrective information that is relatively advantageous to their own side (Shin and Thorson 2017). Such ‘social correction’ of misinformation is most effective when users have a mutual following-follower relationship (Margolin et al. 2018) or when an expert source is attached to the information (Vraga and Bode 2018).

Platforms can also play a role in facilitating the correction of misinformation. When a correction was recommended by Facebook as a ‘related story’ presented under a post containing misinformation about GMOs, it substantially reduced users’ misperceptions about GMOs (Bode and Vraga 2015). The effectiveness of this ‘related story’ function, however, was not seen in the context of vaccine-autism misperception, in which people’s attitudes were held much more strongly (Bode and Vraga 2015). Further, flagging a false story’s publisher as an ‘self-identified source of humor’ can reduce misperceptions even more effectively than flags generated by fact-checkers or other social media users, especially among those most predisposed to believe the misinformation (Garrett and Poulsen 2019). One reason for the effectiveness of this strategy may be that it not only corrects the misinformation but also provides an ‘alternative narrative’ to explain why it exists in the first place (Hamby et al. 2020).

**Format**

In some circumstances, visuals can increase the effectiveness of corrections. Dixon et al. (2015) found that when correcting vaccine misinformation with messages about scientific consensus,
The effect of corrections

providing a visual exemplar of such consensus, such as a photo of a group of scientists, was more effective in refuting the vaccine-autism link than text-only corrections, especially among those who had pre-existing unfavourable views of science. Figures that present numeric information (e.g. the number of attacks in the Iraq War, the growth in jobs, the change in global average temperature) can also reduce misperceptions more effectively than equivalent texts (Nyhan and Reifler 2019). In addition, ‘truth scales’ that visualise the ruling on the veracity of the claim can also be more effective than a text-only correction (Amazeen et al. 2018). Finally, putting fact-checks in a video format increases the likelihood that viewers will pay attention, understand the issue, and accept the correction (Young et al. 2018).

Not all images are equally effective. Hameleers et al. (2020) found that ‘multimodal’ fact-checks on social media (i.e. those that include images) were no better than text-only ones at reducing misperceptions. However, the images used in their experimental manipulation were not directly relevant to the content of the fact-check. This finding suggests that in order for an image or video to increase a fact-check’s effectiveness, it should directly reinforce the central message of the correction.

Wording

Researchers have studied several different aspects of how wording might shape effectiveness. We focus here on two central questions. The first is whether it helps or hurts to repeat the initial misinformation as part of the correction. The second is whether it is useful to provide an ‘alternative narrative’ as part of the correction that provides more detail about the true story.

When a piece of misinformation is repeated, it starts to take on the ‘illusion of truth’, seeming more familiar and thus more true (Berinsky 2017). For example, people who are exposed to ‘fake news’ articles even once are more likely to later say that they are accurate (Pennycook et al. 2018). The worry, then, is that if corrections repeat the initial misinformation, they may have the unintended effect of making the misinformation seem more plausible. However, the empirical evidence for this effect is limited: while corrections that repeat the initial misinformation are slightly less effective than those that do not, they do not fully backfire (Swire, Ecker, et al. 2017). Indeed, recent evidence has even shown that corrections that explicitly repeated the misinformation can be more effective in improving belief accuracy than corrections that avoided repeating the misinformation, possibly due to the increased salience of the falsity of the misinformation (Ecker et al. 2017).

The amount of detail provided in the correction also matters for its success. A full explanation of why a piece of misinformation is wrong (a refutation) is more effective than a simple statement that a piece of misinformation is wrong (a retraction), and these effects persist for weeks after seeing the initial misinformation and correction (Swire, Ecker, et al. 2017). Similarly, Sangalang et al. (2019) found that corrections embedded within a narrative story were effective in correcting misperceptions about tobacco products and changing related behavior intentions, potentially because they increased involvement in and attention to the content of the correction.

The correction worked – what next?

Even when a correction is fully accepted – in other words, when a person believes the correction and understands that the misinformation is false – its effects can linger. This section focuses on how successfully corrected misinformation can shape beliefs, attitudes, and trust.
The effect of misinformation on attitudes can persist

Even when corrective information successfully updates people’s beliefs, it may not be equally effective at influencing attitudes that were influenced by the misinformation. Thorson (2016) refers to these lingering attitudinal effects as ‘belief echoes’, presenting experimental evidence that even when a correction successfully reverts beliefs back to pre-misinformation levels, attitudes are still swayed by the retracted misinformation (Thorson 2016). The ‘belief echoes’ effect has also been replicated in contexts where accurate belief updating failed to produce downstream effects on favourability towards Donald Trump (Nyhan et al. 2019) or attitudes towards immigration (Hopkins et al. 2019).

Corrected misinformation also continues to shape attitudes and opinions in non-political contexts. It affects behavior, memory, and opinion through a process sometimes called the ‘continued influence effect’ (Ecker, Lewandowsky, and Apai 2011; Ecker, Lewandowsky, Swire, et al. 2011). Across a wide range of topics and outcomes, the results are consistent: even when someone accepts a correction, it does not always fully ‘un-ring the bell’ of misinformation.

There are several mechanisms for this continued influence effect. First, people sometimes engage in attitude-congruent motivated reasoning. Even when their misperceptions are corrected, they ‘explain away’ this new information by looking for other arguments that help align the uncongenial information with their preferred worldview. One such strategy is biased attribution of blame: even when partisans accept facts about the changes in economic conditions, they rationalise the information by blaming the opposing party for the worsening conditions and praise their own party for the improvements (Bisgaard 2019). People may also explain away the uncongenial information with biased credibility judgments, expressing their displeasure by concluding that the source of the uncongenial information is not credible after all (Khanna and Sood 2018).

Second, sometimes the misinformation becomes part of a person’s ‘mental model’ of a particular event and thus becomes more difficult to dislodge (Ecker et al. 2015). For example, when retracting misinformation about a fictitious warehouse fire (that it was due to negligence with volatile materials), the retraction was not effective in reducing reliance on the misinformation unless it offered an alternative explanation (evidence of arson was found elsewhere) (Johnson and Seifert 1994). Because alternative explanations help people to revise their mental models, they can be more effective than simple retractions at reducing the continued influence effect (Walter and Murphy 2018).

Finally, a piece of misinformation can carry an ‘affective charge’ that shapes a person’s emotional reactions to the object of the misinformation in a way that even a successful correction cannot fully eliminate (Sherman and Kim 2002). Lodge and Taber (2005) call this ‘hot cognition’ and found that the affective charge attached to a socio-political concept can be activated within milliseconds of exposure, much faster than the cognitive evaluation of the concept. Further, people had a more difficult time processing affectively incongruent information (e.g. cockroach – delightful) than affectively congruent information (e.g. cockroach – disgusting), which implies that corrections that run counter to one’s automatic affective responses may be less effective.

Unintended consequences of the media’s focus on corrections

Finally, it is worth noting that the intense media focus on fact-checking and misinformation may have additional unintended consequences. When people are repeatedly exposed to corrected misinformation, they may infer a larger lesson: that the information environment is a
The effect of corrections

dangerous and difficult to navigate place (Wenzel 2019). Some empirical evidence gives credence to this concern. For example, an intervention such as a general warning about misinformation on social media, despite reducing people’s beliefs in false news headlines, also reduced their beliefs in true headlines (Clayton et al. 2019). This spillover effect can be also seen in similar interventions such as providing tips on detecting misinformation (Guess et al. 2019b), although in both types of interventions, the effect size of reducing beliefs in true headlines is substantially smaller than the effect size of reducing beliefs in false ones. Further, reading corrections in which journalists adjudicate factual disputes can also reduce the confidence in one’s ability to find truth in politics among those who are less interested in the topic under dispute, raising normative concerns over the unintended effect on political efficacy (Pingree et al. 2014).

Conclusion

On average, corrections are successful at moving people closer to the truth. When confronted with a piece of corrective information, at least some people will change their beliefs in the direction of the correction. However, ‘some’ is often not good enough, especially when it comes to misinformation with direct health consequences (for example, vaccines). The literature reviewed in this chapter offers a larger framework for how to think about what makes a correction successful, as well as specific, evidence-based advice for those looking to design corrections that work.

When evaluating whether a correction is successful, it is important to first offer a clear definition of success. Even when a correction is fully accepted, it may not have downstream effects on related attitudes. Normatively, this is not always a problem. To give a specific example, we would neither expect nor want a single correction of a politician’s falsehood to turn someone from an ardent supporter to a fierce opponent. Attitudes, especially central ones like those related to politics and health, are not so easily changed by a single piece of information (or its correction) – nor should they be.

Many of the empirical findings about what makes corrections successful are to some extent common sense. Most importantly, corrections are more successful when a person is inclined to believe them anyway (i.e. when a correction confirms what they already wish was true). However, much of the time, we are concerned about how to make corrections work for people who are not inclined to believe them. A decade ago, to answer this question we would have had to rely on intuition and/or theories imported from other contexts. Today, we can draw on hundreds of empirical investigations of how to correct misinformation across a range of contexts.

While these investigations offer a nuanced view of how and when corrections succeed, a few major lessons stand out. Corrections work better when they come from a trusted source. Corrections work better when they offer the reader (either with visuals or with narrative) a clear, compelling story. And finally, they work better when are met with curiosity, desire for accuracy, and the ability to navigate the increasingly complex media environment.

References


The effect of corrections


