

The backfire effect

Does it exist? And does it matter for factcheckers?

Amy Sippitt

Full Fact is the UK's independent factchecking charity.

About this paper

Full Fact's Research Manager, Amy Sippitt, studies what existing academic evidence tells factcheckers about how our work can be most effective. This briefing looks at the common question posed to factcheckers of whether there is a backfire effect.

We would like to extend our warmest thanks to Brendan Nyhan, Ethan Porter and Briony Swire-Thompson for their comments.

Published by Full Fact, March 2019

Full Fact
2 Carlton Gardens
London
SW1Y 5AA

team@fullfact.org
+44 (0)20 3397 5140
@FullFact
fullfact.org

Registered Charity number 1158683

Published under the Creative Commons Attribution-ShareAlike 4.0 International License

Contents

Quick summary	4
Introduction	5
How we've selected the studies, and some questions you might consider	8
So, where has backfire been observed?	9
Other studies that have found some form of backfire	11
Recent attempts to replicate have found no backfire	13
Other studies	20
Studies in the UK	20
Studies looking at broader issues	20
What does this all mean for factcheckers?	21

Quick summary

- A “backfire effect” refers to the effect that, **when a factual claim reinforces someone’s ideological beliefs**, telling them that the claim is wrong (“debunking” it) can actually make them believe the claim *more strongly* rather than less.
- This supposed effect is often interpreted as meaning that factchecking is ineffective, or even counterproductive.
- This briefing looks at seven major experimental studies that have examined supposed backfire effects, mostly in the United States.
- It finds that while backfire **may occur** in some cases, **the evidence now suggests it is rare rather than the norm**, and that generally **debunking can make people’s beliefs in specific claims more accurate**.
- **Two studies**, from 2010 and 2012, found some evidence of a backfire effect in certain circumstances.
- **None of the five more recent studies** looked at (from 2015, 2017, 2018 and 2019) have found any evidence of the effect.
- The cases where backfire effects were found tended to be particularly contentious topics, or where the factual claim being asked about was ambiguous.
- Our **worldview may still affect** the extent to which a debunk might be effective.
- The impact of debunks on our behaviour is more complicated. This briefing does not look at this in detail; it focuses on how debunks affect belief in the accuracy of the claim in question.
- This review suggests factchecking does help to inform citizens and backfire effects are rare rather than the norm. We still need more evidence to understand how factchecking content can be most effective.

Introduction

A backfire effect, as originally conceived, is the idea that when a factual claim aligns with someone’s ideological beliefs and they are told that claim is wrong, they believe it more strongly than if you hadn’t said anything. This type of backfire is specifically a “worldview” backfire effect, and is also sometimes referred to as a boomerang effect.¹

There is some evidence that this type of backfire effect may occur in some cases—such as with particularly contentious topics, or where the claim being asked about is ambiguous or lacks clarity.² But there is a growing body of evidence that in many cases it doesn’t. Recent studies have concluded that “worldview backfire effects are not the norm and may occur under very specific circumstances”, and that “citizens can accept factual corrections of misstatements even when they are made by one’s preferred candidate during a presidential election”.³

Much of the evidence we have on backfire comes from [laboratory and survey experiments](#), mostly in the United States. The most recent studies now suggest that generally debunks (see definitions box below) can make beliefs in specific claims more accurate.

That said, we’ve still got a long way to go to understand why beliefs in inaccurate claims persist and what can be done to address them.

‘The most recent studies now suggest that generally debunks can make beliefs in specific claims more accurate.’

1 Other types of backfire include for example the “familiarity backfire effect” (the effect of repeating the disputed claim within a debunk) where the evidence is similarly complicated. See: skepticalscience.com/Debunking-Handbook-Part-2-Familiarity-Backfire-Effect.html and Swire B, Ecker UKH, Lewandowsky S, 2017. “The role of familiarity in correcting inaccurate information”, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43 (12), 1948-1961.

2 See Wood T, Porter E, 2017. “The elusive backfire effect: mass attitudes’ steadfast factual adherence”, forthcoming at *Political Behavior*.

3 Swire B, Berinsky AJ, Lewandowsky S, Ecker UKH, 2017. “Processing political misinformation: comprehending the Trump phenomenon”, *Royal Society Open Science*, 4 (3); Nyhan B, Porter E, Reifler J, Wood T, 2019. “Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability”, forthcoming at *Political Behavior*.

The [research also shows](#) that our attitudes can still affect the extent to which a debunk might be effective—the likelihood that someone will be persuaded by it. And there are questions around how to ensure that debunks have lasting effects on beliefs in claims.

Beyond experimental settings, we know that both in the UK and worldwide there are specific inaccurate claims which are still believed by many, despite widespread statements to the contrary.

For example, [Ipsos MORI's 2017 Perils of Perception series](#) found about 20% of people across 38 nations believed it's true that there is a link between some vaccines and autism in healthy children, and a further 38% said they didn't know.

This is despite a large body of evidence⁴ that, in the words of the UK's [National Autistic Society](#), demonstrates “research has comprehensively shown that there is no link between autism and vaccines”. And there is [evidence](#) from the United States that after President Obama published his birth certificate, misperceptions about his nationality initially fell, but then in time increased again.

This briefing primarily focuses on belief in the accuracy of the specific claim in question. Some of the studies also look at the extent to which debunks change people's attitudes or behaviour—to use the earlier example, you might be persuaded that there is no link between vaccines and autism, but at the same time become more convinced you don't want your child to be vaccinated.⁵

These are also [sometimes referred to](#) as “backfire effects”, but we leave them out of our definition of backfire. This is a more complex area, and there is limited research addressing how debunking changes underlying attitudes or behaviour. For factcheckers there is also a question of how to interpret the meaning of this research for our work, since factcheckers' primary focus tends to be on promoting accurate information and not on how individuals choose to make decisions on the basis of that information.

4 See for example: Stratton K, Ford A, Rusch E, & Wright Clayton E, 2011. *Adverse Effects of Vaccines: Evidence and Causality*, (Washington: National Academies Press); Taylor LE, Swerdfeger AL & Eslick GD, 2014, “Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies”, *Vaccine*, 32 (29), 3623-9.

5 See for example: Nyhan B, Reifler J, Richey S, & Freed GL, 2014. “Effective messages in vaccine promotion: A randomized trial”, *Pediatrics*, 133 (4), 835-842.

Definitions

Factual claims: We use this to refer to statements about the state of the world which Full Fact/factcheckers could factcheck, as distinct from statements of opinion.

Corrections: The academic literature on this topic tends to talk about statements which debunk other statements as “corrections”. At Full Fact we talk about “corrections” as public acknowledgements that factual claims were inaccurate (such as a politician correcting the record, or a newspaper correcting an article), so in this briefing we’re just using the terms ‘debunking’ and ‘debunking messages’, except where directly quoting from academic studies.

Debunking: Factual messages which seek to rebut inaccurate factual claims. Here we will use this as a catch-all term, which includes debunks which might not directly say a claim is wrong (for example they might simply provide the accurate information, but not directly call out the claim), or content where the rebuttal is only one part of a broader journalistic report on the topic.

Factchecks: We will use this term specifically for published articles in which the primary purpose of the content is evaluating and explaining the accuracy of a claim (such as might come from dedicated factchecking organisations).

Attitude-inconsistent information: We use this for text containing factual information that is inconsistent with someone’s beliefs, but does not seek to directly address specific factual claims quoted in the same place.

How we've selected the studies, and some questions you might consider

We have focused this briefing on the studies that we have seen referenced the most by academics in relation to backfire effects⁶, and specifically on studies that examine the effects of debunks of factual claims (rather than simply the effects of attitude-inconsistent information).⁷

We have also focused on experimental studies where respondents are randomly assigned to treatment and control groups.

We look in detail at seven studies, and reference the findings of three further studies and three literature reviews. This briefing explores the studies from a practical perspective. We've looked at things like: what format of debunk has been tested (such as within news articles or standalone conclusions)? How clear is the debunk? What types of claims and claimants have been tested—highly controversial ones, or less contentious topics? Who has been sampled and how large is the sample?

Many of the studies tend to be conducted using participants on Amazon's Mechanical Turk (MTurk). MTurk is an online marketplace where participants get paid for participating in surveys. While its use is common in academic studies—it's cheaper than nationally representative sample providers—it has some significant caveats to be aware of.

For one, people who choose to join the marketplace are known to differ in at least some ways to those who don't. In the United States, Conservatives are **known to be under-represented** on the platform for example. And secondly, people who are on MTurk can fill out a lot of surveys, including on the same topic, and **there is evidence** that at least some of them start to 'game' the system according to what they think is expected or not expected of them.

Some of the studies below have attempted to check for differences, comparing MTurk samples to nationally representative samples—and have generally found no significant differences.

6 For example in Flynn DJ, Nyhan B, and Reifler J, 2017. "The nature and origins of misperceptions: understanding false and unsupported beliefs about politics", *Political Psychology* 38 (S1), 127-150; Lewandowsky S, Ecker UKH, Cook J, 2017. "Beyond misinformation: understanding and coping with the "post-truth" era", *Journal of Applied Research in Memory and Cognition* 6 (4), 353-369; Nyhan B, Porter E, Reifler J, Wood T, 2019. "Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability", forthcoming at *Political Behavior*.

7 For this reason we have excluded Hart PS, Nisbet EC, 2011. "Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies", *Communication Research*, 39 (6), 701-723, and Schaffner B and Roche C, 2016. "Misinformation and motivated reasoning: responses to economic news in a politicized environment", *Public Opinion Quarterly* 81 (1), 86-110.

So, where has backfire been observed?

There are two main studies that [tend to](#) be pointed to as evidence of factual backfire taking place under some circumstances (leaving out the studies on vaccine effects). Both of these focus on short debunks as they might appear in news articles, rather than full factcheck articles explaining the analysis. They also test generally less conclusive debunking messages compared to more assertive conclusions tested in recent studies discussed below.

“When corrections fail: the persistence of political misperceptions”

Brendan Nyhan and Jason Reifler (published in *Political Behavior*, 2010)

The idea of a backfire effect first became particularly well known from a [study](#) by Brendan Nyhan and Jason Reifler in 2010, where they found several instances of where “corrections actually increase misperceptions among the group in question”.

Their theory was that this happens because when people are shown information that goes against their political beliefs, they will “counterargue” it in their minds strongly enough that then they’ll end up with stronger beliefs in the original inaccurate claim than if they weren’t shown the debunking information.

Ever since, journalists and others have [repeatedly referenced this study](#) (or related studies) to say that attempts to correct people cause backfire effects. (One [recent example](#), from New York Times columnist Nicholas Kristof: “One challenge is that fact-checking doesn’t work very well... This is called the ‘backfire effect.’”) However, even on this study the [authors say](#) the effects were “overstated and oversold” and that they never actually found that it *always* happened.

They conducted two studies, one in autumn 2005 and one in spring 2006, testing five claims in total among samples of less than 200 US undergraduates. They found evidence of backfire in response to two claims they defined as inaccurate, both made by ex-US President George W. Bush.

The first and most well-known claim, tested in the first study, was about whether Iraq had weapons of mass destruction (WMD) that were destroyed or hidden before US forces arrived in the US invasion. Participants were shown a news article about comments made by George Bush defending the decision to invade Iraq and were then asked the extent to which they

agreed with the statement “Immediately before the U.S. invasion, Iraq had an active weapons of mass destruction program, the ability to produce these weapons, and large stockpiles of WMD, but Saddam Hussein was able to hide or destroy these weapons right before U.S. forces arrived”.

Some had been shown the news article on its own, and some saw the news article with a debunk within the article saying that the Central Intelligence Agency had released a report to say no weapons had been stockpiled. Conservatives who saw the news article with the debunk ended up agreeing with the statement more strongly than those who saw the news article without the debunking element—a “backfire effect”.

The second claim displaying evidence of backfire, tested in the second study, was about whether George Bush’s tax cuts paid for themselves in increased government revenue.

There was no significant evidence of backfire in a further three claims tested in the second study, including a simplified claim about WMD.⁸ There was some evidence of debunks having a neutral effect.

So what can we make from this? The experiment purposefully covered a highly controversial topic in American politics where people would have prior beliefs, to differ from hypothetical scenarios tested in previous research. So it’s arguably unsurprising that individuals were unpersuaded by a single news item. A [recent study](#), which we’ll come on to, tested a simplified version of the claim and found no backfire—which could either reflect the lower intensity of the issue given it was tested a few years later, or less ambiguity and nuance in the wording of the claim.

The authors also said they purposefully tested debunks which did not come down hard on what the truth was, so they would appear like they do in news articles—in this instance, saying that the Central Intelligence Agency had released a report to say no weapons had been stockpiled, rather than explicitly saying the claim was false. It’s possible this made the debunk less convincing.

This study also had a small sample of fewer than 200 people—which means it [might have been](#) a mistaken finding—and was conducted among undergraduates rather than a national sample.

More recent evidence has suggested that these findings may have been due to the ambiguity of the statement tested, and due to this smaller sample size. We’ll come back to this later.

⁸ This includes an additional claim mentioned in a footnote in the paper, where the authors say they tested corrections to a claim made by Michael Moore (in the movie “Fahrenheit 9/11”) that the war in Afghanistan was motivated by oil company Unocal’s desire to build a natural gas pipeline through the country. They found no relevant statistically significant effects.

“The hazards of correcting myths about health care reform”

Brendan Nyhan, Jason Reifler and Peter Ubel (published in *Medical Care*, 2012)

In this larger study of around 900 US participants from an opt-in internet panel, Nyhan and Reifler worked with Peter Ubel to test for backfire in response to a claim made by former Alaskan Governor Sarah Palin. They found backfire in just one subgroup: among Palin supporters with high political knowledge.

Participants were randomly assigned to either a control group in which they read an article on Sarah Palin’s claims about “death panels” or an intervention group in which the article also contained corrective information refuting Palin. The claim related to President Obama’s health reforms which Palin claimed would create a “death panel” in which bureaucrats decide whether to continue a person’s health care.

For people with high political knowledge and who supported Sarah Palin, they found they clung more strongly to the beliefs they held. This was the only group where backfire effects were found in this study.

The study used a similar debunking style as the 2010 study where they quoted someone else as saying the claim was inaccurate, using the text: “However, non-partisan health care experts have concluded that Palin is wrong”.

Brendan Nyhan and Jason Reifler have [updated their beliefs](#) since these studies. Nyhan [now says](#) that facts can change minds “in some cases”—the question is when and how durably they can be changed. He points to macro public opinion data which he says is less encouraging, as we can see from surveys like Ipsos MORI’s Perils of Perception series.

Other studies that have found some form of backfire

Some other studies have found evidence of backfire when it comes to underlying behavioural intentions, but not for specific belief in the claim debunked. For example, as referenced earlier, two studies⁹ testing the effects of debunks addressing specific fears about vaccines found the debunks were

9 Nyhan B, Reifler J, Richey S, & Freed GL, 2014. “Effective messages in vaccine promotion: A randomized trial”, *Pediatrics*, 133 (4), 835-842; Nyhan B, Reifler J, 2015. “Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information”. *Vaccine*, 33(3), 459-64.

effective at reducing beliefs in the claim they were debunking, but still decreased the intention to vaccinate among those with high levels of concern about vaccine side effects.

This raises the question of what effect debunks or factchecks can have (and should seek to have) on underlying attitudes and behaviour—especially when the influence of our attitudes on our factual beliefs is often hard to disentangle from the extent to which our factual beliefs cause our attitudes.¹⁰

¹⁰ Duffy B, 2018. *The Perils of Perception: Why we're wrong about nearly everything* (London: Atlantic Books).

Recent attempts to replicate have found no backfire

Recent attempts to directly test for the backfire effect, generally among much larger samples than previously used, have found little evidence of the effect.

These studies have tested claims and debunks appearing on their own, as well as claims and debunks appearing in a mock news article.

“The elusive backfire effect: mass attitudes’ steadfast factual adherence”

Thomas Wood and Ethan Porter (published in *Political Behavior*, 2018)

A recent study by Thomas Wood and Ethan Porter conducted five studies testing 52 claims with more than 10,100 people in the US.

They selected prominent issues, ranging from the incidence of gun violence through to undocumented immigrants’ criminal activities. They tested claims and debunks both as separate pieces of text, and within mock news articles.

They conclude that backfire is “stubbornly difficult to induce” and found: “The average subject exposed to the correction subsequently expressed attitudes more in line with the facts.” This includes with Nyhan and Reifler’s WMD text, where they found no backfire.

That’s not to say there was no effect of partisanship—in the majority of cases they still found that people were more or less likely to believe a claim was true or false depending on their political beliefs. But they found that debunking did move individuals towards being more accurate on average, including among both liberals and conservatives.

They say their results do “not lead us to conclude that backfire is categorically impossible. Certain issues and certain questions—perhaps asked at moment when ideology or partisanship, or both, are particularly salient—might plausibly trigger factual backfire”.

The types of debunks they tested varied from fairly non-confrontational factual statements as might appear in news articles such as “in fact, according to public records, homicides

of law enforcement officers have been declining for decades”, to more direct debunks, for instance stating that a claim was “plainly false”.

What makes their study particularly interesting is that they also tested to see whether the complexity of the claims participants were asked to agree or disagree with, and the relevance of the debunk to the claims, made any difference to the effect of the debunk.

Nyhan and Reifler’s original study asked respondents if it was true that “Immediately before the U.S. invasion, Iraq had an active weapons of mass destruction program, the ability to produce these weapons, and large stockpiles of WMD, but Saddam Hussein was able to hide or destroy these weapons right before U.S. forces arrived.” To see what effect the ambiguity of this statement had, Wood and Porter tested this and a simpler version: “Following the US invasion of Iraq in 2003, US forces did not find weapons of mass destruction”.

They found no backfire from either version. However, the simpler version was the only version of the question to successfully elicit a change in responses from participants—with liberals presented with the simpler version adopting the debunk and conservatives showing no effect. Wood and Porter suggest this might be because the original wording offered more room for dispute, and was also just potentially too complex for respondents.

When they varied the complexity of six other claims, they found no level of complexity caused backfire, but more complex statements did decrease the effect of the debunk among conservative respondents.

Lastly, they asked respondents to rate the ‘relevance’ of debunking statements to the original claim they were debunking. Surprisingly, they found no significant differences in the extent to which debunks changed people’s beliefs depending on whether the debunk was seen as closely related or tenuous.

All but the last of their studies were conducted with participants from MTurk. Their last test compared results between 1,000 Mechanical Turk participants and 1,000 participants from a nationally representative sample. They found generally no significant differences in responses to corrections, although people identifying in the centre of the political spectrum on MTurk were found to be slightly more responsive to new factual information than those in the national sample.

“Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability”

Brendan Nyhan, Ethan Porter, Jason Reifler and Thomas Wood (forthcoming in *Political Behavior*, 2019)

Wood and Porter have since collaborated with Nyhan and Reifler on a study at the height of the 2016 US Presidential elections, and found even at the peak of election season they could reduce misperceptions about trends in crime and unemployment among both Republicans and Democrats, based on claims made by Donald Trump.

They tested two claims: one made in Donald Trump’s nomination acceptance speech that violent crime had increased substantially, with the study conducted several weeks after; and one made by Donald Trump in the first 2016 presidential debate that jobs were moving from the US to Mexico, with the study conducted that evening. Misperceptions decreased among both Clinton and Trump supporters, compared to those who saw the claim without any corrections.

No Hillary Clinton claims were tested, so we only have evidence of possible partisan reactions by Trump supporters.

The debunks they tested were similar to the fairly non-confrontational debunks tested in previous studies and were based on media debunks at the time. The first one started by saying, “According to the FBI’s Bureau of Justice Statistics, the violent crime rate has fallen dramatically and consistently over time...”, and the second by saying, “In fact, according to the Bureau of Labor Statistics, unemployment has fallen in both states...”.

In the first study they tested to see what happens when you include quotes from elites trying to undermine the debunking information—in one instance where former Trump campaign manager Paul Manafort questioned the validity of the statistics, and in another where he suggested the FBI was biased to the Clinton campaign. They found some evidence that these reduced the effect of the debunk on Trump supporters.

They also asked respondents about how accurate they thought the crime statistics were, and how fair and unbiased or biased the article was. Despite their beliefs becoming more accurate, Trump supporters were more likely to view the article as less accurate and fair when there was a debunk. They were also less likely to view the statistics as accurate when they were used in a debunk, especially when the experiment said they had been questioned by a Trump staffer.

Finally, both studies also tested to see if the debunk affected citizens' attitudes towards Trump and found it didn't. This is arguably unsurprising. As the authors comment, the study only looks at the effect of a single debunk. Further research is needed to see what the effect is on attitudes when individuals see multiple debunks of claims by the same politician, and when shown debunks of claims made by opposition politicians.

The first study was tested on around 3,000 Mechanical Turk participants, and a further 1,200 from a nationally representative sample. No significant differences were found in the results—with both showing a significant decrease in misperceptions among both Clinton and Trump supporters. The second study was conducted with 1,500 MTurk participants.

“Processing political misinformation: comprehending the Trump phenomenon”

Briony Swire, Adam Berinsky, Stephan Lewandowsky, and Ullrich Ecker (published in *Royal Society Open Science*, 2017)

This study, a collaboration between Briony Swire, Adam Berinsky, Stephan Lewandowsky and Ullrich Ecker, provides further evidence of belief change after debunks. They tested reactions to statements made by Donald Trump in the run-up to the 2016 US Presidential election.

Around 1,800 US participants from MTurk took part in a first study in November 2015, and around 1,000 US participants from a national online sample took part in a second study in July 2016.

They were primarily focused on how the source of the claim influences the effectiveness of debunks. While we will not go into these findings in detail, overall they found “no evidence for a worldview backfire effect in either experiment”.

Participants were shown statements made by Donald Trump on the campaign trail, half of which were accurate and the other half inaccurate. Debunks/affirmations to the claims were then presented, with a statement about the accuracy of the claim (e.g. “This is false” or “This is true”) and an explanation of this assessment with an explicit reference to a reputable non-partisan source.

Unlike the other studies explored here, participants were asked how accurate they thought the statement was before they saw the debunk/affirmation, then given a reminder of their original evaluation when they were asked again, either straight away or after a week, about their belief in the statement. Despite this they found both Democrats and Republicans showed a “substantial amount” of belief change—although with some decline after a week.

They found no change in participants' voting preferences after the debunks.

“Does truth matter to voters? The effects of correcting political misinformation in an Australian sample”

Michael Aird, Ullrich Ecker, Briony Swire, Adam Berinsky and Stephan Lewandowsky (published in Royal Society Open Science, 2018)

Building on the results of the previous research, this study sought further evidence of whether debunks change factual beliefs. Where the previous study used four accurate and four inaccurate claims, this study looked at whether debunks of disproportionate numbers of inaccurate and accurate claims by a politician affect support for that politician.

It has a much smaller sample than the previous studies discussed, with 370 Australian participants, including around 80 undergraduate students and the rest recruited online. But it provides some useful preliminary evidence for possible responses to debunks with a non-US sample.

Participants were shown short conclusions assessing the accuracy of statements made by politicians from opposing parties: Bill Shorten from the Labor party, and Malcolm Turnbull from the Liberal party.

They found debunks of inaccurate claims reduced beliefs to equally low levels across “left-wing” and “right-wing” participants, and evaluations of accurate claims increased the accuracy of beliefs across left-wing and right-wing participants. They also found debunks of myths actually reduced beliefs more strongly if the myths came from a favoured politician, since participants were more likely to believe those myths in the first place and therefore their beliefs had “further to fall” (this was mainly among left-wing participants).

The authors said it's possible these results were affected by the claims covered in the study—which they said were unlikely to have challenged strong beliefs, and which may have differed in importance between the two politicians.

When altering whether participants were shown an equal number of accurate and inaccurate claims by a politician, or four false statements and only one true statement, they found the latter caused significant decline in reported feelings towards the politician being factchecked—at least within this experimental setting.

The study reports the findings of a follow-up study conducted in the USA, not published yet, which found Americans' feelings towards politicians barely shifted in the disproportionate condition—unlike the Australian sample where there was a significant reduction in beliefs. This highlights the importance of carrying out more research on this topic across different cultures and country contexts.

“Do people actually learn from fact-checking? Evidence from a long-term experiment during the 2014 campaign”

Brendan Nyhan and Jason Reifler (not peer reviewed)

The largest study of real-life factchecks is by Brendan Nyhan and Jason Reifler, who looked at the effect of factchecks produced by the US factcheckers PolitiFact in 2014.

In the study, which hasn't gone through peer review, they randomly exposed a representative panel of 1,000 Americans to either factchecks or unrelated press releases over a series of surveys during the 2014 mid-term election campaign. These included true, false and half-true Democrat and Republican claims.

They found factchecks significantly improved the accuracy of people's beliefs, with overall little difference by partisan beliefs. This effect was strongest for those with high political knowledge—with the proportion of correct answers increasing by 17 percentage points from those who saw no factchecks to those who saw the factchecks. For those with low political knowledge, these proportions increased by 11 percentage points.

Where the factchecks were very long, the study put together shorter versions—including the introductory text, the final conclusion and the truth rating. However, they provide the results of a smaller experiment which tested the full-length versions and found similar effects. (Those who saw the full-length versions were more likely to say there was too much detail.)

The authors say these findings are “relatively large” considering participants were asked about the accuracy of the claims a while after—participants completed their first survey in September 2014 and then were shown the factchecks in a series of three mini surveys between then and the final survey in November 2014.

More research is needed to understand what might be contributing to this. One possibility is this is due to participants reading the factchecks unusually carefully, since the survey tested if participants had read each factcheck and they had

to read the factcheck again if they answered a question about it incorrectly. While their study didn't compare factchecks to the types of debunks embedded within news articles, they also say that the findings could suggest the less ambiguous style of factchecking may be more effective.

Other factors may have influenced it too. They say the topics may have been less important than the controversial topics explored in previous experiments. The factchecks shown to participants also included claims from lesser-known state-level politicians rather than national-level politicians, so participants may have had fewer prior attitudes before reading the factcheck.

Lastly, the factchecks tested included a truth rating, which Full Fact doesn't use. We see our job as filling in the shades of grey when campaigners often talk in black and white. We don't know how much difference rating scales make—a [separate study](#), looking at the effect of truth scales using made-up factchecks of fictional claims and politicians, found scales made little difference for political claims, but we need more evidence on this topic.

Other studies

Studies in the UK

There has also been [one peer-reviewed study](#) on this broad topic conducted in the UK in 2013. This looks at the effect of showing people information on the unemployment rate and the growth rate on beliefs about recent economic performance. They found providing this information generally made individuals' beliefs about recent economic performance more accurate, regardless of their prior beliefs.

Studies looking at broader issues

This briefing has focused specifically on evidence relating to the existence or not of backfire. There is a whole range of further research still to explore on factors that improve the receptivity of factchecks—for example, studies [have suggested](#) that replacing inaccurate information with new information is better than simply stating something is incorrect, and that the formats of factchecks can make them more effective—for example, [videos](#) and [graphs](#) have been found to be more effective than text.

What does this all mean for factcheckers?

This review suggests factchecking does help to inform citizens and backfire effects are rare not the norm.

We still need more evidence to understand how factchecking content can be most effective. A [review](#) of psychological studies testing out methods to counter misinformation concluded in 2017 that “Mounting evidence suggests that the process of correcting misinformation is complex and remains incompletely understood”. That is especially the case outside the United States, which is where the majority of these studies have been conducted.

The studies discussed here have shown that political beliefs can affect the extent to which debunking can be successful. We need more evidence to understand this in the United Kingdom and in other countries outside the US. Evidence on the effect of factchecks is promising, and needs exploring further as to how we can make factchecks particularly effective.

‘We still need more evidence to understand how factchecking content can be most effective.’

One of the areas we are also interested in exploring more is under what circumstances someone with deeply entrenched inaccurate factual beliefs might be persuaded to change them. For example, what difference does it make if the person or organisation putting out the message is trusted by the individual—and how might factchecking organisations gain that trust? The paper by Briony Swire and others is one of a number of papers building evidence on the former of these questions.

Full Fact’s work aims to empower the public by holding those in power to account for the information they use—to help prevent inaccurate claims from being made and from being spread before they become widespread misperceptions. We do this by seeking corrections from politicians and journalists to stop the spread of claims, and by seeking systemic changes that improve the accuracy of public debate. We are working to gather more evidence on the effectiveness of this in the coming years.

Full Fact
2 Carlton Gardens
London
SW1Y 5AA

team@fullfact.org
+44 (0)20 3397 5140
@FullFact
fullfact.org

