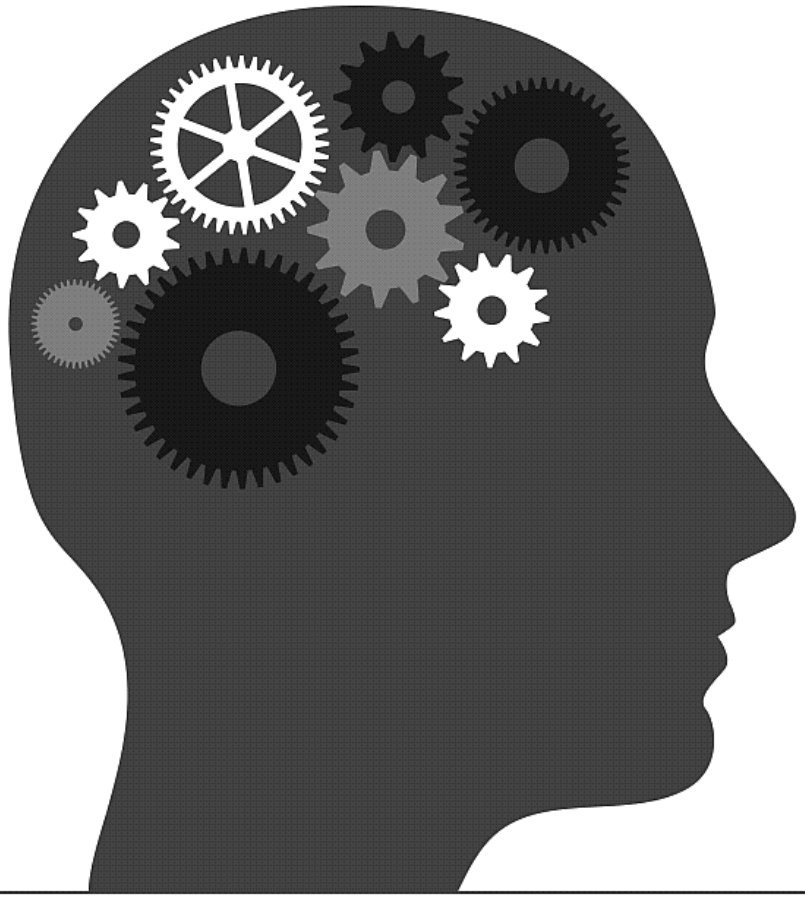
the DEBUNKING handbook



John Cook

Stephan Lewandowsky

Written by:

John Cook, Global Change Institute, University of Queensland Stephan Lewandowsky, School of Psychology, University of Western Australia

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

gciQ
Global Change Institute

THE UNIVERSITY OF WESTERN AUSTRALIA

Achieve International Excellence

First published in November 2011.
Version 2 published on 23 January 2012.
For more information, visit http://sks.to/debunk

Cite as:
Cook, J., Lewandowsky, S. (2011), The Debunking Handbook. St. Lucia, Australia: University of Queensland. November 5. ISBN 978-0-646-56812-6. [http://sks.to/debunk]

Debunking myths is problematic. Unless great care is taken, any effort to debunk misinformation can inadvertently reinforce the very myths one seeks to correct. To avoid these “backfire effects”, an effective debunking requires three major elements. First, the refutation must focus on core facts rather than the myth to avoid the misinformation becoming more familiar. Second, any mention of a myth should be preceded by

explicit warnings to notify the reader that the upcoming information is false. Finally, the refutation should include an alternative explanation that accounts for important qualities in the original misinformation.

Debunking the first myth about debunking

It's self-evident that democratic societies should base their decisions on accurate information. On many issues, however, misinformation can become entrenched in parts of the community, particularly when vested interests are involved.[1,2] Reducing the influence of misinformation is a difficult and complex challenge.

A common misconception about myths is the notion that removing its influence is as simple as packing more information into people's heads. This approach assumes that public misperceptions are due to a lack of knowledge and that the solution is more information - in science communication, it's known as the "information deficit model". But

that model is wrong: people don't process information as simply as a hard drive downloading data.

Refuting misinformation involves dealing with complex cognitive processes. To successfully impart knowledge, communicators need to understand how people process information, how they modify their existing knowledge and how worldviews affect their ability to think rationally. It's not just what people think that matters, but how they think.

First, let's be clear about what we mean by the label "misinformation" - we use it to refer to any information that people have acquired that turns out to be incorrect, irrespective of why and how that information was acquired in the first place. We are concerned with the cognitive processes that govern how people process corrections to information they have already acquired - if you find out that something you believe is wrong, how do you update your knowledge and memory?

Once people receive misinformation, it's quite difficult to remove its influence. This was demonstrated in a 1994 experiment where people were exposed to misinformation about a fictitious warehouse fire, then given a correction clarifying the parts of the story that were incorrect.[3] Despite remembering and accepting the correction, people still showed a lingering effect, referring to the misinformation when answering questions about the story.

Is it possible to completely eliminate the influence of misinformation? The evidence indicates that no matter how vigorously and repeatedly we correct the misinformation, for example by repeating the correction over and over again, the influence remains detectable.[4] The old saying got it right - mud sticks.

There is also an added complication. Not only is misinformation difficult to remove, debunking a myth can actually strengthen it in people's minds. Several different "backfire effects" have been observed, arising

from making myths more familiar,[5,6] from providing too many arguments,[7] or from providing evidence that threatens one's worldview.[8]

The last thing you want to do when debunking misinformation is blunder in and make matters worse. So this handbook has a specific focus - providing practical tips to effectively debunk misinformation and avoid the various backfire effects. To achieve this, an understanding of the relevant cognitive processes is necessary. We explain some of the interesting psychological research in this area and finish with an example of an effective rebuttal of a common myth.

1

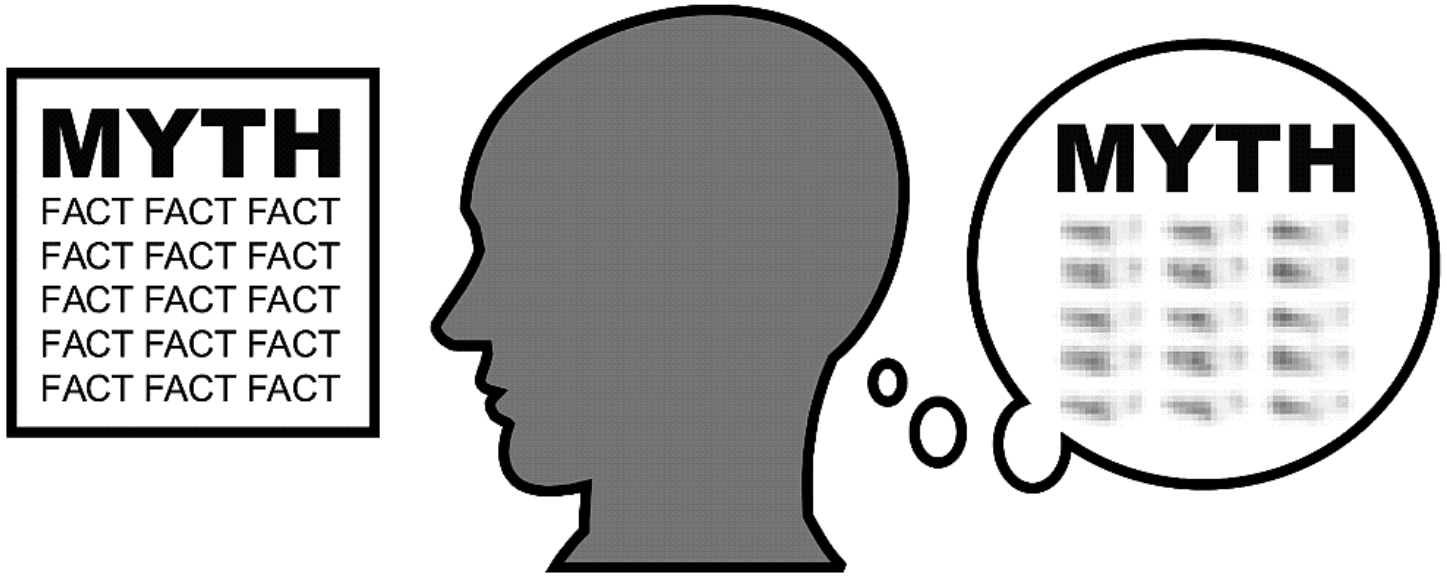
The Familiarity Backfire Effect

To debunk a myth, you often have to mention it - otherwise, how will people know what you're talking about? However, this makes people more familiar

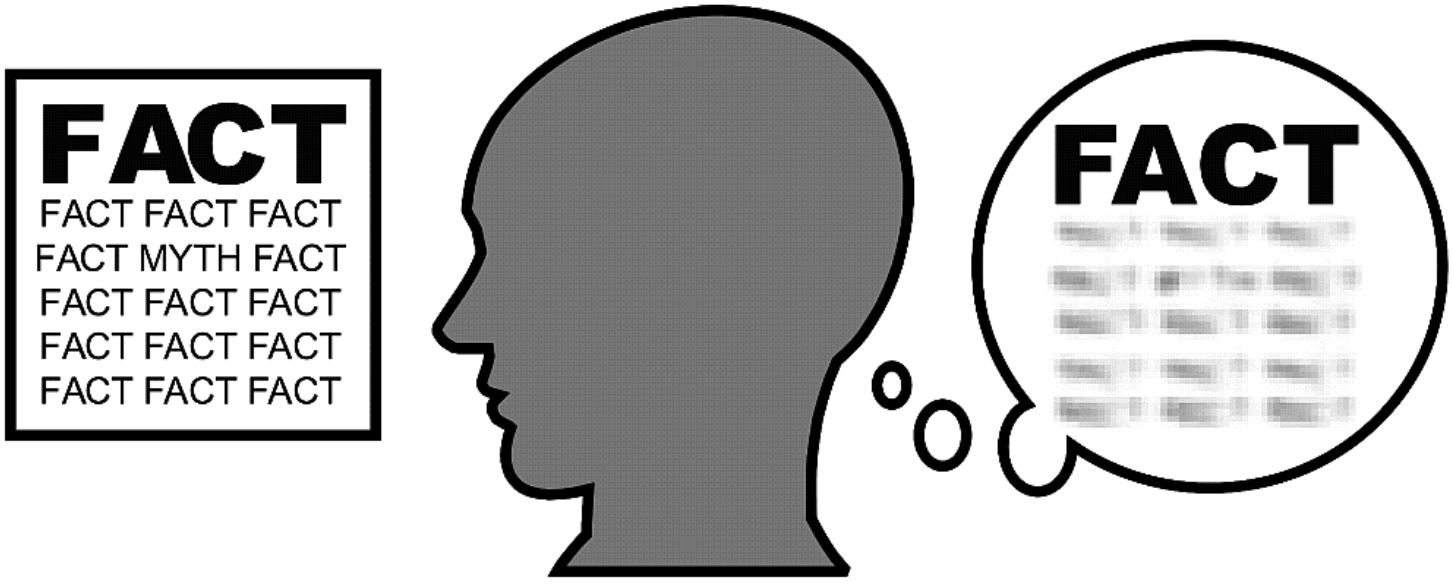
with the myth and hence more likely to accept it as true. Does this mean debunking a myth might actually reinforce it in people's minds?

To test for this backfire effect, people were shown a flyer that debunked common myths about flu vaccines.[5] Afterwards, they were asked to separate the myths from the facts. When asked immediately after reading the flyer, people successfully identified the myths. However, when queried 30 minutes after reading the flyer, some people actually scored worse after reading the flyer. The debunking reinforced the myths.

Hence the backfire effect is real. The driving force is the fact that familiarity increases the chances of accepting information as true. Immediately after reading the flyer, people remembered the details that debunked the myth and successfully identified the myths. As time passed, however, the memory of the details faded and all people remembered was the myth without the "tag" that identified it as false. This effect is particularly strong in older adults because their memories are more vulnerable to forgetting of details.



How does one avoid causing the Familiarity Backfire Effect? Ideally, avoid mentioning the myth altogether while correcting it. When seeking to counter misinformation, the best approach is to focus on the facts you wish to communicate.



Not mentioning the myth is sometimes not a practical option. In this case, the emphasis of the

debunking should be on the facts. The often-seen technique of headlining your debunking with the myth in big, bold letters is the last thing you want to do. Instead, communicate your core fact in the headline. Your debunking should begin with emphasis on the facts, not the myth. Your goal is to increase people's familiarity with the facts.

Example of debunking a climate myth

Sun and climate are going in opposite directions

Core fact emphasised in headline

Over the last few decades of global warming, the sun has shown a slight cooling trend. Sun and climate are going in opposite directions. This has led a number of scientists to independently conclude that the sun cannot be the cause of recent global warming.

Core facts reinforced in initial text

One of the most common and persistent climate myths is that the sun is the cause of global warming.

Myth

This myth cherry picks the data - showing past periods when sun and climate move together but ignoring the last few decades when the two diverge.

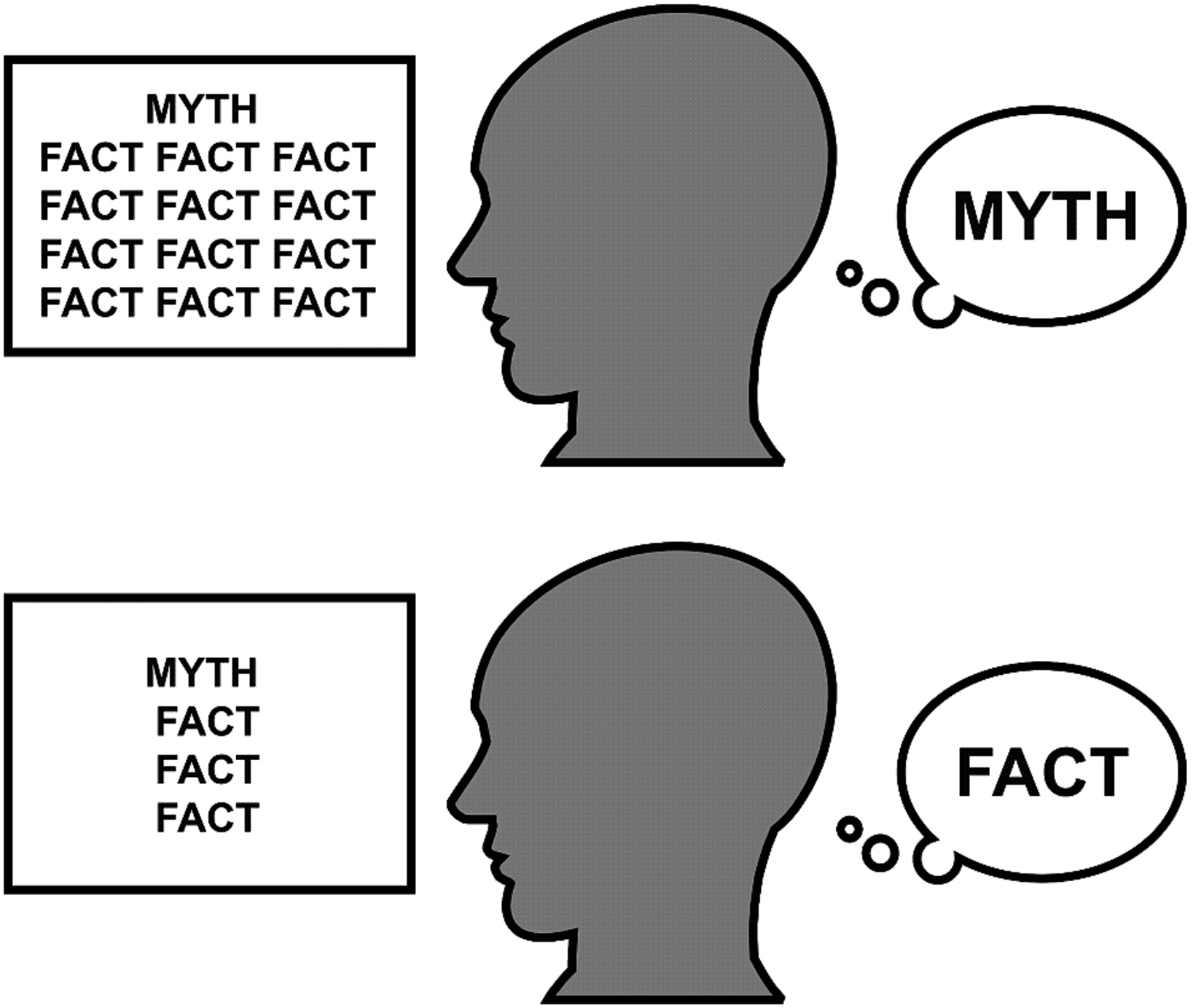
Explaining how the myth misleads (alternative explanation, see Page 5)

The Overkill Backfire Effect

One principle that science communicators often fail to follow is making their content easy to process. That means easy to read, easy to understand and succinct. Information that is easy to process is more likely to be accepted as true.[7] Merely enhancing the colour contrast of a printed font so it is easier to read, for example, can increase people's acceptance of the truth of a statement.[9]

Common wisdom is that the more counter-arguments you provide, the more successful

you'll be in debunking a myth. It turns out that the opposite can be true. When it comes to refuting misinformation, less can be more. Generating three arguments, for example, can be more successful in reducing misperceptions than generating twelve arguments, which can end up reinforcing the initial misperception.[7]



The Overkill Backfire Effect occurs because processing many arguments takes more effort than just considering a few. A simple myth is more cognitively attractive than an over-complicated correction.

The solution is to keep your content lean, mean and easy to read. Making your content easy to process means using every tool available. Use simple language, short sentences, subheadings and paragraphs. Avoid dramatic language and derogatory comments that alienate people. Stick to the facts.

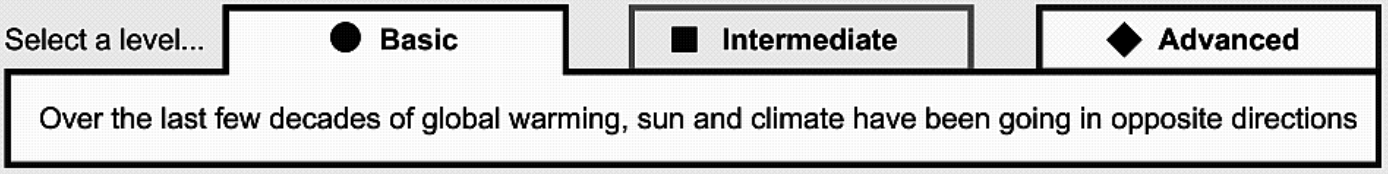
End on a strong and simple message that people will remember and tweet to their friends, such as “97 out of 100 climate scientists agree that humans

are causing global warning"; or "Study shows that MMR vaccines are safe." Use graphics wherever possible to illustrate your points.

Scientists have long followed the principles of the Information Deficit Model, which suggests that people hold erroneous views because they don't have all the information. But too much information can backfire. Adhere instead to the KISS principle: Keep It Simple, Stupid!

Having your cake and eating it too

Writing at a simple level runs the risk of sacrificing the complexities and nuances of the concepts you wish to communicate. At Skeptical Science, we gain the best of both worlds by publishing rebuttals at several levels. Basic versions are written using short, plain English text and simplified graphics. More technical Intermediate and Advanced versions are also available with more technical language and detailed explanations. The icons used on ski runs are used as visual cues to denote the technical level of each rebuttal.



3

The Worldview Backfire Effect

The third and arguably most potent backfire effect occurs with topics that tie in with people's worldviews and sense of cultural identity. Several cognitive processes can cause people to unconsciously process information in a biased way. For those who are strongly fixed in their views, being confronted with counter-arguments can cause their views to be strengthened.

One cognitive process that contributes to this effect is Confirmation Bias, where people selectively seek out information that bolsters their view. In one experiment, people were offered information on hot-button issues like gun control or affirmative action. Each parcel of information was labelled by its source, clearly indicating whether the information would be pro or con (e.g., the National Rifle Association vs. Citizens Against Handguns). Although instructed to be even-handed, people opted for sources that matched their pre-existing views. The study found that even

when people are presented with a balanced set of facts, they reinforce their pre-existing views by gravitating towards information they already agree with. The polarisation was greatest among those with strongly held views.[10]

What happens when you remove that element of choice and present someone with arguments that run counter to their worldview? In this case, the cognitive process that comes to the fore is Disconfirmation Bias, the flipside of Confirmation Bias. This is where people spend significantly more time and thought actively arguing against opposing arguments.[8]

This was demonstrated when Republicans who believed Saddam Hussein was linked to the 9/11 terrorist attacks were provided with evidence that there was no link between the two, including a direct quote from President George Bush.[11] Only 2% of participants changed their mind (although interestingly, 14% denied that they believed the link in the first place). The vast majority clung to the link between Iraq and 9/11, employing a range

of arguments to brush aside the evidence. The most common response was attitude bolstering - bringing supporting facts to mind while ignoring any contrary facts. The process of bringing to the fore supporting facts resulted in strengthening people's erroneous belief.

If facts cannot dissuade a person from their pre-existing beliefs - and can sometimes make things worse - how can we possibly reduce the effect of misinformation? There are two sources of hope.

First, the Worldview Backfire Effect is strongest among those already fixed in their views. You therefore stand a greater chance of correcting misinformation among those not as firmly decided about hot-button issues. This suggests that outreaches should be directed towards the undecided majority rather than the unswayable minority.

Second, messages can be presented in ways that reduce the usual psychological resistance. For example,

when worldview-threatening messages are coupled with so-called self-affirmation, people become more balanced in considering pro and con information.[12,13]

Self-affirmation can be achieved by asking people to write a few sentences about a time when they felt good about themselves because they acted on a value that was important to them. People then become more receptive to messages that otherwise might threaten their worldviews, compared to people who received no self-affirmation. Interestingly, the "self-affirmation effect" is strongest among those whose ideology was central to their sense of self-worth.

Another way in which information can be made more acceptable is by "framing" it in a way that is less threatening to a person's worldview. For example, Republicans are far more likely to accept an otherwise identical charge as a "carbon offset" than as a "tax", whereas the wording has little effect on Democrats or Independents—because

their values are not challenged by the word “tax”.[14]

Self-affirmation and framing aren’t about manipulating people. They give the facts a fighting chance.

Filling the gap with an alternative explanation

Assuming you successfully negotiate the various backfire effects, what is the most effective way to debunk a myth? The challenge is that once misinformation gets into a person’s mind, it’s very difficult to remove. This is the case even when people remember and accept a correction.

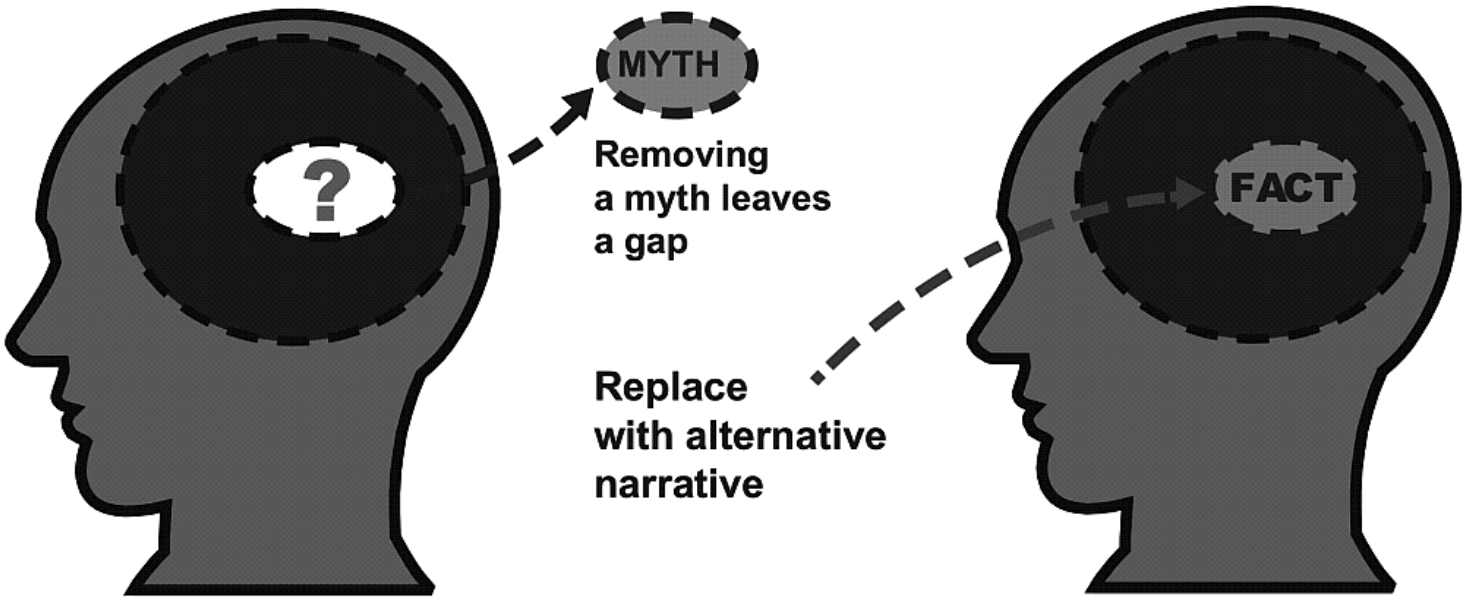
This was demonstrated in an experiment in which people read a fictitious account of a warehouse fire.[15,16,3] Mention was made of paint and gas cans along with explosions. Later in the story, it was clarified that paint and cans were not present at the fire. Even when people

remembered and accepted this correction, they still cited the paint or cans when asked questions about the fire. When asked, "Why do you think there was so much smoke?", people routinely invoked the oil paint despite having just acknowledged it as not being present.

When people hear misinformation, they build a mental model, with the myth providing an explanation. When the myth is debunked, a gap is left in their mental model. To deal with this dilemma, people prefer an incorrect model over an incomplete model. In the absence of a better explanation, they opt for the wrong explanation.[17]

In the warehouse fire experiment, when an alternative explanation involving lighter fluid and accelerant was provided, people were less likely to cite the paint and gas cans when queried about the fire. The most effective way to reduce the effect of misinformation is to provide an

alternative explanation for the events covered by the misinformation.



This strategy is illustrated particularly clearly in fictional murder trials. Accusing an alternative suspect greatly reduced the number of guilty verdicts from participants who acted as jurors, compared to defences that merely explained why the defendant wasn't guilty.[18]

For the alternative to be accepted, it must be plausible and explain all observed features of the event.[19,15] When you debunk a myth, you create a gap in the person's mind. To be effective, your debunking must fill that gap.

One gap that may require filling is explaining why the myth is wrong. This can be achieved by exposing the rhetorical

techniques used to misinform. A handy reference of techniques common to many movements that deny a scientific consensus is found in Denialism: what is it and how should scientists respond?[20] The techniques include cherry picking, conspiracy theories and fake experts.

Another alternative narrative might be to explain why the misinformer promoted the myth. Arousing suspicion of the source of misinformation has been shown to further reduce the influence of misinformation.[21,22]

Another key element to effective rebuttal is using an explicit warning ("watch out, you might be misled") before mentioning the myth. Experimentation with different rebuttal structures found the most effective combination included an alternative explanation and an explicit warning.[17]

Graphics are also an important part of the

debunker's toolbox and are significantly more effective than text in reducing misconceptions. When people read a refutation that conflicts with their beliefs, they seize on ambiguities to construct an alternative interpretation. Graphics provide more clarity and less opportunity for misinterpretation. When self-identified Republicans were surveyed about their global warming beliefs, a significantly greater number accepted global warming when shown a graph of temperature trends compared to those who were given a written description.[13]

Another survey found that when shown data points representing surface temperature, people correctly judged a warming trend irrespective of their views towards global warming.[23] If your content can be expressed visually, always opt for a graphic in your debunking.

Anatomy of an effective debunking

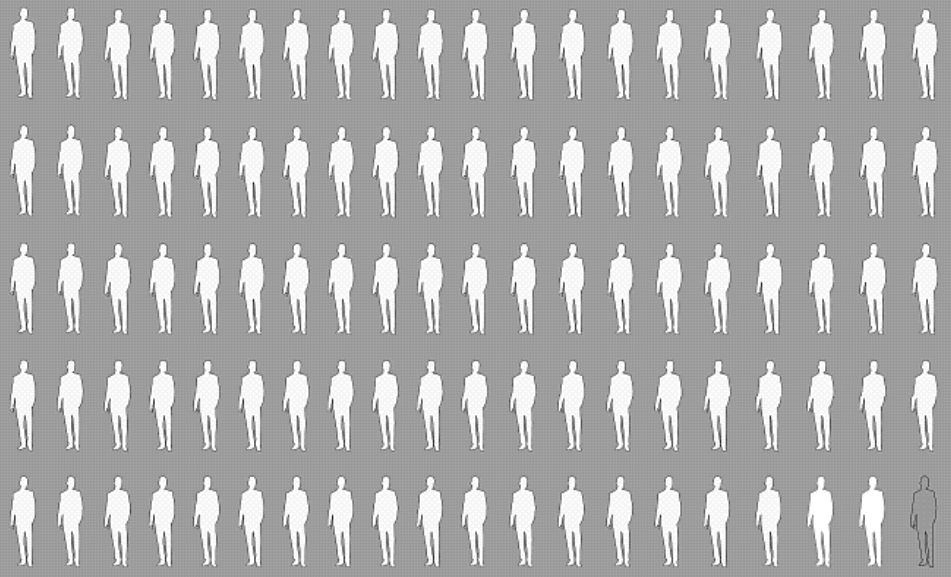
Bringing all the different threads together, an effective debunking requires:

- **Core facts**—a refutation should emphasise the facts, not the myth. Present only key facts to avoid an Overkill Backfire Effect;
- **Explicit warnings**—before any mention of a myth, text or visual cues should warn that the upcoming information is false;
- **Alternative explanation**—any gaps left by the debunking need to be filled. This may be achieved by providing an alternative causal explanation for why the myth is wrong and, optionally, why the misinformers promoted the myth in the first place;
- **Graphics** – core facts should be displayed graphically if possible.

The following example debunks the myth that there is no scientific consensus about man-made global warming, because 31,000 scientists signed a petition stating there is no evidence that human activity can disrupt climate.

97 out of 100 climate experts agree humans are causing global warming.

Several independent surveys find 97% of climate scientists who are actively publishing peer-reviewed climate research agree that humans are causing global warming.



On top of this overwhelming consensus, National Academies of Science from all over the world also endorse the consensus view of human caused global warming, as expressed by the Intergovernmental Panel on Climate Change (IPCC).

However, movements that deny a scientific consensus have always sought to cast doubt on the fact that a consensus exists. One technique is the use of fake experts, citing scientists who have little to no expertise in the particular field of science.

For example, the OISM Petition Project claims 31,000 scientists disagree with the scientific consensus on global warming.

However, around 99.9% of the scientists listed in the Petition Project are not climate scientists. The petition is open to anyone with a Bachelor of Science or higher and includes medical doctors, mechanical engineers and computer scientists.

Core fact communicated in headline

Core fact reinforced in opening paragraph, fleshed out with additional details.

Core fact reinforced with infographic

Explicit warning cueing reader that misinformation is coming and indicating the nature of the misinformation.

The myth

The gap created by this debunking is how can there be a consensus if 31,000 scientists dissent? This gap is filled by explaining that almost all the 31,000 scientists are not climate scientists.

Bibliography

1. Jacques, P. J., & Dunlap, R. E. (2008). The organisation of denial: Conservative think tanks and environmental skepticism. *Environmental Politics*, 17, 349-385.
2. Oreskes, N., & Conway, E. M. (2010). Merchants of doubt. Bloomsbury Publishing.
3. Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When discredited information in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (6), 1420-1436.
4. Ecker, U. K., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, 18, 570-578.
5. Skurnik, I., Yoon, C., Park, D., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, 31, 713-724.
6. Weaver, K., Garcia, S. M., Schwarz, N., & Miller, D. T. (2007). Inferring the popularity of an opinion from its familiarity: A repetitive voice sounds like a chorus. *Journal of Personality and Social Psychology*, 92, 821-833.
7. Schwarz, N., Sanna, L., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight:Implications for debiasing and public information campaigns. *Advances in Experimental*

Social Psychology, 39, 127-161.

8. Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, 32, 303-330.
9. Reber, R., Schwarz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth, *Consciousness and Cognition*, 8, 338-3426.
10. Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50, 755–69.
11. Prasad, M., Perrin, A. J., Bezila, K., Hoffman, S. G., Kindleberger, K., Manturuk, K., et al. (2009). "There Must Be a Reason": Osama, Saddam, and Inferred Justification. *Sociological Inquiry*, 79, 142-162.
12. Cohen, G. L., Sherman, D. K., Bastardi, A., Hsu, L., & McGoey, M. (2007). Bridging the Partisan Divide: Self-Affirmation Reduces Ideological Closed-Mindedness and Inflexibility in Negotiation. *Personality & Soc. Psych.*, 93, 415-430.
13. Nyhan, B., & Reifler, J. (2011). Opening the Political Mind? The effects of self-affirmation and graphical information on factual misperceptions. In press.
14. Hardisty, D. J., Johnson, E. J. & Weber, E. U. (2010). A Dirty Word or a Dirty World?: Attribute Framing, Political Affiliation, and Query Theory, *Psychological Science*, 21, 86-92
15. Seifert, C. M. (2002). The continued influence of misinformation in memory: What makes a correction effective? *The Psychology of Learning and Motivation*,

41, 265-292.

16. Wilkes, A. L.; Leatherbarrow, M. (1988). Editing episodic memory following the identification of error, *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 40A, 361-387.
17. Ecker, U. K., Lewandowsky, S., & Tang, D. T. (2011). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38, 1087-1100.
18. Tenney, E. R., Cleary, H. M., & Spellman, B. A. (2009). Unpacking the doubt in "Beyond a reasonable doubt:" Plausible alternative stories increase not guilty verdicts. *Basic and Applied Social Psychology*, 31, 1-8.
19. Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, 35, 2019-2032.
20. Diethelm, P., & McKee, M. (2009). Denialism: what is it and how should scientists respond? *European Journal of Public Health*, 19, 2-4.
21. Lewandowsky, S., Stritzke, W. G., Oberauer, K., & Morales, M. (2005). Memory for fact, fiction and misinformation: The Iraq War 2003. *Psychological Science*, 16, 190-195.
22. Lewandowsky, S., & Stritzke, W. G. K., Oberauer, K., & Morales, M. (2009). Misinformation and the 'War on Terror': When memory turns fiction into fact. In W. G. K. Stritzke, S. Lewandowsky, D. Denemark, J.

Clare, & F. Morgan (Eds.), *Terrorism and torture: An interdisciplinary perspective* (pp. 179-203). Cambridge, UK: Cambridge University Press.

23. Lewandowsky, S. (2011). Popular consensus: Climate change set to continue. *Psychological Science*, 22, 460-463.