

SANDER VAN DER LINDEN

A PSYCHOLOGICAL VACCINE AGAINST MISINFORMATION

This article is an edited excerpt from Sander's book, 'Foolproof: Why We Fall for Misinformation and How to Build Immunity'.

Our team (the University of Cambridge's Social Decision-Making Lab) headed off to New York to collaborate with Google Jigsaw, Google's technology incubator. We met up with Beth Goldberg, Director of Research & Development and together with our colleague Stephan Lewandowsky, we started to have regular conversations with Google about how to inoculate people against extremism on social media.

Just like a vaccine introduces your body to a weakened version of a harmful virus, it turns out that the mind can be inoculated against harmful misinformation by exposing people to—and persuasively refuting—weakened doses of misinformation. The process of psychological inoculation works by: (a) forewarning people of an impending manipulation attempt, and (b) arming people in advance with the arguments and cognitive tools they need to counter-argue and resist exposure to persuasive misinformation (known as a 'prebunk'). Or, as one BBC journalist writing about our research put it: "Like Han Solo, you shoot first."

66 The mind can be inoculated against harmful misinformation by exposing people to—and persuasively refuting—weakened doses of misinformation.

Beth was very interested in scaling our inoculation approach en masse via YouTube (owned by Google). One of the common techniques Beth identified is the use of 'false dichotomies'. A false dichotomy is a manipulation technique designed to make you think that you only have two options to choose between, while in reality, there are many more. Because YouTube doesn't really deal in headlines or social media posts, the issue here is that these more subtle rhetorical techniques are often being used — persuasively — by political Guru's in YouTube videos. From rants that spread fake news about Covid-19 and climate change to attempts to recruit people into QAnon and ISIS.

For example, one ISIS recruitment video explicitly aimed at Western Muslims was titled, "There is no life without Jihad" — a clear example of a false dichotomy: either you join 'jihad' or you cannot lead a meaningful life.

To produce the vaccine, we needed to synthesise weakened doses, so we started to make our own animated videos. The videos follow the inoculation format closely and start with an immediate warning that you (the viewer) might be targeted with an attempt to manipulate your opinion. We then show people how to spot and refute misinformation that explicitly makes use of these techniques by exposing them to a series of weakened examples (the microdose) so that people can easily identify and resist them in the future.

For example, in the video that inoculated people against the 'false dichotomy' technique, we pulled material from Star Wars III – Revenge of the Sith. We show the climactic confrontation between Anakin Skywalker, soon to become Darth Vader, and his mentor, Obi-Wan Kenobi. Obi-Wan says, "My allegiance is to the Republic, to Democracy!" to which Skywalker replies: "If you're not with me, then you're my enemy." This is clearly a false dichotomy. We explain to the viewer that Obi-Wan is simply trying to prevent Anakin from joining the dark side; just because he disagrees with Anakin doesn't automatically make them enemies. Obi-Wan points out the fallacy in his reply: "Only a Sith deals in absolutes."

We ran several large randomised controlled trials where we either exposed people to one of our short videos or a control video about 'freezer burn'. We then asked people how manipulative they found a series of arguments and how willing they were to share them with others. An example of the false dichotomy quiz would be the following post: "Why give illegal immigrants access to social services? Why should we help illegals when we should be helping homeless Americans instead?"

In the experiment, we asked people to rate many such misleading items in the hope that their ability to discern manipulative from non-manipulative content would improve.

This is exactly what we found. Unlike the control group, the inoculated groups became much better at identifying which posts contained a specific manipulation strategy and were

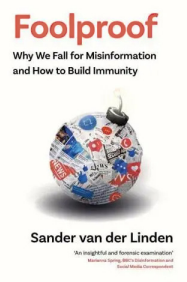
subsequently less likely to want to share this type of content with others in their social network. Beth's idea was that we could implement and scale these videos on YouTube by inserting them in the 'non-skippable ad space' (you know, when you're trying to watch a video on YouTube and you get stuck watching an annoying ad that you can't skip? That's where our inoculation video would be placed).

We leveraged YouTube's ad platform to upload and target millions of U.S. users who are known to watch political content with either our inoculation or the control videos. Beth then got YouTube to agree to allow us to customise their 'brand lift' survey (which usually polls people on whether they recognise a brand) for a scientific experiment. Within twenty-four hours, they would be presented with a quiz in the ad space evaluating their ability to spot the misinformation technique they had been inoculated against (for example, the use of false dichotomies, emotional manipulation, scapegoating etc). We were able to reach about 5 million 'impressions' (views) with a single campaign. After watching the 90-second inoculation video, we boosted people's ability to spot misleading content by about 5–10 per cent. That might not seem much at the individual level, but this is in a realistic setting for a single dose of a short video clip that can be scaled across potentially hundreds of millions of people.

66 Can enough people be inoculated to achieve psychological herd immunity?

Of course, the work is far from finished. There is much to explore about how applying the principles of psychological inoculation can empower individuals and policymakers to address societal challenges effectively. For example, we discovered that the vaccine wears off over time so campaigns ideally need to feature "booster shots" and feedback to enhance longer-term learning. In Google's latest prebunking campaigns—which reached the majority of social media users in Poland, Czechia, and Slovakia—they uncovered substantial cultural variation in the effectiveness of inoculation, suggesting that careful local tailoring is important. The ultimate question, of course, is whether enough people can be inoculated to achieve psychological herd immunity. Only then will misinformation no longer have a chance to spread. This will likely necessitate integrating inoculation against misinformation in our national educational curricula.

This article is an edited excerpt from Foolproof: Why We Fall for Misinformation and How to Build Immunity. Copyright © 2023 by Sander van der Linden. Published by 4th Estate. Used with permission of the publisher, 4th Estate. All rights reserved. Sander van der Linden is Professor of Social Psychology in Society and Director of the Cambridge Social Decision-Making Lab in the Department of Psychology at the University of Cambridge.

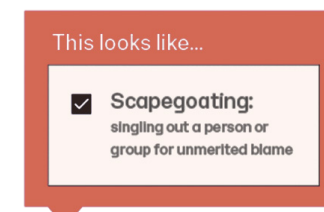


1. Emotional warning



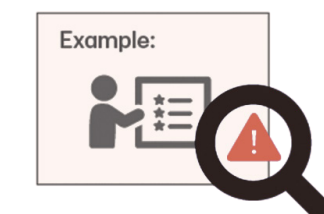
Users are alerted that there are impending "attacks" to manipulate them.

2. Refute the attack



Users are equipped to spot and refute a manipulative message.

3. Micro-dose



Users see example(s) of manipulative message to identify it in the future.