

# microbiologist



## > INSIDE

### Assuring the safety of vaccines

The plastisphere

Mandatory childhood vaccination

Latin anyone?

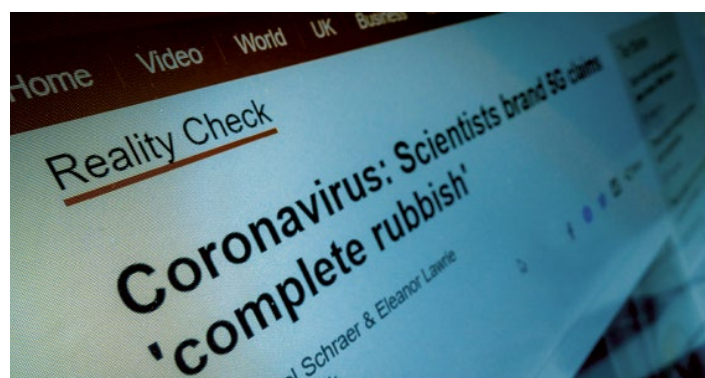
# A psychological ‘vaccine’ against fake news

**Rakoen Maertens and Sander van der Linden**

*University of Cambridge, UK*

As governments across the world are rolling out COVID-19 vaccination policies, they are not only facing challenges around vaccine logistics but are also fighting an uphill battle against the onslaught of ‘fake news’. In fact, commentators have argued that misinformation might be the most contagious thing about the virus. The consequences of belief in fake news can be dangerous, whether it concerns fake cures that lead people to ingest harmful substances or vandalism of the 5G network infrastructure because of false assertions that there is a link between radiation and COVID-19. In our recent research, we found clear evidence that belief in common misinformation about the virus is strongly associated with reduced intentions to get vaccinated and a lower likelihood to recommend the vaccine to others. We therefore need an effective way to counteract and tackle the spread of misinformation in society.

Despite years of research in cognitive and behavioural science on how to curb the impact of misinformation, a magic bullet solution for the problem has not been found. The classical method is known as *debunking* and entails issuing a correction after people have already been exposed to a falsehood. This method leaves much to be desired because it often requires repeating the myth in an effort to debunk it. We know that the more someone is exposed to a myth, the more familiar the myth feels, and the more weight the brain will give the claim in memory retrieval: an effect known as the *illusory truth effect*. In other words, when fact-checking misinformation, there is always a risk that the myth becomes more readily available in one’s memory, potentially leading to a negative effect that reinforces the myth while people easily forget about the correction. Sometimes the benefits of debunking outweigh the adverse effects of potentially increasing familiarity through repetition but even when debunking is effective, it does not solve the problem: the spread of



misinformation—especially on social media—outpaces the rate at which we can fact-check and so people are repeatedly fooled by manipulative information. In other words, debunking is never agile enough: even if it works, the damage is often already done, and the myth continues to spread.

This leads to the natural question of whether we can prevent misinformation from taking root in the first place? In the 1960s, American psychologist William McGuire developed a framework known as *inoculation theory*, which closely follows the biomedical analogy. In brief, McGuire posited that—similar to a biomedical vaccine—the ‘cognitive immune system’ needs to become familiar with a weakened version of the ‘virus’ (the manipulation attempt) in order for it to develop ‘mental antibodies’. At the time, McGuire was thinking about people’s susceptibility to propaganda during the Cold War and whether it was possible to produce a ‘vaccine against brainwash’. Although he had some initial success with his experiments, showing that people can become more immune to persuasion when they are forewarned and exposed to a severely weakened version of the ‘persuasive attack’, he never quite tested his ideas in the context of propaganda, and the theory slowly faded into history.

Sixty years later, at the Cambridge Social Decision-Making Lab, we decided to renew focus on inoculation theory. One insight that McGuire may not have foreseen is that we can

now actually borrow models from epidemiology to study the spread of information pathogens. Misinformation is highly contagious as it can spread from one person to another without the need for physical contact. The vaccine analogy has, therefore, never been more apt.

To develop a scalable vaccine that provides long-term cognitive protection against a wide range of misinformation pathogens, we further developed the prophylactic framework of inoculation theory—otherwise known as '*prebunking*'. To illustrate how this approach works in a controlled laboratory environment, we started

by evaluating whether we could distil a sufficiently weakened dose out of a specific myth: misinformation about climate change. In particular, we used a screenshot of a real website that hosts a bogus petition allegedly signed by thousands of scientists claiming that global warming is a hoax. In the experiment, we forewarned people about the petition (without naming it) and provided them with a pre-emptive refutation (a weakened dose that contains the '*prebunk*'). For example, the forewarning message contained an explanation of the flaws and fallacies utilised in the misinformation (e.g. the use of fake experts, including false signatories such as

# PSSST....!







*Graffiti on a Manchester motorway bridge about COVID-19 being a hoax*

Charles Darwin and members of the Spice Girls). The warning and weakened dose are meant to trigger people's vigilance and attention (to start the production of mental antibodies), and the message offers people concrete ways to resist the misinformation. After people were inoculated, participants were exposed to a full dose of the misinformation. Our findings showed that while those who received a placebo treatment were negatively impacted by the misinformation, inoculated participants were substantially less likely to be fooled by it. This experiment was replicated three times and extended to show that even one week after the intervention, inoculated individuals were still protected against the misinformation attack.

After some initial success with an isolated issue in one context, we successfully developed the first vaccine of the second generation: an intervention to inoculate people against a broad range of misinformation tactics. Rather than trying to pre-empt every single myth, we deemed it more efficient to develop a broader-spectrum vaccination that targets the very building blocks of the misinformation virus itself (its nucleic acid). After a year's worth of investigation, we identified the techniques common to nearly all online misinformation; 'The Six Degrees of Manipulation': discrediting, appealing to emotion, group polarisation, impersonation (e.g. fake experts), conspiracy theories and trolling. An important second theoretical innovation was that we wanted to simulate a social media feed to increase the validity of our testing environment. This also allowed us to examine the notion of 'active inoculation' or the idea that instead of passively providing people with the facts beforehand, you let people generate their own intellectual antibodies in an interactive learning setting.

Accordingly, in collaboration with the Dutch media literacy organisation *DROG*, we designed the inoculation-based game *Bad News* ([getbadnews.com](http://getbadnews.com)). In the game, participants take on the role of misinformation producer

and aim to gain followers (whilst maintaining their credibility) by creating and sharing manipulative news headlines according to one of the six degrees of manipulation. During gameplay, players are forewarned about the dangers of fake news and are exposed to weakened doses of the six manipulation techniques in a controlled environment, often using humour. Importantly, these doses are strong enough to trigger people's motivation to learn how to protect themselves but not so strong as to actually dupe them (or infect people with the virus). Players typically participate in a quiz beforehand where they are asked to rate how reliable they find a large series of posts that are either credible news items or items that contain one of the key misinformation techniques. Notably, the test items are different from the training items used in the game. Both big data samples (of people who voluntarily play the game) as well as randomised controlled trials on the *Bad News* game have shown that it effectively reduces people's susceptibility to fake news headlines, that the game boosts people's confidence in their own ability to discern manipulative from credible news and reduces self-reported readiness to share fake news.

After these promising findings, we decided to further test and evaluate our intervention with the UK government. The government helped translate the game into 20 languages around the world, which enabled us to do additional large-scale cross-cultural replications with positive results. The game has now been played by over a million people around the world and is implemented in high school and university curricula in several countries.

Following the recent declaration by the World Health Organization (WHO) of a worldwide 'infodemic' and the success of our second-generation vaccine, we set out to develop a similar inoculation game to help protect people against the harmful impact of COVID-19 misinformation specifically. Together with the UK Cabinet Office and with support from the WHO and United Nations *Verified*

campaign, we launched *GoViral! The GoViral! Game* prebunks three common techniques used to spread misinformation about the coronavirus: fearmongering, the use of fake experts and conspiracy theories.

Although these advances are promising, we needed to know more about the long-term efficacy of the cognitive vaccine. Similar to how it is imperative to determine the long-term effectiveness of COVID-19 vaccines, research on inoculation-based interventions will need to investigate how to increase the interventions' long-term success. In general, the literature shows that the benefits of cognitive inoculation remain intact for up to two months, but that its effectiveness starts to dissipate after a few weeks. Importantly, research has found evidence for the potential of 'booster sessions'. Just like for the Pfizer vaccine—where a booster jab is needed three weeks after the initial injection—cognitive immunity can be prolonged with an additional 'booster shot'. In an experiment with *Bad News*, we found evidence for inoculation effect decay after two months when we excluded follow-ups, but full retention of the protective effects for at least three months after three booster sessions.

## FURTHER READING

Basol M, Roozenbeek J, van der Linden S. Good news about Bad News: gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of Cognition* 2020; 3(1), 1–9 <https://doi.org/10.5334/joc.91>

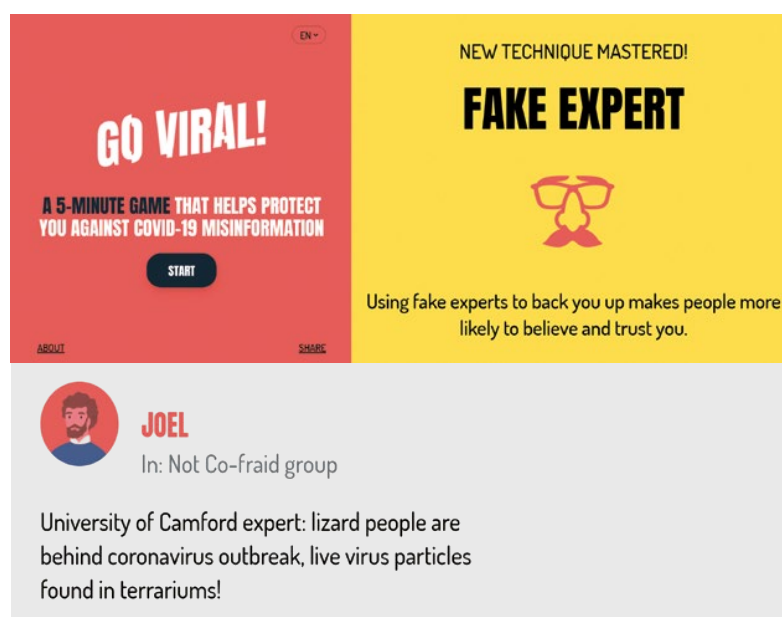
Maertens R, Anseel F, van der Linden S. Combatting climate change misinformation: evidence for longevity of inoculation and consensus messaging effects. *Journal of Environmental Psychology* 2020; 70, 101455 <https://doi.org/10.1016/j.jenvp.2020.101455>

Maertens R, Roozenbeek J, Basol M, van der Linden S. Long-term effectiveness of inoculation against misinformation: three longitudinal experiments. *Journal of Experimental Psychology: Applied* 2020. Advance online publication. <https://doi.org/10.1037/xap0000315>

Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G et al. Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science* 2020; 7(10), 201199. <https://doi.org/10.1098/rsos.201199>

Roozenbeek J, van der Linden S. Fake news game confers psychological resistance against online misinformation. *Palgrave Communications* 2019; 5(1), 65 <https://doi.org/10.1057/s41599-019-0279-9>

van der Linden S, Roozenbeek J, Compton J. Inoculating against fake news about COVID-19. *Frontiers in Psychology* 2020; 11, 566790 <https://doi.org/10.3389/fpsyg.2020.566790>



Screenshot of *GoViral!*, a game illustrating the fake expert technique ([www.goviralgame.com](http://www.goviralgame.com)).

Ultimately, we aim to follow the vaccination analogy to its logical conclusion: herd immunity. Psychological vaccines do not work in exactly the same manner as biological vaccines of course, their efficacy is nowhere near 95% immunity and they wear off over time if not boosted regularly. Having said this, we are currently running computer simulations to try to estimate what percentage of a population needs to be vaccinated at a sufficient rate within a given (online) community to effectively contain the spread of misinformation. To help ensure that misinformation-induced vaccine hesitancy does not spread further, governments, schools, technology companies and civil society can therefore help spread and scale the vaccine. For example, some social media companies, such as Twitter, have already started experimenting with prebunking on their platform. We are currently collaborating with Google, who through their research and innovation hub Jigsaw, is helping us to develop short inoculation videos, which could be embedded as an advertisement on, for example, YouTube (before people are exposed to misinformation). However, at the end of the day, we need a multi-layered defence system. We need to prebunk first where possible, but also continue to rely on real-time fact-checking as well as debunking after the fact if needed. Only if these combined pre- and debunking measures are implemented quickly, widely and in various settings, will we be able to contain the spread of misinformation in society.

At present, we do not yet know at what rate the misinformation virus evolves, but we know it is getting more sophisticated. As with any virus, we have to stay alert and be ready to update our vaccines whenever a new variant of misinformation arises. In the words of the renowned Defence against the Dark Arts Professor, Severus Snape, 'our defences must be as flexible and inventive as the arts we seek to undo'.