A Practical Guide to Prebunking *Misinformation*

A collaboration between:

University of Cambridge
BBC Media Action
Jigsaw

ABOUT

Prebunking is a technique gaining prominence as a means to build preemptive resilience to misinformation. This guide was developed for practitioners interested in defending against misleading and manipulative information. It documents the foundations of prebunking, aiming to translate academic research into a practical how-to guide that enables groups and individuals with no prior knowledge of behavioral psychology to deploy their own prebunking interventions.

This work is a collaborative effort between the University of Cambridge, Jigsaw (Google) and BBC Media Action. The University of Cambridge's Social Decision-Making Lab has been at the forefront of developing prebunking approaches, based on inoculation theory, designed to build people's resilience to mis- and disinformation.

Jigsaw, a team at Google, has partnered with leading universities around the world, including the University of Cambridge, to test prebunking in a variety of settings in order to understand the advantages and limitations of this approach.

BBC Media Action, the BBC's international development charity, is adapting and testing the use of prebunking approaches as one of its strategies to tackle information disorder in the various countries where it works.

This guide was written by the following people (listed in alphabetical order by organization): Mikey Biddlestone, Trisha Harjani, Sander van der Linden, and Jon Roozenbeek (University of Cambridge), Alasdair Stuart (BBC Media Action), Beth Goldberg, Meghan Graham, Mari Iwahara, Bomo Piri, Peter Weigand, and Rachel Xu (Jigsaw).

If you have any questions or concerns related to the research in this guide, please contact Jon Roozenbeek at the University of Cambridge's Social Decision-Making Lab.

If you would like more information on BBC Media Action's work tackling information disorder (including prebunking approaches) or you have any other enquiries for BBC Media Action, please email Alasdair Stuart.

Cite as: Harjani, T., Roozenbeek, J., Biddlestone, M., van der Linden, S., Stuart, A., Iwahara, M., Piri, B., Xu, R., Goldberg, B., & Graham, M. (2022). A Practical Guide to Prebunking Misinformation.

UNIVERSITY OF CAMBRIDGE BBC MEDIA ACTION JIGSAW

Table of *Contents*

Part 1: *Why* Prebunking

Part 2: *How to* Prebunk

01:
Why Prebunking

1.1

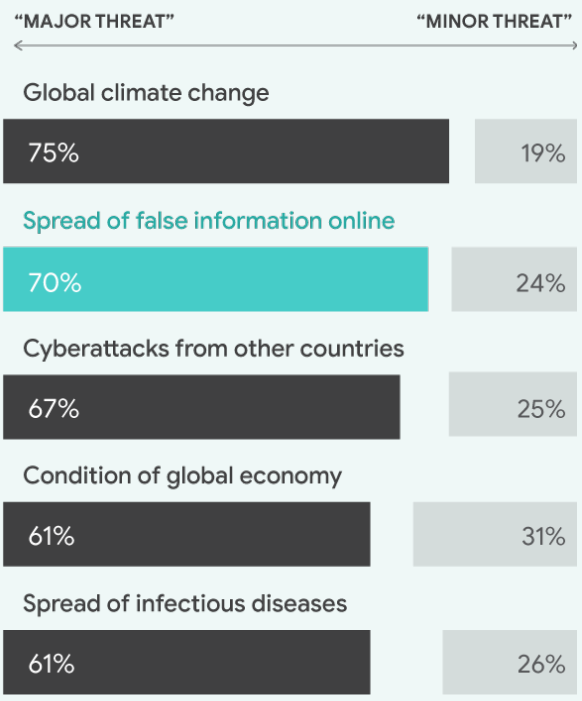
The landscape

The proliferation of misinformation online is a serious threat to public safety and modern democracy.

The real life consequences are serious — regions where COVID-19 disinformation thrived experienced higher death rates from the virus despite vaccine availability compared to neighboring regions.[1]

TOP PERCEIVED GLOBAL THREATS

Data Source: Pew Research Center



Meanwhile, a 2022 Pew Research Center poll across 19 countries found that 70% of respondents cited misinformation as a major threat to their country, second only to climate change as a global threat.[2]

Fighting back against misinformation is a challenge. A number of interventions have been designed to help minimize the spread and consumption of misinformation and disinformation,[3] including but not limited to debunking, nudges, automated labels, and information literacy boosts.[4] But there are many difficulties — both practical and conceptual — that hinder success at scale. One prominent approach — commonly known as debunking — targets misinformation after it has already spread. While corrective measures shown after seeing misinformation (such as fact checks) can be effective, they are often time consuming, expensive, and tricky to deploy with necessary speed. Misinformation can be quite "sticky," in the sense that individuals often continue to rely on it after they have been exposed to it, even after corrections have been made.[5] Moreover, fact checks are challenging as they have not historically received much engagement: research on over 50,000 debunking posts on Facebook found that very few audiences exposed to misinformation actually interacted with fact-checking posts.[6]

As a result, researchers have tried to find ways to prevent misinformation before it has taken hold in the first place. Pre-emptive approaches occur before people are exposed to misinformation and are commonly referred to as pre-emptive debunking or "prebunking." While there are many different types of prebunking interventions, they are often based on inoculation theory. Prebunking messages build mental defenses for misinformation by providing a warning and counterarguments before people encounter it. Note that while inoculation is usually most effective when an individual is reached beforehand, it's still possible to inoculate someone after they have been exposed to misinformation but haven't yet been persuaded (discussed further in 2.2: Define your audience).

1.2

How prebunking works

Prebunking focuses on how people are commonly manipulated and misled online, rather than directly challenging falsehoods or telling people what they need to believe.

Given the difficulty of dislodging beliefs based on misinformation, there is a growing field of research into helping people resist persuasion by misinformation in the first place. One approach borrows from biomedical science. Inoculation protects people against misinformation by teaching them to spot and refute a misleading claim via pre-exposure to a weakened dose. Prebunking, (or "attitudinal inoculation") is a way to teach people to spot and resist manipulative messages — before they happen. Prebunking has been demonstrably effective at helping a wide range of people build resilience to misleading information, including those across the political spectrum.[7] This technique focuses on *how* people are commonly manipulated and misled online, rather than directly challenging falsehoods or telling people what they should believe. As such, it can resonate with a wide audience because it is generally educational, non-judgmental, and non-accusatory in tone. It often focuses on the higher-order techniques and narratives being shared, seeking to empower individuals to spot *how* they are being manipulated. Prebunking assumes no prior capabilities or knowledge of a topic, making it widely usable across age groups and settings. For example, the first-ever prebunking game, *Bad News*, was designed to be used by educators to teach young people in schools how to spot the techniques used by malicious actors.

EXAMPLE: PREBUNKING MANIPULATION TECHNIQUES (FALSE DICHOTOMIES)

One example of prebunking reveals the common trick of "false dichotomies" in misinformation — providing a choice between only two options, even though there are many in reality. *View video >*

PREBUNKING ADVANTAGES	OUTCOME
Grounded in a large evidence base since the 1960s	Well-tested and shown to be effective in many scenarios
Proactively addresses persistent misleading narratives or techniques that are relevant over time and can be deployed across multiple topics and domains	Scales easier than fighting individual claims
Requires no pre-existing knowledge or capabilities on behalf of the viewer	Can be effective across ages and education levels
Non-accusatory in tone, invites non-judgmental learning, and taps into audience's innate desire to not be manipulated	Makes audiences more open to this type of preventative intervention
Can be effective while being apolitical by addressing misleading narratives or techniques rather than specific claims	Can be effective across the political spectrum, and in at least one study it was effective among those with conspiratorial beliefs[8]

Inoculation theory

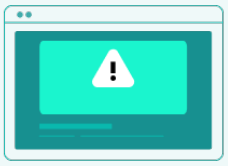
Prebunking is built on inoculation theory, which was developed in the 1960s by social psychologist William McGuire, and designed to be used as a psychological "vaccine for brainwash."[9]

Much like how medical vaccines confer physiological resistance against future infection, psychological inoculations confer resistance against future attempts of attitudinal manipulation (akin to the immunity provided by antibodies).

Studies over the past 60 years have shown inoculation to be effective across cultures and on a wide range of subjects including the environment, public health, crisis management, and animal rights, among others.[10,11,12,13,14]

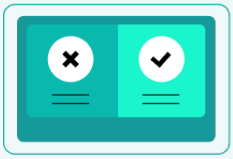
More recently, academics have demonstrated how inoculation messages can reduce the influence of misinformation and extremist propaganda online.[15]

Practically, inoculation involves two parts:



1. Forewarning

A warning activates the viewer's mental defenses against unwanted attempts to persuade them by alerting them that they are likely to encounter misleading messages in the near future.



2. Preemptive refutation

An effective rebuttal provides the viewer with tools to counter misleading information they may see in the future. In addition to equipping them with counter-arguments in advance, it helps to include a "micro-dose"or weakened example of the misinformation, so that they can more easily recognize it in the future.

Prebunking messages grounded in this foundational structure can strengthen the viewer's mental resilience to persuasive attacks in the future.[16] The limitations of this resilience is discussed further in 1.4 Limitations of prebunking.

What kind of information can be prebunked?

There are two predominant forms of prebunking that address misinformation at a higher level beyond specific misinformation claims. They both address different types of misinformation:

1. Misinformation narratives

Misinformation encountered online often comes in the form of claims or opinions about a particular topic. However, individual misinformation claims can often feed into broader narratives. Issue-based prebunking tackles the broader, persistent narratives of misinformation beyond specific claims.

Tackling individual misinformation claims is time-consuming and reactive, while prebunking broader narratives can dismantle the foundations of multiple claims at once and be much more effective at building resilience to new claims that share this false foundation.

EXAMPLE: HUMANS AND CLIMATE CHANGE

Consider the following statement:

"31,000 scientists have signed a petition: many climate scientists disagree over whether human release of greenhouse gasses are harming the Earth's climate."

This claim is one of many falsehoods that are part of the broader, misleading narrative that there is no scientific consensus on human-caused climate change and that climate change is instead part of Earth's natural cycle. Prebunking can address this broader narrative, warning people to be skeptical of those who seek to cast doubt on the scientific consensus that humans are contributing to climate change, without necessarily debating the facts of this specific claim about a petition.

2. Misinformation techniques

Technique-based prebunking focuses on the tactics used to spread misinformation. While the information that is used to manipulate and influence individuals online can widely vary, the techniques that are used to mislead are often repeated across topics and over time. Some commonly used tactics are outlined below.

EXAMPLE: COMMON MISINFORMATION TECHNIQUES

Some commonly used techniques are outlined in detail on the next page.

→

COMMON MISINFORMATION TECHNIQUES

TECHNIQUE	EXAMPLE
Impersonation Spreading information as another person or organization in order to appear more trustworthy and credible.	*"NASA admitted that climate change occurs naturally as a result of changes in Earth's solar orbit and not anthropogenic factors."* **EXPLANATION:** This example uses NASA as a way to increase the credibility of the statement, even though NASA has never made such a claim.
Emotional manipulation Using language that leverages strong emotional language to spark reactions — including fear or outrage.	*"What this airline did for its passengers will make you tear up — SO heartwarming."* **EXPLANATION:** This example shows how information can be presented to deliberately spark an emotional reaction to promote clicking and sharing and reduce critical evaluation.
Polarization Exaggerating existing differences between two groups to create a sense of hostility towards another party, such as using "us" versus "them" language. This is sometimes leveraged between political groups but can be used in many contexts.	*"People's Party: Don't believe the Worker Party liars. They said they would abolish student debt yet more people today are in debt than ever."* **EXPLANATION:** This example uses hostile "othering" language by describing another party as liars.
Conspiratorial ideation Explaining events from traditional news using alternative explanations that give weight to the idea that a small set of individuals, usually a secretive, malicious, elite group, are controlling these events.	*"Vaccines are just a way for billionaires to track us with their microchip vaccines! Who's really in control of our bodies here?"* **EXPLANATION:** This example encourages conspiratorial ideation by claiming people are not in control, referring to a mysterious group who is, in this case billionaires, and making unsubstantiated claims.
Ad hominem attack Ad hominems, Latin for "to the person," target the individual making an argument to take attention away from the argument's substance and shift it toward personal details. While such details might be relevant (e.g. if they show the person is not credible), they can also be entirely irrelevant and used as a distraction tactic.	*"Barbara has an uncontrollable temper and apparently a personality disorder too! We can't have someone crazy in power."* **EXPLANATION:** This example attacks characteristics of the leader, instead of discussing their policies or leadership decisions.

COMMON MISINFORMATION TECHNIQUES (CONT.)

TECHNIQUE	EXAMPLE
False dichotomy This is a type of logical fallacy that makes it appear as if there are only two sides or choices in a debate or situation, when in reality there are many more.	*"Either you support the energy protests or you don't believe in justice."* **EXPLANATION:** This example positions two ideas as opposite sides of a spectrum — making "supporting energy protests" and "believing in justice" as opposites — when it is possible to support both or neither at the same time, as well as many other positions someone may take.
False balance Presenting a debate as having two relatively balanced viewpoints that oppose each other when in fact, one argument has much more evidence to support it.	*"Experts debate the shape of the earth. While scientist Reece Chow has found the earth is spherical, expert Rene Paul argues that the earth is flat."* **EXPLANATION:** In this example, despite consensus amongst scientists that the earth is round, the placement of an "expert" that supports a flat-earth theory gives the argument more apparent support than it really has.

1.3

Formats and technical considerations

Prebunking interventions typically are either active — meaning the person interacts with questions or prompts to learn about the process of crafting misinformation — or passive — meaning the person observes a prebunking message.

Each of these approaches has advantages and disadvantages in terms of scalability, effectiveness, longevity, cost, and online engagement. Broadly speaking, the longer and more involved the viewer is in an intervention, the higher the effect size and the better the longevity of the inoculation effect.

MATCH YOUR CONTENT AND PLATFORM

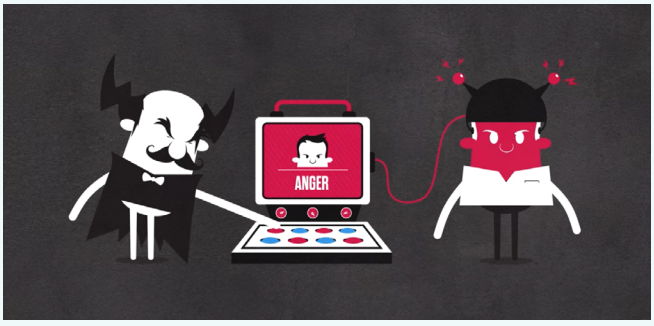
Content designed for one platform (ie: YouTube, TikTok, website) may not always be easily shared across other platforms, so it's important to think about where your content will live when choosing a format.

"Passive" prebunking

These interventions provide viewers with all the information needed to resist misinformation, without requiring them to actively engage beyond processing the information. For example, a video explaining how a technique is manipulative is a passive approach. Passive formats researched to date have included **text**, **graphics**, and **videos**.[17,18,19,20]

Passive prebunking interventions can be simpler from a production standpoint. For instance, a text-based prebunking intervention — like a series of pop-up messages — is relatively easy to implement at scale on social media. However, it is less immersive and interactive, which is likely to yield a smaller and shorter lasting impact than a more engaging — or active — format like a game.[21]

EXAMPLES



VIDEO EXAMPLE: FALSE DICHOTOMIES

This video example — produced by Jigsaw and Cambridge University — uses culturally relevant examples to help viewers understand and recognize the use of false dichotomies in the spread of misinformation. *View video >*

INFOGRAPHIC EXAMPLE: COVID-19 CONSPIRACY THEORIES

This UNESCO infographic explains conspiracy theories by using COVID-19 as an example.[22]

“Active” prebunking

Alternatively, active prebunking interventions require the individual to take action, making choices that help them retain information and engage more deeply with the content they see. The primary active approach researched to date is **games**.[23,24]

While games are more immersive and allow individuals to be inoculated against multiple manipulation techniques commonly used in misinformation, they require a larger investment from the viewer in terms of time and focus, which may reduce the number of people engaging with it. They are also a larger investment to produce, though some high-impact games have been implemented on a large-scale — like *Go Viral* (below).

Audio-based prebunking, such as broadcasting prebunking messages via radio or through chat apps (e.g. WhatsApp), is an underexplored medium that would benefit from further research (see 1.5 Future areas for exploration for more).

SELECTING A FORMAT

Each prebunking format has advantages and disadvantages in terms of scalability, effectiveness, longevity, cost, and engagement. Broadly speaking, the longer and more engaged the person is with the prebunk, the greater the size and duration of the prebunking effect.

GAMING EXAMPLES:

BAD NEWS

This was the first-ever prebunking game. It is a choice-based browser game created by DROG and the University of Cambridge in which players take on the role of a fake news producer and learn to identify and mimic six misinformation techniques (e.g. trolling, conspiratorial reasoning, impersonation) over six levels. Since then, several other games with similar premises have been designed. *View game >*

HARMONY SQUARE

Set in a peaceful community known for its pond swan and annual Pineapple Pizza Festival, this game appoints the player as the “Chief Disinformation Officer,” tasked with polarizing the people of Harmony Square and using trolling campaigns during political elections.
View game >

GO VIRAL!

This game similarly simulates the player’s descent into an online echo chamber where misinformation about the Covid-19 pandemic is common. Over three levels, players learn about the use of emotionally manipulative language, the use of fake experts to lend credibility to misinformation and the use of conspiratorial thinking to sow doubt. So far the game has had over 200 million impressions.[25] *View game >*

1.4

Limitations of prebunking

While prebunking has proven particularly effective at protecting individuals against manipulation attempts, there are some known limitations — and others that require more investigation to fully understand.

Scalability

Prebunking has proven effective with a wide variety of audiences, but practitioners should proceed with caution and pilot test when sharing messages across different types of misinformation, audiences, and platforms. Scaling to too broad an audience without appropriate specificity or local context can lead to lower engagement or oversimplification that can reduce efficacy. On the flip side, prebunking a single narrative or issue can narrow the relevant audience for that message and limit scalability (e.g. targeting vaccine hesitant audiences with a message prebunking vaccine misinformation).

CONTENT LIMITATIONS

Not all prebunks are equally scalable. Some narratives, even if they are made up of multiple claims, are still highly specific to a topic or area of misinformation. Since technique-based prebunks can be used across many topics, it maybe more scalable across many types of misinformation compared to issue-based prebunks.

Issue-based prebunks, however, are likely to provide deeper protection against specific topics and narratives. Being aware of the pros and cons of each approach is important when selecting an approach.

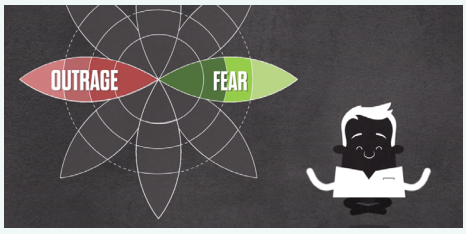
RISK OF OVERSIMPLIFICATION

A major challenge with scaling prebunking stems from the way users interact with content online. For people to engage with content on social media, information must be shortened to deliver information to the user as concisely as possible. This is increasingly the case with the rise of new media platforms, which makes it difficult to capture the nuance required to be effective. Presenting the three components of prebunking in a short, engaging way can be particularly challenging. Oversimplifying your message may make it ineffective, cause confusion, and even risk spreading misinformation further.

PLATFORM LIMITATIONS

Different platforms encourage different audience interactions, and using the same creative format across multiple channels can limit efficacy. Social media platforms are designed for specific content formats that may not perform as well on other platforms. Additionally, different platforms may host different misinformation narratives and use different types of messengers, such as influencers, so it can be challenging to optimize a message for more than one platform.

EXAMPLE: TRUTH LAB SERIES



For example, Roozenbeek et al. developed five short, animated videos that were presented to participants as 30 or 90-second advertisements on YouTube videos.[26] They found that the videos improved people's detection of manipulation attempts, discernment of trustworthy versus untrustworthy content, and the user's decision to share misinformation. View videos >

Length of effects

It is typical for learning from educational interventions to fade over time. Research has shown that this decay can be countered by administering "booster shot" interventions, or short reminders that prebunk the misinformation again at a later date. This may involve a repeat of the original prebunk or a shortened version summarizing key points.[27,28]

EXAMPLE: BOOSTER VIDEOS

Researchers from Jigsaw and the Universities of Cambridge and Bristol created booster videos to remind people of what they saw in previous, longer prebunking videos — analogous to a digital "booster shot." The experiment found that prebunking videos were able to protect individuals initially for around 10 days, and a 30-second booster video at Day 10 was a useful reminder that extended the protection to at least 30 days. View videos >

Unintended effects

When crafting prebunking interventions, practitioners should be vigilant and consider potential negative reactions to the message. Although backfire effects (meaning interventions inadvertently increasing people's belief in misinformation) do not appear to be a significant cause for concern,[29] some individuals are likely to resist any intervention. In the case of prebunking, for example, people particularly resistant to attempts to influence and alter their attitudes may not appreciate prebunking messages.

For example, one study found that messages prebunking white supremacist narratives had no effect on people with extreme right-wing beliefs, which suggests a resistance to this type of message for those with hardened views.[30] It is important to consider the influence of outliers in the audience when designing and analyzing prebunking messages.

1.5

Future areas for exploration

While inoculation interventions have existed since the 1960s, prebunking interventions in the digital age are still actively being researched and developed. More investment, research, and testing is needed to fully understand how to best prebunk on a global scale.

Global understanding

Despite misinformation being a global issue, much of the research on prebunking has been conducted in the Global North, such as the US, UK, and Europe. More research is needed to understand how prebunking can best be applied and contextualized in other countries around the world.

Factors such as language, demographics, geography, and cultural diversity can all play into the success or failure of scaling an approach like prebunking and need to be further understood in context.

Tackling closed applications

It is especially challenging to understand the spread of misinformation in closed messaging platforms such as WhatsApp and Telegram. When the technology is specifically designed to be private, it is inherently difficult to understand trends and habits. To date there has been limited research on how to apply prebunking to target this information space.

It would be valuable to test what types of prebunking content best engage users of closed chat apps, what formats they may choose to share with others (to multiply the impact of the intervention), and what effect this has on the impact and spread of misinformation in closed messaging spaces (e.g. can inoculation theory content reduce user belief in misleading or false information shared by friends or family, or reduce the likelihood of users sharing such content with their own contacts?).

EXAMPLE: BAD NEWS IN INDIA

A recent study found that the *Bad News* game was effective at prebunking individuals in India, with participants rating false news as less reliable after playing the game.[31]

BBC Media Action is working to adapt and distribute prebunking videos via existing social media channels with high reach in North Africa. The effectiveness of the campaign will subsequently be evaluated, with results expected to be shared in early 2023.

Formats and message lengths

Prebunking research to date has predominantly focused on text, videos, and interactive games. But there are many other formats in which humans consume information — more research is needed to understand how prebunking might be effectively adapted to different formats such as audio or memes.

AUDIO-BASED INTERVENTIONS

In some contexts, audiences still primarily rely on audio formats to receive and communicate information (e.g. in some rural communities in Africa, where community radio remains the primary source of information, or where high data costs mean people prefer to use audio content in WhatsApp groups rather than video content).

Developing audio-based prebunking approaches, exploring and testing dissemination of such approaches through radio programming or chat apps, is an underexplored area which could have great benefit in these contexts.

BITE-SIZE INFORMATION

While prebunking interventions using online games and short animated videos (approximately two minutes each) designed for digital distribution have been proven impactful, some digital audiences are more likely to engage with shorter length digital content (e.g. 30 seconds or less), and/or are moving to platforms that favor such content (e.g. TikTok). While some early work has demonstrated effectiveness of 30-second prebunking videos, further work is needed to explore whether and how prebunking can be adapted to this kind of "bite size" digital media content.[32]

LONGER-FORM NARRATIVE MEDIA

Longform programming such as TV or radio drama, or reality-shows, are designed to reach mass audiences. There is a compelling body of evidence, including from BBC Media Action's work, demonstrating that locally crafted and well-researched narrative-led media outputs can engage audiences at scale, bringing about social and behavior changes. The use of and evidence for the power of storytelling to address development issues at scale in low resource settings is ever widening and includes: HIV/AIDS, gender-based violence, gender norms, social cohesion, sanitation, contraceptive use, and child survival.[33,34,35,36,37,38,39]

BBC Media Action's experience has demonstrated that storytelling formats can be very useful in raising sensitive issues in a non-confrontational way, which is critical in societies where key power holders may be directly contributing to the spread of misinformation. However, to date, there has been no attempt to integrate prebunking approaches into such content. It would be innovative to test whether storylines within a drama could be used to convey a prebunking message to the drama's audience, such that they experience (through what happens to the characters in the drama) a prebunking warning. Such approaches have the potential to reach much larger audiences, and importantly may also engage more vulnerable populations who are unlikely to use online gaming or see digital inoculation theory content (e.g. older people who use social media less often).

Role of the messenger

Much of the research surrounding prebunking to date has explored the content and format of the prebunking message (the foundations of which have been distilled in this document), and how effectiveness varies based on these different levers. However, very limited, if any, research to date has considered how prebunking effectiveness may vary depending on the messenger or speaker of the prebunk.

Humans react differently to information from different sources – expertise, authoritativeness, trust, and bias, can all play a role in how we perceive and internalize messages from a messenger. More recent reviews of the inoculation literature have begun to examine the role of source credibility in achieving attitudinal resistance.[40] More research is required to understand which actors – e.g. social media influencers, public figures, authoritative organizations, news announcers etc. – are more effective messengers of prebunking information, in which contexts, and to which audiences.

Additional areas of inquiry

Prebunking as a field is growing rapidly to keep pace with an ever-evolving information environment. As the research advances, so do the actors who seek to spread misinformation, who adapt and evolve to find new ways to manipulate. Additional areas of inquiry will naturally emerge in concert, and researchers and practitioners alike must constantly push the frontier of knowledge to understand how to better protect our society against misinformation.

02:
How to Prebunk

2.1

When and who should do it

Prebunking works best when narratives and manipulation techniques are not fully understood by the audience, or the audience's position on the topic is dynamic. Once beliefs about a topic are solidified or polarized, it can be challenging to prebunk.

When contemplating prebunking as an approach to tackle misinformation, it's helpful to check that the following conditions apply:

When narratives or techniques can be anticipated

Misinformation narratives and techniques are often repeated over time, and across different topics. With thoughtful analysis of these trends, application of these narratives and techniques to new misinformation can often be anticipated. For example, recurring moments such as election cycles, health crises, and environmental disasters are often ripe for misinformation, and some techniques or narratives that occur at these times can be repeated.

EXAMPLE: SMALLPOX AND COVID

Vaccines are a perennial topic of misinformation. They have been accused of being "unnatural" since their invention and the false claims made about them are often recycled.

For example, in the 1800s, the smallpox vaccine was rumored to turn people into "human-cow hybrids" due to its cowpox-derived formula. Today, COVID-19 vaccines are similarly alleged to "alter your DNA."[41] This narrative was reasonably predictable ahead of time, and therefore may have been an effective candidate for prebunking.

Before audiences have been convinced

Audience receptivity is key when designing a prebunking intervention. Ideally, the intervention will reach audiences before they buy-in to misinformation. While there is some evidence to suggest that prebunking can still work after exposure to misinformation (known as "therapeutic inoculation"), it is more effective when audiences have not yet been fully convinced of the claim or narrative.[42] When designing a prebunking intervention, consider who your audience is, the degree to which they already believe the misinformation you are aiming to prebunk, and the current media and/or political landscape to determine the appropriateness of a prebunking intervention.

EXAMPLE: SCIENTIFIC RACISM

Research by Jigsaw and American University found that prebunking white supremacist narratives among Americans was effective in reducing support for white supremacist messengers and their narratives among the vast majority of those tested. However, the prebunking videos had no effect for those who showed strong pre-existing white supremacist beliefs (as measured by surveys like the right-wing authoritarianism scale and social dominance orientation scale).[43]

Who should do it

Due to the increasing distrust in online information, it is important that you have a strong foundation of trust and credibility with your audience when prebunking.

Ensure that your organization has the following:

☑ Expertise to speak authoritatively on the topic

The information space is oversaturated with advice and disputes over accuracy. Before embarking on prebunking, ensure that you have the necessary and sufficient expertise to credibly address the misinformation in question. If needed, partnering with respected experts, scholars, and authoritative bodies can be a great way to demonstrate expertise.

☑ Trust and good will with your audience

Audiences are more likely to trust the content of a message if they trust the source sharing it. If you have a strong relationship with the audience you're trying to reach, or feel that they have a positive affinity to you and/or your brand, you may be well positioned to prebunk misinformation. If you are concerned about the level of trust an audience has in you, consider partnering with a group or creator that has a stronger relationship with that audience.

☑ The capacity to engage

Prebunking should not be a one way conversation. Plan to have resources available to monitor, iterate, and measure your efforts. It is also important to maintain humility to engage in a dialogue with your audience after sharing messages that tackle misinformation.

2.2

Getting started

Here are five steps and considerations to keep in mind as you create your prebunking material:

STEP 1:

Choose your subject:

What misinformation do you seek to prebunk?

The subject of your intervention is based on the misinformation that you wish to target and may range anywhere from global crises such as climate change and pandemics to more individual-level issues like perceptions surrounding mental health.

As noted above in 2.1 When and who should do it, consider the following when choosing a subject:

- Make sure you have relevant expertise on the misinformation and audience that you are targeting, or are working with subject matter experts who do.
- Do your research on the misinformation landscape to identify prominent and burgeoning narratives and techniques your audience encounters.

STEP 2:

Choose your audience:

Who are you trying to reach with your prebunk efforts?

Consider the audience for your intervention and try to understand their current relationship to the information you are trying to share and what they may be interested in hearing from you.

As noted above in 2.1 When and who should do it, consider the following when choosing your audience:

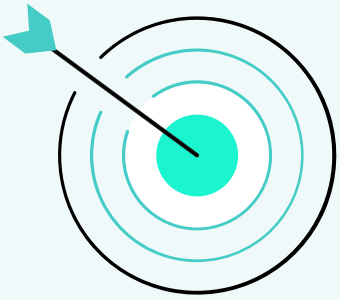
- Can you anticipate some of the techniques/narratives before they become widespread? Can you anticipate new techniques/narratives as the information landscape evolves?
- Has your audience already engaged with the technique and/or narrative that you are trying to dislodge? How ingrained are their beliefs?

2.2
GETTING STARTED (CONT.)

STEP 3:

Define your goal(s):

Specify the goals of your intervention: What outcomes do you hope to achieve after your prebunking intervention has been shared?



Prebunking interventions can achieve a range of outcomes that fall into three categories:

1. Knowledge or skills

Prebunking can teach audiences new knowledge (e.g. accurate statistics) or skills (e.g. ability to discern misinformation) to combat misinformation and build resilience to future manipulation.

2. Attitudes

Prebunking can shift audiences' attitudes about their own capabilities to defend themselves from misinformation or change their perceptions of an actor spreading misinformation (e.g. trustworthiness of a misinformation source).

3. Behaviors

Prebunking can change audiences' behaviors in the way they interact with, consume, or respond to misinformation (e.g. reducing sharing of misinformation).

OUTCOMES AND OBJECTIVES

These are not comprehensive, and there may be other goals you hope to achieve. Be sure to define them clearly and early, so that your organization is aligned on the objective of your intervention.

The outcomes you seek will influence how you design your prebunking intervention (see 2.2 Getting started), as well as 2.3 Measuring success.

2.2

GETTING STARTED (CONT.)

STEP 4:

Choose an approach: Issue vs. Technique-based

Select your approach: Do you want to prebunk an issue or a technique?

Issue-based approach

Issue-based (also known as narrative-based) prebunking targets broader, persistent narratives of misinformation, beyond specific claims. This allows you to tackle the foundation of many claims, enabling you to more effectively dismantle misinformation instead of fact-checking individual claims. More information on misinformation narratives can be found in 1.3 Misinformation narratives.

WHEN IS ISSUE-BASED PREBUNKING APPROPRIATE?

If the misinformation you're tackling requires a refutation grounded in specific facts and explanations of a topic, narrative prebunking could be a great approach.

Technique-based approach

Technique-based prebunking reveals commonly used techniques and tactics that are prevalent across multiple claims and misinformation narratives. This approach helps audiences understand *how* they may be manipulated, rather than disputing the content of the manipulation. More information on manipulation tactics used to spread misinformation can be found in 1.3 Misinformation techniques.

WHEN IS TECHNIQUE-BASED PREBUNKING APPROPRIATE?

If there are techniques that are commonly deployed across multiple claims and narratives, technique-based prebunking could be an effective way of providing broad resistance across multiple encounters of misinformation. Technique-based prebunking, because it is not tied to specific misinformation claims or narratives, makes it easier for your intervention to be more apolitical, which can be useful for more politicized misinformation topics.

2.2
GETTING STARTED (CONT.)

STEP 5:

Choose a format:

What medium is best to deliver your prebunking message?

Prebunking messages can be delivered in a variety of formats — to date, the literature has explored prebunks in the following formats: text, audio, visual, video, and games.

Each of these have their respective advantages and disadvantages in terms of scalability, audience engagement, effect size, long-term effectiveness, and cost. These are outlined in 1.3 Formats and technical considerations.

In general, more "active' approaches may yield deeper manipulation resilience. However, more engaging formats (like video games), often require more time and effort, and require significant buy-in from the audience to engage. “Passive” approaches can be quicker to develop and scale — however, they need to be designed and deployed thoughtfully in order to have a lasting effect. Note that these are generalizations based on the literature to date — effect sizes may vary depending on the intervention.

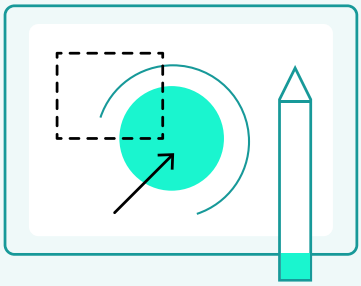
Questions to consider when deciding on a format:

- What media platforms and formats is the intended audience already engaging with?
- How much time and effort (or money) do you have to invest in production?
- Do you have the necessary design capacity to develop visually engaging messages, like infographics, videos, or games?
- What scalability and degree of online engagement are you hoping to achieve? Will your format keep your audience’s attention?
- Will your message be evergreen or require more resources to update periodically?

STEP 6:

Design your intervention

What components should you keep in mind?

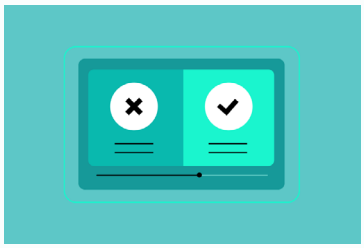


Inoculation messages can build up people's resistance or "mental antibodies" to encountering misinformation in the future, in the same way vaccines create antibodies that fight against future infection. But there are certain criteria that need to be met for an intervention to successfully qualify as a prebunk.

These are three key components of successful prebunking messages:

1. Warning

Alert users of attempts to manipulate them



2. Preemptive refutation

Explain the narrative/ technique and how it is manipulative

3. Microdose

A weakened or practical example of misinformation that is harmless (e.g. will not radicalize or distress your audience or repeat the misinformation)

BE HUMBLE

Sometimes, the information landscape changes quickly, especially during times of crisis (e.g. a new virus). Acknowledge the limitations to your explanations and counterarguments where possible, and be transparent about where information is still evolving.

You can take creative license when designing your intervention, but retaining these key components is important for maintaining scientific integrity. For additional creative guidance, see section 2.4 Creative considerations.

2.3

Measuring success

Once you have designed your prebunk, how will you know if it succeeds at achieving your goals?

It is helpful to have a measurement plan in place so to understand whether and how your intervention achieves your intended goal(s). Measuring the impact of your intervention provides useful feedback for future prebunking efforts and helps other practitioners. In order to measure success, there are three foundational steps:

STEP 1:

Define your key metrics

The metrics you choose should be directly tied to the goal(s) you hope to achieve.

As outlined in 2.2 Define your goals, common goals may involve changing an audience's **knowledge/skills, attitudes, and/or behaviors**. Some common metrics corresponding to these outcomes include:

Knowledge- or skill-based outcomes

- Ability to identify a misinformation technique
- Ability to discern a misinformation narrative
- Ability to distinguish between true and false information

Attitude-based outcomes

- Confidence in their own abilities to detect misinformation
- Trust in the reliability of a source
- Mood as a result of seeing a piece of misinformation (e.g. anger, fear)
- Tendency toward conspiracy theories

Behavior-based outcomes

- Consumption of misinformation (e.g. time spent on misinformation sources)
- Engagement with misinformation (e.g. comments)
- Sharing of misinformation
- Support for misinformation (e.g. likes)

You may choose to design your own metrics for your intervention — whatever metrics you decide to use, ensure that they adequately and accurately measure the goal(s) you set out to achieve. It is recommended to use a combination of metrics to measure your goal(s).

MATCH YOUR METRICS TO THE MESSAGE

If your prebunk is issue-based, the metrics should include questions on the same subject (or issue) featured in the prebunking message.

Similarly, technique-based interventions should be measured using questions that test viewers' knowledge, attitudes, or behaviors for the same technique(s) featured in the prebunk.

STEP 2:

Collect data

Once you have your metrics, what data do you need to measure these outcomes? For example, if you seek to inoculate someone on how to spot a false dichotomy, what information will tell you whether they have learned what a false dichotomy is?

In the literature, researchers have often used one or a combination of three ways to collect data to measure your desired outcomes:

1. Tasks

Tasks are used to test knowledge, skills, or characteristics of a person who has been exposed to your intervention. This could be as simple as a survey question that asks them, for example, to identify the correct manipulation tactic present in an example.

2. Self-reported responses

Self-reported responses are collected using surveys, asking questions of the person before and/or after they have interacted with your intervention. This may be about an attitude or intent that they may have after being exposed to your intervention. For example, a self-reported response to measure a change in attitudes could be to rate the trustworthiness, reliability, accuracy, etc.[44] of an example social media on a Likert scale from 1 ("not at all reliable") to 7 ("very reliable").[45,46,47,48,49,50,51]

3. Behavioral observation

Behavioral observation is when you collect data that records a person's behavior before, during and/or after they have interacted with your intervention. For example, you could collect data from a specific social media platform and assess how much misinformation was shared by a set of users. While behavioral data is the most direct measure of real-world impact, behavioral data can be difficult to obtain — it usually requires data access from the platform on which you are running your study, or a heavy computational effort to scrape data from the platform.

Due to the limited access to behavioral data, a lot of academic research instead uses **self-reported surveys as a proxy for behavioral data**, by asking for self-reported behavioral intent or judgements as a proxy for behavior. For example, to measure behavioral intent, you could ask the person to self-report whether they would share a piece of information.

DATA ACCESS

How you collect this data depends on the platform where you deploy your prebunking intervention.

For example, if you use a social media platform, you could collect data through a follow-up survey (if available).

If you choose to use your own platform, you might have access to behavioral data (e.g. whether or not someone clicks on a link to misinformation).

STEP 3:

Analyze your data

Once you have your data, how do you know if your prebunking intervention impacted your key metric?

Data analysis can happen at different levels of sophistication:

1. Measure after (easiest)

Collect data on your desired metrics after they have engaged with your prebunking intervention. This may tell you how resilient your audience is to misinformation at the time of measuring, but does not tell you whether this is due to your prebunking efforts.

2. Compare before and after

Collect data on your audience's performance on key metrics before and after exposure to your intervention. This allows you to observe the change in their performance after seeing your intervention, which may give some insight into the effectiveness of your prebunk. However, there may be other factors influencing the change in outcomes. Without a control group, you will not be able to say for certain if your prebunking was the primary driver of any change in their knowledge/skills, attitudes, or behaviors.

Conduct a randomized control trial (scientifically robust)

This is the most rigorous and scientific way to measure the effectiveness of your intervention. An introductory guide to RCTs can be found here. It is worth noting: the only way to truly know whether your prebunking intervention is definitively effective is by conducting a proper randomized control trial and statistical analysis on the data. However, noting that many organizations may not have the capacity to conduct full-scale statistical analyses, we have provided lighter-touch alternatives in this guide.

2.4

Creative considerations

Tone

When writing your prebunking message, determine the tone that is appropriate for your message and audience (e.g. serious, humorous, formal, casual, educational, etc.). The right tone depends on the relationship between your organization and your target audience, as well as the subject matter you are tackling.

When deciding on a tone, consider what will keep your audience's attention, and the best way to effectively convey your message.

Many prebunking interventions have used entertaining explanations and examples dispersed with humor throughout to maintain an audience's attention.[52,53] Some interventions have deployed narrative storytelling to better explain the components of a prebunk. But there are many topics that can be inappropriately matched with humor (for example, events that involve human suffering). Use your judgment and understanding of your audience to determine which tone will best connect with them.

EFFECTIVE VS. APPROPRIATE

Be very thoughtful about what tone is appropriate for your creative and which is effective. If you're not, something could go viral for the wrong reason.

EXAMPLE: LIGHT-HEARTED, ANIMATED PREBUNK

This prebunking video designed by Jigsaw and University of Cambridge to define the tactic of ad hominem attacks uses cartoon villains to help viewers detect manipulations online. *View video >*

EXAMPLE: SERIOUS PREBUNK

This more serious video produced by Jigsaw and Demagog shows actors depicting friends discussing real-life scenarios and narratives about Ukrainian refugees. *View video >*

CREATIVE CONSIDERATIONS (CONT.)

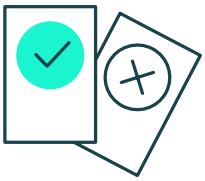
Before deploying your message, some final considerations:

Verify your sources

Check your sources. Make sure to be transparent about where your information comes from and avoid leaving audiences to guess your intentions.

Context

When creating a prebunking message, it is important to consider providing the viewer with access to more context on the subject and next steps. What will your viewers do after they see your message? Is there somewhere to direct them to learn more or help them get involved in spreading the word on preempting this misinformation? Consider a landing page with more information or create a call to action that helps amplify your message.



Creative testing

There are variables that cannot be anticipated, even given the best intentions and creative process. We recommend testing your creative (videos, infographics, games) with focus groups representative of your target audience to understand how audiences may respond before sharing with large groups online.

2.5

Watchouts

Prebunking is not a one-size-fits-all solution to ending misinformation as we know it. There are limitations and changing trends that will impact your efforts. A few are listed below.

ONLY ONE PART OF YOUR TOOLBOX

Prebunking is not definitively better than all other interventions — but it is a good first line of defense. It's also useful to bear in mind that individual-level interventions (including prebunking) work in tandem with system-level interventions (for example combatting polarization and organized disinformation).[54] All misinformation interventions have pros and cons, and many can be effective under the right circumstances. Remember prebunking is only one part of the toolbox — it can even be used in combination with other tools.

BE AWARE OF HARDENED VIEWPOINTS

As stated earlier in 2.1 When and who should do it, prebunking has proven effective when audiences are not fully bought into the misinformation. Once people's viewpoints have solidified, due to politicization or radicalization on a topic, prebunking is less likely to be effective and audiences may respond poorly. It can be difficult to gauge when this has taken place. Depending on your audience, you may choose different messages/channels for different audiences.

DON'T PATRONIZE YOUR AUDIENCE

When trying to share information or teach an audience a new skill, there is a risk of making your audience feel patronized. Online audiences are smart and digest large quantities of information very quickly. Avoid speaking down to them or treating them like children. Always maintain intellectual humility and a non-judgmental tone.

AVOID OVERSIMPLIFYING YOUR MESSAGE

When sharing information online, content creators often have to condense their messages into engaging bite-sized pieces in order to hold their audiences' attention — this is increasingly the trend on newer social media platforms. While shorter information may be easier to scale, it's harder to communicate nuance this way. Make sure that you are not oversimplifying your message to the point of rendering it ineffective. If the platform or medium you are using does not support complicated messaging, consider how you might direct them to a more detailed source so that viewers can dive deeper if interested.

POSSIBILITY FOR MISINTERPRETATION

Regardless of efforts to communicate effectively, efficiently, and deeply, it is possible for audiences to misinterpret your message. Plan for misinterpretation by including a link to where audiences can get more context on your efforts and goals.

2.6

Prebunking Checklist

Designing your intervention

- [ ] **Choose your subject**
  What misinformation are you seeking to prebunk?
- [ ] **Choose your audience**
  Who are you targeting with your intervention?
- [ ] **Define your goals**
  What outcomes do you hope to achieve?
- [ ] **Choose an approach**
  Will your intervention be tackling the content of the misinformation or the tactics?
- [ ] **Choose a format**
  What format best fits your intervention? (text, infographic, video, etc.)
- [ ] **Design your message**
  Build your intervention based on cultural, tactical and audience cues.
- [ ] **Deploy your messsage**
  Share your creative on designated platforms
- [ ] **Measure success**
  What metrics align with your intended outcome(s) and how will you measure results?

References

1 Leonardo Bursztyn, Aakaash Rao, Christopher P. Roth and David H. Yanagizawa-Drott, "Misinformation during a Pandemic," National Bureau of Economic Research, June 2020. https://www.nber.org/papers/w27417

2 Jacob Poushter, Moira Fagan, and Sneha Gubbala, "Climate Change Remains Top Global Threat Across 19-Country Survey," Pew Research Center's Global Attitudes Project (blog), August 31, 2022. https://www.pewresearch.org/global/2022/08/31/climate-change-remains-top-global-threat-across-19-country-survey/

3 We define misinformation here as information that is false, misleading, and/or fallacious and that can intentionally or unintentionally result in harmful consequences; disinformation is misinformation that is produced intentionally, for example as part of an organized campaign. For ease of reference, throughout this document, we will refer to all false or misleading information as misinformation, including anything that may have originated from disinformation or malinformation.

4 Jon Roozenbeek, Jane Suiter, and Eileen Culloty, "Countering Misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions," European Psychologist (September 20, 2022) advance online publication. https://doi.org/10.31234/osf.io/b52um

5 Stephan Lewandowsky et al., "Misinformation and Its Correction: Continued Influence and Successful Debiasing," Psychological Science in the Public Interest 13, no. 3 (December 2012): 106–31. https://doi.org/10.1177/1529100612451018

6 Fabiana Zollo et al., "Debunking in a World of Tribes," ed. Jose Javier Ramasco, PLOS ONE 12, no. 7 (July 24, 2017): e0181821. https://doi.org/10.1371/journal.pone.0181821

7 Sander van der Linden et al., "Inoculating the Public against Misinformation About Climate Change," Global Challenges 1, no. 2 (February 2017): 1600008. https://doi.org/10.1002/gch2.201600008

8 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience against Misinformation on Social Media," Science Advances 8, no. 34 (August 26, 2022): eabo6254. https://doi.org/10.1126/sciadv.abo6254

9 W. J. McGuire, "Resistance to Persuasion Conferred by Active and Passive Prior Refutation of the Same and Alternative Counterarguments," The Journal of Abnormal and Social Psychology 63, no. 2 (September 1961): 326–32. https://doi.org/10.1037/h0048344

10 Jon Roozenbeek, Sander van der Linden, and Thomas Nygren, "Prebunking Interventions Based on the Psychological Theory of 'Inoculation' Can Reduce Susceptibility to Misinformation across Cultures.," Harvard Kennedy School Misinformation Review, (February 3, 2020). https://doi.org/10.37016//mr-2020-008

11 Sander van der Linden et al., "Inoculating the Public against Misinformation About Climate Change," Global Challenges 1, no. 2 (February 2017): 1600008. https://doi.org/10.1002/gch2.201600008

12 Josh Compton, Ben Jackson, and James A. Dimmock, "Persuading Others to Avoid Persuasion: Inoculation Theory and Resistant Health Attitudes," Frontiers in Psychology 7 (February 9, 2016). https://doi.org/10.3389/fpsyg.2016.00122

13 Bobi Ivanov et al., "Using an Inoculation Message Approach to Promote Public Confidence in Protective Agencies," Journal of Applied Communication Research 44, no. 4 (October 2016): 381–98. https://doi.org/10.1080/00909882.2016.1225165

14 Robin L. Nabi, "'Feeling' Resistance: Exploring the Role of Emotionally Evocative Visuals in Inducing Inoculation," Media Psychology 5, no. 2 (May 2003): 199–223. https://doi.org/10.1207/S1532785XMEP0502_4

15 John Cook, Stephan Lewandowsky, and Ullrich K. H. Ecker, "Neutralizing Misinformation through Inoculation: Exposing Misleading Argumentation Techniques Reduces Their Influence," ed. Emmanuel Manalo, PLOS ONE 12, no. 5 (May 5, 2017): e0175799. https://doi.org/10.1371/journal.pone.0175799

16 Cecilie S. Traberg, Jon Roozenbeek, and Sander van der Linden, "Psychological Inoculation against Misinformation: Current Evidence and Future Directions," The ANNALS of the American Academy of Political and Social Science 700, no. 1 (March 2022): 136–51. https://doi.org/10.1177/00027162221087936

17 Sander van der Linden et al., "Inoculating the Public Against Misinformation About Climate Change," Global Challenges 1, no. 2 (February 2017): 1600008. https://doi.org/10.1002/gch2.201600008

18 John Cook, Stephan Lewandowsky, and Ullrich K. H. Ecker, "Neutralizing Misinformation Through Inoculation: Exposing Misleading Argumentation Techniques Reduces Their Influence," ed. Emmanuel Manalo, PLOS ONE 12, no. 5 (May 5, 2017): e0175799. https://doi.org/10.1371/journal.pone.0175799

19 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience against Misinformation on Social Media," Science Advances 8, no. 34 (August 26, 2022): eabo6254. https://doi.org/10.1126/sciadv.abo6254

20 Stephan Lewandowsky and Muhsin Yesilada, "Inoculating against the Spread of Islamophobic and Radical-Islamist Disinformation," Cognitive Research: Principles and Implications 6, no. 1 (December 2021): 57. https://doi.org/10.1186/s41235-021-00323-z

21 Jon Roozenbeek and Sander van der Linden, "How to Combat Health Misinformation: A Psychological Approach," American Journal of Health Promotion 36, no. 3 (March 2022): 569–75. https://doi.org/10.1177/08901171211070958

22 Melisa Basol et al., "Towards Psychological Herd Immunity: Cross-Cultural Evidence for Two Prebunking Interventions against COVID-19 Misinformation," Big Data & Society 8, no. 1 (January 2021): 205395172110138. https://doi.org/10.1177/20539517211013868

References (cont.)

23 Melisa Basol, Jon Roozenbeek, and Sander Van der Linden, "Good News About Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News," Journal of Cognition 3, no. 1 (January 10, 2020): 2. https://doi.org/10.5334/joc.91

24 John Cook et al., "The Cranky Uncle Game—Combining Humor and Gamification to Build Student Resilience Against Climate Misinformation," Environmental Education Research (June 14, 2022), 1–17. https://doi.org/10.1080/13504622.2022.2085671

25 "GCS International Joins the Fight against Health Misinformation Worldwide," Government Communication Service of the United Kingdom (February 18, 2021). https://gcs.civilservice.gov.uk/news/gcs-international-joins-the-fight-against-health-misinformation-worldwide/

26 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience against Misinformation on Social Media," Science Advances 8, no. 34 (August 26, 2022): eabo6254. https://doi.org/10.1126/sciadv.abo6254

27 Rakoen Maertens et al., "Long-Term Effectiveness of Inoculation against Misinformation: Three Longitudinal Experiments," Journal of Experimental Psychology: Applied 27, no. 1 (March 2021): 1–16. https://doi.org/10.1037/xap0000315

28 Bobi Ivanov, Kimberly A. Parker, and Lindsay L. Dillingham, "Testing the Limits of Inoculation-Generated Resistance," Western Journal of Communication 82, no. 5 (October 20, 2018): 648–65. https://doi.org/10.1080/10570314.2018.1454600

29 Briony Swire-Thompson, Joseph DeGutis, and David Lazer, "Searching for the Backfire Effect: Measurement and Design Considerations," Journal of Applied Research in Memory and Cognition 9, no. 3 (September 2020): 286-299. https://doi.org/10.1016/j.jarmac.2020.06.006

30 Brian Hughes, Kurt Braddock, Cynthia Miller-Idriss, Beth Goldberg, Meili Criezis, Pasha Dashtgard, and Kesa White, "Inoculating Against Persuasion by Scientific Racism Propaganda: The Moderating Roles of Propaganda Form and Subtlety," SocArXiv. (July 31, 2021). https://doi:10.31235/osf.io/ecqn4

31 Ananya Iyengar, Poorvi Gupta, and Nidhi Priya, "Inoculation Against Conspiracy Theories: A Consumer Side Approach to India's Fake News Problem," Applied Cognitive Psychology (September 14, 2022) acp.3995. https://doi.org/10.1002/acp.3995

32 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience against Misinformation on Social Media," Science Advances 8, no. 34 (August 26, 2022): eabo6254. https://doi.org/10.1126/sciadv.abo6254

33 Abhijit Banerjee, Eliana La Ferrara, and Victor Orozco-Olvera, "The Entertaining Way to Behavioral Change: Fighting HIV with MTV," Cambridge, MA: National Bureau of Economic Research (July 2019). https://doi.org/10.3386/w26096

34 S. Usdin et al., "Achieving Social Change on Gender-Based Violence: A Report on the Impact Evaluation of Soul City's Fourth Series," Social Science & Medicine 61, no. 11 (December 2005): 2434–45. https://doi.org/10.1016/j.socscimed.2005.04.035

35 UNICEF, "Technical Note on Gender Norms" (United Nations, n.d.). https://www.unicef.org/media/65381/file/GP-2020-Technical-Note-Gender-Norms.pdf

36 Ada Sonnenfeld et al., "Strengthening Intergroup Social Cohesion in Fragile Situations," 3ie Systematic Review 46 (2021). https://www.3ieimpact.org/evidence-hub/publications/systematic-reviews/strengthening-intergroup-social-cohesion-fragile

37 "Creatively Tackling Sanitation in India," BBC Media Action (September 2020). https://www.bbc.co.uk/mediaaction/publications-and-resources/research/summaries/executive-summary-navarangi-re-sept-2020/

38 Rachel Glennerster, Joanna Murray, and Victor Pouliquen, "The Media or the Message? Experimental Evidence on Mass Media and Contraception in Burkina Faso," August 21, 2022. https://www.povertyactionlab.org/sites/default/files/research-paper/working-paper_3835_Mass-Media-and-Contraception_Burkina-Faso_Aug2022.pdf

39 Danielle A. Naugle and Robert C. Hornik, "Systematic Review of the Effectiveness of Mass Media Interventions for Child Survival in Low- and Middle-Income Countries," Journal of Health Communication 19, no. sup1 (May 6, 2014): 190–215. https://doi.org/10.1080/10810730.2014.918217

40 Josh Compton, Sander van der Linden, John Cook, and Melisa Basol, "Inoculation Theory in the Post-Truth Era: Extant Findings and New Frontiers for Contested Science, Misinformation, and Conspiracy Theories," Compass (May 5, 2021). https://doi.org/10.1111/spc3.12602

41 Renee DiResta, "'Prebunking' Health Misinformation Tropes Can Stop Their Spread," Wired (August 28, 2021). https://www.wired.com/story/prebunking-health-misinformation-tropes-can-stop-their-spread/

42 Josh Compton et al., "Inoculation Theory in the Post-Truth Era: Extant Findings and New Frontiers for Contested Science, Misinformation, and Conspiracy Theories," Social and Personality Psychology Compass 15, no. 6 (June 2021). https://doi.org/10.1111/spc3.12602

43 Brian Hughes et al., "Inoculating against Persuasion by Scientific Racism Propaganda: The Moderating Roles of Propaganda Form and Subtlety," preprint: SocArXiv (July 31, 2021). https://doi.org/10.31235/osf.io/ecqn4

44 Jon Roozenbeek, Sander Van der Linden, Rakoen Maertens, Stefan M. Herzog, Michael Geers. Ralf Kurvers, and Mubashir Sultan, "Susceptibility to Misinformation Is Consistent Across Question Framings and Response Modes and Better Explained by Myside Bias and Partisanship than Analytical Thinking," Judgment and Decision Making, Vol. 17, No. 3 (May 2022): pp. 547–573. https://journal.sjdm.org/22/220228/jdm220228.pdf

45 Rakoen Maertens et al., "Long-Term Effectiveness of Inoculation against Misinformation: Three Longitudinal

References (cont.)

Experiments," Journal of Experimental Psychology: Applied 27, no. 1 (March 2021): 1–16. https://doi.org/10.1037/xap0000315

46 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience Against Misinformation on Social Media," Science Advances 8, no. 34 (August 26, 2022): eabo6254. https://doi.org/10.1126/sciadv.abo6254

47 Melisa Basol et al., "Towards Psychological Herd Immunity: Cross-Cultural Evidence for Two Prebunking Interventions against COVID-19 Misinformation," Big Data & Society 8, no. 1 (January 2021): 205395172110138. https://doi.org/10.1177/20539517211013868

48 Stephan Lewandowsky and Muhsin Yesilada, "Inoculating Against the Spread of Islamophobic and Radical-Islamist Disinformation," Cognitive Research: Principles and Implications 6, no. 1 (December 2021): 57. https://doi.org/10.1186/s41235-021-00323-z

49 Melisa Basol, Jon Roozenbeek, and Sander Van der Linden, "Good News About Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News," Journal of Cognition 3, no. 1 (January 10, 2020): 2. https://doi.org/10.5334/joc.91

50 Jon Roozenbeek and Sander van der Linden, "How to Combat Health Misinformation: A Psychological Approach," American Journal of Health Promotion 36, no. 3 (March 2022): 569–75. https://doi.org/10.1177/08901171211070958

51 Jon Roozenbeek and Sander van der Linden, "The Fake News Game: Actively Inoculating against the Risk of Misinformation," Journal of Risk Research 22, no. 5 (May 4, 2019): 570–80. https://doi.org/10.1080/13669877.2018.1443491

52 John Cook et al., "The Cranky Uncle Game—Combining Humor and Gamification to Build Student Resilience Against Climate Misinformation," Environmental Education Research, June 14, 2022, 1–17. https://doi.org/10.1080/13504622.2022.2085671

53 Jody C. Baumgartner and Amy Becker, eds., "Political Humor in a Changing Media Landscape: A New Generation of Research," Lexington Studies in Political Communication (Lanham: Lexington Books, 2018).

54 Jon Roozenbeek, Jane Suiter, and Eileen Culloty, "Countering Misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions," European Psychologist (20 September 2022) advance online publication). https://doi.org/10.31234/osf.io/b52um

A Practical Guide to Prebunking Misinformation

UNIVERSITY OF CAMBRIDGE

BBC MEDIA ACTION

JIGSAW