Government
Communication
Service

RESIST 2

Counter-disinformation toolkit

Foreword

"An unexciting truth may be eclipsed by a thrilling falsehood"

Aldous Huxley, 1958.

During the COVID-19 pandemic, technology and social media have been used to keep people safe, informed, productive and connected. However, this same technology is also enabling misinformation and disinformation to spread, becoming increasingly sophisticated since the publication of the original RESIST toolkit. The pandemic has brought to bear the terrible impact of misleading information, as people look for truth, comfort and guidance for themselves and their loved ones in distressing times. The pandemic has taught us many lessons, but especially for us as communicators, it has shown us that impactful communications can save lives.

The UK government has been working and learning alongside its international partners to take action against disinformation. Through multiple trainings and collaborative partnerships with government, multilateral and civil society communicators, the original toolkit has been used extensively around the world. The RESIST toolkit has been used to ensure we uphold the democratic values we hold dear and that communicators - as the deliverers of important and often life-changing messages - speak to their audiences effectively. The proliferation of false information severely undermines trust in democratic institutions, processes and cohesive societies. As communicators we all shoulder the responsibility of delivering truth to our citizens to maintain the fabric of our institutions and uphold our democratic values.

This refreshed version of the RESIST toolkit reflects the new realities of the threat that mis- and disinformation poses today. It explores new techniques and tactics, and how organisations can effectively identify, monitor and respond. The toolkit takes a systematic, evidence-based approach for helping organisations build societal and individual resilience to disinformation.

Alex Aiken

Executive Director
of Government
Communications

RESIST 2 – What's New?

Since we developed the original RESIST framework in 2018, hundreds of communications professionals and civil servants from the UK and around the world have provided feedback about how they use the toolkit, and what they would like to see in future iterations. This updated version reflects both the changing demands of the communication profession, and the evolving information environment.

Key new developments:

Threats

While the key state actors remain the same, new techniques and approaches emerged during the Covid-19 pandemic, and the threat posed by misinformation has become more urgent. We have therefore developed the scope of the toolkit to cover "false and misleading information", and greatly expanded the section on "Recognise" to better reflect current threat vectors.

Audiences

Vulnerabilities within the audiences targeted by mis- and disinformation have changed. For example, those vulnerable to mis- and disinformation during the Covid-19 pandemic represented far larger audiences than previous campaigns. We have expanded the toolkit to include a greater variety of communication tools to reach these audiences, and additional structured analysis techniques to weigh up risk and impact.

Partnerships

The community working to counter the negative effects of mis- and disinformation has grown. National and international stakeholders, including independent media, civil society, digital platforms, and academic institutions, perform crucial roles in a whole of society approach to maintaining a healthy information environment. Effective collaboration with partners, including policy levers outside of traditional communication tools, is now better represented in the toolkit.

Outcomes

Given the range of counter mis- and disinformation initiatives internationally, we recognise a need to better evaluate the effectiveness of communication interventions. The Tracking Outcomes section of the toolkit has evolved into a more comprehensive section on Tracking Effectiveness, focused on assessment of impact metrics.

Case studies

RESIST training has always drawn heavily on case studies and examples from participants' own experiences. For this updated version, we have integrated many examples both from the UK and international partners into the toolkit, to help better elucidate techniques and provide inspiration.

Contents

Executive Summary

Recognise

Recognise provides an overview of the information environment as it is today. It highlights the sheer volume of information shared online and the challenges this poses for communicators. It explains the differences between misinformation, disinformation and malinformation and why these categorisations are useful to understand the impact on audiences, including threatening public safety, fracturing communities and undermining trust in public institutions and the media. It also provides a checklist that can be used by communicators to determine whether a piece of information is likely to be false.

Early Warning

Early Warning begins with an overview of the available tools that can be used to monitor the media environment. It explains how to identify areas of your organisation - and its priorities - that are vulnerable to mis- and disinformation. It will support communicators to focus on the monitoring of key vulnerabilities by mapping out the objectives, audiences, brands and information and determining the extent to which they are at risk from mis- and disinformation.

Situational Insight

Situational Insight explains how communicators can turn information into actionable insight. It notes that insight can be produced on an ongoing basis (daily, weekly or monthly reports), or in an ad hoc manner to respond to emerging threats and issues. It helps communicators to create insight that is accessible for officials, including ways to explain insight clearly and avoid the use of jargon.

Impact Analysis

Impact Analysis explains how communicators can use 'structural analysis techniques' to predict the likely impact of a piece of mis- or disinformation. It stresses the importance of establishing and following clearly defined processes to ensure that assessments are objective and do not rely on the instincts - or 'gut feeling' - of practitioners.

Strategic Communications

Strategic Communications outlines the key skills communicators should utilise when developing and implementing communication strategies. It explains how communicators can make their content more engaging and impactful, including by identifying the best channels to reach target audiences, and using 'friendly voices' to increase the credibility and reach of communication activity. It outlines the different ways communication activity can be undertaken - proactive, reactive or active - and details how communicators can assess and implement these different types of approaches in diverse scenarios.

Tracking Effectiveness

Tracking Effectiveness outlines the importance of measuring the extent to which strategic communication activities have been effective. It sets out the importance of distinguishing between outputs and outcomes, and offers examples of metrics that can be used to determine the effectiveness of communications against pre-defined objectives.

Why Do We Need RESIST?

The amount of information on the internet seems endless. Each minute of every day, WhatsApp users send 40 million messages; Facebook users share 150,000 posts, YouTube users upload 500 hours of content, there are 350,000 new tweets, and nearly half a million engagements on Reddit[1]. Add to this traditional print and broadcasting media, email and telephones, and it is clear that the information environment is more contested than ever.

Communications must compete in this crowded information space. Your job as a communications specialist is to understand your audiences, to earn their attention and trust, so that you can supply them with important - and sometimes vital - information. This means developing an understanding of threats to the information environment, particularly those that stand between you and your audiences.

Sometimes people accidently share false information. Perhaps they didn't read it properly, or they misunderstood or misremembered what they had read, or they were given the wrong information to begin with. This is known as misinformation. **Misinformation refers to verifiably false information that is shared without an intent to mislead.** The effects of misinformation can still be harmful.

People also deliberately spread false or manipulated information. Usually, it is because the individuals and organisations that create it have something to gain from deception. **Disinformation refers to verifiably false information that is shared with an intent to deceive and mislead.** It can often have harmful consequences.

Sometimes true or partially true information is used in such a way that it has similar effects to disinformation. For example, facts such as statistics can be twisted or taken out of context to support false interpretations. This is known as malinformation. **Malinformation deliberately misleads by twisting the meaning of truthful information.**

These different definitions are good to know, but real-life instances of false and misleading information tend to be less clear-cut. Often, you won't know somebody's intention or goals. Sometimes people spread disinformation because they think it's true. This means that it technically becomes misinformation. Malinformation can be challenging to contest because it is difficult to inject nuance into highly polarised debates. In general, we recommend not getting too hung up on definitions and to retain focus on your priorities, such as supplying accurate and credible information, and where necessary protecting the public from harm.

This toolkit enables you to develop a response that can be used when dealing with all types of manipulated, false, and misleading information. You can use the definitions of mis-, dis- and malinformation to help think about who is spreading it and why, but when developing and delivering a response to false information, it's most important to **focus on the harm it can do**.

Manipulated, false and misleading information can:

- threaten public safety
- fracture community cohesion
- reduce trust in institutions and the media
- undermine public acceptance of science's role in informing policy development and implementation

[1] https://web-assets.domo.com/blog/wp-content/uploads/2020/08/20-data-never-sleeps-8-final-01-Resize.jpg

- damage our economic prosperity and our global influence; and
- undermine the integrity of government, the constitution and our democratic processes.

The aim of this toolkit is to reduce the impact of false, misleading, and manipulated information on UK society and our national interests, in line with democratic values. **Our primary objective is to give the public confidence in assessing the veracity of information themselves, so they are equipped to make their own decisions.**[2]

Online Harms

The UK Government's Online Harms White Paper[2] identifies several harms from online activity and false information. The most relevant for this toolkit are:

- terrorist and extremist content
- abuse and harassment
- online mis- and disinformation
- online manipulation
- abuse of public figures

You can find out more about how to prioritise the harms most relevant to your work in Section 2: Early Warning and Section 4: Impact Analysis.

This toolkit provides a consistent and effective approach to identifying and tackling a range of manipulated, false, and misleading information that government and public sector communicators may experience. The RESIST model is divided into components that can be used independently or tailored depending on the kind of organisation and the threats it faces. Sometimes you will need to respond before you know what something is or while you are investigating it - i.e., there won't always be time to do the steps in order.

Communications departments play a central role in recognising and responding to mis- and disinformation. You will often be the first to see it. This toolkit helps you develop routines to make informed assessments of risk and to share your insights and work with other parts of your organisation. It helps you to formulate recommendations and responses, and to evaluate your actions. The approach set out in this toolkit will contribute to a robust early warning system for recognising and responding to threats and emerging trends in the information environment.

[2] https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

Recognise Mis- and Disinformation

While the distinctions between mis-, dis- and malinformation are useful from a theoretical perspective, in practice they can be hard to distinguish. In many instances, it won't be obvious whether a questionable social media post is entirely false or whether there is malign intent behind it. It also won't be clear if it is an isolated incident, or indicative of sustained malicious intent. This section will introduce you to the important things to look for so that you know how to recognise misleading or manipulated information. In addition to introducing the basics of how to recognise mis- and disinformation, this section also serves as an introduction to the RESIST method.

In this section, you will learn:

- how to identify the problematic components of misleading or manipulated messages
- some of the ways that messages fit within and support problematic narratives
- how to better understand the brand values and identities of those who spread problematic messages and narratives
- how to weigh up the intent of those who spread the messages and narratives
- how to weigh up the possible or likely impact of techniques used

1.1 Investigate the messages

The most common way to first notice mis- and disinformation is when you encounter **messages** that draw your attention and raise concerns. A message is a form of communication aimed at a group of recipients. It may for example take the form of a social media post, tweet, meme, or comment, or a letter, flyer, poster, or slogan. Is the message an opinion? Opinions are usually subjective, which means that they cannot be verifiably false. **If the message is simply a statement of opinion, you should not treat it as disinformation**. However, if the opinion is based on verifiably false, deceptive, or manipulated information that has the potential to cause harm, it may be worth investigating further.

First, you should look out for five of the most common components of mis- and disinformation. We call these the **FIRST indicators**, because they are almost certainly the first things that will draw your attention. Note that recognising deceptive techniques is only the first stage of discovery. You should work through the following stages in order to better recognise the scope of the problem.

Fabrication

Is there any manipulated content? E.g., a forged document, manipulated image, or deliberately twisted citation.

Identity

Does anything point to a disguised or misleading source, or false claims about someone else's identity? E.g., a fake social media account, claiming that a person or organisation is something they are not, or behaviour that doesn't match the way the account presents itself.

Rhetoric

Is there use of an aggravating tone or false arguments? E.g., trolling, whataboutism, strawman, social proof, and ad hominem argumentation.

Symbolism

Are data, issues or events exploited to achieve an unrelated communicative goal? E.g. historical examples taken out of context, unconnected facts used to justify conspiracy theories, misuse of statistics, or conclusions that are far removed from what data reasonably supports.

Technology

Do the communicative techniques exploit technology in order to trick or mislead? E.g. off-platform coordination, bots amplifying messages, or machine-generated text, audio and visual content.

Fabrication case study:

In 2021, hackers altered the text of two Polish government websites to falsely claim that there had been a leak from a nuclear power station in neighbouring Lithuania. The hackers then took control of the Facebook and Twitter accounts of prominent political figures to spread the content.

Identity case study:

In 2018, Pro-Kremlin groups created fake Instagram pages designed to look like legitimate news aggregators. Targeting an upcoming election in Ukraine, they linked to stories containing heavy political biases and false information, and the pages accumulated 100,000 followers prior to their removal.

Rhetoric case study:

According to a March 2021 survey, more than a quarter of people in the UK read information about COVID-19 that could be considered false or misleading. People from minority ethnic backgrounds were particularly targeted with provocative rhetoric that implied that coronavirus vaccines contain pork or monkey animal parts, that the vaccine will alter human DNA, and that it is part of a plan to implant microchips. Respondents from minority ethnic backgrounds were around twice as likely to say that the claim made them think twice about getting vaccinated.

Symbolism case study:

To support its illegal annexation of Crimea in 2014, pro-Kremlin media spread narratives that engaged in historical revisionism, including the false claims that Crimea is a natural part of Russia, that Ukrainian control over Crimea was a historical anomaly, and that Ukraine has never been a genuinely sovereign country.

Technology case study:

In 2020, a day before Indian Legislative Assembly elections, a political party released deepfake videos of a candidate criticising a competitor in a language he did not speak. 5,800 WhatsApp groups shared the manipulated content, which reached up to 15 million people.

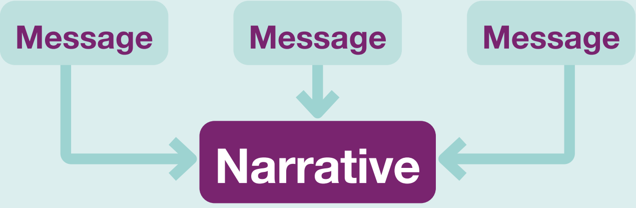
See Annex A for a glossary of common terms related to mis- and disinformation.

1.2 Unpack the narratives

Messages are the building blocks of **narratives**. Narratives are a form of storytelling that helps to explain and shape perceptions of an issue. They are stories that are designed to influence a target audience. If you see lots of messages on a topic, it is likely that you will be able to identify one or more of the narratives that they fit into or help to construct.

Narratives are generally made up of seemingly disparate messages or statements, brought together to tell a particular story. These stories are then more relatable to a broader audience, and can unify groups with different beliefs and interests, making them convenient vehicles for spreading misleading or deceptive ideas. **As a communicator, you should familiarise yourself with the misleading narratives that affect your principal areas of work and responsibility**. *You can find out more about how to do this in Section 2: Early Warning.*

Narratives are in essence **simple stories that give shortcuts to understanding complex issues**. They often express things about identity, community, and purpose. They are often not literally true, but rather they carry the aggregated, distilled beliefs of a community built up over time by many people across many statements. If you identify a series of messages that fit within, or help to construct a misleading narrative, that can be an indicator of mis- or disinformation.



Case study:

In late 2020, memes began to circulate suggesting that the AstraZeneca vaccine, which uses a modified virus derived from viruses found in chimpanzees, would turn people into monkeys. Drawing upon symbolism from the Planet of the Apes movies as well as the principle of reverse evolution (the vaccine devolving humans to apes), the memes were widely spread on pro-Kremlin media and social media. Many of the memes featured manipulated images of politicians and researchers, among others, with ape or chimpanzee features replacing their human features. The intention behind each of these messages appear to have been to sow fear and doubt around the AstraZeneca vaccine with the aim of boosting confidence in the Russian Sputnik V vaccine. Each message therefore took the form of a simple joke that nonetheless contributed to a more significant narrative that had geopolitical objectives.

1.3 Consider the brand and who it speaks to

If you're able to identify messages and narratives based on verifiably false, deceptive, misleading, or manipulated information, the next step is to get a sense of who is spreading the information. **The aim is to better understand the values, identities and beliefs that drive these narratives, and with whom they have credibility**. We call this the brand. A brand is what people think and feel about somebody or something, the characteristics and qualities that distinguish it from everything else. It is not about the person behind the account, but rather about the **persona that the account attempts to project to resonate with certain target audiences**.

Narrative
Narrative
Narrative
Brand

Case study:

During 2018-19, a regional Indonesian independence movement conducted widespread protests against the government. A series of pro-independence social media pages and accounts appeared that identified themselves as being news providers based in that region. Despite at first appearing sympathetic to the independence movement, they gradually shared dramatic anti-protester rhetoric, such as referring to the protesters as extremists, criminals, and anarchists, and these pages also promoted pro-government narratives. Their 'brand', in other words, was deceptive. An investigation found that more than 70 social media accounts and pages had been created as part of a coordinated campaign against the independence movement by a public relations company.

On social media it can sometimes be hard to understand whether an account represents a real person, advocacy movement, business, or a troll designed to spread certain messages. The account may make claims to being a certain type of person or organisation (based on the image and description), or it may give very little information. Since it may not be possible for you to accurately attribute a social media account or set of posts to its owner, it is better to focus on two aspects of its brand identity: **what** the account claims to represent, and **who** it claims to represent.

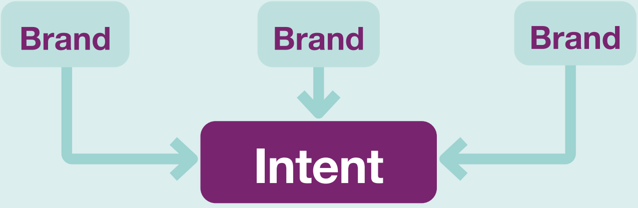
Narratives serve to unify people and groups with very different interests. **Understanding the different brands that participate in these narratives and comparing what you can learn about the brands with examples of behaviour should help you to better understand where different participants fit into the debate**. If mis- and disinformation are strongly connected to the brand identity, that can be an indication of something worth digging into further. You can find out more about how to share and analyse information about false or misleading narratives in Section 3: Situational Insight. *You can learn more about how to weigh up the impact of mis- and disinformation on your brand in Section 4.5*

1.4 Weigh up the intent

Closely connected to the brand is the intent. **Intent is notoriously difficult to determine**. However, if you weigh up the kinds of brands that share problematic messages and narratives, a picture should begin to emerge. **Often, the position of the account within its community – i.e. how its brand fits with the other brands it engages with – can help to reveal intent**.

If the account sharing misleading content is willing to discuss, delete or correct false statements, for example, it is a strong indicator that there is no intent to mislead. If, however, there are signs that the account is deliberately spreading false content, trolling, or attacking individuals and organisations, the risk is higher that there may be an intent to cause harm.

Intent may vary quite widely even among those who produce the same kinds of messages, who adhere to similar narratives, and whose brand values overlap. Below are several examples of the most common intentions that motivate individuals and organisations to spread false, misleading or manipulated information.



Because I believe

Sometimes people spread misleading or manipulated information because they genuinely believe something, because they want to persuade others, or because they want it to be true:

- "I am inclined to support any statements that resonate with my strongly-held beliefs."
- "Even if this specific case isn't true, it's an example of the kind of thing that goes on."
- "Our objectives are so important that bending the rules doesn't matter: the ends justify the means."

Because I have grievances

Groups and individuals sometimes use dubious communication tactics because they believe it is the best way for their voices to be heard:

- "I feel disenfranchised, and this community is giving me a voice."
- "I blame this social group for my problems, and I want to damage them."
- "I'm angry and want others to feel the way I do."

Because I can

Sometimes people use the anonymity of being online to do things they would never do in real life, which can confer status within a hierarchical online community:

- "Here I can say the things I really want to say without consequences."
- "People really respect and listen to me, so I give them more of what they want."
- "I wanted to see if it was possible to do this and get away with it."

Because I want to discredit somebody or something

Often, people spread misleading information that is aimed at negatively affecting the credibility, trust and reputation of a target person or organisation:

- "By discrediting this organisation, we will get more support for our goals."
- "This person is bad. We need to tell the world the truth."
- "Maybe this exact case isn't true, but it's an example of the sort of thing they would do."

Because I want to contribute to polarisation

Sometimes the intent is to contribute to existing tensions by aggravating them, thereby eroding the middle ground:

- "There are only two sides in this debate and you must choose one or the other."
- "I am here to promote my beliefs and help to crowd out opinions I do not like."
- "I will use any means available to prove my point and disprove yours."

Because I can make money

Sometimes people spread misleading or manipulated information because they can make money from it:

- “Click on these links that seem to align with your views, and I will sell your data, generate some ad revenue, and try to steal your passwords.”
- “I am an influencer for hire. For a fee I will support your cause using all methods available.”
- “By spreading false information and preying on uncertainty, I am able to sell products that I know do not work.”

Because I am part of an Information Influence Operation

On rare occasions, hostile state and nonstate actors conduct espionage designed to undermine the prosperity and security of another country. Such activities can have an online component that takes advantage of sensitive debates and issues. These activities are usually hidden in plain sight, among genuine people with legitimate worries and grievances:

- “I have gained the trust of this group by creating material that they share and by supporting their posts.”
- “I am cultivating groups and individuals to make them more extreme.”
- “My online activities are coordinated with the work of several other agents to achieve a specific effect on a country’s decision-making, social cohesion, and alliances.”

1.5 Consider the impact of the techniques

We began by analysing messages using the FIRST Indicators to establish the use of false, misleading or manipulated information. We then connected these messages to narratives that may be problematic to your areas of responsibility, as well as to the integrity of public debate more generally. Third, we considered how the brands spreading the messages and narratives present themselves, how they behave in their communities, and who they are targeting. Drawing this together, we then weighed up the likely motivations and intentions of those creating and spreading the false, misleading or manipulated information.

The final step in recognising mis- and disinformation is to consider the impact of the techniques that are being used. This involves understanding how messages, narratives, brands, and intent fit together so that they create an impact. [3]

[3] https://www.cardiff.ac.uk/news/view/2547048-high-profile-western-media-outlets-repeatedly-infiltrated-by-pro-kremlin-trolls

Case study:

The impact of anti-vaccine narratives is an example of a type of cumulative harm that many different but interconnected messages can cause. Antivax narratives predate Covid-19 and have previously been falsely associated with the risk of autism in children alongside many other conspiracies. False and misleading information targeting Covid-19 has for example suggested microchip implants, magnetising of people's arms, effects on fertility, that Christian Eriksen's heart attack during Euro 2020 was caused by the Pfizer vaccine, and of course the monkey memes.

Weighing up impact is challenging, but in this case overall vaccination figures would provide an indication of how the public is responding to the contested information environment. In other cases, it may be possible to obtain polling or survey data. However, in many cases the only available data will be the comments and discussions that you see online in conjunction with the mis- and disinformation. Since these debates can easily be manipulated – indeed, comments sections of newspapers and blogs can be the targets of disinformation campaigns just as easily as comments on social media posts[3] – understanding the impact often comes down to focusing upon your organisational priorities and responsibilities.

You can learn more about how to weigh up the possible risks and impact of mis- and disinformation in Section 4: Impact Assessment.

Summary

This section has covered:

- the differences between **mis-**, **dis-** and **malinformation.**
- the **messages** that are spread and whether they are false, misleading or manipulated.
- the ways in which messages fit within and construct **narratives**
- the **brand** identities of those spreading the messages and narratives
- some indications of the **intent** of the accounts that spread the messages and narratives.
- a sense of the possible or likely **impact** of the techniques, messages and narratives

The remaining components of RESIST 2 will support you in deciding whether to respond, and if so how to craft that response.

Early Warning

This section will help you to answer the following questions:

- how do I focus digital monitoring on **risks**?
- how can I use digital monitoring to **protect my priority issues, tasks and audiences**?

2.1 Focus your monitoring on risk

Monitoring of traditional and digital media has improved greatly since the days of cutting out interesting stories from newspapers with scissors and storing them in a binder. There are now multiple monitoring services available, which are managed by specialised government teams, purchased from analytics companies, or gathered by yourself. They range from very simple data (number of likes, shares and comments) to advanced (big data, sentiment analysis and network analysis).

You will already have regular access to some monitoring products. **However, many will not be focused specifically on the risks posed by mis- and disinformation**.

The existing monitoring you receive should give a good sense of your key audiences, influencers and of the broader debates that relate to your priority policy areas. You should use this monitoring to gain an understanding of:

- digital debates that are taking place in relation to your organisation and its work;
- the main attitudes held by key influencers and audiences;
- how influencers and segmented audiences engage on digital platforms with your organisation and its work
- changes in trends over time

The value of this knowledge is that it enables you to improve your preparedness for handling mis- and disinformation. It can offer early warnings of potential threats and risks, and give a sense of what is normal and what might involve manipulation of debates according to **FIRST principles**. Monitoring can help you to better understand where to look, and what to look for. You can supplement it with your own specialist insights using tools such as:

Platform analytics

Each social media platform has an analytics function that provides data on accounts or pages that you own. Platforms that you own pages on are an important source of insight for understanding how people engage with your content.

Google Trends

Shows how frequently terms are searched for on Google. The results can be broken down by time, country, and related queries to focus attention on a specific timeframe, location, and/or topic. This is useful for revealing spikes in interest and can help guide your attention to specific days, locations or topics where interest in a debate has changed.

TweetDeck

Create a Twitter dashboard to follow multiple timelines, accounts and search terms in real time. Note that you can monitor accounts and keywords in Tweetdeck without being a follower. Available at tweetdeck.twitter.com.

Browser extensions

There are a number of apps that can be added to your browser to speed up or even automate functions such as translation, image searches and taking screenshots. This is especially useful for speeding up simple tasks that you need to do often.

Lots more tools are available for a wide number of platforms and in multiple languages.

2.2 Protect your priorities from mis- and disinformation

The next step is to develop contingency planning around the risks of mis- and disinformation. The below grid will help you to figure out some of the important areas where misleading or manipulated information can provide significant risks to your organisation and the people depending on its services. Use monitoring data to answer the below questions. You can also work with colleagues to brainstorm or "red team" the risks. **This information can help you to increase your organisation's resilience to mis- and disinformation**, by for example improving campaign planning, better directing monitoring, raising awareness of vulnerabilities, and providing early warning of potential threats.

	Our priorities	Areas of risk
Objectives to protect	What are our **priority policy areas and responsibilities**?	What are the prevailing attitudes in these areas that could be harnessed for mis- and disinformation? What types of mis- or disinformation could be particularly harmful to our priorities and our audiences?
Information to protect	What are our **key messages and narratives**?	What misleading or manipulated information is being spread? What are the main messages and narratives we should be aware of? What is untrue or misleading about them?
Brands to protect	What are the **core values** that we stand for?	What values and interests do we and our partners wish to project? What types of mis- or disinformation could undermine our credibility, engagement, or ability to deliver results?
Audiences to protect	Who are the **key stakeholders and audiences** affecting or dependent on our policy areas?	What are their values and interests? Who do they communicate with and listen to? Which parts of their relationship with my organisation are susceptible to mis- and disinformation?

Summary

Digital monitoring should be focused on your key priorities, influencers and audiences. There are a number of government units that provide analysis and insight products that may be relevant to you. There are also many free and paid tools that can be used to support analysis. You should use combinations of these tools to create a monitoring toolkit that suits your needs.

The purpose of digital monitoring in relation to disinformation is ultimately to help you to **reduce vulnerabilities, plan for risk, and protect your priorities**. This kind of focused planning can help give you an early warning if disinformation appears within your priority policy areas or among key influencers and audiences.

The knowledge that you develop in these steps should be operationalised in the next step: creation of **insight**.

Situational Insight

This section will help you answer the following question:

- what is insight in the context of mis- and disinformation and how should it be used to support a timely response?

By the end of this section, you will be familiar with the basic steps required to understand an insight report containing relevant information for your organisation.

3.1 Turning monitoring into insight

Monitoring becomes valuable when it is turned into **insight**. Insight is a form of analysis that turns **interesting data** into **actionable data**. It answers the question, “So what?”. At its core, insight is about understanding audiences to support communication planning. Insight should be used to:

- baseline/benchmark over time to show change
- identify emerging trends and provide early warning of threats
- understand how mis- and disinformation is distributed to key audiences
- generate recommendations
- provide support for developing and targeting messages and campaigns, including preclearance of lines

Insight usually takes the form of reports that are circulated daily, weekly or ad hoc depending on need. Much of the data can be drawn automatically from the monitoring toolbox or dashboard that you developed in the previous section. A good insight report can be as short as one or two pages: **put the most important information at the top and get to the “so what” quickly**.

Bear in mind that your insight product might be the first time that people in your organisation are exposed to digital monitoring data as a basis for analysing mis- or disinformation. It should be usable as a briefing for special advisers, policy advisers, senior staff and ministers, so explain things clearly by avoiding jargon and using images where possible.

A disinformation insight product should at a minimum include:

- **Key insights and takeouts.** A top line summary including a short commentary explaining the ‘so what’ and setting out your recommendations for action; and
- **Sections on key themes and issues covering**:
  - relevant outputs from your department on priority issues, for example a ministerial announcement
  - examples of disinformation relating to these outputs, including where and how it is circulating
  - key interactions and engagements, for example is the disinformation being dealt with organically, is it being picked up by journalists and influencers and if so which ones?
  - trends and changes in attitudes (and influencers and audiences) over time (this can be combined with any polling data you have)
  - your commentary and recommendations for a response

Insight is important, but you should always weigh it up against known risks (section 2.2) and your own knowledge of what is normal debate in your area of expertise. **Be wary of exaggerating the impact of single examples of misleading or manipulated content, or of falling into the trap of feeling that every false post should be corrected**. It is also worth remembering that niche debates on social media rarely represent broader public opinion. Insight is most valuable when you consider it as an indicator of emerging trends that you use alongside other sources to create an overall assessment of impact. The following section will discuss this further.

Summary

The goal of a disinformation insight product is to share the early warning signals you have gleaned from digital media monitoring with the people who need a situational briefing. As with all monitoring, it can be used in long-term planning, for example in an annual report or as part of a campaign evaluation, for ad hoc issues, or produced on a weekly or even daily basis. They should be short, clear and to the point. *A basic example of an insight product is given in Annex B*.

Impact Analysis

This section will help you answer the following questions:

- how do you determine levels of **confidence** in your assessment of mis- and disinformation?
- what is the likely **impact**?
- how does the mis- or disinformation affect your **areas of responsibility, or ability to achieve your organisation's objectives**?
- how does mis- and disinformation affect your **communication with the public**?
- how does the mis- or disinformation affect your **brand or organisation's reputation**?
- how should I **prioritise**?

The following section provides a number of structured analysis techniques which cover a series of questions that can help to guide your assessment of the goals, impact and reach of potential disinformation you have identified through monitoring and insight. This can be used to help you decide whether to act and, if so, how. The following steps will lead you through these more detailed questions, in ways that help to unpack decision-making and avoid "gut feeling" responses.

	Our priorities	Key questions
Objectives to protect	**Policy areas and responsibilities**	1. Is the mis, dis- or malinformation harmful to our priorities? In what ways?
Information to protect	**Key messages and narratives**	2. What problematic communication techniques and behaviours have you identified? 3. What misleading or manipulated content is being spread? What are the main messages and narratives we should be aware of? What is untrue or misleading about them?
Brands to protect	**Core values**	4. What values and interests do the accounts spreading mis- or disinformation wish to project, and to which target groups?
Audiences to protect	**Key stakeholders and audiences**	5. How widely is mis- or disinformation spreading, and to which target groups?

4.1 What is the degree of confidence?

Structured analysis techniques are a well-established means of standardising assessments and decision making. They are mainly used in assessments where analysts look at different parts of a puzzle and need to share the same process and language. **We draw upon simplified versions of these techniques here because handling mis- and disinformation should not be based on a gut feeling. You need to follow a structured, coherent process using a common language that leads to consistent decisions**.

While it may be tempting to expect clear yes/no answers to many situations, analysis, and the subsequent response to mis- and disinformation, often involves uncertainty. This is because situations are ongoing and there is rarely enough information to be sure of who is behind activities, what their goals are, and what impact their efforts will have. **What we see on social media or in insight reports should be treated as indicators of possible trends rather than firm evidence of public opinion**. Rather, we must consider these activities in terms of **risk and likelihood**.

Careful use of language is crucial to give a nuanced assessment of what appears to be happening. For example, you may have high confidence that a piece of social media content is disinformation, but low confidence in who is ultimately behind it and why they are spreading it. If you are sure, say so. If not, it is helpful to add a small indicator of the likelihood that something is true. It is enough to place a letter in parentheses at the end of a proposition; low confidence [L], medium confidence [M], or high confidence [H]. Note that there needs to be collective agreement before any attribution is made. More advanced users may wish to use additional structured analysis techniques together with the PHIA confidence yardstick to add greater nuance to their assessments; *see Annex C for more details*.

Confidence levels

High confidence [H]: the evidence currently available is sufficient to reach a reasonable conclusion.

- "Digital platforms and researchers have linked this group to a previous information influence operation [H]."

Medium confidence [M]: it is possible to reach a reasonable conclusion based on the available evidence, but additional evidence could easily sway that conclusion.

- "Based on the identity of this account, the networks it belongs to and its previous behaviour, there does not appear to be an intent to mislead [M]."

Low confidence [L]: there is some relevant evidence, but it is taken in isolation or without corroboration.

- "Many of the disinformation posts appear to have been translated from a foreign language or use linguistic idioms that suggest the network is based in a particular foreign country [L]."

4.2 How does mis- or disinformation affect your areas of responsibility?

Government departments provide services to the public according to their areas of responsibility. One method of determining whether to respond to mis- or disinformation is to consider how it impacts your department's responsibilities.

	Our priorities	Key questions
Objectives to protect	**Policy areas and responsibilities**	1. Is the mis-, dis- or malinformation harmful to our priorities? In what ways?

Below is an example of how you might answer these questions by turning your priorities into a matrix. Develop a similar matrix of your own that covers the areas relevant to your priorities. You can then mark which of the priorities are affected by the misleading or manipulated information you have encountered. Include confidence levels to mark how likely a scenario appears to be.

Does it affect the ability of your organisation to do its job?	Does it affect the people who depend upon your services? (be specific; see 4.5)	Does it pose a significant risk to the general public?
Ability to deliver services	Key stakeholders	National security
Reputation (see 4.4)	Key audiences	Public safety
Policy areas/goals	Niche audiences	Public health
Individual staff/staff safety	Vulnerable audiences	Climate of debate (see 4.3)

For example:

Public health

- “Accounts are spreading false information that encourages individuals not to use this potentially life-saving service. There is some evidence of moderate impact among users of the service’s Facebook page [M].”

Reputation

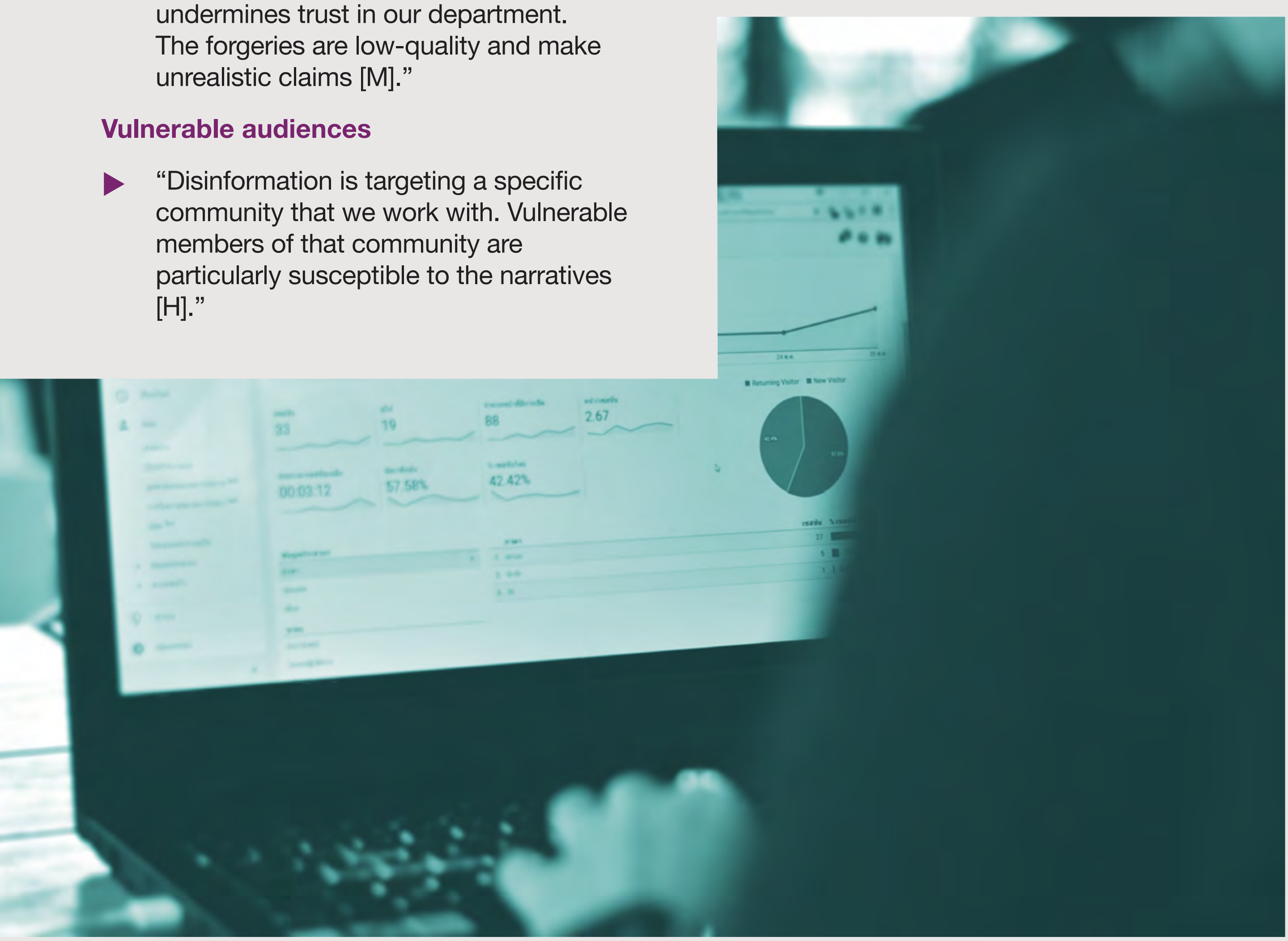
- “A small group of users are spreading forged material that potentially undermines trust in our department. The forgeries are low-quality and make unrealistic claims [M].”

Vulnerable audiences

- “Disinformation is targeting a specific community that we work with. Vulnerable members of that community are particularly susceptible to the narratives [H].”

4.3 How does the mis- or disinformation affect your communication with the public?

The integrity of the information exchanged between governments and members of the public is crucial to a well-functioning society. It may be appropriate to engage with mis- and disinformation if it affects your communication with the public to a degree that you are no longer able to communicate effectively with them, for example to deliver crucial public services.



	Our priorities	Key questions
Information to protect	**Key messages and narratives**	2. What problematic communication techniques and behaviours have you identified? 3. What misleading or manipulated content is being spread? What are the main messages and narratives we should be aware of? What is untrue or misleading about them?

Problematic communication techniques, behaviours and content are easily assessed using the **FIRST indicators**. The steps under **Recognise** will guide you through this information collection process and help you to develop succinct answers to these questions.

Examples:

- "Here are multiple sockpuppet accounts being used that have previously been active spreaders of disinformation [M]."
- "Narratives use symbolism and rhetoric, including historical revisioniandm and whataboutism, to agitate left-leaning groups [H]."
- "Much of the trolling seems to be coordinated off-platform to drown out specific voices at specific moments [H]."

4.4 How does the mis- or disinformation affect your brand?

In many cases, misleading or manipulated information will assert values that compete with those of government departments for credibility. For example, those spreading the misleading or manipulated information may claim to be protecting the public, fighting for a good cause, or have access to unique information sources. They will often assert this identity in order to target it at specific audiences. **It will not always be possible to know who is behind an account or network, but it should be possible to assess the account to better understand their values, interests and patterns of behaviour.**

	Our priorities	Key questions
Brands to protect	**Core values**	4. What values and interests do the accounts spreading mis- or disinformation wish to project, and to which target groups?

Use the answers from the previous questions to form an informed opinion about how the mis- or disinformation relates to your organisation, your values and priorities. Then check the information you have collected to see if you can answer the following about the accounts and networks creating or sharing the mis- or disinformation:

- who do they claim to be?
- what values do they claim to stand for?
- who are their target groups?
- is their behaviour consistent with their profile?
- what is their track record of liking, sharing and commenting on posts?
- do they have a stake in the issue?
- how do they engage with other accounts, particularly those that disagree with them?
- do they seek accurate information and correct mistakes?

4.5 What is the likely reach of the mis- or disinformation?

You should make an assessment of how extensively you believe the mis- and disinformation will be engaged with. Is it likely to disappear within a few hours or does it have the potential to become tomorrow's headlines?

	Our priorities	Key questions
Audiences to protect	**Key stakeholders and audiences**	5. How widely is mis- or disinformation spreading, and to which target groups?

Exposure/reach	Likelihood
Little interest: very limited circulation and engagement	
Filter bubble: some engagement within niche audiences with similar worldview / automated circulation	
Trending: some discussion online, may include open debate and rebuttals	
Minor story: some reporting on mainstream media	
Headline story: affecting day-to-day operations	

Examples:

- “a small number of accounts are particularly influential in this network. At least one is deliberately spreading false information, others are more balanced but still share false information [H]”
- “members of this group have posted that they will ignore all government guidance on this topic. They are using hashtags that reach large audiences, which means that interesting content could generate traction [M]”
- “a number of public figures have now commented on this content. It is likely to reach mainstream news [H]”

4.6 How should I prioritise the mis- and disinformation?

Once the previous steps are completed, you should be able to assign a priority level to the mis- or disinformation. Is misinformation at risk of causing harm to the public, or are there indications that it will be ignored? Is disinformation likely to become part of a major international crisis, like the Salisbury poisoning, or endanger life, such as misinformation around COVID-19 vaccines, or is it enough simply to monitor developments?

Below are three example priorities: high, medium and low. You may need to develop your own criteria for prioritising disinformation based on your specific needs and experiences. The principle is that the goal, impact and reach should inform how urgently you prioritise the case.

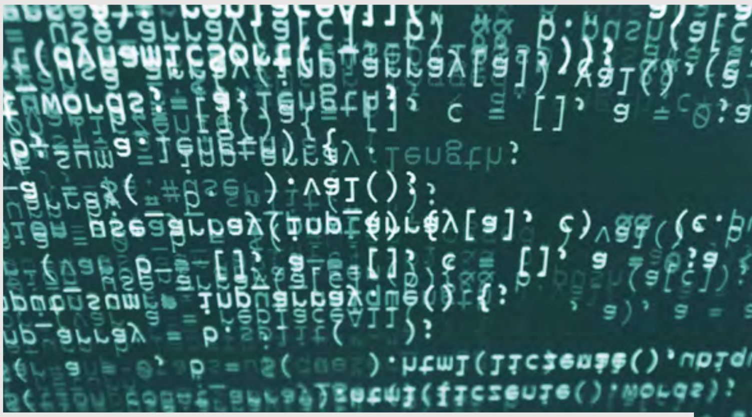
Keep your assessment outcome-focused, i.e. does what you are seeing represent a significant obstacle to achieving your priorities? If not, it should be lower priority. **The role of government is not to respond to every piece of false or misleading information**. You should not take on the role of arbiter of truth or moderator of public debate. **A prioritised response is one in which there is a clear and compelling need to protect government objectives, information, brands and/or audiences**.

	Description	Actions	Internal audiences	Tools
High	Significant risk to the public, e.g. health or national security and has a high likelihood of making headlines. Much of the evidence is high confidence and builds a clear picture. It requires immediate attention and escalation.	Make senior staff, SpAds/ policy advisers and other parts of government aware of issue and its priority. Share insight and analysis. Prepare quickly for a cross-government response.	Senior staff Wider government	Share insight Briefings Prioritise short-term comms

Example:

- Following the poisoning of two UK residents in Salisbury, Russian news sources began a campaign of harassment of the investigating services and UK government. Early warnings from digital media enabled the production of briefings for senior staff across government to prepare for a disinformation crisis.

	Description	Actions	Internal audiences	Tools
Medium	Negative effect on a policy area, departmental reputation or a large stakeholder group and is trending online. The evidence indicates a potential for harm if left unchallenged. It requires a response.	Make senior staff and SpAds/policy advisers aware of issue. Share insight and analysis within department. Investigate the issue and prepare press lines based on known facts.	Senior staff Policy advisers Monitoring and analysis teams	Insight Briefings Press lines Prioritise short and medium-term comms



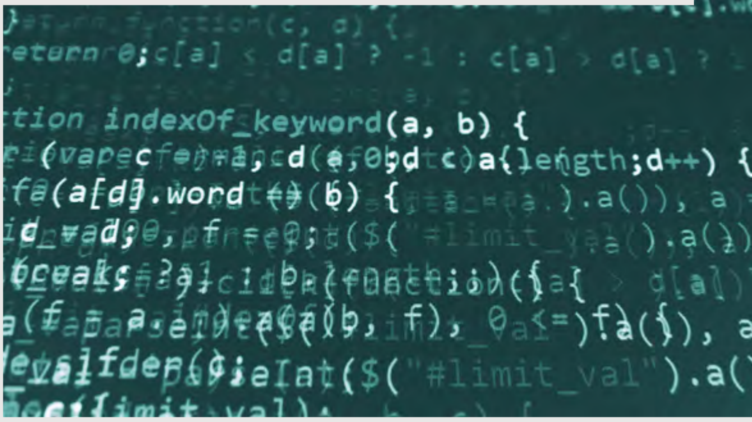
Example:

- A trade press with limited circulation misleadingly claims that a recent parliamentary vote determined that animals have no feelings. Early warning assessment highlights a risk that the narrative may be picked up by mainstream press. Insight, briefings and press lines are prepared either to proactively correct the story or to prepare for possible mainstream interest in the policy area.

	Description	Actions	Internal audiences	Tools
Low	Potential to affect the climate of debate about e.g. a department's work and has limited circulation. The evidence is of mixed quality. The debate should be routinely followed but intervention is unnecessary/ undesirable.	Share insight and analysis with in media department. Investigate the issue and prepare press lines/ narratives based on known facts. Conduct a baseline analysis of debate and track any changes.	Comms officers Monitoring and analysis teams	Insight Press lines Baseline analysis Prioritise medium and long-term comms

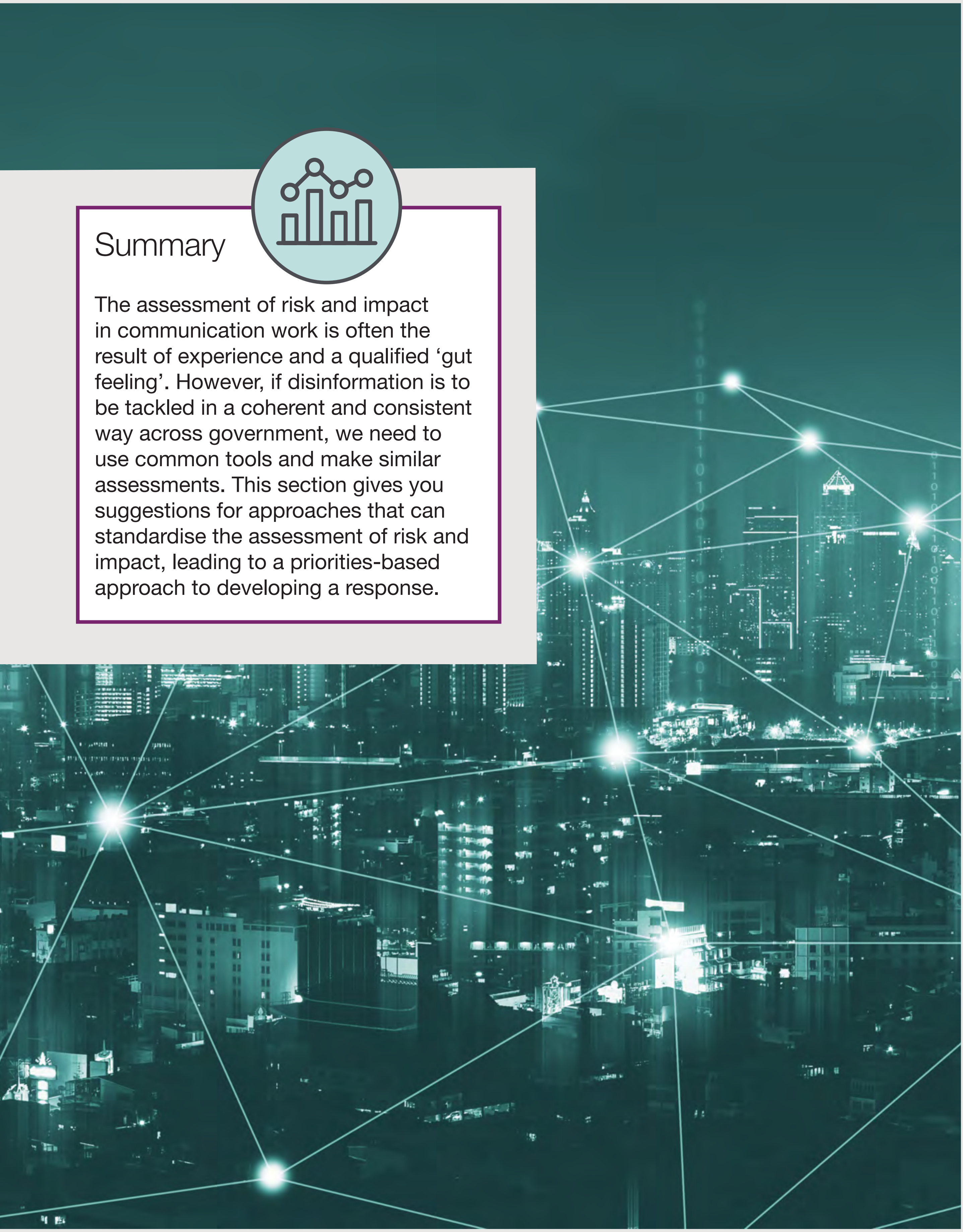
Example:

- A conspiracy theory has emerged holding the UK government responsible for a major public safety incident. The theory is only being circulated by fringe groups known for anti-government sentiment, and runs counter to current mainstream debates. Insight and press lines are prepared, but no response is made for the time being. The area is monitored and baseline analysis is used to spot any sudden changes in the climate of debate.



Summary

The assessment of risk and impact in communication work is often the result of experience and a qualified ‘gut feeling’. However, if disinformation is to be tackled in a coherent and consistent way across government, we need to use common tools and make similar assessments. This section gives you suggestions for approaches that can standardise the assessment of risk and impact, leading to a priorities-based approach to developing a response.



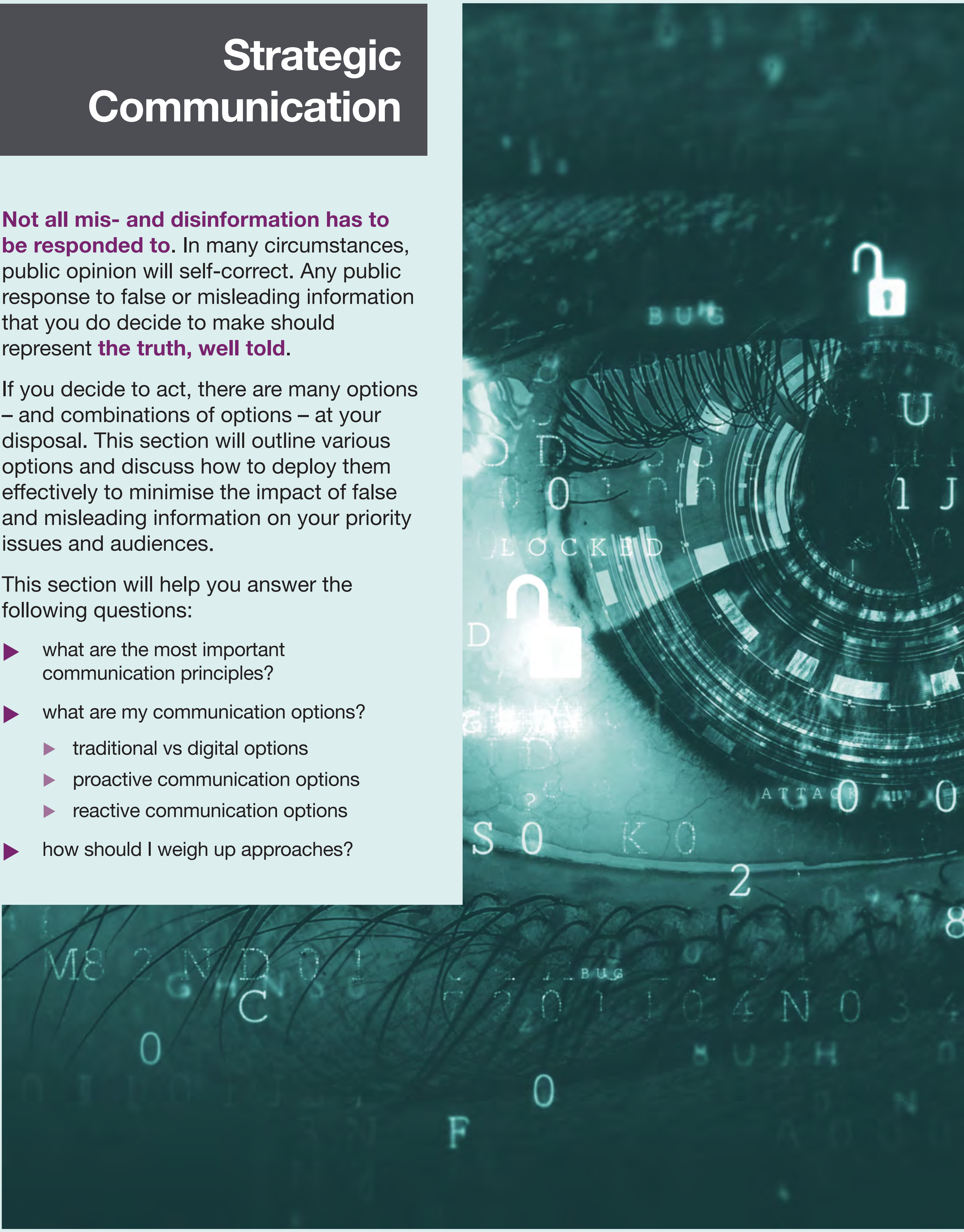
Strategic Communication

Not all mis- and disinformation has to be responded to. In many circumstances, public opinion will self-correct. Any public response to false or misleading information that you do decide to make should represent **the truth, well told**.

If you decide to act, there are many options – and combinations of options – at your disposal. This section will outline various options and discuss how to deploy them effectively to minimise the impact of false and misleading information on your priority issues and audiences.

This section will help you answer the following questions:

- what are the most important communication principles?
- what are my communication options?
  - traditional vs digital options
  - proactive communication options
  - reactive communication options
- how should I weigh up approaches?



5.1 Follow communication best practice

The Organisation for Economic Cooperation and Development (OECD) has used the collective expertise and learnings of its members and partners to develop good practice principles to help address mis- and disinformation.

The resulting draft *OECD Principles of Good Practice for Public Communication Responses to Mis- and Disinformation* has been devised to:

- inform government policies and communication that resonate with citizens' needs and leverage stakeholders as part of a whole-of-society approach
- empower communicators through institutionalised approaches that are public-interest driven and evidence-based
- mitigate the spread and effects of mis- and disinformation through building capacity for timely, preventive and forward-looking efforts to respond to problematic content

The OECD Principles of Good Practice for Public Communication Responses to Help Counter Mis- and Disinformation

Transparency

Governments strive to communicate in an honest, clear and open manner, with institutions comprehensively disclosing information, decisions, processes and data within the limitations of relevant legislation and regulation. Transparency, including about assumptions and uncertainty, can reduce the scope for rumours and falsehoods to take root, as well as enable public scrutiny of official information and open government data.

Inclusiveness

Interventions are designed and diversified to reach all groups in society. Official information strives to be relevant and easily understood, with messages tailored for diverse publics. Channels and messages are appropriate for intended audiences, and communication initiatives are conducted with respect for cultural and linguistic differences and with attention paid to reaching disengaged vulnerable, underrepresented or marginalised groups.

Responsiveness

Governments develop interventions and communications around the needs and concerns of citizens. Adequate resources and efforts are dedicated to understanding and listening to their questions and expectations to develop informed and tailored messages. Responsive approaches facilitate two-way dialogue, including with vulnerable, underrepresented and marginalised groups, and enable an avenue for public participation in policy decisions.

Whole-of-society

Government efforts to counteract information disorders are integrated within a whole-of-society approach, in collaboration with relevant stakeholders, including the media, private sector, civil society, academia and individuals. Governments broadly promote the public's resilience to mis- and disinformation, as well as an environment conducive to accessing, sharing and facilitating constructive public engagement around information and data. Where relevant, public institutions coordinate and engage with non-governmental partners with the aim of building trust across society and all parts of the country.

Public-interest driven

Public communication should strive to be independent from politicization in implementing interventions to counteract mis- and disinformation. Public communication is conducted as separate and distinct from partisan and electoral communication, with the introduction of measures to ensure clear authorship, impartiality, accountability and objectivity.

Institutionalisation

Governments consolidate interventions into coherent approaches guided by official communication and data policies, standards and guidelines. Public communication offices benefit from adequate human and financial resources, a well-coordinated cross-government approach at national and sub-national levels, and dedicated, trained and professional staff.

Evidence based

Government interventions are designed and informed by trustworthy and reliable data, testing, behavioural insights and build on the monitoring and evaluation of relevant activities. Research, analysis and learnings are continuously gathered and feed into improved approaches and practices. Governments focus on recognising emerging narratives, behaviours, and characteristics to understand the context in which they are communicating and responding.

Timeliness

Public institutions develop mechanisms to act in a timely manner by identifying and responding to emerging narratives, recognising the speed at which false information can travel. Communicators work to build preparedness and rapid responses by establishing coordination and approval mechanisms to intervene quickly with accurate, relevant and compelling content.

Prevention

Government interventions are designed to pre-empt rumours, falsehoods, and conspiracies to stop potentially harmful information from gaining traction. A focus on prevention requires governments to identify, monitor and track problematic content and its sources; recognise and proactively fill information and data gaps to reduce susceptibility to speculation and rumours; understand and anticipate common disinformation tactics, vulnerabilities and risks; and identify appropriate responses, such as “pre-bunking”.

Future-proof

Public institutions invest in innovative research and use strategic foresight to anticipate the evolution of technology and information ecosystems and prepare for likely threats. Counter-misinformation interventions are designed to be open, adaptable and matched with efforts to build civil servants’ capacity to respond to evolving challenges.

5.2 What are my communication options?

The following sections will outline some of the most common communication options that are available to communications teams when you choose to respond to mis- or disinformation. First is a short discussion about **traditional** versus **digital communication** methods, followed by **proactive communication options**, which is where you attempt to push back on false or misleading information before it has become widespread; and **reactive communication options**, which are about counteracting a potentially harmful message, narrative, actor, or objective.

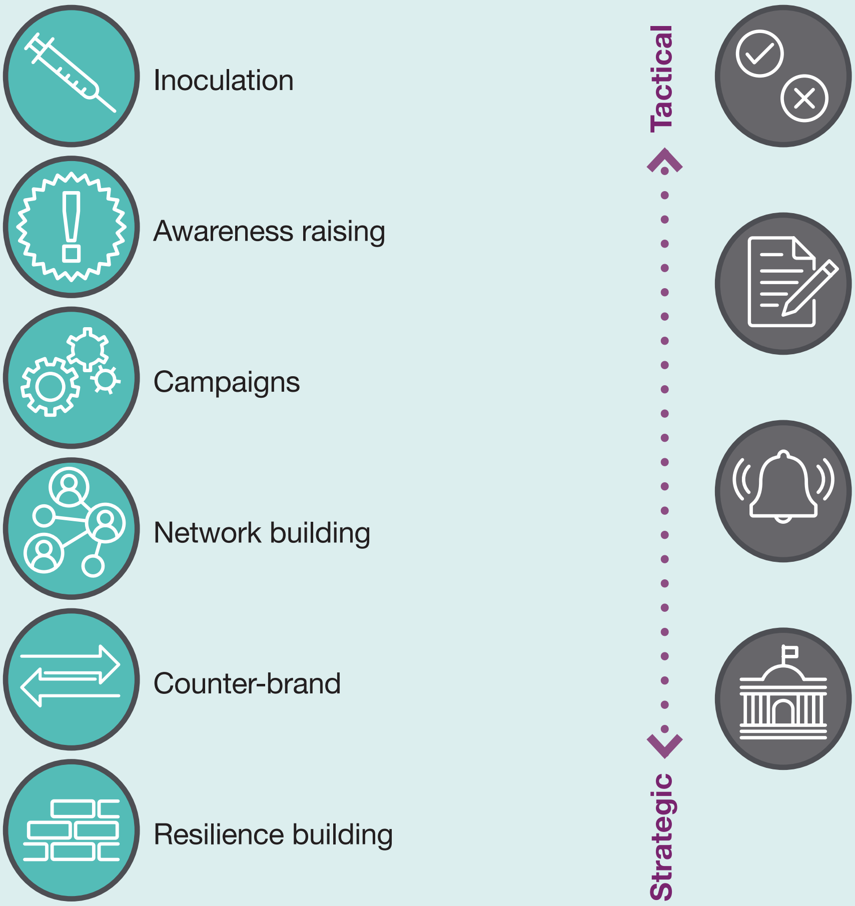
Communication channels

Efforts to inform and engage the public

- **Face-to-face engagement**
- **Media relations**
- **Social media engagement**

Proactive efforts to pre-bunk, raise awareness, and shape the information environment

Reactive efforts to debunk, counter, and restore the information environment



5.3 Using traditional and digital communication channels

Many of the communication options for mitigating the impact of mis- and disinformation are found in general approaches to government communication, such as media relations and social media engagement. This work can be both proactive and reactive, insofar as it shapes the information environment and sometimes includes responses and corrections to false or misleading interpretations of information. Common to these approaches is a reliance on the OECD Principles to ensure clear and credible communications capable of informing the public so that they have access to correct information, as well as changing perceptions and behaviours where appropriate.

Face-to-face engagement

Media relations

Social media engagement

Face-to face engagement

When should it be used? In some cases, the most credible communication method is simply to talk. However, it is also difficult to do this at scale. It is most relevant when for example a key stakeholder or influencer requires a tailored briefing, or a group or community needs to see boots on the ground.

How does it work? Communication via media or social media allows for the scaling up of what is in essence the direct communication between you and a member of the public. Face-to-face means returning to the core principles of building trust by relating directly at a human level (even if, in some cases it may not literally be face-to-face, e.g. a phone call). Note that such meetings can also be used as content for traditional and social media. Techniques include:

- briefing: speaking to individual or small groups in order to explain the context of an issue, can be on or off-the-record
- stakeholder meetings: engaging with key stakeholders so that they can reach out to their communities, for example by briefing a religious leader or teacher
- community outreach: a site visit, including for example a spoken statement or town hall style meeting to answer common questions

What is its impact? In-person communication is often considered the most credible because it removes the sense of distance, as well as layers of mediation by e.g. journalists, from communication. It also allows for the translation of at times quite abstract policy issues into everyday language. Even if it can only be deployed selectively, content from face-to-face meetings can be used to support other communication techniques at scale, such as sharing clips of meetings on social media.

Media relations

When should it be used? Media relations is one of the traditional approaches to communications work. It is often most relevant when issues affect mainstream audiences, such as readers of newspapers and consumers of television news. This category also includes specialist outlets (in areas such as current affairs, economics, and trade magazines) who may be particularly influential to different target groups.

How does it work? Many of the tools listed below are designed to ensure that the government can get its message across so that the public has access to the right information. Media act as mediators, adding their own interpretations before delivering stories to their audiences. Tools include:

- press release - a written statement, sometimes including interviews and data, that can be used as a primary source of information by journalists
- statement - an on-the-record statement, for example by an elected official, senior government official, or other credible voice depending on the issue
- earned media coverage - offering speakers to media to encourage media coverage about a specific topic, announcement or issue
- background briefing - speaking to journalists or other stakeholders to explain the context of an issue, can be on or off-the-record
- promoting friendly voice - third parties can be a valuable means of building bridges to sceptical audiences, particularly if they are seen as an objective source of credible information

What is its impact? Media relations aim to shape media coverage on a day-to-day level. Effective media handling is part of typical communications work, and many of these tools will be valuable for ensuring the correct information is available to the public in the first place, in order to reduce areas of doubt where false and misleading narratives can develop.

Social media engagement

When should it be used? Some audiences are most effectively and directly reached through social media. If the decision has been made to develop a department’s presence on social media, for example through a Facebook page or Twitter account, this can provide additional opportunities to understand and engage with the public using targeted communications. Owned channels on social media allow for more detailed knowledge of a target audience and more control over the message compared to media outreach. If you want to speak directly with the public, social media channels can offer a rapid, cost-effective means of engaging.

How does it work? Analysis of digital media channels, including social media, can reveal for example that certain target audiences are searching for reliable information sources or are spreading false or misleading materials. Engagement can involve creating relevant content and targeting it to those audiences targeted by or most impacted by false information through owned social media channels. Content can be tailored and directed to selected audiences. The UK Government’s Rapid Response Unit (RRU) has created a simple process “FACT” for conducting this kind of work:

- Find - use media monitoring sources, including previous steps of RESIST, to identify false and misleading information
- Assess - weigh up the risk of false or misleading messages and narratives, and consider how they are spreading
- Create - develop engaging content
- Target - weigh up whether to respond directly to posts, contact the author/publisher, use targeted advertising, and/or develop segmented content

What is its impact? Effective social media engagement can quickly identify trends and allow direct contact with highly targeted audiences without additional mediation through, for example, journalists. This enables a variety of positive outcomes including attitude and behaviour change, building resilience, shaping the information environment, showing a willingness to engage in debate, as well as a credible rapid response, all focused on using insights from data to reach "eyeballs". *For more guidance see FACT (see Annex D) and Section 6 on Tracking Outcomes*.

Who will approve content?

Who absolutely needs to review and sign off on your content? For example your Head of News and/or Head of Communications, followed by the relevant Special Adviser. You should secure contacts for each of these positions who will be able to respond on their behalf if your regular contacts are absent.

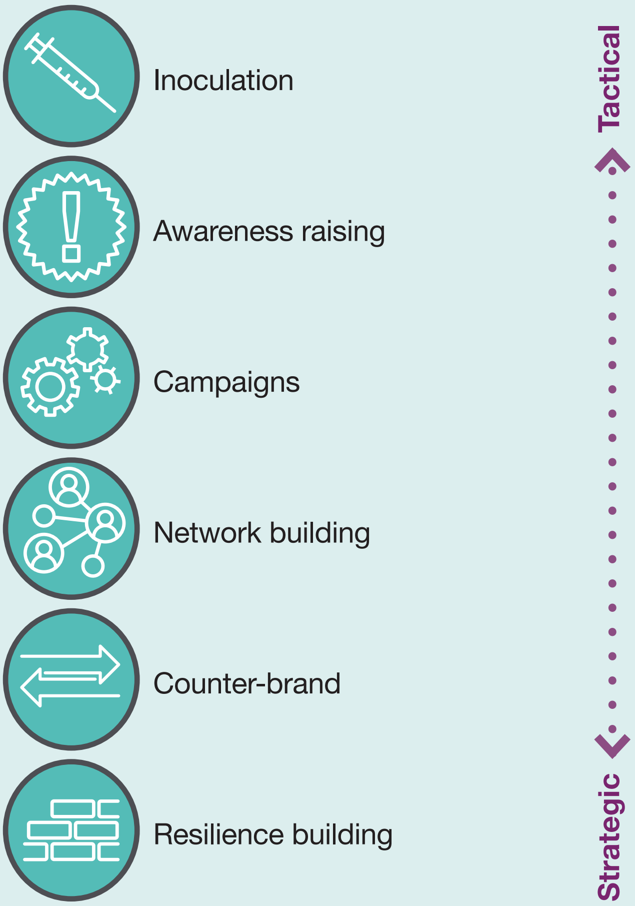
If you have been creating and sharing situational insight in the form of monitoring reports – as set out in the situational insight section – this will help people to understand the context in advance. They will already have an understanding of the disinformation affecting your organisation or its policy areas, which will help when you need to build a case to respond and when you want to clear content quickly.

Can lines be pre-cleared? If insight is already available into an emerging disinformation trend, it may be possible to preclear some press lines before they are needed. For example, some government departments have weekly routines to pre-clear lines via their subject experts several days before an event is likely to make the news, in case of negative reporting.

It is important to note that traditional and social media are not mutually exclusive. News media articles are often widely shared, discussed and debated on social media. Social media posts and debates are sometimes reported on, or integrated into stories, by traditional media. Digitisation of news media means that many news platforms integrate social media functions, such as comment fields, into their formats; similarly, sharing news is part of most social media platforms. The approaches referred to here are therefore cross-cutting tools that can be applied in many situations, including as part of the response to false and misleading information.

5.4 Proactive communication options

Proactive communications is about taking steps to push back on false or misleading information before it has become widespread. It includes efforts to pre-bunk, raise awareness, and shape the information environment in order to minimise risk of harm to the public.



Inoculation

When should it be used? When you want to proactively counteract ("pre-bunk") false messaging before it has become widely spread.

How does it work? The PROVE[4] framework has been developed and tested by researchers to ensure effective evidence-based communication. It can be used to develop clear, informative messaging using scientific evidence to explain nuance and uncertainty around complex issues. It is not supposed to advocate or persuade, only inform.

- Pre-bunk: anticipate mis- and disinformation through media monitoring and risk assessment and prepare to pre-emptively warn the public
- Reliably Inform: trust is built by informing openly rather than persuading. This means ensuring that information reflects expertise, honesty, and good intentions
- Offer balance: do not skew or ignore evidence, but rather ensure balance in how evidence is presented
- Verify quality: be open about the quality of the underlying evidence so that the credibility of the information is clear
- Explain uncertainty: disclose any uncertainties, gaps and risks with the current evidence

What is its impact? Inoculation helps to strengthen resilience by informing the public about an issue at risk of false or misleading information, and preparing and empowering their ability to engage with that content. Experiments show that target audiences find PROVE content to be more engaging, more trustworthy, and more interesting to read than other content. It also elicited less negative emotional responses (such as anger), and less negative cognitive responses[5].

[4] https://www.nature.com/articles/d41586-020-03189-1

[5] According to experiments conducted on behalf of HMG by Cambridge University

Awareness raising

When should it be used? When you want to proactively shape public debate about issues likely to be subjected to mis- and disinformation.

How does it work? Information and awareness campaigns use proactive communications to nudge, advocate, influence, and persuade target groups to behave in a manner that ultimately benefits society. They may involve slogans ("Stay Home. Protect the NHS. Save Lives"), instructions ("Don't drink and drive"), or simple processes ("See it. Say it. Sorted"), as well as more traditional public affairs and advocacy work. They often draw upon rudimentary storytelling and establish narratives. The approach can involve:

- awareness raising and public information campaigns
- brand and narrative development
- storytelling around ongoing trends, threats and risks
- specific warnings, advice and guidance to the public
- use of audience research to develop targeted nudge, advocacy, influence and persuasion tactics
- publishing of evidence or research to support a wider advocacy effort

What is its impact? Awareness raising efforts are capable of shaping public debate about an issue in order to change behaviour. The principles of public information campaigns are well-established in for example OASIS (*see Annex E*).

Campaigns

When should it be used? All communications should be viewed in the context of a wider campaign, for example, what do we want to achieve and where does it fit with other activities[6].

How does it work? A campaign is a planned sequence of communications and interactions that uses compelling narratives over time to deliver a defined and measurable outcome, such as behaviour change.

- Objectives: Start with the policy aim and develop communications objectives that will deliver this
- Audience/Insight: Use data to understand your target audiences
- Strategy/Ideas: Plan the campaign strategy including messaging, channels, and partners/influencers
- Implementation: Plan the campaign implementation by allocating resources and setting out timescales for delivery
- Scoring/Evaluation: Monitor outputs, outtakes and outcomes throughout your campaign

What is its impact? Effective communication planning, coordination, and measurement delivers added value to all communications. While not mis- or disinformation specific, it is important to continue to use OASIS in such issues as a matter of best practice.

[6] See https://gcs.civilservice.gov.uk/guidance/marketing/delivering-government-campaigns/guide-to-campaign-planning-oasis/ for further guidance

Network building

When should it be used? If an issue is likely to persist in the medium or long term, it is important to develop networks capable of shaping an effective response over time.

How does it work? Networks of like minded allies and organisations provide a safe space for solving problems together. Each party can work from its relative strength; e.g. governments can use their convening power and policy/legislative capabilities; NGOs their credibility and subject expertise; researchers their ability to generate evidence and knowledge; journalists can use their investigative capabilities and connections to audiences. Networks can be within a government, within a country embracing cross-sector stakeholders, and internationally. They can:

- build and maintain networks of experts and policymakers
- catalogue known mis- and disinformation vectors and actors
- support the creation of independently-verified repositories of factual information
- provide specialist training and workshops
- shape consensus among key stakeholders around problems and solutions
- support development of long-term, sustainable relationships with target audiences

What is its impact? Networks should ultimately share knowledge and expertise in order to strengthen the community against a threat. For example, the ability of the UK government to work with partner governments, researchers, NGOs, and journalists during the aftermath of the Salisbury poisoning led to coordinated sanctions, published independent research into the associated disinformation campaigns, and a credible independent exposé of the perpetrators. Such impact would not be possible without strong networks.

Counter-brand

When should it be used? When you want to expose the behaviour of a persistently hostile actor who spreads false or misleading claims. Note that this may require political approval.

How does it work? Counter-brand refers to a range of communicative activities that collectively seek to ensure a reputational cost to actors who persistently spread false, misleading and harmful information. It ensures that the target audiences of false messages and narratives are better informed about the interests, values and behaviour of the sources behind mis- and disinformation. Techniques include:

- use branding to explain and project your identity and values, and to explain the negative behaviour of persistently hostile actors
- expose contradictions in the values, identity, interests, and behaviour of the sources of false and misleading narratives
- use satire where appropriate for target audiences
- address audience hopes/needs and provide an alternative vision
- make a technical or political attribution
- work with credible partners to attribute disinformation sources and/or deliver counter-brand content

What is its impact? Counter-brand communications represent a strategic effort to engage with a threat actor and hence are not to be taken lightly. Efforts to counter-brand Daesh involved revealing the realities of life under a brutal regime, thereby exposing the contradictions between the disinformation and the reality. Attributions, for example in the statement by the then Prime Minister Theresa May that Russia was responsible for the Salisbury poisoning, lead to geopolitical consequences such as sanctions. Ultimately, the anticipated impact is to raise the costs of spreading false and misleading information for a specific actor, and to increase public resilience to the sources of false information.

Resilience building

When should it be used? For long term efforts aimed at increasing the ability of target audiences to critically-engage with false or manipulated information.

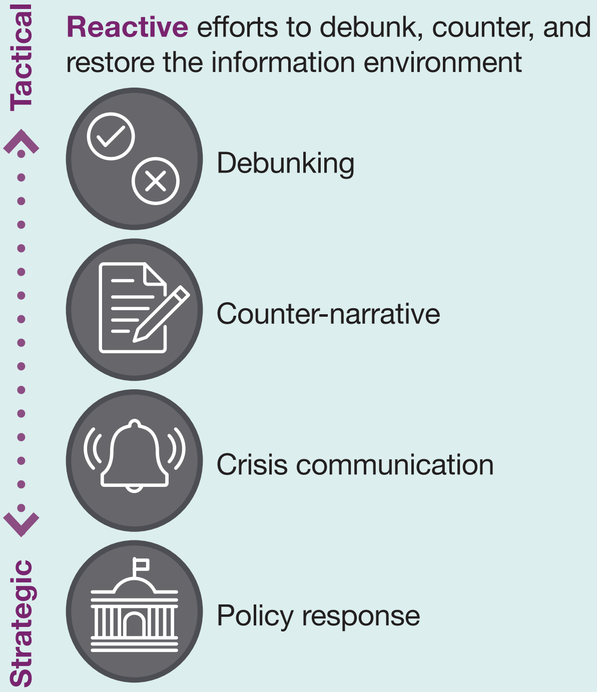
How does it work? The aim of resilience building and media literacy initiatives is to empower people to better understand how false information can be spread on and offline, so that they can more effectively engage with what they see, read, and hear. Targeted education and training can develop techniques such as source criticism, identifying bias, using logic and argumentation, and interpretation of data.

- media literacy education
- training to detect flawed and fallacious argumentation
- interactive training content
- participatory campaigns
- gamification
- use of satire to encourage critical thinking
- outreach to schools and vulnerable groups
- engagement with digital platforms to include media literacy into their product design

What is its impact? Improving the resilience of the population to manipulation through false or misleading content. This makes it harder for misinformation to spread, by for example encouraging people to check sources before they share social media posts. A more informed and media-savvy population is less likely to be drawn in to disinformation campaigns.

5.5 Reactive communication options

Reactive communications is about counteracting false or misleading information so that the public is empowered to make an informed choice. These options are specifically designed to counteract a potentially harmful message, narrative, actor, or objective.



Reactive efforts to debunk, counter, and restore the information environment

Debunking

When should it be used? When false or manipulated information is circulating and you wish to counteract the impact of the false information by asserting the truth.

How does it work? The principle is that false or manipulated information should not be left unchallenged. Counter-messaging involves engaging at the level of individual messages can include efforts to for example correct the record, debunk a myth, or fact check. Traditional media handling often involves correcting false or misleading information, however in the mis- and disinformation field there are specific best practices that have been established by leading experts of health and climate change misinformation[7]. These include a simple process for constructing messages to as to maximise clarity and impact:

- Fact: lead with the truth
- Myth: point to false information
- Explain fallacy: why is it false?
- Fact: state the truth again

What is its impact? Fact checking and debunking are widespread practices that perform an important role in ensuring that false information does not go unanswered. Until recently, we have emphasised the risk that repeating misinformation may backfire since it helps to reinforce the false messages. Recent studies show that this risk is generally lower than of not responding at all[8]. A problem with this approach is however that it is time-consuming and deals with individual messages, which means it can be challenging to use at scale and can feel like whack-a-mole. Some NGOs, nonprofits and media houses have created initiatives that cover different issues and markets, and it can be useful to collaborate with them to increase the impact of governmental counter-messaging.

- Stick to the subject. False or misleading messages are often designed to draw you into wider narratives that are off-topic and you are unlikely to want to debate. There may be a “kernel of truth” to false claims that are deeply personal and will be difficult to separate from the issue at hand. Be aware of how messages fit within narratives and stick to the issue that is connected to your priorities.
- Use facts, examples and evidence wherever possible. Independent evidence confirmed by multiple credible sources is ideal. However, be aware that some legitimate sources will be discredited from the perspective of certain audiences.
- Don’t engage trolls. Watch out for combinations of rhetorical devices such as social proof, strawman, whataboutism, ad hominem, claims of no evidence, etc. If somebody repeatedly uses these techniques in their online engagements, they are likely not interested in correcting false or misleading information.

Counter-narrative

When should it be used? When false or misleading narratives, stories, conspiracies, or myths develop into shorthand, or a delivery mechanism, for mis- and disinformation content.

How does it work? Countering narratives involves exposing falsehoods and contradictions in how important issues are explained to different

[7] https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf

[8] https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf

audiences and where possible replacing them with a more truthful narrative. Narratives are sometimes seen as being in competition, hence the notion of a "battle of narratives".

- Focus on asserting your preferred narratives and ensure consistency across communication activities
- Dispute incorrect assumptions in false narratives, including using debunking where appropriate
- Tell your story using evidence to support core elements
- Tell their story your way

What is its impact? Countering mis- and disinformation narratives helps to shape the understanding of complex issues, which can make audiences more resilient to individual false and misleading messages. For example, false conspiratorial narratives around the causes of Covid-19 provide a delivery mechanism for a range of other conspiracies about vaccines that can, by extension, cause public harm. Countering the narratives at source impacts upon the viability of associated mis- and disinformation messages, rather like pulling up the roots of a weed rather than simply removing the leaves.

- Be wary of the battle of narratives. While different narratives may be in competition for attention, the biggest threat to a narrative is the "say-do gap", also known as the difference between the stories we tell ourselves, and our realities. Often, the "winning" narrative is not the "best" story, but that which most compellingly explains the realities of target audiences.

Crisis communication

When should it be used? Crisis communication is about dealing with misconceptions and crises as they occur. The truth is not always immediately clear. These approaches are not mis- and disinformation specific, but are nonetheless reliable and well-established means of ensuring that accurate information reaches target audiences as it becomes available.

How does it work? Communicators have a range of tools to deal with the majority of crises that become newsworthy or impact lives. The tools mainly involve providing information to target audiences via key influencers, such as journalists, editors, thought-leaders, community leaders, experts, and other stakeholders.

- Holding statement: acknowledgement of an emerging issue with the promise that more information will be on its way
- Statements and interviews: on-the-record information that can be used as a primary source for media or target audiences
- Stakeholder engagement and briefings: speaking to journalists or other stakeholders in order to explain the context of an issue, can be on or off-the-record
- On the ground activity: participating in community meetings, providing visible assistance, or distributing posters, newsletters, etc
- Demarche: a dossier of information from different sources offering in-depth background on an issue
- Q&A: providing frequently asked questions on own websites/social media accounts, or town hall style public meetings to answer common questions

- Risk mitigation comms: communications aimed at preparing and mitigating continued risks
- Paid advertisement: in some cases, it may be appropriate to advertise on media or social media (note that you may need political authorisation to do this)
- Search engine optimisation (SEO): in some cases, it may be appropriate to promote government content on search engines
- Investigation/final conclusion: in some cases, it may be appropriate to publish the results of an investigation to offer closure on an issue

What is its impact? Crisis communication aims to protect the public during immediate short term crises. It is part of normal communications work, and many of these tools will be valuable for correcting mis- and disinformation that has reached mainstream audiences during a crisis or that presents a risk to the public. Note that it is advisable to draw upon the OECD Principles when conducting transparent and credible crisis communication.

Policy response

When should it be used? When you need additional policy options to counteract the strategic intent of a persistently hostile actor.

How does it work? Using policy levers to counter malign intent means developing detailed knowledge about the capabilities, will, and opportunities of a persistently hostile actor in the information environment, such as in the case of hostile state threats with a disinformation component. Aspects of this knowledge can support a communication response, other parts should be shared more widely within government and with trusted partners.

- You may wish to ignore messages and narratives and focus instead efforts on revealing and countering the strategic intent or effect on key target audiences
- It may be appropriate under certain circumstances and with the appropriate permissions to directly, or together with partners such as social media platforms, to seek to block, disrupt, remove, or expose the sources of harmful content
- In conjunction with colleagues working on policy, it may be appropriate under certain circumstances to work with the communicative dimensions of policy decisions or actions that government makes
- In some cases, the activities may warrant an official government response aimed at punishing previous actions, deterring current operations, or raising the costs of future actions.

What is its impact? Countering the intent of an actor means to a certain degree going head-to-head with them, so the impact should involve a significant cost to their capabilities. For example, during the Salisbury poisoning, the UK government limited its work on countering the dozens of individual pro-Kremlin disinformation messages and narratives and instead focused on revealing the overall intent of the wider disinformation campaign. Some of this strategic work takes place at policy level but requires input from communication specialists. Other parts are about the communication strategy itself, and where the focus should lay when considering the various communications options mentioned here.

- More than communication of policy. If the government has decided to take policy-level actions against a persistently hostile actor, it may be valuable to think of all related communication activities in terms of their signalling value, for example as part of a deterrence posture.

5.6 How to weigh up approaches?

When you have conducted your risk assessment, you will have reached a conclusion about the priority of the false or misleading information. This will enable you to consider a range of communicative tools which you can then tailor to relevant target groups. Generally, the higher the priority, the more focus should be placed on short-term reactive responses, at least initially. Note that a combination of short, medium and long-term approaches may be necessary, depending on the priority of the issue. You should use the OASIS model to plan your communication activities (*see annex E*).

	Our priorities	Key questions
Objectives to protect	**Policy areas and responsibilities**	**1. Is the mis- or disinformation harmful to our priorities? In what ways?** *Media relations is well suited to high priority issues that have reached or are about to reach mainstream media and broad audiences, or that are of general interest to the population or specialist group. Social media is useful for direct communications, potentially with highly segmented audiences.* *Proactive approaches are well suited for emerging issues where there are medium-to-long term opportunities to shape the debate and build public resilience* *Reactive approaches are well suited to ongoing medium-and-high priority issues particularly on social media or other niche forums where there is a need for specialist counter-mis- and disinformation tactics* *Note that a combination of approaches may be necessary in order to coordinate, for example, mainstream resilience-building public information alongside specific countermeasures on social media.*
Information to protect	**Key messages and narratives**	**2. What problematic communication techniques and behaviours have you identified?** *You may need to draw on a variety of approaches to inform the public, and colleagues, about the behaviours (Proactive) or directly counteract them (Reactive).* **3. What misleading or manipulated content is being spread? What are the main messages and narratives we should be aware of? What is untrue or misleading about them?** *Consider the balance between correcting false messages contra narratives. Proactive approaches are useful for shaping mainstream media coverage by informing the public about broader issues, as well as for building longer-term resilience. Reactive for directly counteracting individual messages and narratives at source.*

	Our priorities	Key questions
Brands to protect	**Core values**	**4. What values and interests do the accounts spreading mis- or disinformation wish to project, and to which target groups?** *Proactive approaches include a range of tools for actively informing and engaging the public about current and emerging issues. Reactive approaches are suited to tackling the sources of false and misleading information directly. Discrepancies between what accounts say and do can help to reveal intent, which could open for counter-brand responses.*
Audiences to protect	**Key stakeholders and audiences**	**5. How widely is mis- or disinformation spreading, and to which target groups?** *Understanding audience segmentation, and the influencers relevant to specific audience, can help to prioritise the cases according to types of harm. This is a crucial step for determining which channels and audiences to target, and whether to primarily counter or seek to build longer-term resilience.*

Case Study 1: Russian disinformation on the Salisbury poisonings

Following the poisoning of Sergei and Yulia Skripal in March 2018, the UK government was able to successfully counter a determined attempt by Russia to deliberately mislead the public in the UK and overseas. By the end of the first week, a dozen ‘alternative’ explanations for the poisonings had appeared on Russian domestic media, through official government statements, and via digital media. Combinations of disinformation techniques were used to seed and then amplify these narratives. The EU East StratCom Task Force published a version of its Disinformation Review that collected evidence of these examples demonstrating the use of fabrications, malign rhetoric, memes and trolling. The FCDO also shared disinformation updates across government, with international partners, as well as civil society and researchers.

On 12 March 2018, the then Prime Minister Theresa May stated to Parliament that it was “highly likely” that Russia was behind the poisoning. In total more than 40 narratives were seeded by predominantly Russian sources and shared by a variety of influencers. During these early phases, the UK and its allies successfully pushed back on the disinformation strategy by ignoring individual narratives and simply collecting and exposing the overall approach Russia was taking. Meanwhile, efforts to share information with international allies behind the scenes supported significant policy responses eventually leading to an array of sanctions.

	Description	Channels	Proactive	Reactive
Short-term Responses	The disinformation requires an immediate response. Use rapid communications to rebut, correct or counter disinformation in accordance with the established facts.	Prioritise media relations and face-to-face contact with trusted allies, journalists and researchers	Use a counter-brand approach to push back on adversary. Use official statements to release intelligence assessments and public health information. Pre-bunk emerging risks by exposing adversary strategy. Activate network and friendly stakeholders.	Weigh up resources and immediate threats to determine balance between counter-message, counter-narrative and crisis handling. Focus reactive efforts on on public safety and assurance

Case Study 2: Counter-Daesh Global Coalition communications

Established in 2015, the UK Counter-Daesh Global Coalition Communications Cell employed an international, cross-government strategic communications approach to countering Daesh propaganda. By agreeing its overarching principles with 83 partners (across governments and civil society) at the strategic level, the Cell had the flexibility to achieve its objectives in an agile manner. Its daily messaging pack reaches 1,000 officials and has developed into a key resource for identifying and sharing trends, providing rapid response, and creating consensus across the Coalition.

Communications originally focused on countering Daesh's propaganda by rebutting and refuting their claims, but it soon became clear that it was more effective to expose false narratives of life under Daesh. Moving into a proactive posture, the Cell launched whole-of-coalition campaigns like 'Take Daesh Down' and increased the positive messaging focused on life after Daesh, illuminating important international and grassroots stabilisation efforts in Iraq and Syria.

YOU CAN TAKE DAESH DOWN ON YouTube

How to report a video that supports ISIS:

1. Start by clicking the ⋮ icon near the content.

2. Select REPORT and follow the prompts.

3. Done. Easy. Effective. Confidential.

	Description	Channels	Proactive	Reactive
Medium-term Responses	The disinformation requires a considered response. Use a combination of communications to assert your own values/brands. Tie proactive measures with your normal everyday communications and work with stakeholders/ influencers to create consensus around your position.	Prioritise contact with trusted allies, journalists and researchers. Use the same channels as those used by Daesh, and draw on local partners wherever possible	Create campaigns that can be tailored to local conditions Develop counter-brand response revealing differences between propaganda and life under Daesh. Develop campaigns about life after Daesh	Keep up to date on messaging trends and rebuttals with daily info pack. Work with allies to identify and remove Daesh content. Provide content to empower local partners to debunk as they see fit

Case Study 3: Stop the Spread

In 2020, in response to widespread misconceptions about Covid-19 vaccines, the UK partnered with the World Health Organisation (WHO) to raise awareness of the problem. Campaign content included social media infographics explaining the safety of vaccines, positive messaging that debunked common misinformation themes, and the "Stop the Spread" campaign in partnership with BBC World News and BBC.com aimed at encouraging people to double-check information. Many of the campaign assets were made available to governments to support their own efforts.

In addition, the partners created a campaign called "Reporting Misinformation", which raised awareness of how to report potentially wrong or misleading information on different digital platforms. This worked in conjunction with Stop the Spread and the "WHO Mythbusters" misinformation database to educate both in identification and reporting of misinformation. Finally, the campaign also included the game "GoViral!", which provided an interactive means of learning about mis- and disinformation tactics.

	Description	Channels	Proactive	Reactive
Long-term Responses	The disinformation requires a coherent, sustained response to create long-term change. Develop and assert strategic narratives in relation to an issue by shaping the information space to promote your own position and deter others (raising the threshold).	Social media and BBC channels	Credible partnership between WHO and UK, with WHO and BBC as key platforms to reach global audiences. Awareness raising to shape perceptions of the misinformation problem and link to resilience building	Establish database of known mis- and disinformation narratives and debunk.

The communication options outlined here are not exhaustive, but they should contribute to the idea that there are a range of communicative tools at your disposal to deal with false and misleading information.

Summary

This section has covered some of the most important principles of government communication and outlined in some detail a range of communication options available to you. These include communication channels (face-to-face, traditional media, social media), as well as several proactive and reactive methods that you can draw upon. Finally, the section offers some short case studies to help explore how different approaches can be weighed up to shape effective responses to mis- and disinformation.

Tracking Effectiveness

This section will help you answer the following question:

- what can I track to ensure my strategic communications are effective and will help me learn for future disinformation incidents?

By the end of this section, you will be familiar with the basic steps required to track output and outcome metrics to keep your organisation engaged in a continuous learning process when responding to disinformation.

6.1 The importance of tracking strategic communications

Tracking the effectiveness of strategic communication will enable you to determine whether your strategy for countering a specific piece of mis- or disinformation has been successful or not.

Where a strategy has been deemed to have been successful, you can replicate it and use it to counter other pieces of disinformation.

Where it has been deemed to have been unsuccessful, you can modify your strategy, or tailor your content, to ensure that it better matches the interests and preferences of your target audience. Doing so will increase the likelihood that your audience engages meaningfully with your content, thereby reducing vulnerability to disinformation.

In cases where you adapt your strategy and implement new strategic communications, it is vital that you track effectiveness once more. Not doing so could result in a subsequent strategy also failing to meet its intended aims.

Baselines are a critical part of measurement at the outset of your campaign. Measuring baselines enables you to clearly understand whether your strategic communications are successful in creating change, as measured against the baseline established at the beginning. Tracking the effectiveness of your strategic communications is therefore an iterative and ongoing process, carried out at regular intervals for the duration of your campaign. Doing so will enable you to tailor and modify your communications as the campaign progresses, thereby increasing its effectiveness.

6.2 How to measure the effectiveness of strategic communication

Before you begin to measure the effectiveness of your strategic communications you need to first determine which of your objectives are outputs, and which are outcomes, in order to understand the type of results your intervention is having on your audience. In short:

1. **Outputs** can be understood as the pieces of information that you create and disseminate. Tracking the effectiveness of outputs means measuring the extent to which your strategic communications have *reached* and *engaged* the target audience.

2. **Outcomes** can be understood as the impact your strategic communications have had in the real world. Tracking the effectiveness of outcomes means measuring the extent to which your strategic communications have directly contributed towards your target audience *thinking* or *behaving* differently.

Once you understand the differences between outputs and outcomes you can begin to map out your objectives. These will differ depending on the specifics of what you are trying to achieve. However, in principle, counter disinformation campaigns should aim to achieve the five objectives set out below:

- reach the audience vulnerable to a specific piece of mis- or disinformation (output)

- present that information in a manner which engages the audience and captures their attention (output)

- direct the the audience to an alternative source of information that is both legitimate and credible (output)
- increase the proportion of reporting or online activity that references your communications or information from your communications (output)
- build their resilience whilst enhancing their ability to think critically about the information they encounter (outcome)
- change their attitudes, perceptions and behaviours towards a particular issue or topic (outcome)

Tracking effectiveness therefore requires you to determine the extent to which your strategic communications have been successful in achieving those five objectives.

Examples of metrics that can be used to measure the effectiveness of your *outputs* include:

- demographic information on the types of people accessing and engaging with your content
- the number of people who go on to click on a link to a more detailed source of legitimate and credible information (and conversely, the number of people continuing to access disinformation)
- the amount of time people spend engaging with the information you are disseminating . Examples include “dwell time”, which means the amount of time people spend consuming your information
- the extent to which your communications are accessible and available in public discourse. For example, amongst the articles covering an issue or topic, how many reference your message/ communications. This is also known as “share of voice”

You can also use broader metrics which focus on the outcomes your strategic communications have contributed towards.

Examples of metrics that can be used to measure *outcomes* include:

- survey and polling data on the views and attitudes of your target audience from reputable market and social research companies
- statistics on the numbers of people undertaking desired actions, for example the number of people getting vaccinated for COVID-19

It should be noted, however, that it is not often possible to definitively prove a direct link between your communications and the attitudes and behaviours of demographic groups (or society more broadly). This is because there are numerous factors at play beyond your activities.

That being said, measuring broader shifts in the attitudes and preferences of your target audience - and their behaviours - can be useful in giving you an indication of whether your activities are contributing towards overall positive trends and desired outcomes. As well as helping you to assess the long-term impact of mis- and disinformation narratives and the extent to which they remain a threat.

Objective	Metrics	Effectiveness of output or outcome
Reach the audience vulnerable to a specific piece of disinformation	Demographic information on the types of people engaging with your content	Output
Present them with engaging information that captures their attention	The amount of time people spend viewing information (dwell time), the extent to which they engage with that information (likes, comments, shares)	Output
Direct them to a more in-depth source of legitimate information	The click-through rate (the number of people who click on a link to more in-depth information), the amount of time people spend viewing that information (dwell time)	Output
Increase the proportion of reporting or online activity that references your communications and messages	The extent to which your communications are accessible and available in public discourse - for example, of the articles covering an issue or topic, how many reference your message/ communications?	Output
Build audience resilience to a specific piece of disinformation and enhance their ability to think critically	The number of people engaging with a specific piece of disinformation (is it decreasing or increasing?)	Outcome
Change audience views, attitudes and perceptions about a particular issue or topic	Survey and polling data on the views, attitudes and perceptions of your target audience	Outcome
Change the way in which the audience behaves	The number of people amongst your target audience who are changing their behaviour (examples include quitting smoking, getting vaccinated, applying to become a teacher)	Outcome

GoViral! case study

Extensive academic research into 'inoculation theory'[9] has shown that giving people a taste of the techniques used to spread fake news on social media increases their ability to identify and disregard misinformation in the future. GoViral! is an online game designed to help the public understand and discern three of the most common Covid-19 mis- and disinformation tactics used by online actors, so they can better protect themselves.

The objectives of GoViral! are to reach audiences vulnerable to health misinformation, and present information in an engaging gamified format to direct them to credible sources of information i.e, the World Health Organisation. The ultimate aim is to build citizen resilience to disinformation and change perceptions of vaccines in a positive way.

9

Assessing outputs:

Data on outputs and their effectiveness is collected using a number of datapoints from Google Analytics:

Badges: part completion of the game	Scores: full completion of the game	Actions: shares of the GoViral! game on social media	Outbounds: visits to the WHO website

[9] https://www.nature.com/articles/d41586-020-03189-1

This data provides continuous insight into how effective the game has been at reaching and engaging audiences.

Assessing outcomes:
Those falling foul of health misinformation are less likely to engage in positive health behaviours, such as taking the COVID-19 vaccine. However, pinpointing whether the game has increased vaccine confidence directly is impossible due to the number of factors involved in people's views, attitudes and behaviours. Whilst experiments show the effectiveness of the game in lab settings, and output data shows that the game engages target audiences, this does not tell us about the long-term outcomes of the game.

To address this, further development is underway by the University of Cambridge to establish a consolidated metric for measuring the vulnerability of audiences to mis- and disinformation. This metric could be used in the collection of survey data to assess the impact of communication interventions before and after they are implemented and illustrate their direct impact on attitudes and behaviours.

Summary

After this final section, you now have a sense of:

- why **tracking the impact** of your strategic communications is important
- the difference between **output and outcome** measures
- **examples of metrics** that can be used to track your impact

Tracking effectiveness brings us full circle to better understand how we can constantly improve our strategic communications in response to disinformation.

About the author

Dr James Pamment is associate professor at the Department of Strategic Communication at Lund University.

The Government Communication Service (GCS) is the professional body for people working in communication roles across government. Our aim is to deliver world-class communications that support Ministers' priorities, improve people's lives and enable the effective operation of our public service.

gcs.civilservice.gov.uk | @ukgovcomms

Designed by the Global Communications Academy