Database guide: categories, criteria & use of information

Introduction 2

**1] Criteria & evidence used to identify the type of mis/disinformation** 4

1. The way in which the claim was ‘wrong’– from unproven to false 4
2. Misinformation or disinformation 7
3. The type/s of information effect – cognitive to affective 10
4. Topic or topics of the claim 12
5. Whether perceived credibility could be affected by ‘worldview’ 17
6. Repetition of the claim 20
7. Specific claims, broad narratives or meta-misinformation 22
8. What broad false narrative – if part of one 25
9. Degree of distortion – from completely false to broadly true 35
10. Format/s in which claim spread 38

**2] Criteria & evidence of origins, channels or settings, & drivers** 40

1. Type/s of originator &/or promoter of the claim 40
2. Perceived credibility of source: (i) ‘authoritative sources’ (ii) community network 44
3. Country or countries to which it related 46
4. Country or countries in which it was observed 46
5. Channels & settings in which it spread 46
6. Factors that drove creation &/or responses to misinformation 50
7. Factors that drove creation &/or deliberate distribution of disinformation 54

**3****] Criteria & evidence of effects, consequences & factors that shape them** 57

1. Factors shaping audience reception of the credibility of the claim 58
2. Observed & potential changes in understanding, attitude, behaviour caused 61
3. Whether (i) had effect w/out action (ii) audience had capacity & motivation 65  
   to act on false aspect
4. **Whether specific substantive consequences observed - & if so what field** 68
5. Whether observed consequences that occurred: direct, conditional, cumulative 79
6. Scale of observed consequences 82
7. Severity of observed consequences 84
8. Duration of observed consequences 88
9. **Whether claim had substantial potential to cause (further) specific** 90 **consequences - & what field**
10. If potential consequences – would they be direct, conditional or cumulative 138
11. If potential consequences – what scale 141
12. If potential consequences – what severity 144
13. If potential consequences – what duration 148

**For bibliographic references:** see bibliography in **Fake News – What’s the Harm?**

Introduction

*Fake News – What’s the Harm?*, published by University of Westminster Press in 2025, is based on (i) close examination of a sample of false and misleading information set out in the database described here and (ii) a range of existing scholarly evidence of information and misinformation effects set out in the study. It aims to provide:

a critique of the dominant narratives and approaches shown in academic research and adopted in key counter-misinformation approaches

a theory of factors that predict the potential for and nature of the consequences which misinformation and disinformation may cause or contribute to

To verify the accuracy of the findings set out in the database, and described in the database, colleagues are invited to review (i) the criteria by which the assessments are made, (ii) the quality of evidence used and (iii) the findings reached.

The database is made up of 250 entries – each entry representing information identified as in one or other way false or misleading by one of six independent fact-checking organisations operating in sub-Saharan Africa 1 July-31 December 2019: AFPFactcheck; Africa Check; BBC Reality Check; Congo Check; Fact Hub/IFCN and Ghana Fact. In addition to the 250 examples of misinformation examined in detail, the database includes 42 examples of other fact-checks of the same claim and 110 examples of factchecks of information similar to one or more of the 250 main entries examined.

The organisations whose work is used were all chosen for having been independently assessed as adhering to an internationally recognised editorial code – (i) the International Fact-Checking Network (IFCN) Code of Principles,[[1]](#footnote-1) or (ii) the BBC editorial guidelines.[[2]](#footnote-2) I set out in the study full details of the process by which the fact-checking organisations were selected, how they select the claims they check and how they carry out that work.

Each entry in the database starts with a specific claim that were assessed by one or more fact-checking organisation as being in some way false or misleading. As explained in the study, I define ‘claim’ as a factual assertion made by a person or organisation, explicitly or implicitly, whether made by the spoken or written word or through presentation of an image.[[3]](#footnote-3) I include in the database a link to the fact-check or fact-checks that assessed the accuracy of each claim so that colleagues can review this evidence for yourselves.

To attempt to better understand the (i) type, (ii) drivers, and (iii) observed or potential contribution of these clams to specific consequences, I have set out information in 30 different categories on each and every entry – from the way in which the information was said to be misleading to the evidence of apparent or potential effects.

Note: as I address, in *Fake News – what’s the harm?*, not only the observed or potential contribution of these claims to specific consequences, but also types and drivers of false information, the database includes a wider range of findings than those detailed in the guidelines for the 2024 trial described in the book which focused purely on identifying potential for consequences.

I set out below both the sources of information and a detailed explanation of the criteria used. And I explain how the information from the database is used in the study. While the finding made in some categories is based on simple, obvious evidence and requires little explanation, some does involve complex judgement, and I set out the basis for these findings here, and I seek to make all judgements consistent across the database.

The criteria used in the version of the model rolled out to fact-checkers in future will iterate, as and when needed, based on review and evidence-based feedback from colleagues.

1] Criteria & evidence used to identify the type of mis/disinformation   
The first goal of the study is to establish useful ways to identify the type or types of information identified in the study as being in some way false or misleading, from the way in which the information is or was false or misleading to whether it was a specific false claim or fed into a broader false narrative and the format or formats in which it spread. I set out in the first five columns of the database: (A) the claim made, (B) the organisation that checked it, (C) the date the fact-check was published, (D) a link to the factcheck(E) a summary of claim and reality. I then set out the evidence drawn from the fact-check in 10 key categories about the nature of the information. I set out information about a further 7 categories on the claim origins and settings and 13 categories related to potential for harm.

Ten factors related to the nature of the information or claim.

The way in which the claim was ‘wrong’ – from unproven to false

The evidence in the database shows, the majority of claims identified by fact-checkers as ‘misinformation’ are not wholly false but contain at least *some* element of accurate information creating, in some cases, a false or misleading understanding and, in other cases, an understanding that bears an element of inaccuracy but is broadly true. This matters, since legislation on ‘false news’ commonly identifies information in a binary fashion as either ‘true’ or ‘false’ (Cunliffe-Jones et al, 2021). Furthermore, for media literacy to enable individuals to identify and dismiss information which is false or misleading, requires that students understand the way in, and extent to which, information falsifies or misleads understanding. And a fact-check that recognises the element in the false information that is accurate is more likely to persuade those who believe the information that it is misleading information.

I have used 10 different classifications developed by (i) different fact-checking organisations adhering to the International Fact-Checking Network Code of Principles and/or (ii) academics in the field. I have made the assessments based on evidence set out in the fact-checks – about (i) the claim and (ii) the reality of the situation described. The nature of the evidence provided in the fact-checks varies considerably depending on the nature of the claim. I explain the nature of this evidence in the database.

|  |  |  |
| --- | --- | --- |
| **Type of distortion** | **Definition [[4]](#footnote-4)** | **Example** (Database entry in brackets) |
| **unproven** | A factual claim for which no definitive proof is publicly available | e.g. a claim suggested a group of named Nigerian politicians had been barred from the United States for election rigging. The US declined to comment. The politicians denied it. There was no publicly available evidence either way – hence I rated it unproven. See (Entry 34). |
| **satire thought of as true** | A factual claim that originated as satire but is understood as true | e.g. a Facebook post mocking attitudes towards northern Nigerians in the southeast of the country was clear satire, using evident humour, but understood by many readers as true. See: (Entry 39). |
| **mislabelled, misattributed or misidentified information** | A still or moving image that is labelled, attributed or identified in a false or misleading way[[5]](#footnote-5) | e.g. a photo of a group of Christians threatened with execution by militants in Iraq in 2003, was published mislabelled as occurring in Afghanistan in 2019. The photo was real but did not show what was stated. See: (Entry 40). |
| **misleading information** | A factual claim that bears some truth but lacks crucial context or detail in way that reframes meaning | e.g. a video of an SA minister making a remark in 2017 was published and claimed to be filmed in 2019 after a wave of xenophobic violence. The implication was an SA minister the attacks. The lack of context reframed the meaning. See: (Entry 49). |
| **an overstated / understated claim** | A factual claim that overstates or understates the level or scale of a state or situation | e.g. an online post claimed that Chad had the world’s highest internet charges. The charges are not the world’s highest but high (top five of 181 countries in one index). The post is thus not wholly wrong but overstates the costs. See: (Entry 130). |
| **conflated claim** | A claim that conflates two or more facts in a way that reframes meaning | This Africa Check report in 2015, showed official statements that conflated evidence of peaceful protests and violent public order incidents in a way that exaggerated the scale of unrest. <https://africacheck.org/fact-checks/blog/comment-politicians-and-police-are-misusing-unrest-figures>. |
| **False** | A factual claim that is wholly false | e.g. An online post falsely claimed that a ship carrying food from Nigeria to Cameroon had sunk. No ship had sunk. Moreover, the implication made in the post about what this showed about the situation in Cameroon was also untrue. See: (Entry 134). |
| **fabricated or manipulated information** | Content, presented as real, all or part of which is fabricated or manipulated to create a false meaning. This may involve the fabrication of a quotation, document or image, or the manipulation of an image to reframe its meaning. | e.g. a Facebook page was set up in the name of the education ministry in South Africa – and use to announce faked plans to hand out condoms to schoolchildren to curb teen pregnancies. Both the page and the claim made were fabrications. See: (Entry 92) |
| **imposter content** | Content deliberately falsely presented as work of another individual or organisation. | e.g. online posts claimed an image showed a screenshot of a Tweet by the then then US President criticising protests against the Ethiopian prime minister. See: (Entry 89) |
| **coordinated inauthentic behaviour** | The creation and coordination of a disinformation operation such as a network of fake accounts to create a false or misleading understanding. | I identify an operation that cloned an independent Eritrean news organisation – and published false information – as an example of this behaviour.  See: (Entry 108) |

As the examples above show, the ways in which information may be false or misleading vary considerably. I use this evidence to argue in the study that the way in which, and extent to which, information is false or misleading matters for both the research approaches taken to this field and the efficacy of measures taken to counter misinformation.

1. Misinformation, disinformation or ‘fake news’

As the researchers Emily Vraga and Leticia Bode noted in 2020 and others since,[[6]](#footnote-6) research on misinformation has grown in volume and scope in recent years, but “defining ‘misinformation’ in a consistent and coherent way has been a challenge for the field”. In the study I review and critique the definitions of ‘misinformation’, ‘disinformation’, ‘fake news’, and the evolving academic and policy debate about their meaning.   
In the study I propose definitions of each of these phenomena based on this critique and a close review of the evidence from the database of the ways in which the claims examined   
(i) were false or misleading, (ii) the factors that contributed to the creation and spread of the false or misleading claim (iii) the evidence available of intentionality and (iv) evidence of how accuracy of certain statements does or may change over time.

|  |  |  |
| --- | --- | --- |
| **Misinformation or disinformation** | **Criteria used to determine whether a claim is mis- or disinformation** | **Examples** |
| **Misinformation** Information that, according to the best evidence and expert opinion publicly available at the time, is false or misleading but may not have been known to be so by those who created or spread it. (See discussion of this definition in the study). | * The information is false or misleading * There is no hard evidence available that proves the information was known to be false or misleading when created or shared | e.g. Sky News misidentified a stow away on a plane from Kenya. It apologised for the error. No evidence indicated that the Sky knew the name was false when it broadcast its report. See: (Entry 107) |
| **Disinformation**  Information that, according to the best evidence and expert opinion publicly available at the time, is false or misleading, is or was known to be so by those who created or spread it, for a social or political purpose.  (See discussion of this definition in the study). | * Hard evidence of intent to mislead e.g. fabrication of information &/or the information is part of a clearly organised campaign to use false information promote a particular view or goal. | e.g. a letter criticising Nigeria’s president, purporting to be from the German foreign minister, was fabricated See: (Entry 43). It was part of a long-term disinformation campaign by separatists. |
| * Evidence indicates that the information or image wasfabricated or manipulated (e.g. manufactured quotes, fabricated or manipulated image) – not simply a false interpretation or misunderstanding | e.g. a post online cut and presented an image from a video game as a photograph of a French military helicopter crash. This is deliberate fabrication and therefore disinformation. See:  (Entry 157) |
| * **Note:** I include imposter content as fabricated or manipulate content – and therefore disinformation. * I also include hoaxes and scams that use fabricated information to draw in audience for financial or other reasons. * I include ineffective or unproven medical treatments, **if** the information provided is found to be fabricated or known to be false   **Note: I do not include**   * Satire, where the satirical or parodical intent is or was declared or the humour is clear * Ineffective or unproven medical treatments, even if promoted commercially, if there is no hard evidence the information was fabricated or known to be false or misleading | e.g. A fake newspaper front page said a Nigerian politician was on a UAE ‘terror watchlist’. Fabrication is deliberate. See: (Entry 168)  e.g. A series of hoaxes providing false information to seek charitable donations. See: (Entry 97)  e.g. promoters of sales of an untested anti-AIDS medication claimed it was in Phase 2 trials in France. No such trials had taken place. See: (Entry 123)  e.g. a satirical report about a restaurant with a license for cannibalism. The website had a clear disclaimer its articles were satire. See: (Entry 99)  e.g. Individuals promoted claims a tisane was a cure for malaria & bilharzia. There was no evidence they knew the claim to be false. See: (Entry 125) |
| **‘fake news’** i.e. Fabricated information, presented as true, produced in a pseudo-journalistic style to deceive about its meaning.  See Egelhofer & Lecheler 2019.[[7]](#footnote-7) | * The information is presented as true, and published in a traditional news format * Evidence shows the fabrication or manipulation of the content | e.g. a post presented in ‘news report’ style, fabricated an image to suggest a bomb blast had taken place when no such event had occurred. See: (Entry 56) |
| **Misinformation that originated in disinformation**  In the study I consider an argument of researchers Bontcheva and Posetti that the term misinformation is itself misleading as “it is evident that the content at hand often owes its origin to others’ deliberate acts of disinforming citizens with harmful intent" [[8]](#footnote-8). I examine examples I classify as misinformation for evidence they originated in disinformation. | * Information that is false or misleading but not proven to be known as such by those who published or promoted it at the time * The publication or promotion of the misinformation followed from a piece of disinformation, known by its promoters to be false | Police in Ghana falsified a child’s age in court papers to secure his conviction as an adult. They knew his real age. This was disinformation. False reports subsequently emerged online that the child was detained in Nigeria. There was no hard evidence those who claimed this knew he was in Ghana. See: (Entry 103) |

The distinct nature of these different forms of false information affects the efficacy of counter measures taken. Other things being equal, an individual who has made a genuine mistake is, more likely to correct his or her statement than one who made the false claim deliberately, while individuals motivated by financial reward are more likely to respond to a restriction on their financial earnings than an appeal to ethical principles.

The type of information effect – cognitive to affective

The effect information has, or may have, on audience attitudes, behaviour and understanding is another factor in shaping its potential consequences. Based on information set out in the fact-checks, and evidence of reader responses, I have assessed all entries for their actual or potential information effect: (i) **affective information** (shaping attitudes or expressing or evoking emotions), (ii) **behavioural information** (stimulating or enabling behaviour), and/or (iii) **cognitive information** (creating a false understanding).

|  |  |  |
| --- | --- | --- |
| **Type of effect** | **Criteria & evidence used** | **Example** |
| **Solely or primarily cognitive effect**   * The claim changed, or had potential to change, the audience’s understanding of the status of a person, group, institution, event or state of affairs or medical or scientific reality. (See Perse, 2001 for more discussion of cognitive effect. Reference at foot of page). | * the claim sets out primarily practical, technical information or information of academic interest * the claim does not explicitly express or evoke emotion * the claim does not explicitly exhort or enable behaviour * reader responses show effect on understanding | e.g. - false data about the popularity of politicians in Kenya was published in 2019. It did not evoke emotion and, three years before the next election, due in 2022, there was no opportunity for behaviour. Reader responses showed the information created and enforced a false understanding. See: (Entry 5) |
| **Solely or primarily affective**   * The claim expressed and influenced, or had potential to influence, mood, feelings and attitudes, including positive or negative evaluations. (See Perse, 2001 for details). | * the claim included emotive language or imagery * the claim evoke emotional responses from readers or those affected * the claim did not explicitly exhort or enable behaviour | e.g. an online post showed an image of what was claimed to be a South African government employee playing solitaire online while the public waited. This provoked anger but did not substantively affect understanding or behaviour. See: (Entry 4) |
| **Solely or primarily behavioural effect**   * The claim explicitly stimulated or enabled observable action or behaviour, via behavioural information, or had potential to do so. | * the claim includes content that explicitly exhorts or enables behaviour * evidence from reader responses shows a behavioural effect or intention | e.g. an online message claimed to be a code of behaviour Nigerian military required of the public during a security operation. The claim had both affective and cognitive effect – but was primarily behavioural. See: (Entry 81). |
| **A combination of two or more of the above**. | * The claim included a combination of types e.g. both explicit emotional language and an explicit exhortation to behaviour * Evidence from reader responses shows two or more types of effect | e.g. – False reports foreign nationals were kidnapping children from schools in South Africa created a false understanding of a situation (cognitive effect), were expressed in emotive language and evoked fear (affective responses) and called for people to come to the schools (a behavioural effect). See: (Entry 54) |

Topic or topics of the claim

I argue in the study that the topic or topics of a claim may help predict both the potential for and the nature of any consequences information may cause contribute to. In terms, first, of potential for consequences, we know from much research (Lewandowsky et al[[9]](#footnote-9)) that, other things being equal, a claim, true or false, that relates to a question on which an individual has a strongly-held pre-existing view is less likely to cause a change in understanding, attitude or behaviour than a claim on a topic on which they have no such fixed opinion, such as whether or not a particular celebrity has died or a dispute over a practical topic.

Second, I argue that, other things being equal, the topic of the claim will shape the field in which consequences that do or may occur: misinformation about medical issues being more likely to have effects on public health than on climate change, for example. Third, I argue the *precise* topic of the claim is likely to help shape the scale, severity and duration of the effects. For example, a false claim about health topic is more likely to cause widespread but mild consequences if it relates to treatment for a common but mild condition than if it relates to a rare and more severe condition.

Using evidence set out in the fact-check, I have identified content related to 20 broad topics among the claims examined. The list of topics is, of course, non-exhaustive. It reflects both the nature of events taking place in the period and the claim selection process of the fact-checkers – which I explain in detail in the study. Other samples taken at different periods, in different countries, by different methods, would likely identify misinformation on a number of other topics. I also note that many claims relate to more than one topic and where this happens I reference the primary topics. For example I identified a claim about a rape case involving two well-known people as being related to (i) gender (ii) justice (ii) celebrities.

I set out a table of topics and topics from the database below. See the reference notes for examples.

|  |  |
| --- | --- |
| **TOPICS of Misinformation - Identified in the study** | **SUB-TOPICS -  Identified in the study** |
| **ACCIDENTS & DISASTERS** | **Three sub-topics**. (i) occurrence, threat of accidents, disasters;[[10]](#footnote-10) (ii) causes & responses to accidents, disasters;[[11]](#footnote-11) (iii) hoax appeals for those in need re accidents, disasters.[[12]](#footnote-12) |
| **BUSINESS & ECONOMY** | **Two sub-topics.** (i) operations of a particular business or business sector;[[13]](#footnote-13) (ii) the major economic indicators.[[14]](#footnote-14) |
| **CELEBRITIES** | **One sub-topic.** (i)activities, views &/or health of celebrities & the famous.[[15]](#footnote-15) |
| **CRIME & JUSTICE** | **Seven sub-topics.** (i) type of people responsible for particular crimes &/or crimes in general;[[16]](#footnote-16) (ii) rates & risks of particular types of crime &/or in general;[[17]](#footnote-17) (iii) particular crimes or court cases;[[18]](#footnote-18) (iv) particular laws, standards or rules of operation;[[19]](#footnote-19) (v) standards, behaviour and performance of the police & justice system;[[20]](#footnote-20) (vi) prison numbers & conditions;[[21]](#footnote-21) (vii) abuse, crime and risks of abuse and crimes against children.[[22]](#footnote-22) |
| **DEVELOPMENT & POPULATION (Development infrastructure & population)** | **Three sub-topics.** (i) size of population;[[23]](#footnote-23) (ii) state of infrastructure (road, rail, electricity, key buildings etc);[[24]](#footnote-24) (iii) the status of key human development indicators.[[25]](#footnote-25) |
| **EDUCATION** | **Four sub-topics.** (i) standards & outcomes of teaching in schools & universities;[[26]](#footnote-26) (ii) attendance in schools, universities;[[27]](#footnote-27) (iii) government spending on education & standard of facilities;[[28]](#footnote-28) (iv) operation of the exam system.[[29]](#footnote-29) |
| **ENVIRONMENT & CLIMATE** | **Four sub-topics.** (i)evidence, causes, costs &/or ways to limit climate change;[[30]](#footnote-30) (ii) the energy sector, renewables;[[31]](#footnote-31) (iii) wildlife, biodiversity and wildlife protection,[[32]](#footnote-32) (iv) use and/or waste of natural resources.[[33]](#footnote-33) |
| **FINANCIAL & OTHER OPPORTUNITIES** | **One sub-topic.** (i)company give aways, financial opportunities;[[34]](#footnote-34) |
| **GENDER** | **Five sub-topics.** (i) attributes & behaviour of the genders;[[35]](#footnote-35) (ii) codes of dress & behaviour for women;[[36]](#footnote-36) (iii) the economic & workplace status of genders;[[37]](#footnote-37) (iv) gender abuse & violence – includes prostitution & sex trafficking;[[38]](#footnote-38) (v) marriage laws and practices.[[39]](#footnote-39) |
| **GOVERNANCE** | **Three sub-topics.** (i)standards of behaviour in government;[[40]](#footnote-40) (ii) occurrence of corruption;[[41]](#footnote-41) (iii) official salaries & other costs of government.[[42]](#footnote-42) |
| **HEALTH** | **Seven sub-topics.** (i) the state of public health;[[43]](#footnote-43) (ii) the prevalence, nature & causes of mental ill-health;[[44]](#footnote-44) (iii) access, cost of access & quality of health services;[[45]](#footnote-45) (iv) the effectiveness &/or risks of different health treatments;[[46]](#footnote-46) (v) the causes, effects, symptoms, means of spread, prevalence &/or susceptibility to health conditions;[[47]](#footnote-47) (vi) the actions & effects on public health of health practitioners, researchers, charities, companies & policymakers;[[48]](#footnote-48) (vii) the effects on health of particular activities, &/or non-medicinal products.[[49]](#footnote-49) |
| **JOB OPPORTUNITIES** | **One sub-topic.** (i)claims firms are hiring staff when they are not.[[50]](#footnote-50) |
| **MEDIA** | **Two sub-topics.** (i) the accuracy, impartiality and or regulation of the medias;[[51]](#footnote-51) (ii) imposter claims to be mainstream media coverage.[[52]](#footnote-52) |
| **MIGRATION & COMMUNITIES** | **Five sub-topics.** (i) the level of migration &/or legal status of immigrant or other social groups;[[53]](#footnote-53) (ii) the actions or views of an ethnic, national, immigrant or other social group;[[54]](#footnote-54) (iii) the socio-economic status of &/or discrimination against immigrant or other social groups;[[55]](#footnote-55) (iv) a particular social group having ‘special powers’;[[56]](#footnote-56) (v) the history of slavery, discrimination, exploitation.[[57]](#footnote-57) |
| **MISCELLANEOUS** | **Three sub-topics.** (i) the natural, supernatural &/or spiritual world;[[58]](#footnote-58) (ii) unusual feats and records;[[59]](#footnote-59) (iii) random topics.[[60]](#footnote-60) |
| **OTHER COUNTRIES** | **Four sub-topics**. (i) a foreign government’s actions or views;[[61]](#footnote-61) (ii) government or a public figure’s relationship with a foreign country;[[62]](#footnote-62) (iii) the actions or views of an international organisation;[[63]](#footnote-63) (iv) comparisons between countries.[[64]](#footnote-64) |
| **POLITICS** | **Eight sub-topics.** (i) cost of or access to government services;[[65]](#footnote-65) (ii) nature &/or results of politicians’ policies, actions, competence;[[66]](#footnote-66) (iii) nature &/or results of government spending;[[67]](#footnote-67) (iv) practical election processes &/or risks of participating in elections;[[68]](#footnote-68) (v) facts &/or fairness of actual or expected results of elections;[[69]](#footnote-69) (vi) actions or views of politicians & their allies;[[70]](#footnote-70) (vii) public opinion &/or support for politicians or policies;[[71]](#footnote-71) (viii) political appointments.[[72]](#footnote-72) |
| **PRIVATE INDIVIDUALS** | **Two sub-topics.** (i)identity of a private individual;[[73]](#footnote-73) (ii) reputation of a private individual.[[74]](#footnote-74) |
| **SEX, SEXUAL ANATOMY & SEXUALITY** | **Three sub-topics.** (i) sexual behaviour by age, gender or in general;[[75]](#footnote-75) (ii) facts of sexual anatomy or sexual health;[[76]](#footnote-76) (iii) sex with animals, aliens.[[77]](#footnote-77) |
| **VIOLENT UNREST, WAR & CONFLICT** | **Three sub-topics.** (i) politically linked violence, including terrorism;[[78]](#footnote-78) (iii) state security or military issues;[[79]](#footnote-79) (iv) violent civil unrest (social, ethnic, religious).[[80]](#footnote-80) |

Whether perceived credibility could be affected by ‘worldview’

As I have already noted, we know that, other things being equal, a claim, whether true or false, that relates to an issue central to an individuals’ existing understanding of the world – their ‘worldview’, or *weltanschauung* – is less likely to cause a change in understanding, attitude or behaviour, than a claim on a topic of purely practical interest or on which they have no firm view (Lewandowsky et al, 2012. Duffy 2018).[[81]](#footnote-81)

This does not, of course, mean that a claim on such a topic is more or less likely to be perceived as credible by all who see it. It indicates instead that, other things being equal, the claim is more likely to be perceived as credible by some and less likely by others. Where, by contrast, the claim relates to a topic on which individuals have no particular pre-existing view, its perceived credibility is more likely to be affected by the persuasive or unpersuasive nature of the information itself.

While I do not and cannot know the view of everyone who was exposed to the information, the views of many individuals exposed to online claims could be identified either in their responses to the claims or other information it was possible to find about them and their affiliations or beliefs. For example, many of those who responded to a false claim exaggerating the number of Christians killed by Muslims in Nigeria in 2019 can be seen from their online profiles to be committed religious activists, adhering to a particular Christian worldview.

Given, however, that it was not always possible to identify the views those exposed to the information, I identify in the database whether entries were about topics that either (i) related or (ii) had the potential to relate to, individuals’ worldview, and were thus more likely to change those individuals’ attitude, understanding or behaviour – or, did not have this potential. And I use this information to indicate whether the claim was, thus, more or less likely to be perceived as credible and/or affect an existing view, if such a view existed.

|  |  |  |
| --- | --- | --- |
| **Relation to ‘worldview’** | **The topic related to** | **Examples** |
| **Related to ‘worldview’** | (i) political loyalties or ideology | e.g. a false claim about media coverage of former SA president Jacob Zuma. See: (Entry 19).  Many in South Africa were fiercely loyal to the former president and saw media as conspiring against him. Those with this view would be more likely to see the claim as true. |
| (ii) understanding of human nature | e.g. a false claim about SA women seeking welfare payments.  [See:](https://africacheck.org/fact-checks/meta-programme-fact-checks/no-evidence-south-african-women-top-list-africas-heavy) (Entry 171).  Many in South Africa and elsewhere view the nature of low-income women as being to act irresponsibly to seek benefits. Those with this view would be more likely to see the claim as true. |
| (iii) issues of religious identity &/or religious/moral matters | e.g. a false claim about the number of Christians killed in Nigeria. See: (Entry 96).  Many in Nigeria, and elsewhere, view the world as more hostile than it is to people of their religious faith. Those Christians with this view would be more likely to see the claim as true. |
| (iv) national, cultural or ethnic loyalties or belief systems | e.g. a false claim about that a man in Nigeria was made to ‘bark’ and killed by being ‘cursed’. See: (Entry 14).  Many across Africa and the world hold to a belief in psychic or supernatural effects. Those with this belief would be more likely to see this claim as true. |
| (v) matters that are subject of conspiracy theories. | e.g. a false claim about the effects of vaccines on children. See: (Entry 186).  Many around the world hold to a view that vaccines are dangerous, and – in some cases – see claims made for vaccines as part of a conspiracy. Those with this belief would be more likely to see this claim as true. |
| **Unrelated to ‘worldview’** | (i) practical matters, not topics related to identity or belief | e.g. a false claim about a wealthy businessman running a promotional giveaway. See: (Entry 95).  The claim related to whether or not a well-known businessman was running a giveaway promotion that month. This was not a topic on which individuals would have had a pre-existing view. |
| (ii) topics about which most people had or would be likely to have no strong prior view | e.g. a false claim about the medical effects of getting up in the night. See: (Entry 76)  This was a topic on which most individuals lacking basic medical training and education would have, or be likely to have, no strong prior view. |

Repetition of the claim

Repetition is another factor that helps shape a false claim’s potential effects. We know from numerous studies (Allport & Lepkin, 1945; Swire-Thomson et al, 2017; Pennycook et al, 2018; Fazio et al, 2019; Lacassagne et al 2022) that, other things being equal, the repetition of information increases its perceived credibility – even where the claim is highly implausible (Lacassagne et al, 2022). And, other things being equal, a claim’s repetition can thus affect potential for changes to attitudes, behaviour or understanding, and thus its potential to cause or contribute to substantive consequences.

For this reason, I assess in the database evidence that a claim or close variants of the claim were repeated over time – and use this information in my model for predicting potential consequences.

To do this I identify entries as either (i) appearing from only one source, or channel, on one occasion – as identified in the factcheck; (ii) repeated on one or more sources or channels over a period of under two months – as identified in the factcheck; (iii) repeated on one or more sources or channels over a period of two months or over – as identified in the factcheck; or (iv) note that the fact-check does not provide this information.

I note in the study that many fact-checks do not report on whether the claims appeared in other sources, and none do so exhaustive, and may thus understate the repetition.

|  |
| --- |
| **The claim appeared on only one occasion** – **as reported in the fact-check**. I.e. It may have appeared elsewhere but this was not recorded.  e.g. a false claim that the Ugandan finance ministry was seeking applications for grants, appeared on one platform, on one occasion – on a faked version of the ministry’s social media page. See: (Entry 85). |
| **The claim has been repeated in similar form from several sources or channels** **over a short period** [i.e. *< two months*] i.e. appearing in more than one channel or setting over a recent period of up to two months thus showing a certain level of current ubiquity]  e.g. a false claim that individuals had attacked the Nigerian embassy in Benin spread on several channels over a short period following the close of the border to trade. It was a time-specific claim. See: (Entry 79) |
| **The claim been repeated in similar form from several sources or channels over extended period** [i.e. *> two months*] i.e. it shows both a certain level of ubiquity and repetition over time.  e.g. a false claim that Islamic militants were going door-to-door posing as medics and injecting people with HIV started in India in 2017 and spread on different platforms and countries through 2018-2019.  [See:](https://factcheck.afp.com/false-police-alert-about-islamic-state-jihadists-posing-medics-has-spread-country-country) (Entry 84) |
| **The fact-check provides no details of the possible repetition of the claim**  e.g. the fact-check of a false claim which led to the death of two medical researchers, and injury of another in Ethiopia did not provide clear evidence of whether the claims had been made previously. See: (Entry 241) |

Specific claims, broad narratives or meta-misinformation

Some of the claims studied relate to specific, narrow issues or events, and – even if repeated – do not form a broad false narrative as such. I identify these in the database as ‘specific false claims’. Others do feed into just such a broad false narrative; a series of interlocking claims that come together in an overarching narrative – the potential for effect greater than any single claim (Festinger et al, 1957. Lewandowski et al, 2012. Jolley et al, 2017 and more). We also know from the studies such as that by Hameleers et al, in 2020, hat the perceived prevalence of false information in society has an effect or potential for effect on public trust in information in general. This is what I term *meta-misinformation*: information that creates belief that misinformation is more prevalent than it really is, leading to the erosion of public trust in information more broadly.

It is beyond the scope of this study to measure the prevalence of, or exposure to, misinformation on the continent overall. However, I have assessed whether the entries in the database were (i) a specific false claim, (ii) part of an existing broad false narrative, or (iii) related to or affecting perceptions of the accuracy of information or a source of information in general: *meta-misinformation*. I have done this on the criteria shown below.

|  |  |  |
| --- | --- | --- |
| **From specific to meta false claims** | **Criteria** | **Notes & examples** |
| **Specific false claims** | (i) The claim is false and is related to a specific situation, event, individual or organisation  (ii) The claim does not feed into a broad false narrative |  |
| **Criteria in more detail**  The claim itself is false or misleading | All entries in the database meet this criteria. |
| The claim itself might be a one-off or repeated – but does not feed into a broader narrative.  [*Being repeated does not necessarily make it part of a broader, overarching narrative*]  Though it is not, itself, repeated, other similar claims are made. However, these claims also do not feed into an overarching false narrative as such. | e.g. A false claim certain workers were banned from working in a Johannesburg suburb was repeated – but there was no broader false narrative about such bans.  [See:](https://factcheck.afp.com/no-johannesburg-has-not-banned-informal-waste-collectors) (Entry 3)  e.g., the study identifies many examples of fake job adverts. These form a type of false information, but not a broad false narrative. See: (Entry 98) |
| The claim is false and does feed into a broader narrative – however the narrative is itself broadly accurate – not false. | E.g., a specific claim about a shooting in South Africa was false but the wider narrative that crime rates were high was broadly accurate See: (Entry 18) |
| **Claims that form part of a broad false narrative** | (i) The claim is false and is related to a specific situation, event, individual or organisation  (ii) The claim feeds into a broad false narrative |  |
| **Criteria in more detail**  The claim itself has to be false or misleading | All entries in the database meet this criteria. |
| The claim does not merely repeat the point of similar claims but combines with other claims to form a broader, overarching narrative | e.g. a false claim that health workers harvest organs illegally while patients are alive fed into a broader narrative that health workers seek to harm the public. See: (Entry 167) |
| Evidence shows the broader narrative is itself false | e.g. a claim about a hotel cleaner making herself pregnant by a billionaire then demanding support fed into a false narrative about women making false claims of abuse by men. See: (Entry 93) |
| **meta-misinformation** | (i) Mis/disinformation that inundates an information space  (ii) Misinformation that purports to come from a source of accurate information (iii) A false claim that accurate information is misinformation, or misidentifies a reliable source as a source of false information. |  |
| Misinformation that inundates an information space – i.e. an effect related to prevalence, not to specific claims | Note: I did not measure for prevalence or perceived prevalence. This is addressed by other studies. |
| Information that misrepresents itself as coming from a source of accurate information – and thus sows doubt &/or misunderstanding about an accurate source | e.g. An imposter account sows doubt about the accuracy of a source of accurate information on Eritrea. See: (Entry 108) |
| Information that falsely claims that accurate information is inaccurate or that a source of accurate information is a source of misinformation | e.g. A false claim that SA media were censoring reports of a politically charged trial, thus sowing a false understanding of the accuracy of reporting on the issue. See: (Entry 19) |

What broad false narrative – if part of one

I have set out below a list of the false narratives I identified in the database, setting out both (i) examples of a claim which fed into them and (ii) evidence of the nature of the narrative.

I have done this both (i) to assist verification by colleagues reviewing the database – providing evidence of the existence of these narratives, and how and why I identified them as false – and (ii) to understand the nature of the effects they may have – the consequences for society of broad false narratives about vaccine effects being different to those about celebrity behaviour for example. From the entries completed so far, I have identified claims as feeding into the following false narratives – set out according to the topics identified in section four of the database. As you will see below, I have – so far – identified no evidence of claims that feed into false narratives on certain topics, such as accidents and disasters, and evidence of claims feeding multiple different false narratives on other topics, such as health or politics. While my sample is, of course, non-exhaustive, I will discuss in the study the implications of this for counter misinformation approaches.

|  |  |
| --- | --- |
| **ACCIDENTS & DISASTERS** | No broad false narratives related to accidents & disasters identified among database entries |
| **BUSINESS & ECONOMY** | **Evidence of broad false narratives** |
| **False narrative about the activities of a particular business or its staff**  e.g. a false claim the mobile phone company MTN was giving away data to customers. See: (Entry 75) | I identified a large set of false claims in the study about South African mobile phone company MTN which – taken as a whole – would cast undue doubt on its credibility. This included claims that the firm was abruptly stopping its ‘mobile money’ operations in Cameroon, and Ghana, prosecuting loan defaulters, running scams claiming it was giving away data or 15,000 phones – which affected those who believed the claims,naming a new, untested MD and moving into condom production. |
| **CELEBRITIES** | No broad false narratives related to celebrities identified among database entries |
| **CRIME & JUSTICE** | **Evidence of broad false narratives** |
| **False narrative that exaggerates the scale and risk of crime in South Africa**  e.g. a false claim re-a “gang shooting” in Cape Town that was actually taken from a film set. Gang violence is real in Cape Town but not as prevalent as the film suggested. See (Entry 18) | Violent crime is high in South Africa – but not as high as widely perceived. This is due, in part, to a broad false narrative commonly found in traditional and social media exaggerating the scale of crime (Duffy, 2018). See: (Entry 68) |
| **False narrative that exaggerates level of political bias, corruption in judiciary in South Africa**  e.g. false claims about political bias of a leading prosecutor (See: Entry 71) and corruption of six prominent judges (See: Entry 223) | Bias and corruption exists in the South African judiciary – but not to the extent suggested. The claims made against the judiciary served, however, to undermine public confidence in prosecutions and judicial sentences. |
| **DEVELOPMENT & POPULATION (Pop’n, infrastructure, development)** | **Evidence of broad false narratives** |
| **False narrative that exaggerates the known population of Lagos and/or Nigeria**  e.g. a false claim that Lagos had a known population of 21 million. See: (Entry 194). This and similar claims | The population of Nigeria, and its major cities like Lagos, has been a subject of false narratives for decades – federal and provincial policymakers presenting different figures for financial and political reasons. |
| **False narrative about the state of development in African countries in the 19th, 20th century**  e.g. a false claim about a photo taken in DR Congo in 1952 and said to show the situation in Nigeria, Kenya or other countries. See: (Entry 110) | The state of development of Africa, and African countries, during the 19th and 20th centuries has long been a topic of true and false historical narratives. The claims identified in the example shown left, fed into false narratives about the development of Nigeria and Kenya, portraying the population as less worldly than it was at the time. |
| **EDUCATION** | No broad false narratives related to education identified among database entries |
| **ENVIRONMENT & CLIMATE** | **Evidence of broad false narratives** |
| **False narrative that warnings of the effects of climate change are exaggerated or a hoax**  e.g. a false claim that icebergs melting will not contribute to sea levels rising. See: (Entry 189) | Opponents of action to counter climate change have sought, for decades, to dismiss concerns by arguing, variously, that climate change is either not real, or not manmade, or that its effects will not be as severe as evidence suggests. The claim shown left is part of this broad false narrative. |
| **False narrative that understates activity that contributes to climate change**  e.g. a false claim spread by Russian media that forest fires in the Amazon basin caused less damage than those in central Africa. See: (Entry 135) | Supporters of energy companies and the logging industry have for decades promoted a false narrative underplaying the role of their activity in contributing to climate change. The claim shown left fed into the narrative that what was happening in Brazil was normal and of little effect on climate. |
| **FINANCIAL OPPORTUNITIES** | No broad false narratives related to financial opportunities identified in database entries. |
| **GENDER** | **Evidence of broad false narratives** |
| **False narrative that women make up claims of sexual abuse, exploitation by men**  e.g. a false claim that a woman, who was raped, had apologised and admitted making up her claim. See: (Entry 20) | Women have been known, on rare occasions, to make false accusations of sexual abuse and rape. A common false narrative suggests that this is frequently done. See ‘“She Lied”: Social construction, rape myth prevalence in social media, and sexual assault policy’ (Stabile et al, 2019), on this false narrative. As a related type of claim see: (Entry 93) |
| **False narrative that exaggerates support for polygamy**  e.g. a false claim that the government in Eritrea had made polygamy obligatory by law. See: (Entry 124) | Studies have long shown that woman generally do not support polygamy. Nevertheless, a false narrative can be found widely in west Africa and elsewhere that women generally support the practice.  See as another example: <https://pubmed.ncbi.nlm.nih.gov/12291820/> |
| **GOVERNANCE** | No broad false narratives on governance identified among database entries |
| **HEALTH** | **Evidence of broad false narratives** |
| **False narrative that exaggerates the risks to health from mobile phones**  e.g. a false claim that 5G phone masts are so dangerous those installing it have to wear hazmat suits. See: (Entry 203) | Mobile phones have negative effects on health for some. However a false narrative that old-style phone masts to 4G, 5G and the phones themselves cause a wide range of negative effects to health is false. See report on the history of phone-related conspiracy theories <https://fullfact.org/online/5g-and-coronavirus-conspiracy-theories-came/> |
| **False narrative that a cure has been found for HIV/AIDS &/or it is no longer a threat to public health**  e.g. a false claim that a cure has been found for HIV so there is no need for people to use condoms. See: (Entry 229) | Since the emergence of HIV/AIDS in the 1980s, it has been subject of many false narratives, from the claim that HIV does not cause AIDS (a claim which affected public policy in SA 2000-2005) to claims AIDS was not real, only affected White people, and/or that various fake treatments were ‘cures’ |
| **False narrative that vaccines are dangerous to health**  e.g. a false claim that 75% of children who received vaccines in a Mexican town died or were hospitalised. See: (Entry 186) | Most vaccines do or can cause some medical side effects, to some recipients. However, a broad false narrative exaggerates the risks that vaccines pose to health. This study is one of numerous studies of this false narrative. See: Pertwee et al, 2022. <https://www.nature.com/articles/s41591-022-01728-z> |
| **False narrative that ‘natural’, ‘indigenous’ health treatments can be presumed more effective than ‘western’ ones**  e.g. a false claim that a tisane of artemisia is effective against malaria –in part because it is natural. See: (Entry 124) | Much natural medicine is known to be effective. However, a broad false narrative exists that natural, indigenous or traditional treatments be presumed to be effective health remedies because they are natural.  See studies of this narrative e.g. Rozin et al: <https://doi.org/10.1016/j.appet.2004.03.005> |
| **False narrative that ‘eastern’ health treatments can be presumed more effective than other treatments**  e.g. a false claim that ‘Chinese medicine’ recommends pricking a person’s ears, fingers with a needle to treat a stroke. See: (Entry 30) | Much traditional eastern medicine is proven effective. However, a broad false narrative exists that traditional ‘eastern’ or ‘Chinese’ treatments can be presumed more effective than western ones for this reason alone. |
| **False narrative that health workers seek to cause harm to patients, the public**  e.g. a false claim that doctors do not anaesthetise but rather paralyse and then remove and sell the organs of organ donors. See : (Entry 167) | It happens but is very rare that health workers deliberately harm their patients. However, a false narrative that this happens is common and has caused or correlated with attacks on health workers in Sierra Leone and Liberia, (2014) Ethiopia, (2018) the DR Congo, (2019)across South Asia (2022)and among anti-vaxxers in the United States (2022). |
| **MEDIA** | **Evidence of broad false narratives** |
| **False narrative that exaggerates the extent of political bias, censorship in the media**  e.g. a false claim that national and international media conspired to cover up the absence of Nigerian President Buhari from a summit. See: (Entry 47) | Many media in many countries have been shown to censor and/or display political biases. Nevertheless, a false narrative is common that media have censor news which they have in fact covered.  This claim from South Africa was another such a case. See: (Entry 19) |
| **MIGRATION & COMMUNITIES** | **Evidence of broad false narratives** |
| **False narrative about the type of people responsible for crime, other negative behaviour**  e.g. foreign nationals wrongly blamed for burning down a historic building in South Africa. See: (Entry 32) | Through history, marginalised communities have often been falsely blamed for crime and other actual or alleged negative behaviour – from the witchcraft trials of 16th century Europe to racial profiling around crime today. This study found many claims wrongly held ‘others’ held responsible for crime & other negative behaviour. |
| **False narrative that exaggerates the number of migrants, foreign nationals in the country**  e.g. false claim that 800,000 Nigerians live in South Africa (the total is far lower), among a total of 13 million illegal immigrants (total is far lower). See: (Entry 206) | Many countries see false claims about the number of migrants &/or foreign nationals who live in their country. Allied with false narratives about (i) the behaviour (ii) treatment of these foreign residents these narratives have potential to increase social tensions and may play a role in civil unrest. |
| **False narrative that exaggerated the scale, nature of xenophobic violence in South Africa 2019**  e.g. a false claim that photos showed Nigerians attacked and killed in the xenophobic violence taking place in South Africa. See: (Entry 53) | A wave of xenophobic violence broke out in South Africa in September 2019. Subsequent studies by a University of the Witwatersrand group found 18 people killed, 22 physically assaulted and 865 temporarily displaced. Those known killed were Congolese South Africans, and Zimbabweans. Others may have been. Many reports at the time misidentified the nationality of victims and the scale of the violence. |
| **False narrative that exaggerates the level of discrimination against Muslims**  e.g. a false claim that Angola had banned Islam. See: (Entry 150) | Studies show that Muslims in many countries face discrimination on grounds of their religion. At the same time, a false narrative exists that exaggerates the scale and degree of discrimination faced. This claim – that Angola had ‘banned’ Islam – is an example of this false narrative. |
| **False narrative that exaggerates the extent and threat of harm, violence by Muslims; oppression of Christians**  e.g. a false claim that Nigerian Muslim herders killed 6,000 Christian farmers in six months. See: (Entry 96) | It is also the case that, while Muslims have played a part in violence against Christians and others in Africa and elsewhere, the scale and threat of this violence is often exaggerated. This 2019 claim that Muslims were planning a ‘war against Christians’ in Nigeria was one small part of this broad false narrative. See: (Entry 174). |
| **False narrative that exaggerates the extent to which the Nigerian government discriminates against Igbos/Biafrans**  e.g. a false claim that Nigerian authorities were forcing the country’s first local carmaker to move out of the southeast. See: (Entry 227) | Igbos have often suffered from discrimination in Nigeria since the early days of independence. However, the extent of discrimination is often exaggerated by false claims that stir grievances. These claims bolster support for independence for the region – based, in part at least, on a false factual understanding. |
| **False narrative that understates the level of racial injustice and inequality in post-apartheid South Africa**  e.g. a false claim that white people own ‘only’ 22 percent of land today in South Africa. See: (Entry 16) | Many supporters of the former regime in South Africa promote a narrative that understates the level of racial injustice and inequality persisting in the country. A repeated false claim that 400,000 whites live in poverty, for e.g., exaggerated the extent of poverty among white South Africans. See others in the study. |
| **False narrative that black people are given unfair advantages in post-apartheid South Africa**  e.g. a false claim that the government in South Africa planned to give high school certificates to black high school dropouts to help them get jobs. See: (Entry 27) | Linked to the above false narrative, is a false narrative that black people are given unfair advantages in post-apartheid South Africa.  Thus, for example, a white supremacist claimed that a black airline employee who mishandled airline baggage was protected by employment laws which unfairly protected black South Africans. See: (Entry 113). |
| **MISCELLANEOUS** | **Evidence of broad false narratives** |
| **False narrative that humans are capable of impregnating animals**  e.g. a false claim that a zookeeper in Indonesia had sex with and impregnated a female orangutan. See: (Entry 25) | Among the many false reports online, one sub-field feeds a false narrative that humans are capable of impregnating animals – something humans are scientifically unable to do. The reports that feed the narrative draw audiences for junk news sites. |
| **False narrative that seemingly miraculous events or effects are proof of religious beliefs**  e.g. a false claim that a girl who disrespected the Koran was turned into a monkey. See: (Entry 179) | Other reports feed a false narrative that seemingly miraculous events or effects – from rocks floating in mid-air to different waters refusing to mix – have occurred and are proof of the believers’ religious beliefs. |
| **False narrative that some people have supernatural powers that can do harm to others**  e.g. a false claim that certain women were ‘bird witches’ who had come to do harm to others through supernatural powers. See: (Entry 132) | A further false narrative exists that goes beyond expressions of religious faith to argue that some individuals have supernatural powers that are or could be used to harm others. Such claims can be used, as noted in the entry 132 shown left, to justify abuse of and violence the often marginalised people accused of such powers. |
| **OTHER COUNTRIES** | **Evidence of broad false narratives** |
| **False narrative that other countries responded strongly to the xenophobic violence in South Africa in 2019**  e.g. a false claim that Rwandan President Paul Kagame had called for South Africa to be expelled from the African Union. See: (Entry 65). | After xenophobic violence in South Africa in September 2019, numerous claims emerged in media and online that other named countries were taking a strong stance against South Africa – boycotting a summit meeting held there or threatening sanctions. These claims built into an overarching narrative – and put pressure on other countries to do the same. |
| **False narrative that exaggerates the extent to which 'political correctness' shapes the policies of international organisations**  e.g. a false claim that the WHO was defines the inability to find a sexual partner a ‘disability’. See: (Entry 11) | International organisations to use diplomatic, rights-focused language. This tendency is often exaggerated or distorted by critics as organisations bowing to what they identify as ‘political correctness’ or excessive concern for ‘rights’ over ‘responsibilities. The claim shown left fed into this narrative. |
| **False narrative that foreign powers support an insurgency or bid for independence**  e.g. a false claim that the German foreign minister criticised Nigeria for violation of the rights of Biafrans. See: (Entry 43) | Through history, secessionists have made false claims that foreign powers supported their bid for independence (See history of the Nigerian civil war/Biafran war; 1967-70). In the database I identified numerous such false claims by supporters of current independence campaigns – in Biafra, Cameroon and elsewhere. |
| **False narrative that Western powers plot to reduce the size of African &/or Muslim population**  e.g. a false claim that a drug given to African men who have a circumcision to reduce risk of HIV infection “reacts” after 10 years and is killing men – part of a Westen plot to kill Africans. See: (Entry 235) | Western powers have often called for efforts to reduce the rate of population growth in Africa as a means of improving living standards.  These calls, and wider hostility to the West have led to a series of false claims of secret effort to reduce the African &/or African Muslim population by spreading different forms of harmful medical treatments. |
| **False narrative that France supports Jihadists in west Africa &/or harms/exploits region**  e.g. a false claim that the Cameroonian army had detained French soldiers fighting for Islamic militia Boko Haram. See: (Entry 240) | France has long intervened in events in francophone West Africa – both during and since colonialism. A narrative widespread online suggests France today supports jihadist movements battling governments in the region. The narrative is false but the claims appear frequently. |
| **False narrative that exaggerates the role of the United States in the region**  e.g. a false claim that the United States had sent 3,500 troops to Cameroon to take control of the Anglophone region. See: (Entry 117) | The United States also has a long history of intervening in African countries – during and since the Cold War. A false narrative exists nevertheless that exaggerates the role the United States currently plays in politics and security in the region. There are several examples in the study. |
| **False narrative that the United Nations supported the rebel movements in DR Congo**  e.g. a false claim that UN troops participated alongside rebel forces in a massacre of Congolese. See: (Entry 160) | The United Nations has maintained a peacekeeping operation, MONUSCO, in DR Congo since 1999. It has long been accused by pro-government forces of favouring and event supporting the rebel movements that have been fighting the government. There has never been evidence to support the false claims. |
| **False narrative that exaggerates the interference by Rwanda in DR Congo**  e.g. a false claim that Rwanda controls phone lines in DR Congo and can listen in on conversations. See: (Entry 138) | Rwanda has actively intervened in its neighbour, DR Congo, since the 1994 genocide, sending in troops and intervening in other ways. However, a narrative exists in DR Congo that exaggerates the extent and scale of this interference. The entry shown left is an example of this false narrative. |
| **False narrative that US President Donald Trump supported the Ethiopian PM and opposed his critics**  e.g. a false claim that US President Donald Trump had condemned protestors in Ethiopia, calling them extremists. See: (Entry 86) | Supporters of the Ethiopian government, led in 2019 by Prime Minister Abiy Ahmed, used multiple false claims online to promote a false narrative that then US President Donald Trump supported Ahmed and opposed his critics. Trump made no such claims and there was no evidence he took a view on the actions of the protestors. |
| **POLITICS** | **Evidence of broad false narratives** |
| **False narrative that ‘all*’* politicians are untrustworthy and fail to keep their promises**  e.g. a false claim that a community threw a politician in a river after he broke an election promise to build a bridge to the village. See: (Entry 67) | Repeated surveys show low levels of trust in politicians in general. While many politicians fail to state the truth and keep their promises, this is not a universal trait.  The narrative that ‘all politicians lie’ is false. |
| **False narrative that Nigeria's president is ill or dead and has been replaced or side-lined**  e.g. a false claim that Nigeria’s president had died, been replaced by a body double from Sudan, and had failed to attend a summit. See: (Entry 47) | Separatists in Nigeria promoted in 2017 a false claim that the Nigerian president had died and been replaced by a “body double”. The claim was subsequently expanded into a broader false narrative – that officials and media colluded in maintaining the fictional president in office – all apparently aimed at undermining governmental legitimacy. |
| **False narrative that exaggerates risks for those participating in an election**  e.g. a false claim that a video of a shooting in South Africa showed ‘thugs’ shooting at a political rally ahead of elections in Nigeria. See: (Entry 91) | Pre-election violence is common in Nigeria <https://www.hrw.org/news/2019/06/10/nigeria-widespread-violence-ushers-presidents-new-term>.  Nevertheless, a broad false narrative common in Nigeria and elsewhere exaggerates the threat of violence in ways that may affect voting. |
| **False narrative that portrays policymakers as more draconian than they are in reality**  e.g. a false claim that a planned Nigerian law would introduce hanging for criticising the government. See: (Entry 101) | Policymakers in Nigeria have often acted in draconian manner to the public – members of the military taking actions that violated both domestic and international human rights laws.   Nevertheless, a false narrative exists which exaggerates the severity of laws passed as shown in this example. |
| **False narrative that the government is more effective than it really is**  e.g. a false claim that the DR Congo military had found a major store of hidden rebel munitions – exaggerating the military success of the new government. See: (Entry 151) | The promotion of a government’s record as more effective (or less effective) than it was in tackling perceived challenges is one of the commonest false narratives in politics effectiveness. The study found numerous examples of this false narrative in different countries. For another e.g. see Nigeria’s president declaring falsely that the country reached ‘food security’ (See: Entry 249) |
| **False narrative about the popularity of particular politicians, in general &/or as candidates in elections**  e.g. a false claim that a vast crowd had come to attend a rally by a political candidate in Kenya. See: (Entry 181).  [Note – Final column in this entry provides links to multiple such claims] | Like the claim that governments or insurgencies have more support from foreign powers than they do in reality, the claim that politics candidates, or policies, are more popular than they really are, is another false narrative common in politics. We see this in the study with claims that falsified opinion poll evidence (See: Entry 5) or falsified evidence of physical support. |
| **False narrative that benefit payments skew the behaviour of recipients in socially negative ways**  e.g. a false claim that pregnant women in South Africa drink heavily in order to harm their babies and then claim benefit payments. See: (Entry 171) | The claim that welfare payments skew public behaviour is another common false narrative – popular with conservatives in the US, Europe and Africa.  Media in South Africa and UK have since at least 2013 reported falsely that women in the country drink to ensure their babies suffer foetal alcohol |
| **False narrative about the role of Bill &/or Hillary Clinton in US & world affairs**  e.g. a false claim that Hillary Clinton had sold 20% of US uranium stocks to Russia in return for payments to the Clinton Foundation. See: (Entry 185) | The Clintons, like many public figures on the left and right, are subject of numerous false claims that build into a false narrative distorting understanding of their influence in world affairs. These false narratives may have potential to influence certain political and social attitudes |
| **False narrative that exaggerates corruption &/or incompetence of South African authorities**  e.g. a false claim that a senior South African police officer wanted to close police stations at 6 p.m. to keep the police safe from criminals | While there many examples of corruption and incompetence can be found in South Africa, studies of public opinion show that many people exaggerate these problems reinforced by &/or due in part to this sort of false narrative. |
| **False narrative that the former public prosecutor in South Africa was politically biased**  e.g. a false claim that former public prosecutor Thuli Madonsela was a tacit supporter of apartheid. See: (Entry 71) | In 2016, British PR firm Bell Pottinger was hired to run a covert information campaign making the case that those supporting the removal from office of former president Jacob Zuma were allied with ‘white monopoly capital’ and tacit supporters of the former apartheid regime. The false claims against the former prosecutor who charged Zuma with corruption were part of that broader false narrative. |
| **PRIVATE INDIVIDUALS** | No broad false narratives related to private individuals identified among database entries |
| **SEX & SEXUALITY** | **Evidence of broad false narratives** |
| **False narrative that exaggerates the sexual promiscuity in a particular demographic group**  e.g. a false claim about a government plan to curb excessive teen pregnancies. See: (Entry 92) | Claims are commonly that exaggerate the sexual promiscuity of particular demographic group – defined by age, ethnicity or gender.  In the study, several claims, feed into a broad false narrative about sexual behaviour by age (Entry 92) or gender (Entry 172) |
| **VIOLENT UNREST, WAR & CONFLICT** | **Evidence of broad false narratives** |
| **False narrative that exaggerates support for Islamic militants from Nigerian authorities, prominent figures**  e.g. a false claim that former Nigerian VP Atiku Abubakar was on an international terrorism watchlist due to his support for jihadists in northern Nigeria. See: (Entry 168) | Unproven claims have circulated in Nigeria for many years – suggesting that particular named or unnamed Nigerian politicians and prominent figures supported the Boko Haram movement &/or other militant movements.   Typically, no evidence is produced. Some claims such as the one shown left can be shown to be false. |
| **False narrative about popular support for one side in an ethnic or national conflict**  e.g. a false claim that exaggerated popular support for secessionists in Cameroon. See : (Entry 163) | A false narrative that exaggerates popular support for one side in a conflict over the other is common occurrence in conflicts.  In this study, I identify such claims in Cameroon, Nigeria, DR Congo and other countries. |
| **False narrative that exaggerates the scale of suffering of Anglophones in Cameroon**  e.g. a false claim that a boat carrying food to Cameroon had sunk, increasing suffering among English-speaking Cameroonians. See: (Entry 134) | Anglophones in Cameroon have suffered from punitive responses by the government since the region’s secessionist insurgency started in 2017.  A false narrative exists, nevertheless, that exaggerates the scale and severity of this suffering. This entry is an example. |
| **False narrative that exaggerates the risk of conflict or war in Nigeria**  e.g. a false claim that former president Olusegun Obasanjo predicted a ‘second civil war’ in Nigeria. See link: (Entry 169) | A false narrative that exaggerates the threat of civil conflict or war in Nigeria has existed for decades since the end of the civil war 1967-1970. Claims that political leaders such as former president Olusegun Obasanjo (see left) or leaders of the United States and United Kingdom have predicted such a war (See: entry 174) feed into this narrative. |

Degree of distortion – from completely false to broadly true

In the study, I seek to assess the actual and potential consequences of misinformation. And, while all the entries in the database are false or misleading *to some degree*, they differ in the extent or degree to which they are false. For this reason, the entries also differ, I argue, in the extent which, if believed to be true, they distort the audience’s understanding of reality compared to what the audience would have believed if presented with information that was entirely accurate.

For example, a photograph of people killed in an ambush on a road in northern Nigeria in 2018 was posted online a year later, presented as a recent image showing the aftermath of an attack on vehicles in the south of the country (See – Entry 17). The image was ‘wrong’ but, as it happened, several travellers *had* in fact been killed on the road in southern Nigeria recently, as described in the post. So while the label given to the photo was indeed wrong, the understanding created was broadly accurate.

Using evidence from the fact-checks of the claim made the reality of the situation, I have categorised entries as either (i) wrong in some aspect but creating an understanding that is broadly accurate; (ii) including an element of truth, but creating an understanding that is misleading; and (iii) completely or mostly false. I do this based on the following criteria:

|  |  |  |
| --- | --- | --- |
| **Finding** | **Criteria** | **Examples** |
| **Wrong in some aspect but the underlying picture is broadly accurate** | This relates to claims where:   * A detail of the claim is inaccurate but this detail is immaterial and the key event or circumstance described as occurred broadly as said. | e.g. A post claimed DR Congo signed a rail contract. The used the wrong name of the firm involved. The name was immaterial. The broad understanding is correct. See: <https://factuel.afp.com/non-la-rdc-na-pas-signe-un-accord-avec-la-firme-russe-russell> |
| * The image used was from another event or place – but the type of event or circumstance described had occurred broadly as described. | e.g. A post claimed a photo showed an SA company’s office torched in Nigeria in response to recent events. The photo was from an earlier event but the broad understanding was correct. Other SA firms had been attacked. See: (Entry 49) |
| * The claim narrowly overstated or understated a real state of affairs | e.g. an article claimed Chad has the world’s highest internet charges. They are, in fact, among the world’s highest. The understanding created was broadly correct. See: (Entry 130) |
| * The underlying claim related to intent and an element of the claim was wrong but the main actors’ intent was as described. | e.g. a video posted online claimed ex-Ivorian president Laurent Gbagbo was attacked; showing public hatred for the former leader. In fact, a crowd attacked a man it mistakenly believed to be Gbagbo. The understanding of popular reactions to him was broadly true. See: (Entry 89). |
| **includes an element of truth but is misleading** | This relates to claims where:   * A substantial element of the claim is accurate but taken out of context making it misleading | e.g. a post claimed that an SA minister made anti-immigrant comments during recent xenophobic violence. He did so two years earlier and in a different context. See.  (Entry 49) |
| * The event claimed occurred but not for the reasons stated – making the claim misleading | e.g. reports claimed three countries had decided to ‘boycott’ a major event in SA in 2019. They had never planned to attend. This was not a ‘boycott’. The claim was misleading. See: (Entry 57). |
| **completely or mostly false** | This relates to claims where:   * The claim is an individual or institution made a remark they did not make and there was and is no evidence these were their thoughts*.* | e.g. reports claimed the US ambassador to DR Congo made remarks disparaging recent elections. The words were actually a journalist’s questions. There is no evidence the words used reflected the ambassador’s thoughts. See: (Entry 154) |
| * The claim is an individual or institution is said to have made a remark they did not make, and the situation was not as described | e.g. a post claimed that a former footballer made a remark claiming France supported jihadists. There was evidence he said this and no evidence the situation was as claimed. (See additional claims in Entry 120). |
| * An event, circumstance or effect is claimed to have occurred. Evidence shows the event, circumstance or effect was in no way as described, nor was it similar what was described. | e.g. posts claimed that banana and honey are a proven cure for TB and other diseases. There is no evidence these have any such effect. See: (Entry 158) |
| * The event, circumstance or effect claimed to have happened did not happen and could not have happened for scientific or practical reasons. | e.g. posts claimed that a zookeeper in Indonesia had had sex with an orangutan and made it pregnant. This did not happen and could not have happened for scientific reasons. See: (Entry 25). |

Format/s in which claim spread

Finally, for this section, I assess the formats in which the claims spread. I argue in the study that, while some research suggests visual information seen as ‘true to life’ has greater persuasive effect than text content (Messaris & Abraham, 2001. Powell et al, 2015. Sundar et al, 2021), other studies find limited differences in persuasive effect attributable to the mode and format of information (Hameleers et al, 2020b. Wittenberg, 2021. Barari et al, 2021). Nevertheless, understanding the formats in which claims spread assists in understanding of the means and tools to be used by those who wish to counter it.

I note, meanwhile, that while in some cases the fact-checkers whose work is used provided details of different channels and formats in which claims they checked appeared, in other cases they did not, and may not have been exhaustive in doing so when they did. I have noted in the database where details of the format were unclear in the fact-check.

Based on evidence set out in the fact-checks, I have identified misinformation as initially appearing as one or more of the following formats:

|  |  |
| --- | --- |
| **FORMAT IN WHICH THE CLAIM APPEARED** | **EXAMPLES** |
| **Books** | e.g. a false claim that Hillary Clinton had sold US uranium to Russia – first appeared in a book. See: (Entry 185) |
| **Broadcast content (TV or radio)** | e.g. a false claim that an anti-FGM activist had the procedure was made initially in a broadcast by a Kenyan TV station. See: (Entry 26) |
| **Community rumours or myths** | e.g. a false claim that medical researchers in Ethiopia were abusing children was made initially peer-to-peer offline. See: (Entry 274) |
| **Email statements** | e.g. a false claim that pricking someone’s ears can cure them of a stroke was made initially in chain email statements. See: (Entry 30) |
| **Faked documents, web pages or sites** | e.g. the false claims made from a faked Eritrean website were made initially through the faked website. See: (Entry 108) |
| **Memes shared online** | e.g. a false claim that doctors were harvesting organs of patients against their will were made initially as memes shared online. See: (Entry 167) |
| **Official statements** | e.g. a false claim about road improvement by the DR Congo government was made initially in an official statement. See: (Entry 152) |
| **Photos or photo captions – shared online** | E.g. a false claim the DR Congo army had found a major rebel weapons depot was made initially in photos posted online. See: (Entry 151) |
| **Product labels** | E.g. a false claim a Gabonese doctor had created an AIDS vaccine were made on product labels – among other formats. See: (123) |
| **Public signs or posters** | e.g. a false claim that informal workers were banned from a suburb of Johannesburg was made initially on a public sign. See: (Entry 3) |
| **Speeches to live audiences** | e.g. a false claim by Nigeria’s president about the effects of climate change on Lake Chad was made initially in a live speech. See: (Entry 73) |
| **Videos or video captions – shared online** | e.g. a false claim Nigeria’s First Lady was locked up in the presidential villa was made initially in a video caption online. See: (Entry 78) |
| **Voice notes or phone messages** | e.g. a false claim that an Ebola vaccine was a poison created by Rwanda to kill DR Congolese people was made initially in a voice note. See: (Entry 243) |
| **Written news articles, written posts online or on messaging apps.** | e.g. the numerous claims in written articles in traditional media and online or in messaging apps – such as a false claim that the UK and US were withdrawing citizens from Nigeria due to a coming war. See: (Entry 174). |

2] Criteria & evidence of origins, channels or settings, & drivers   
Having established the types of information found in the study – from misinformation to disinformation to the ways it misleads, its information effect, topic, repetition and format – it is important – in order to understand (i) the effectiveness of countermeasures, and (ii) the effect of different sources and settings of misinformation on the consequences it has or may have – to understand (i) who or what causes it to occur, (ii) where, (iii) in what channels and settings it occurs (iv) and what drives this process. To do this, I examine the following:

1. Type/s of originator &/or promoter of the claim

In the study, I seek to understand the types of individual or organisation who **originate** and/or **promote** the false claims. Doing so helps us to better understand their (i) motivations, (ii) their actual or potential perceived credibility, and (iii) the means that might be used to counter the effects of the false claims they make – some creators of false information being more prone to correcting or withdrawing their false claims than others.

For this study, I do not attempt to identify all who **share** the particular false claim. That is beyond my means. I have, however, identified by type those who either (i) originated the misinformation studied – i.e. made a new false claim on their own behalf, or (ii) promoted it, taking a false claim from another source, giving it their own imprimatur or meaning, and sharing it to their audiences.

Doing so, I have broken the originators and promoters into three broad groups based on the extent to which they are identifiable and could be approached directly – using the criteria set out below: (i) well-known and identifiable individuals & organisations; (ii) lesser known but identifiable individuals & organisations, & (iii) unknown individuals & organisations.

|  |  |  |
| --- | --- | --- |
| **Well-known, identified individuals and organisations** | | |
| **Type** | **Criteria** | **Example from study** |
| **Academic, expert or thinktank** | Applied to the individual academic, or expert who makes the claim if it is published in their name. It need not necessarily be published in an academic journal but they must be identified as the source. | e.g. A technical expert who made a false claim about control of the phone network in DR Congo. See: (Entry 138) |
| **Business or business leader** | Applied to business, business organisation or person representing a business itself. | e.g. a drugs firm which made false claims about HIV and TB infections in Kenya. See: (Entry 215) |
| **Celebrity or influencer** | Applied to a famous person, notably in the fields of entertainment, fashion or sport, or social media influencer. Does not apply to those prominent for other reasons such as notable politicians or scientists, i.e. well known politicians would be in politicians group and well-known scientists would be in academics). | e.g. a video clip of villagers throwing a politician into a river was promoted by a well-known Kenyan TV personality. See: (Entry 67) |
| **Domestic politicians, political parties** | Applied to elected national and local politicians, candidates seeking election and/or political parties (not political activists)**.** | e.g. A Senegalese opposition leader claimed falsely that an image online showed prison conditions in the country. See: (Entry 135) |
| **Foreign politicians, political parties** | As for domestic, but foreign**.** | e.g.a Nigerian politician made a false claim about remarks made by a South African politician. See: (Entry 49) |
| **International organisations** | Applied to inter-governmental bodies such as African Union, subregional bodies, UN agencies, WHO. | I have not identified claims from this type of organisation among the first entries studied. I will update as or when I do. |
| **Labour leaders** | Applied to labour union officials, statements from labour unions. | I have not identified claims from this type of individual among the first entries studied. I will update as or when I do. |
| **Mainstream domestic media** | Applied to national and local broadcast, print and online media with a professional news-driven editorial structure. | e.g. a mainstream South African news site, published a false claim about heavy drinking by women in South Africa. See: (Entry 171) |
| **Mainstream foreign media.** | Applied to foreign broadcast, print and online media with a professional news-driven editorial structure. | e.g. The Washington Post published a false claim about heavy drinking by women in South Africa. See: (Entry 171). Note: The Post made a correction. |
| **NGOs & civil society organisations** | Applied to domestic and international organisations. | e.g. a suicide prevention charity made a false claim about suicide rates in Nigeria. See: (Entry 173) |
| **Online news sites and blogs** | Applied to online news sites and blogs that adhere to news-driven editorial standards but are not associated with mainstream media. | e.g. a user-survey website published a false claim about the South African city of Pietermaritzburg. See: (Entry 69) |
| **Religious leaders and institutions** | Applied to official leaders within religious organisations. I do not use it for religious activists or lay people with religious views. | e.g. a Nigerian preacher posted a doctored photo of a rock supposedly floating in the air as proof of God’s power. See: (Entry 24) |
| **Satirists** | Applied to satirists and satire sites who define themselves as satirists or parody accounts, save where deliberate intent to mislead can be proven. | e.g. a satire site published a false claim about Nigerian immigration fraud in the US. See: (Entry 118) |
| **State officials** | Applied to non-politicians speaking for state, such as policy, army or civil servants. | e.g. The Nigerian police chief claimed falsely Abuja was one of the safest cities in the world. See: (Entry 277) |
| **Traditional rulers or community leaders** | Applied to those whose authority comes as a traditional or community leader. Distinct from community or ethnic activist. | I have not identified claims from this type of individual among the first entries studied. I will update as or when I do. |
| **Lesser known but identifiable individuals and organisations** | | |
| **Community networks spreading myths** | Applied to an unsubstantiated claim long entrenched as a community belief – whether it is spread online or offline. | e.g. unknown community members promote a false claim that black people cannot get skin cancer E.g.  https://www.bbc.co.uk/ programmes/p07wvv10 |
| **Hyper-partisan news media or forum** | Applied to broadcast, print or online media that promotes exclusively one-sided information as ‘news’, regardless of its accuracy. | e.g. a partisan news site claimed images showed French army helicopters that crashed in Mali. See: (Entry 157) |
| **Identifiable conspiracy theorists** | Applied to identifiable individuals or groups believing and/or promoting a provably false theory that an event or situation is the result of a conspiracy. | e.g.conspiracy theorists in South Africa promoting a conspiracy theory about Hillary Clinton. See: (Entry 185) |
| **Identifiable ethnic or religious activists** | Applied to identifiable individuals or groups frequently promoting an ethnic or religious agenda. | e.g. a white supremacist in South Africa published a claim that a million white South Africans now live in poverty. See: (Entry 187) |
| **Identifiable political or social activists** | Applied to identifiable individuals or groups frequently promoting a political or social agenda - not a politician or party official. | e.g. activists supporting the DR Congo president created and spread a false claim of army success. See: (Entry 151) |
| **Junk news websites or pages** | Applied to an online news site or page purporting to offer news if majority of content comprises false or misleading information. I do not apply it to tabloid-style sites if a majority of what they report is broadly accurate. | e.g. a junk news site falsely claimed the WHO had warned southern Nigerians against eating beef from the north. See: (Entry 178) |
| **Self-styled health practitioner or news site** | Applied to individuals, companies, blogs or news sites that present themselves as offering health services or advice, without recognition or approval from an approved medical authority**.** | e.g. The self-styled health news site called Natural News promoted a vaccine conspiracy theory report. See: (Entry 186) |
| **Unknown and unidentifiable individuals and organisations** | | |
| **Hoaxers and scammers** | Applied to individuals or online operations that originate or promote false information intended to deceive either for amusement or financial reward. It does NOT refer to hoaxers for political or social effect. | e.g. a hoax web page promoted this among many fake job adverts. See: (Entry 98) |
| **Unidentified individual/s offline** | Applied to unidentified individuals spreading false information offline. | e.g. an unknown individual put up a fabricated sign in a public place. See: (Entry 3) |
| **Unidentified individual/s online** | Applied to individuals whose online identify cannot be verified. May have created a fake or imposter online account for political or social effect, or to cause practical harm e.g. through spreading computer virus. | e.g. unidentified individuals – no known background – created and spread a false claim about government employees in South Africa. See (Entry 4) |

Perceived credibility of source: (i) ‘authoritative sources’ &(ii) ‘community networks’

While it is clear from research that the public do not necessarily have to believe the source of information to be credible in order to believe a claim to be true or potentially true, (See: Pennycook & Rand, 2019 on ‘inattention’), the perceived credibility of the source is shown to increase the likelihood an audience will believe a false claim. The type of source, or sender, that is or may be seen as credible varies greatly depending on, among other factors, the audience, the topic and the context. Different audiences assess credibility in different ways including judging the knowledge the source has or may have of the situation or subject, in particular, to whether they are or may be biased in presenting information. To some audiences, information from people with “medical credentials” is likely to be seen as having have strong credibility, while others will put more faith in information from “alternative” health practitioners. In some countries and contexts, information coming from public authorities is more likely to be seen as credible. In other countries and contexts, authorities are distrusted and information from friendship and social networks is seen as more authoritative (See: Banaji et al, 2019, among others). In the database, I have identified misinformation that is: (i) from a source perceived or likely to be perceived as authoritative – either in general or for that type of claim, (ii) a myth, or false or unsupported belief long entrenched in community networks. I have done so on the following evidence and criteria.

|  |  |
| --- | --- |
| **Types of source perceived or likely to be perceived as authoritative** | |
| **The claim’s type of source** | **Examples from the study** |
| a **false scientific or medical claim made** in a scientific or pseudo-scientific journal or institute **by someone perceived or likely to be perceived by their audience as a credible scientist** | e.g. a false claim that a Gabonese man had created an AIDS vaccine were supported by a supposedly respected French institute – the *Institut International pour le Soutien à la Recherche Scientifique Innovante* – persuading policymakers in Congo and DR Congo to allow its sale. See: (Entry 123) |
| a **claim about a general news event** – not a subject of partisan divide – made by a **media organisation the public broadly trust as an authoritative source on traditional news** | e.g. in east Africa, national media in Kenya, Uganda and Tanzania all reported a claim that the Tanzanian president had mistaken the team the national football team had been beaten by when harshly criticising them for their loss. See: (Entry 2) |
| a **false claim made** by an imposter **in the name of an individual or organisation** – **where the audience believes the claim genuinely represents the views** of that individual or group. | e.g. a false claim that an image showed a tweet by the leader of the ANC Women’s League. Responses showed many believed the tweet, which was fake, was genuine and showed her distress at the economic situation in Zimbabwe and hope that South Africa would not do the same. See: (Entry 39) |
| a **false claim on a technical matter** from a **perceived technical or pseudo-technical expert** | e.g. false claim made by a perceived telecoms expert about control of the telecoms network in DR Congo. See: (Entry 138) |
| a **legal claim** from a **perceived legal expert or official report on a legal matter** | e.g. a false claim made in a report about the meaning of a legal phrase in a UN report on Senegal. <https://factuel.afp.com/non-lonu-nexige-pas-la-legalisation-de-lhomosexualite-au-senegal-mais-le-comite-des-droits-de-lhomme> |
| a **statistical or technical claim** from a **politician, official figure** **quoting official statistics** where those statistics are themselves trusted | e.g. a false claim made by a South African minister that the United Nations recommends a particular police-to-population for effective policing. See: (Entry 166) |
| **Types of claim perceived or likely perceived as credible within a community** | |
| **Type of source** | **Examples from the study** |
| A false **claim that has circulated, online or offline in community settings and networks** for many years | e.g. a false claim about treatment for stroke that has circulated in online and offline community networks around the world for years. See: (Entry 30) |
| A specific false **claim which requires acceptance of a false belief that is entrenched in community networks** | e.g. a false claim that two women assaulted by a crowd in Togo were ‘bird witches’ requires belief in witchcraft and therianthropy. See: (Entry 132) |

Country or countries to which it related

In addition to type of source, I have identified in the database the country or countries to which the topic in the claim related.

Together with information on the country in which the information was observed, this helps to assess the capacity of those who saw and believed the claim to act on the information – a factor that helped shape its potential consequences.

Where the claim referenced either a country or people of a particular nationality, all included as countries to which the claim related. For example, I have identified the false claim that foreign nationals were kidnapping children from schools around Johannesburg, as concerning (i) South Africa, where the schools were, and (ii) Somalia and Pakistan, because the false reports suggested Somali and Pakistani nationals were involved. (See: Entry 54). If the claim is not related to any particular country the entry is listed “Not country specific”. If the claim refers to one or more countries but does not identify them (for example a false claim I listed it as concerning “multiple” countries. (See: Entry 52)

Country or countries in which it was observed

For the same reasons, I have also identified the country or countries in which the claim was observed as circulating. The list is not exhaustive as fact-checking organisations do not necessarily monitor and record all countries in which a claim circulated – noting only those in which *they* observed it. This information is then used, together with information on the country or countries to which the claim related, to help assess who might be affected and the capacity of the audience to act on the information – shaping its potential for consequences.[[82]](#footnote-82)

1. Channels & settings in which it spread

I argue in the study that the context in which misinformation spreads – not only but including the channel or setting in which it appeared – plays a major role in its potential to cause or contribute to substantive consequences for individuals or society.

Other things being equal, I argue, information published on a policy issue over which the public have no direct control, (for example, a false claim related to the size of the police) is likely to have fewer consequences for society if a social media platform, than if spread in parliament during a debate on the police budget – where those who see the claim as credible have the power to act.[[83]](#footnote-83)

For this reason, based on evidence set out in the fact-checks, I have identified in the database the channels [i.e. the broadcast, print or online media] or settings [i.e. parliament or other formal political venues, community meetings, religious meetings and others] – in which the misinformation appeared. The listing of channels or settings in which each entry appeared is not exhaustive as, while all fact-checkers identified where *they* saw the information, and many reported its appearance in multiple settings, many did not appear to research its different settings. Where they did record misinformation as appearing in multiple channels and settings, I record all those identified.

|  |  |
| --- | --- |
| **Channel or Setting** | **Notes** |
| **academic or expert journal or conference, or report of a think-tank** | e.g. staff at a French research institute – ‘*L’Institut International pour le Soutien à la Recherche Scientifique Innovante’* published statements in academic journals supporting use of an untested AIDS treatment. See: (Entry 123) |
| **books** | e.g. the false claim that Hillary Clinton sold US uranium to Russia, in return for payments to the Clinton Foundation, first appeared in a book, and later spread online. See (Entry 185). |
| **community settings** | e.g. the rumour that health researchers working in a village in Ethiopia were harming children spread peer-to-peer in an in-person community setting. See: (Entry 241)  e.g., rumours about the disappearance of a child from a church in Nigeria spread peer-to-peer offline also. See: (Entry 121)  In the study, I identify community settings as including religious or community meetings, social gatherings, people meeting in public transport and other such events and venues. |
| **group or chain emails** | e.g. the false claim that pricking the ears of a stroke victim with a needle could save their lives first spread as a chain email in the 1990s. See: (Entry 30)  I also show in the study that groups such as anti-vaxxers concerned about what they identify as censorship on social media have started to spread their claims via email instead. |
| **hyper-partisan or junk news websites or newspaper (e.g. Zambia Observer – was entry 254** | I identify both hyper-partisan and junk news websites as being unconcerned – for different reasons - by the objectivity and accuracy of information they publish. I identify the pro-Biafran news website, the Biafran Star, as hyper partisan. See: (Entry 184). I identified a site called ‘Reports Fun’ as a junk news site. (<https://africacheck.org/fact-checks/meta-programme-fact-checks/iceland-not-paying-immigrants-marry-countrys-women-and-has>). Neither appeared concerned with the accuracy of what was published. |
| **mainstream print, broadcast or online media** | False information appears often in mainstream media.  **Some originates in its own reporting**. See examples of false reports from a TV station in Kenya. See: (Entry 174) and a newspaper in Nigeria. See: (Entry 172)  **Some appears from publishing or broadcasting the statement of public figures without correcting false claims** – e.g. a Congolese TV station broadcast a speech by President Felix Tshisekedi and did not identify his false claims as such. See: (Entry 245) |
| **other online platforms: from blogs to corporate websites** | False information also appears on blogs, subject-specialist and corporate websites.  For example a website that describes itself as a “database of user contributed data about cities and countries” published a false claim about the South African city of Pietermaritzburg. See: (Entry 69). |
| **parliament or other political or official settings** | False claims often emerge in parliament, official or political hearings and other official settings – where decisions are made by those in authority.  e.g. Nigeria’s president made a false claim about climate change’s effects in a speech to the United Nations. See: (Entry 73). And in South Africa, the police minister made a false claim about police numbers in a parliamentary debate on the police budget. See: (Entry 180). |
| **public signs, posters, product labels** | False claims appear regularly on public signs, posters, flyers and product labels in public settings across the continent. For example, a false claim was put on a public sign put up illegally in Johannesburg. See: (Entry 3). For another example, see the claim made on product labels for an untested AIDS treatment. See: (Entry 123) |
| **social media platforms or messaging apps** | Most of the misinformation research carried out since 2016 has focused on how misinformation and disinformation spreads on social media and messaging apps. Much of the work of fact-checking organisations focuses on misinformation on these platforms – featuring in too many entries to mention. |
| **A combination of channels or settings** | Numerous entries in the database appeared in a combination of channels or settings – this repetition and combination of sources reinforcing effects.  For example, a false claim made by the Nigerian President about climate change was first made in an official setting, broadcast unedited and reported in traditional media, and then circulated on social media (See entry 72). |

1. Factors that drove creation &/or responses to misinformation

Misinformation emerges and spreads for many reasons. Having identified where the entries originated and spread, I also set out in the database – based on a combination of evidence in the fact-checks and online and offline responses to the false claims – what can and cannot be determined about (i) those who originated and promoted the false information, and why, and (ii) impersonal factors such as lack of access to or trust in accurate information that contributed to its creation.

I note in the study that the failure, at some level, of individuals to verify information is a factor in the creation and/or spread of *all* misinformation and, given that it is universal I do not record it in the database as a specific factor. For the same reason, while I also discuss in the study the effects of the structures and practices of (i) traditional media, (ii) social media and (iii) political parties and parliaments in the creation and spread of false information, I do not identify this in relation to the spread of individual claims in the database.

|  |  |  |
| --- | --- | --- |
| **Factor** | **Criteria for assessing what drove creation &/or spread of/belief in the misinformation** | **Examples & evidence** |
| **Errors in understanding and/or explaining information** | * The topic could be considered complex, the claim was difficult to verify, and there was no evidence in the fact-check or other research of intent to mislead | e.g. A Nigerian newspaper misreported a medical survey. It used data that was not representative of the general population. No evidence of intent to mislead. See (Entry 172). |
| * The originator of the claim corrected the claim when alerted to the error – and there was no evidence in the factcheck or other research of intent to mislead | e.g. a journalist in DR Congo claimed a bridge opened by the president had collapsed. When her error was pointed out, she corrected herself and apologised. No evidence of intent to mislead. See: (Entry 149) |
| **Lack of (i) access to or (ii) trust in accurate information** | * Accurate information exists but evidence in the factcheck or other research shows it was not trusted by the audience. | e.g. Biafran separatists saw but did not trust official reports and media reports of the Nigerian President’s attendance at a summit. See: (Entry 47). |
| * Accurate information on the topic was not available – according to evidence in the factcheck or other research | e.g. A minister misstated the number of rapes annually in Nigeria. Her officials blamed this on earlier failure to collect data. See: (Entry 251). |
| **Adherence to a pre-existing attitude, belief or understanding, behaviour** | * Evidence in the factcheck or other research shows audiences held strong, pre-existing views on identity &, social issues – nationality, ethnicity, gender, religion, crime etc | e.g. a false claim that the Biafran leader had addressed the EU parliament. Those sharing the claim were supporters of Biafran independence. See: (Entry 66) |
| * Evidence in the factcheck or other research shows audiences held strong allegiance to a political party or political cause | e.g. a false claim that a critic of former SA President Jacob Zuma was an ally of former apartheid leader De Klerk. Those sharing the claim were Zuma supporters. See: (Entry 71). |
| * Evidence in the factcheck or other research shows the audience was hostile to and/or showed lack of trust in those in authority &/or belief in conspiracy theories | e.g. a false claim about a politician being thrown in a river. Reader responses showed hostility to politicians. See: (Entry 67).  e.g. a false claim about jihadists were using doctored vaccines to spread HIV. Responses showed adherence to conspiracy theories. See: (Entry 84). |
| * Evidence in the factcheck or other research shows the audience was hostile to and/or showed lack of trust to foreign powers | e.g. a false claim that Rwanda called for South Africa to be expelled from the African Union. Those sharing were hostile to South Africa. See: (Entry 65) |
| * Evidence in the factcheck or other research shows audience belief in supernatural or religious phenomena | e.g. a false claim a girl had been turned into monkey after throwing away a Koran. Responses showed existing belief in supernatural effects. See: (Entry 179) |
| * Evidence in the factcheck or other research shows the audience had an existing false understanding of the subject field | e.g. a false claim that a woman had had triplets after having sex with a dog. While some recognise the report as a joke – some indicated they believed the claim was true. This required a false understanding of biological science. (See: link listed Entry 25). |
| **Emotional responses to the misinformation** | * Evidence in the factcheck or other research indicates readers responded as a result of, or to express, amusement, relief or empathy | e.g. online comments showed people shared an online post that falsely claimed that an elderly woman trained 65 cats to steal from her neighbours for humour. See: <https://africacheck.org/fact-checks/meta-programme-fact-checks/no-elderly-woman-ohio-did-not-train-65-cats-steal-neighbours>. |
| * Evidence in the factcheck or other research indicates readers responded as a result of, or to express, shock, anger, or fear | e.g. online comments showed people who responded to a post that showed government employee in South Africa ignoring her customers –expressed anger at the supposed behaviour. See: (Entry 4) |
| * Evidence in the factcheck or other research indicates readers responded by sharing information because they thought the information was or might be true and helpful | e.g. online comments showed people who responded to a post that claimed Kenyans had been ordered to stop repaying loans did so to share supposedly helpful information – an “early Xmas” as one said. See: (Entry 15) |
| * Evidence in the factcheck or other research indicated readers showed interest in the lives of others | e.g. online comments showed people who shared rumours that the Nigerian President had remarried did so to speculate about his personal life. See: (Entry 74). |
| * Evidence in the factcheck or other research indicated readers felt the information showed they were ‘in the know’, and/or confirmed their existing view | e.g. a false claim that video of a flooded classroom showed school conditions in Kenya and other countries. Those who shared the claim indicated feeling confirmed in their views. See: (Entry 82). |
| * Evidence in the factcheck or other research indicated the information is seen as information authorities or others don’t want public to know | e.g. reader responses showed many believed that (i) photos showed a French helicopter crash in Mali, (ii) this was information the French did not want the public to see – and shared it for this reason. See: (Entry 157). |
| * Evidence in the factcheck or other research shows the false information was seen as hopeful or positive and shared for this reason | e.g. online comments suggested many readers shared a false claim that a type of seed was a cure for diabetes because they thought it hopeful. See: (Entry 153) |
| **Economic, health or social circumstances** | * Evidence in the factcheck or other research indicates readers responded due to desire for employment and/or financial security | e.g. people who responded to a false advert for a (non-existent) work position indicated they did so out of desire for employment. See: (Entry 98) |
| * Evidence in the factcheck or other research indicates readers responded due to desire for a cheap/effective medical cure | e.g. people who responded to a hoax offer of a cheap medical ‘cure’ for HIV (a drink based on natural remedies) indicated hope that this would be a cheap and effective treatment. See: (Entry 145) |
| * Evidence in the factcheck or other research indicates readers responded due to desire for another social goal – e.g. emigration, travel | e.g. people who responded to a hoax about a Canadian visa lottery for Nigerians did so from a desire to emigrate to Canada. See: <https://africacheck.org/fact-checks/meta-programme-fact-checks/scam-no-canada-doesnt-offer-visa-lottery-nigerians>. |

1. Factors that drove creation &/or deliberate distribution of disinformation

Given the covert nature of disinformation, it is not always possible – even for those who run the global online platforms[[84]](#footnote-84) – to prove conclusively (i) who originated false online information and (ii) their motivation in doing so. In this database, I take a cautious approach to attribution of intent, identifying an entry as disinformation only when there is clear evidence of intent to deceive: either (i) evidence that information that was presented as being true, had been fabricated or manipulated in a manner that was misleading; or (ii) research evidence on those behind the information. Where there is no such clear proof, I note the possibility there was intent to deceive but identify the entry as misinformation.

Where it is clear that the intent of those who originated and/or promoted the claim was to deceive, I have deduced the motivation, where possible, from either (i) statements they made about their wider intents (ii) analysis of who or what would benefit from the public being misled. Where it is not possible to make these deductions, I make this clear.

|  |  |  |
| --- | --- | --- |
| **Observed or potential effect** | **Motivation deduced based on these criteria** | **Example** |
| **To increase support for a political party, politician, policy or cause, or undermine an opponent** | * The claim was made by a **known politician, political or social activist or partisan media** - or **unknown actor where fabrication shows intent** * The false claim was or might have been expected to have **increased their support or undermined an opponent** | e.g. a fake version of a newspaper’s front page was created by Cameroonian separatists – promoting  their political demands with apparent intent to increase their support. See: (Entry 163). |
| **To distort political participation in a way that benefits one party*.*** | * The claim was made by a **known politician, political or social activist or partisan media** - or **unknown actor where fabrication shows intent** * The false claim could have been expected to have **affected participation to the benefit of one party**. | e.g. a video fabricated evidence of election-related violence and warned of the risk of election rigging.  Research quoted in the study suggest such type of claims can affect participation in elections and benefit a party. See: (Entry 91) |
| **To promote a particular social goal or view** | * The claim was made by a **known or unknown organisation or individual where fabrication shows intent** * The false claim **was or might have been expected to have promoted a particular social goal or view** | e.g. a sign was erected in a suburb of Johannesburg falsely declaring informal waste collectors banned from the district.  If believed, such a claim could have had the effect of deterring these workers. See: (Entry 3) |
| **To increase attention, audience and/or revenue** | * The claim was made by **traditional media, satirists, junk media, social media influencers &/or hoaxers or scammers** whose operating model **requires large audiences** * The false claim **increased, and/or could have been expected to have increased, the audience** with **financial or other benefits** for who spread it | e.g. A junk news site posted a video claiming a scammer who was cursed died barking like a dog.  The site made money from advertising, by attracting a large audience. The video drew such an audience. See: (Entry 14) |
| **To boost the reputation of a company or business sector or undermine a business rival** | * The claim was made by a **business leader, business, (including self-styled health providers)** – **or unknown actors where the topic relates to a business** * The false claim **affected, and/or could have been expected to have boosted or undermined the reputation of a business or business sector or its rivals** | e.g. a wave of false reports appeared online about mobile phone company MTN. The origin is unknown but the claims included fabricated details.  If believed, claims that MTN was shutting its money transfer operation could have undermined its business. See: [https://ghanafact.com/ 2019/09/mtn-mobile-money-service-to-shut-down/](https://ghanafact.com/2019/09/mtn-mobile-money-service-to-shut-down/) |
| **To sow fear, division or cause practical harm through malice** | * The claim was made by **an unknown actor** or **a known politician, political or social activist or partisan media** * The false claim increased or could have been expected to have **increased division, distrust** of particular communities, institutions; **fears** related to crime, safety, people’s finances, &/or created **reputational or practical harm** to individuals, policymakers | e.g. a fake police alert that Islamic jihadists are posing as medical workers to give injections spreading HIV was created and spread by unknown actors.  The claim increased distrust of medical workers and fears of threats to peoples’ safety. It had potential to cause or contribute practical harm to individuals’ health. See: (Entry 84). |
| **To speculate about a topic – presenting speculation as true** | * The claim was **speculation** by any sort of actor - known or unknown – and **used fabricated material** to present its claim as confirmed information * The claim **attracted or would have been expected to have attracted financial or other gratification** | e.g. the false presentation of speculation about upcoming ministerial appointments by fabricating a supposed approved list. See: (Entry 28) |

3] Criteria & evidence of effects, consequences & factors that shape them

Like the decades-old debate about mass media effects that preceded it (from Lasswell on, analysed in McQuail, 2010 and elsewhere), the academic debate about the effects of misinformation on individuals and society has grown in volume and vigour since 2016.

Researchers and scholars have, since 2016 in particular, sought to establish evidence of the consequences, or lack of them, of particular forms or topics of false information; focusing primarily on the potential of different forms of misinformation to (i) alter the outcome of elections, (ii) spur civil unrest and/or (iii) affect public health.

What this research has not done, to date, has been to theorize the combination of factors that affect the potential of a particular false claim or narrative to cause or contribute to these or other such consequences, and thereby enable society to counter substantive negative effects of particular types or topics of misinformation without unnecessarily or unduly restricting freedom of speech – including the freedom to make false claims.

In the study, I set out a theory – drawn from a combination of (i) close examination of the sample of false and misleading information set out in the database and (ii) a range of existing scholarly evidence of information and misinformation effects – of three factors that shape the potential of different types of mis/disinformation in different contexts to cause or contribute to substantive consequences for individuals and society.

The three factors I identify as shaping the potential for consequences are:

Whether the claim is substantively false, and its capacity to change understanding, attitude or behaviour based on information that is false;

whether sufficient people consider it to be true to have the particular effect;

the audience’s capacity and motivation to act on the information, and its power to affect others if it does so

**In the third section of the database**, I set out my findings on: (Category 18) The perceived credibility of the claim; (Category 19) The observed & potential changes in understanding, attitude, behaviour caused; (Category 20) Whether this (i) had effect w/out action (ii) audience had capacity & motivation to act on false aspect.

In categories 21-25 I then use evidence from the fact-checks and additional research to seek to determine whether specific substantive consequences, caused in whole or part beyond reasonable doubt by the examples of misinformation studied, had occurred.

And in categories 26-30, I set out this theory of potential consequences in practice, using: (i) evidence from the fact-checks specific to potential consequences (e.g. indications of individuals’ intent to use a fake medication) + (ii) a range of existing scholarly evidence identified from information effects in similar circumstances + (iii) the theory set out above and the findings sets out in the first 20 categories in the database – to come to conclusions. Below I set out how I assess the criteria used to assess these factors and how I use this information in the study.

1. Factors shaping audience perception of the credibility of the claim

I argue in the study that misinformation does not always require being seen as true to have consequences such as effects on the mental health of its subjects[[85]](#footnote-85) or by exposing audiences to a computer virus via clickbait reports[[86]](#footnote-86).

In the majority of cases in the study, however, I argue – as summarised above – that false information needs to be perceived as credible, or potentially credible, by a sufficient number to have effect. Based on a review of relevant studies, I set out six factors believe that affect the likelihood information will be perceived as credible.

|  |
| --- |
| **Factors that increase the likely perceived credibility of false information** |
| 1. **Repetition of the claim**   Numerous studies (Allport & Lepkin, 1945; Swire-Thomson et al, 2017; Pennycook et al, 2018; Fazio et al, 2019; Lacassagne et al 2022) show repetition of consonant information increases its perceived credibility. In the database, I identify evidence that a claim has been repeated frequently over time as increasing the likelihood of perceived credibility – regardless of topic. |
| 1. **Perceived plausibility of the claim and its presentation**   Studies such as Madrid-Morales et al, 2021, show information is also more likely to be perceived as true if (i) the claim and (ii) its presentation are perceived as potentially plausible and credible.  I.e., (i) the claim is perceived to show external plausibility, for example there is evidence of previous examples of the sort of thing claimed having happened previously. (See: Hassan & Hitchen, 2020 – on the effect on the perceived plausibility of claims that the president of Nigeria had died in office, of the fact that an earlier president had died in office).  (ii) the claim shows apparent internal logic and consistency e.g. numbers presented add up, and the presentation is plausible e.g. the layout of the information, presentation of logos, and writing style are consistent with what is expected from that source. |
| 1. **Ease of processing of the concept &/or information in the false claim**   Other studies show that information that is easier to process (e.g. using images and large font), is more quickly & superficially processed and likely to be perceived as credible than hard to process information. (See Alter & Oppenheimer, 2009. Pennycook & Rand, 2019).  This effect may, of course, be in conflict with the presentation style expected for the source. For example, information said to come an academic medical journal would be expected to be written in the style of such a journal – not in simple language with pictures. |
| 1. **Affective and/or aligned with pre-existing views**   Studies also show that, other things being equal, information that is affective – arousing anger, fear or humour etc – and/or aligned with individuals’ pre-existing views is likely to be more quickly and more superficially processed than cognitive information (See: Berger, 2011. Brady et al, 2017) and/or incongruent information (See: Nyhan & Reifler, 2010. Lewandowsky et al, 2012. Mitchell et al, 2018) – thus more likely to be perceived credible. |
| 1. **The source or channel is perceived as credible – at least to some of audience**   Further research shows the perceived credibility of a source (e.g. a respected relative; a respected institution) or a channel (e.g. a trusted media) affects the perceived credibility of the claim. (Pornpitakpan, 2004. Van der Meer and Jin, 2019. Banaji et al, 2019) – the more respected the source the more likely the claim is to be perceived as credible.  (Note: see information in category 12 for more). |
| **The false claim is not easily or automatically disproven**  The more easily, quickly a claim is disproven, the less likely it is to be perceived over time as credible. For example – a claim that a celebrity has died may be perceived to be credible in the short-term but this perceived credibility is unlikely to outlast the first public appearance, alive and well. See: Entry 127 – on false claims of the death of footballer Rigobert Song). Claims related to more complex matters – for example statistical claims about landownership in South Africa (See entry 16) – are harder and slower to disprove. |

In the database, I identify entries as either (i) perceived or likely to be perceived as credible or (ii) perceived or likely to be perceived as not credible. I explore in the study contexts in which the question of credibility is more complex than this implies. I make assessments on the following criteria:

|  |  |  |
| --- | --- | --- |
| **Perceived credibility** | **Criteria** | **Examples** |
| **Perceived as credible or likely to be perceived as credible** | * evidence from the fact-check showed a **sufficient number** of responses that could be reliably identified[[87]](#footnote-87) accepted **the claim as sufficiently credible**.[[88]](#footnote-88) | e.g. Reader responses to a false claim about a flooded classroom showed a substantial share of readers perceived the claim as credible. See (Entry 82) |
| * in the absence of evidence of audience responses, the claim reflected **four or more of the six above factors**: (i) repetition; (ii) perceived plausibility; (iii) ease of processing; (iv) affective information effect; (v) perceived credible source; (vi) not easily disproven. | e.g. a false claim about the suicide rate in Nigeria showed four of six factors that make persuasion likely. (repetition; perceived plausibility; easy to process; perceive credible source). See: (Entry 173) |
| **Perceived as false, or likely to be perceived as false** | * evidence from the fact-check showed an **insufficient number** of responses that could be reliably identified saw the claim as true, and a clear **majority found it false**. | e.g. a large majority of readers of a post parodying attitudes to northern Nigerians recognised it was a parody. See: (Entry 41) |
| * in the absence of evidence of audience responses, the information reflected only **three or fewer of the six above factors** | e.g. a false claim about ministerial appointments in Nigeria showed only two persuasive factors. (perceived plausibility; easy to process) See: (Entry 29) |

Observed & potential change in understanding, attitude, behaviour caused

In 1960, the sociologist and communications theorist Joseph Klapper identified four main forms of change to understanding, attitude or behaviour caused by information:   
(i) a reinforcement (or prevention of erosion) an existing understanding, attitude or behaviour; (ii) a minor change in form or intensity of an existing understanding, attitude or behaviour; (iii) a conversion of one existing understanding, attitude or behaviour to another and (iv) the creation of a new understanding, attitude or behaviour.[[89]](#footnote-89)

Using this model, I identify observed and potential specific substantive changes caused to understanding, attitude or behaviour by the misinformation in the database based on whether evidence showed:

1. readers already had an existing understanding, attitude or behaviour related to the topic;
2. the topic was new and unknown to the audience and evidence shows a new specific substantive understanding, attitude or behaviour was created;
3. if it was known, evidence shows the existing specific, substantive understanding, attitude or behaviour changed either in intensity or from one thing to another

See below the findings, criteria and examples of entries where this was applied

|  |  |  |
| --- | --- | --- |
| **Change** | **Criteria** | **Examples** |
| **No evidence observed of specific substantive change in understanding attitude etc on a false basis** | * No evidence was available – in reader responses or other research material – of the information’s effect on audience understanding, attitude or behaviour | e.g. the fact-check of a claim, made on broadcast TV, provided no evidence of effect on reader understanding. See: (Entry 166) |
|  | * The claim was seen as not credible by most of its audience – so had no substantial potential to cause a specific, substantive change in understanding, attitude, or behaviour on a false basis. | e.g. The audience recognised a claim about a mountain road in Kenya was false and thus it could not change their understanding, attitude or behaviour. See: (Entry 42) |
|  | * The claim was either (i) false in some particular aspect but the overall picture it painted was broadly accurate or (ii) the point on which it was false was insubstantial. It had no substantial potential to cause a specific substantive change based on what was false in the claim. | e.g. The claim that a photo showed a building burned down in Nigeria was wrong but similar buildings had been, so the broad understanding was accurate. See: (Entry 50) |
| **Observed reinforcement, on a false basis, of existing understanding, attitude or behaviour** | * Evidence in the fact-check shows the audience or part of the audience already had an understanding, attitude or behaviour * The claim was substantively false but was perceived as credible by the audience or part of the audience * Online comments and/or other evidence in the factcheck or other research indicate the claim confirmed the existing understanding, attitude or behaviour of these readers | e.g. A post falsely alleged SA media failed to report alleged assassination plot against ex-president Zuma. See: (Entry 19). Evidence in the factcheck shows (i) many of those who responded had an existing attitude to and understanding about Zuma and the media (ii) many saw the false information as credible (iii) their comments online indicated it confirmed their view. |
| **Potential reinforcement, on a false basis, of existing understanding, attitude or behaviour** | * Readers had or could be expected to have had an understanding, attitude or behaviour * The claim was substantively false but was or could have been perceived as credible * There was insufficient evidence of effect on reader understanding, attitude or behaviour | e.g. a false claim WHO warned southern Nigerians about meat in which northerners deliberately cultivated a disease. See: (Entry 178)  I identified the entry as having potential to reinforce an existing attitude because:  (i) Many southern Nigerians believed that northern Nigerians intended them harm. (ii) The claim was false but could have been perceived by many as credible. (iii) Insufficient evidence was available of effect on understanding. |
| **Observed minor change, on a false base, in form, intensity of existing understanding, attitude or behaviour** | * Evidence in the fact-check shows readers already had an understanding, attitude, behaviour * The claim was substantively false but was perceived as credible by a substantial number of readers * Evidence in the fact-check showed a minor change in the existing understanding, the intensity of attitude (e.g. expressions of intense emotions such as outrage, fear) or minor changes in behaviour | e.g. A false claim that foreign nationals were responsible for burning down a historic building. See: (Entry 32)  (i) Evidence showed many readers held existing understandings about and attitudes to foreign nationals. (ii) The claim was false but perceived by many to be credible. (iii) Readers expressed anger (intensification of attitude). |
| **Potential minor change, on a false base, in form, intensity of existing understanding, attitude or behaviour** | * Readers had or could be expected to have had an understanding, attitude or behaviour * The claim was substantively false but was or could have been perceived as credible * There was insufficient evidence of effect on reader understanding, attitude or behaviour | e.g. A minister overstated the number of ‘fraudulent marriages’ in South Africa. See: (Entry 177)  (i) Many who heard this would be expected to have had an understanding of the practice. (ii) The claim was narrowly but substantively false but seen as credible (iii) Any effect on understanding would have been minor. No evidence of an effect was provided. |
| **Observed conversion, on a false basis, of one existing understanding, attitude or behaviour to another** | * Evidence in the fact-check shows the audience had a prior related understanding, attitude or behaviour * The claim was substantively false but was or could have been perceived as credible * Evidence in the fact-check (e.g. written statements, evidence of behaviour) showed a change in understanding, attitude or behaviour | e.g. a false claim banning informal workers from a district of Johannesburg. See: (Entry 3)  Evidence in the factcheck showed: (i) Many residents had prior belief these workers were permitted (ii) the claim was false but seen as credible (iii) conversion to a belief that the workers were banned from the area. |
| **Potential conversion, on a false basis, of one existing understanding, attitude or behaviour to another** | * Readers had or could be expected to have had an understanding, attitude or behaviour * The claim was substantively false but was or could have been perceived as credible * There was insufficient evidence of effect on reader understanding, attitude or behaviour | I have not identified this in the first claims reviewed for the database.   I will update when an example when I do so. |
| **Observed creation of a new understanding, attitude or behaviour** | There is no evidence in the factcheck that the audience did or could have had a prior understanding, attitude or behaviour related to the issue  The claim was substantively false but was perceived as credible  Evidence in the factcheck shows a new understanding, attitude or behaviour. (**Note**: a ‘new understanding’ may be very narrow – limited to believing an event which did not happen had occurred). | e.g. a false claim that pricking a stroke victim’s ears or fingers with a needle is a cure. See: (Entry 30)  This was (i) an issue on which there was no evidence readers had a prior understanding (ii) the claim was substantively false but perceived as credible (iii) evidence in the factcheck showed a false new understanding that this was a credible treatment. |
| **Potential creation of a new understanding, attitude or behaviour** | * There is no evidence in the factcheck that the audience had could had have had a prior understanding, attitude or behaviour related to the issue * The information was substantively false but was perceived as credible * There was insufficient evidence of effect on reader understanding, attitude or behaviour | e.g. a false claim that exaggerated the amount of domestic waste produced each year in Nigeria. See: (Entry 192)  (i) There was no evidence readers had a prior understanding of the issue (ii) the claim was false but perceived as credible (iii) it could have affected understanding but there was no evidence of effect. |

1. Whether (i) had effect w/out action (ii) audience had capacity & motivation to act on false aspect

If misinformation causes or contributes to a false understanding or a change in attitude based on a false understanding but this causes no substantive wider effects to the well-being of individuals or society in itself – I do not identify this as a substantive consequence for individuals or society. For that to happen, requires in most cases that some form of action occur.

False information related to an individual may, of course, affect their mental health and well-being, without any further action required for this to occur. (See: Entry 26). In the case of clickbait that exposes a reader to the automatic release of a computer virus, no action after accessing the false information is required for a harmful effect to take place. (See: Entry 176)*.* And misinformation may also cause harm to individuals’ engagement with others in society – through either a specific effect on trust or through the meta-misinformation effect, where it is inaction, not action, that occurs.

For this reason – while I am not able to identify in this study meta-misinformation effect as caused by the perceived wider prevalence of misinformation – I set out in the database where there was and was not evidence from the factcheck or other sources that the claim caused a “direct substantive negative emotional effect to anyone concerned”.

At the same time, I argue that in the majority of cases in the study, for false information to cause or contribute to specific, substantive negative consequences for individuals or society, either the public or policymakers must have both (i) the capacity (ii) motivation to act on the basis of the information in the claim that was false.

Where they do so, I argue, the scale and severity of the consequences caused will depend on both the nature of the claim and the capacity and motivation of those persuaded by the claim to act in ways that will affect not only themselves but others – as members of the public or people in authority – policymakers or decisions-makers in politics, business, the military, civil society or other positions of authority. And for the purposes of the database, I have assessed these factors based on the criteria set out below.

|  |  |  |
| --- | --- | --- |
| **Whether evidence shows the claim caused a direct effect without further action required** | | |
| **The false claim had or may have had substantive effect with no further action required** (e.g. **effect on mental health or via unleashing computer virus, malware**) | The claim **directly affected the attitudes or emotions of those who observed it** in a way that caused a **substantive effect to their mental well-being** | e.g. a Zambian man, attacked in 2017, was re-traumatised by circulation two years later of images falsely suggesting he was victim of new violence. See: (Entry 58) |
| The claim directly affected the actual or perceived reputation of those who are the subject of the misinformation in a way that **caused substantive effect to their mental well-being**. | e.g. an anti-FGM campaigner in Kenya whose reputation was affected by false claims she had undergone FGM, was severely traumatised by the reputational effect. See: (Entry 26) |
| The act of accessing the information **caused a direct effect e.g. exposure to a computer virus** | e.g. people who clicked on link in false report of the death of a singer were affected by a computer virus. See: Entry 176. |

|  |  |  |
| --- | --- | --- |
| **Capacity & motivation to act on false claim** | **Criteria** | **Examples** |
| **The public had capacity and may have been motivated to act on the information – now or in future**  Determined by a combination of (i) the topic, (ii) context, or opportunity, (iii) prior understanding or attitude, (iv) information effects | The claim promoted or discouraged, then or in the future, **behaviour on the individual’s own behalf,** and they had a plausible opportunity to do so within a timeframe when the false understanding caused would have had a determining effect on actions. | e.g. A false claim that encouraged job seekers to respond to a (fake) job advert. See: (Entry 98)  Or a false claim that discouraged people using a particular health treatment – in this case the general use of vaccines. See: (Entry 186) |
| The claim promoted or discouraged, then or in the future, **behaviour toward others** and they had a plausible opportunity to do so within a timeframe when the false understanding caused would have had a determining effect on actions. | e.g. A false claim that encourages individuals to abuse and engage in conflict with others and they had the capacity at the time to do so. See: (Entry 174) |
| **Policymakers had capacity and – if they believed it, or if they believe the public believed it - may have been motivated to act – now or in future**  Also determined by a combination of (i) the topic, (ii) context, or opportunity, (iii) prior understanding or attitude, (iv) information effects – as above. | The claim promoted or discouraged, then or in the future, **decisions** or **policy change** which policymakers had capacity & might have had motivation to act on – **if they believed it** – and had a plausible opportunity to do so within a timeframe when the false understanding would have had a determining effect. | e.g. A claim that vaccines caused children’s deaths could, if believed by policymakers, have caused them to introduce anti-vaccination policies. See: (Entry 186)  Note: In the study, I show evidence that false claims about vaccine effects contributed to policy changes on vaccines in northern Nigeria in 2002/3 and effects this had. |
| The claim **increased public pressure** for a decision or policy which policymakers had capacity & opportunity to implement within a plausible timeframe. | e.g. false claims three African countries were boycotting a meeting in South Africa built public pressure on Nigeria to follow suit. It had capacity to do so. See: (Entry 57) |
| **Neither public nor policymakers have had capacity and been motivated to act on the false aspect in a way that would cause a substantive consequence** | The false claim **promotes no particular behaviour** nor any **attitudinal or cognitive shift** which would be likely to have effects on behaviour in the longer-term. | e.g. a false claim that a large rock had been found floating in mid-air in Jerusalem created no opportunity to act. See: Entry 24. |
| Those **who have capacity to act on the information are shown or would be expected to know the claim was false or misleading** and thus not be motivated to act on the claim. | e.g. a false claim that Nigeria’s president had remarried. See: (Entry 74). Those who had any capacity to act on the claim (members of the family circle) all knew the claim to be false |
| The public or policymakers have capacity to act but **the claim is only narrowly inaccurate** (i.e. either (i) **only marginally inaccurate** or (ii) **not materially false**) or (iii) **the topic is inconsequential**– so any actions taken would  (i) be indistinguishable from action based on wholly accurate information  (ii) not be based on the element that was false  (iii) be inconsequential | e.g. a false claim misidentified images as showing school conditions in several countries but the real conditions in those countries were similar to those shown. See: (Entry 82) |

1. Whether specific substantive consequences observed - & if so what field

As explained in more detail elsewhere, I distinguish in the study between **changes to understanding or attitude**, and **consequences**. If misinformation causes a change in understanding or attitude – but this change causes no observed lasting, material consequences for individuals or society in itself – I do not identify the change as an observed, specific substantive consequence. I also do not include transitory emotional responses such as anger, or inflated or wounded pride, unless shown to cause lasting negative effects to an individual’s physical or mental well-being.

Where I do identify such consequences, I use the term **specific** to make clear that I am referring to consequences caused by the claim itself, not secondary or indirect effects.[[90]](#footnote-90) I use the term **substantive consequence** to refer to a *material, meaningful consequence for the well-being of the individuals or society*. I do identify include a limited number of *psychological effects* including panic, chronic fear and chronic distress as substantive consequences for the effects these are recognised as having on long-term mental and physical health.[[91]](#footnote-91) I also identify certain *political outcomes* such as the distortion of an election outcome, policy or decision-making that affects individuals or society as a substantive consequence. I recognise that some consequences which occurred or may have occurred as a result of the examples of misinformation examined will not have been recorded and that I do not have the means to establish whether they did. The fact I do not record an observed consequence is not proof one did not occur. I also note that many of consequences I record are or may be the result of multiple contributory factor**s** – not one single factor – and hence I identify the effects of claims as having, in these cases, caused or contributed to the consequences.

In the study I note numerous examples of false information that neither caused nor had any substantial potential to cause substantive consequences.[[92]](#footnote-92) However, drawing on evidence provided in the fact-check and/or in subsequent research, I have found a substantial number of entries can be shown to have caused or contributed to substantive consequences for individuals, groups or wider society. And I identify these as occurring in 10 broad fields. Like others, this list is not exhaustive but reflects the forms of effect or possible effect identified from the sample. This comprises consequences for: (i) democracy & governance; (ii) international relations/cross-border conflicts; (iii) public health; (iv) mental health; (v) discrimination, verbal & physical abuse; (vi) civil society violence, insurgencies; (vii) justice system; (viii) business reputation, economic activity; (ix) the environmental; (x) practical harms e.g. computer virus, loss etc.

I set out below: (i) fields & types of observed consequences identified in the study or other research; (ii) the criteria used; (iii) examples from this study or other research.

|  |  |  |
| --- | --- | --- |
| **Fields & Type of consequence** | **Criteria for finding of observed consequences** | **Examples – consequences observed – in this or other studies** |
| **MISINFORMATION EFFECTS ON DEMOCRACY, GOVERNANCE** | | |
| **Distorted voting preferences during an election campaign** | | |
| Distorted voting preferences during election campaign | * Compelling evidence from academic or other study the misinformation had effect on voter preferences   **Note**  Anecdotal evidence would be insufficient to show substantive effect on election result | Evidence from other studies is complex. Gunther et al, 2018, argued misinformation in the US 2016 campaign may have caused such consequences. Guess et al, 2020 argued effects were limited or nil. I conclude such consequences can occur but rare & in specific context. **No such observed consequences were found in the first entries studied.** See discussion of effects in the study*.* |
| **Distorted participation in elections, political protests** | | |
| Distorted participation in election by creating false belief about risks of participation | * Compelling evidence from academic or other study the misinformation had effects on participation * Direct evidence from credible sources (e.g. electoral officials) the misinformation had effects on participation | Evidence from other studies (Bratton et al, 2013, Norris, 2014, etc) suggests misinformation about the risks of violence relate to elections deters voter participation. I accept this finding.  **No such observed consequences were found in the first entries studied.** See discussion in the study*.* |
| Distorted participation in election by creating false belief about election rules/procedures | * Compelling evidence from academic or other study the misinformation had effects on participation * Direct evidence from credible sources (e.g. electoral officials) the misinformation had effects on participation | Anecdotal evidence from election monitors indicates this effect may occur but I have not identified studies of this specific issue. **No such observed consequences were found in the first entries studied**. See discussion of effects in the study. |
| Distorted participation in election by creating false belief about likely outcomes | * Compelling evidence from academic or other study the misinformation had effects on participation * Direct evidence from credible sources (e.g. electoral officials) the misinformation had effects on participation | Studies from 1980s US (Jackson, 1983) to recent elections in Africa (Shenga & Pereira, 2019) suggest that indicating that an election’s outcome is in some way pre-determined reduces voter participation. **No examples of any such consequence were observed in the first entries studied**. See discussion of effects. |
| Distorted participation in political protests by creating false belief about risks | * Compelling evidence from academic or other study the misinformation had effects on participation * Direct evidence from credible sources (e.g. independent observers with sound methodology) of effects | Anecdotal evidence indicates this effect may occur but I have not identified studies of this specific issue. **No examples of any such effect were observed in the first entries studied**. See discussion of effects in the study. |
| **Distorted public policy – outside election process** | | |
| Distorted policies directly by misleading policymakers about a reality | * Compelling evidence from academic, other study * Direct evidence from policymakers (e.g. an admission by an official) misinformation influenced a policy decision | Many studies show this effect. e.g. the influence of false info on HIV/AIDS policy in South Africa 2000-06. (Chigwedere et al, 2008). From this study, **I argue** **a false claim** that a supposed AIDS treatment was effective and backed by a French institute **helped convince Congolese policymakers to allow its sale.** See: (Entry 123). See discussion in the study. |
| Distorted policies indirectly by creating or undermining political pressure on an issue | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. an admission by an official) that public opinion affected by misinformation influenced policy decisions | Many studies have identified the effect of misinformation on public attitudes to health issues & how this shaped policy (See Ayodele, 2007 re-polio vaccines in Nigeria).  **No examples of any such effect were observed in the first entries studied.** See discussion in study. |
| **MISINFORMATION EFFECTS ON INTERNATIONAL RELATIONS, CROSS-BORDER CONFLICTS** | | |
| **Distorted international relations** | | |
| Distorted inter-governmental relations by misleading the public about a false reality | * Compelling evidence from academic or other study * Direct evidence from credible source that misinformation influenced inter-governmental relations in this way by affecting public opinion. | Evidence from other studies indicate short-term exposure (See studies of role of misinformation in UK Brexit vote: Marshall, H. Drieschova, A. 2020) and  long-term exposure to information (See Foos & Bischof, 2021, on influence of media on Brexit vote) can affect inter-governmental relations directly or via public pressure. **No examples of such consequences were observed in the first entries** **studied.** See discussion in study. |
| Distorted inter-governmental relations directly by misleading the policymakers about a false reality | * Compelling evidence from academic or other study * Direct evidence from credible source that misinformation influenced inter-governmental relations | I have not identified studies of this effect. **No examples of such consequences were observed in the first entries studied**. See discussion of effects in study. |
| Distorted international relations at level of individuals, business | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. travel industry, businesses) misinformation had effects on individuals, business operations | Anecdotal evidence indicates this effect may occur but I have not identified studies of this specific issue. In this study, **I argue that following unproven claims** that Nigerians were killed in a wave of xenophobic violence in South Africa, a **misleading claim** that a chain of shops in Nigeria was South African owned **caused or contributed to the shops being attacked by protestors**. See: <https://www.icirnigeria.org/fact-check-are-shoprite-supermarkets-in-nigeria-franchise-outlets/>. See discussion in the study. |
| **Distorted course of cross-border conflicts** | | |
| Caused or distorted cross-border conflicts by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials) that misinformation affected decisions in a cross-border conflict | Numerous studies have examined the role of false information in causing or distorting the conduct of wars (McNamara, 1995. Levine, 2012 etc). **No examples of any such effect were observed in the first entries studied.** |
| Caused or distorted cross-border conflicts by distorting public support for or against action | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials) that public opinion affected by misinformation influenced decisions in a cross-border conflict. | Numerous studies (See Baum & Groeling, 2010 and others) have examined the role of information in causing or shaping wars or conflicts this way by affecting public support.  Among the first entries in the study I identified one example of such a potential effect: **a** **false claim** UN troops collaborated with Uganda-based rebels in a massacre in Eastern DR Congo which **may have** **contributed to subsequent attacks on a UN base**. See (Entry 160) See the discussion in study. |
| **MISINFORMATION EFFECTS ON PUBLIC HEALTH** | | |
| **Distorted readers’ health** | | |
| Distorted health behaviour in way that caused direct harm to health of self or others | * Compelling evidence from academic or other study * Direct evidence from credible sources health behaviour was affected by misinformation & caused direct harm e.g. taking harmful medication | Many studies of this effect exist (Larson, 2018. Cunliffe-Jones 2020.[[93]](#footnote-93) Motta & Stecula, 2021. & others). **No examples of any such effect were observed in the first entries studied.** See the discussion in the study. |
| Distorted health behaviour in a way that caused indirect harm to health of self or others | * Compelling evidence from academic or other study * Direct evidence from credible sources that misinformation affected health behaviour in a way shown to cause indirect harm to health | Studies have also shown this effect (see Thielman et al 2014, and others).  Among the first entries in this study, several showed this effect.  **I argue in the study**, for example, that a **false claim** that cassava cures prostate cancer, may have **caused some people living with prostate cancer to stop taking actual cancer medication.** See: <https://africacheck.org/fact-checks/spotchecks/no-cassava-doesnt-cure-cancer-and-could-make-it-worse>. See the discussion of effects in the study. |
| Distorted behaviour in a way that increases other physical risks to self or others | * Compelling evidence from academic or other study * Direct evidence from credible sources that behaviour was affected by misinformation in a way shown to cause physical risks to self or others e.g. copycat of risky behaviour | Studies show misinformation causing or contributing – for example during a natural disaster (Torpan et al, 2021 etc) or copycat effect – to other forms of dangerous behaviour.  **No such consequences were observed from the first entries studied**. See the discussion in the study. |
| **Caused or contributed to negative sexual behaviours** | | |
| Caused or contributed to sexual behaviour with negative physical or mental health effects | * Compelling evidence from academic or other study * Direct evidence from credible sources that behaviour was affected by misinformation in a way shown to cause negative physical or mental health effects | Several studies of this effect exist. E.g. see Meel BL, (2003)on the myth that sex with a virgin would cure AIDS.  **No such consequences were observed from the first entries studied**. See discussion of effects in the study. |
| **Distorted public health policy** | | |
| Distorted health policies directly by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources that misinformation affected health policy decisions. | As noted above, many studies (Chigwedere et al, 2008 etc) show misinformation can shape policy.  **I argue in the study that**, among entries examined, a **false claim** that an AIDS treatment was effective **altered public policy, persuading policymakers in Congo & DR Congo to allow its sale.** See: (Entry 123). See discussion in the study. |
| Distorted health policies indirectly by creating or undermining public support for policy | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admissions by officials) that public opinion was affected by misinformation and this affected public health policy | Many studies have identified the effect of misinformation on public attitudes to health issues & how this shaped policy (See Ayodele, 2007 re-polio vaccines in Nigeria).  **Among the first entries studied, I did not identify any examples of this consequence**. See discussion of effects in the study. |
| **MISINFORMATION EFFECTS ON MENTAL HEALTH** | | |
| **Caused or contributed to individual distress or public stress, fear or panic** | | |
| Caused or contributed to chronic fear or distress for the subject(s) of misinformation. **Note** I identify **chronic fear**, **distress** as a harm given their known effects on health. | * Compelling evidence from academic or other study * Direct evidence from credible sources that the subject(s) of the misinformation suffered chronic (>month) fear or distress. | Studies identify distress or fear experienced for >one month as chronic (Shalev, 2010 & others) causing potential long-term harm to physical, mental health. **I argue in this study** that a **false claim** that hurt the reputation of an anti-FGM activist left her distraught for weeks, and thus caused **harm to mental health**. See: (Entry 26). See the discussion of effects in the study. |
| Caused or contributed to fear, stress or panic among general public | * Compelling evidence from academic or other study * Direct evidence from credible sources that misinformation caused serious fear, distress or panic among the public. | Anecdotal evidence suggests false information has similar capacity to induce fear and stress in the general public. **In this study, I argue** that **a false claim** that children were being kidnapped from schools **caused parents stress, fear and public panic**. See: (Entry 54). See discussion in study |
| **Distorted public policy on mental health** | | |
| Distorted policies on mental health directly by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources that misinformation affected mental health policy, or decision-making | I have not identified any well-researched studies of this effect. **No such consequences were observed in the first entries studied.** See discussion of effects in the study. |
| Distorted policies on mental health indirectly by creating or undermining public support for policy | * Compelling evidence from academic or other study * Direct evidence from credible sources that public opinion was affected by misinformation and this affected mental health policy | I have not identified any well-researched studies of this effect.  **No such consequences were observed in the first entries studied.** See discussion of effects in the study. |
| **MISINFORMATION EFFECTS ON ABUSE, DISCRIMINATION** | | |
| **Used to enact abuse & discrimination** | | |
| Misinformation used to enact discrimination, abuse | * Compelling evidence from academic or other study * Direct evidence from credible sources of the use of false claims in acts of abuse of &/or discrimination | Studies show that misinformation is used to enact discrimination and abuse. (See Ellis-Peterson, 2018. Kaye, 2019, & others)  In this study, I argue a **false claim** **was used to enact abuse of the women**. See: (Entry 132). See the discussion in the study. |
| **Distorted public policies related to discrimination, abuse** | | |
| Distorted policies related to abuse & discrimination directly by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admissions by officials) that misinformation affected policy on discrimination, abuse | I have not identified any well-researched studies of this effect.  **No such consequences were observed in the first entries studied.** See discussion of effects in the study. |
| Distorted policies related to abuse & discrimination indirectly by affecting public support for policy | * Compelling evidence from academic or other study * Direct evidence from credible sources that misinfo affected public opinion & this affected policy re abuse discrimination | I have not identified any well-researched studies of this effect.  **No such consequences were observed in the first entries studied.** See discussion of effects in the study. |
| **MISINFORMTION EFFECTS ON VIGILANTE VIOLENCE, CIVIL UNREST, INSURGENCIES** | | |
| **Caused or contributed to violence in civil society** | | |
| Caused or contributed to vigilante violence against an individual or individuals | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. police &/or interviews with participants attributing actions to the claims) that misinformation caused or contributed to vigilante violence | Numerous studies and reports (See Banaji et al, 2019. Nur, 2019) have found evidence that misinformation caused or contributed to vigilante violence.  Among the first entries, I argue **a false claim** that the body of a missing child had been found under a church **caused a violent attack on the church and injuries ‘to many.’** See: (Entry 121). |
| Caused or contributed to violence between political, ethnic or religious groups  [Note, I do not include peaceful protests; demos in this category] | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. police, &/or interviews with participants attributing actions to the claims) that misinformation caused or contributed to violence between political, ethnic or religious groups | Numerous studies and reports (See Ellis-Peterson, 2018. Kaye, 2019. Adegoke, 2018) have found evidence that misinformation caused or contributed to political & ethnic violence.  **No such consequences were observed in the first entries studied.** I will update if any are.  See the discussion in the study. |
| **Distorted course of insurgencies** | | |
| Distorted course of insurgencies by misleading policymakers or insurgents about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission leaders of an insurgency or officials) that misinformation influenced decisions by those waging, fighting an insurgency | I have not identified any well-researched studies of this effect.  **No such consequences were observed in the first entries studied.** See discussion in the study. |
| Distorted course of insurgencies by distorting public support for or against the parties | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by leaders of an insurgency or officials) that misinformation affected public opinion & this affected decisions by those waging, fighting an insurgency | I have not identified any well-researched studies of this effect in insurgencies.  **No such consequences were observed in the first entries studied.** See discussion in the study. |
| **MISINFORMATION EFFECTS ON JUSTICE SYSTEM** | | |
| **Distorted operation of justice system** | | |
| Distorted the course or outcome of particular court cases | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. lawyers, legal analysts) that misinformation in the public domain influenced the course or outcome of a particular case | Numerous studies (Frenda et al, 2011. Dando 2020 etc) have been published of ‘misinformation effect’ in court cases.  **No such consequences were observed in the first entries studied.** See discussion in the study. |
| **Distorted public policies related to justice system** | | |
| Distorted justice policies directly by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. officials) that misinformation influenced justice policy | I have not identified any well-researched studies of this effect. **No such consequences were observed in the entries studied.** |
| Distorted justice policies indirectly by creating or undermining public support for a policy | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials) that public opinion affected by misinformation influenced justice policy | I have not identified any well-researched studies of this effect. **No such consequences were observed in the first entries studied.** |
| **MISINFORMATION EFFECTS ON BUSINESS OPERATIONS, MACRO-ECONOMIC ACTIVITY** | | |
| **Distorted business operations** | | |
| Distorted the operations of a business or business sector | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. credible business analysts) that misinformation distorted business sector operations (not reputation) | Studies (Atkinson, 2019. Katwala, 2019 & more) show effects of misinformation on the operations of the businesses targeted. Other studies show effects of disinformation spread by businesses (Hoggan et al, 2009 – re-tobacco firms. Supran & Oreskes 2021 re-energy firms).  **In this study**, a **false claim** made following xenophobic violence in South Africa that a Nigerian chain of shops was South African-owned **caused the shops to be attacked and damaged by protestors**. See: <https://www.icirnigeria.org/fact-check-are-shoprite-supermarkets-in-nigeria-franchise-outlets/>. See the discussion in the study. |
| **Distorted macro-economic activity** | | |
| Distorted macro-economic activity by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. credible economic analysts, officials) that misinformation distorted macro-economic policy and this affected activity | I have not identified any well-researched studies of this effect.. **No such consequences were observed in the first entries studied.** See the discussion in the study. |
| Distorted macro-economic activity by misleading the public about a reality | * Compelling evidence from academic or other study * Direct evidence from credible economic analysts that misinformation distorted public understanding of economy and this affected macro-economic activity | I have not identified any well-researched studies of this effect. **No such consequences were observed in the first entries studied.** See the discussion in the study. |
| **MISINFORMATION EFFECTS ON THE ENVIRONMENT** | | |
| **Distorted public environment policy** | | |
| Distorted public policy on climate change directly by misleading policymakers about a reality. | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials, or evidence by climate analysts) that misinformation distorted policy directly | Numerous studies have explored the effects on public policy of misinformation about climate change (Tesler, 2018. Supran & Oreskes 2021. Webb, 2021. WEF, 2020). **No such consequences were observed in the first entries studied.** See discussion in study. |
| Distorted public policy on climate change indirectly by misleading public opinion about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials, or evidence by climate analysts) that misinformation distorted policy indirectly | Numerous studies have explored the effects on public policy of misinformation about climate change (Tesler, 2018. Supran & Oreskes 2021. Webb, 2021. WEF, 2020). **No such consequences were observed in the first entries studied.** See the discussion in the study. |
| Distorted public policy on wildlife, biodiversity directly by misleading policymakers about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials, or evidence by environmental analysts) that misinformation distorted policy directly | I have not identified any well-researched studies of this effect. **No such consequences were observed in the first entries studied.** See the discussion in the study. |
| Distorted public policy on wildlife, biodiversity indirectly by misleading the public about a reality | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. admission by officials, or evidence by environmental analysts) that misinformation distorted policy indirectly | I have not identified any well-researched studies of this effect. **No such consequences were observed in the first entries studied.** See the discussion in the study. |
| **MISINFORMTION EFFECTS – FINANCIAL & ID LOSS, DIGITAL HARMS** | | |
| Caused financial loss, identity theft &/or digital harms to individuals, organisations | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. news reports, the victims) misinformation caused these harms | A limited number of studies (Cable et al, 2020, and others) have examined the effects of misinformation causing financial loss, digital harms to individuals. Among the entries In this study, I argue several online hoaxes **caused financial loss for those who responded** (See: Entry 97) and examples of clickbait **caused digital harm** (See: Entry 176). |
| Distorted an individual’s professional earnings, career | * Compelling evidence from academic or other study * Direct evidence from credible sources (e.g. news reports, victims) these harms caused | I have not identified any studies of this effect. Among the first entries studied, I argue an example of this effect was the anti-FGM campaigner forced by **false claims** that she underwent FGM **to give up that work**. See: (Entry 26). |

1. Whether observed consequences that occurred were direct, conditional, cumulative

To assess the effectiveness of counter misinformation approaches, it is necessary to understand whether the approaches respond to the speed and context in which such changes occur. Where I have found false claims to have observed consequences, I have identified those changes as (i) **direct**, (ii) **conditional**, (iii) **cumulative** or (iv) a combination of the above. I have done so, based on the following criteria.

|  |  |  |
| --- | --- | --- |
| **Type of change** | **This happens when misinformation is** | **Examples of consequences caused by these changes** |
| **Direct changes**  i.e.   * Immediate * dependent on neither particular conditions nor cumulation to cause the consequence | * **primarily affective** information – **affects the mood of the subject themselves directly on its publication** | e.g. the **immediate distress caused** to an anti-FGM campaigner by a false claim made about her behaviour, negatively affecting her reputation. See: (Entry 26) |
| * **primarily affective** or **behavioural** information –**may affect anyone**, is **not dependent** on particular conditions or cumulative effects for the consequence to occur | e.g. **an emotional reaction to**, or **behaviour in response to,** a false claim about (i) **identity** (national, ethnic, religious, gender etc. issue that might trigger particular behaviour); (ii) **health** (e.g. effects of medication that could cause health behaviours); (iii) **events or circumstances inducing stress for public**; (e.g. faked news of a calamity); **opportunity or need; clickbait** (e.g. hoax appeals for help; hoax job alerts; clickbait stories).  For example, hoax appeals for help for people with medical problems. Medical problems are plausible at any time. No particular conditions were required for the hoax to work. See: (Entry 97) |
| **Conditional changes**  i.e. occurs only   * to particular people &/or * in particular circumstances | * **affective, behavioural** or **cognitive** information – consequence only occurs to **particular people** | For example, **the panic** caused by a false rumour that foreign nationals were kidnapping children from schools in South Africa – was **only felt by parents of children at the schools concerned.** See: (Entry 54) |
| * **affective, behavioural** or **cognitive** information –depends for effect on occurrence of a **political event or circumstance** such as an election or a political protest | For example, **emotional reactions of** (i) **pride or outrage about** the supposed actions, character of a politician affecting voters choices; (ii) **motivation to vote** in an election due to false fear or hope, affecting turnout; (iii) **fear at the prospect of violence** at an election or a political or social protest, affecting turnout (iv) **behavioural responses to misdirection about** the rules or process of voting, affecting voting, affecting the success of voting. |
| * **affective, behavioural** or **cognitive** information – consequences depend on occurrence of **particular circumstances (not just politics)** | e.g. Consequences occur as a response to false claims about a **health crisis** or **natural disaster**.  For example, **panic caused** by false claims that an Ebola outbreak had arrived in Kinshasa **depended for effect on the fact that** (i) **Ebola was prevalent** in the country and (ii) earlier **experience of Ebola created fear**. See: (Entry 131) |
| * **affective, behavioural or cognitive** information – consequences **depend on audience’s adherence to a particular belief** or beliefs to occur | For example, the abuse of two women in Togo accused of being ‘bird witches’ depended on the audience’s pre-existing belief in the supernatural. See: (Entry 132) |
| **Cumulative changes**  i.e. occurs   * through changes to understanding attitude or behaviour over time * to reinforce or change existing views or shape new views | * **Primarily affective or cognitive** information –**repeated over time &/or part of a broad false narrative** – related to a new topic on which audiences may form a view over time | While many of the claims among the first entries studied related to such topics and were often repeated, it was beyond the scope of this study to prove such long-term changes.   See the discussion of the potential for such changes in the section on potential changes. |
| * **Primarily affective or cognitive** information –**repeated over time &/or part of a broad false narrative** – related to a topic on which individuals have a pre-existing view – which may change subject to repetition | While many of the claims among the first entries studied related to such topics and were often repeated, it was beyond the scope of this study to prove such long-term changes.   See the discussion of the potential for such effects in the section on potential changes. |

1. Scale of observed consequences

To properly assess the consequences for society of different types of misinformation, and thus understand the proportionality of any responses, it is also necessary to consider the scale, severity and duration they have or may have: whether (i) they affect individuals, groups or the whole of society, (ii) to mild, moderate or severe effect, and (iii) over the short- or long-term. Where I found consequences to be observed, I have categorised the scale of the consequences, i.e. the number of those affected, as:

1. a specific individual or individuals
2. non-specific individuals (i.e. not specific to a particular person, type or group)
3. a specific group or set of groups, a business or organisation
4. society as a whole

In doing this, it is important to note that save where the finding indicates a single individual, this category does not necessarily indicate **the number** of those affected. Misinformation that affects the behaviour of individuals across society may affect more people than if it were to affect the behaviour of a specific group. **The difference between the categories relates to the specific focus of the effect.** To assess this I have used the following criteria.

|  |  |  |
| --- | --- | --- |
| **Scale at which they occur** | **This happens when misinformation** | **Examples of scale of consequences** |
| **Consequences affect a specific individual or individuals** | * relates to the health, activities, views or character of a specific individual or individuals in a way that harms their reputation, mental wellbeing[[94]](#footnote-94) | e.g. the consequences of a false claim that two specific women acted as ‘bird-witches’ affected, first and foremost, those two specific women. See: (Entry 132). |
| **Consequences affect individuals in general – not a specific set** | * relates to (i) an issue, incident or event that concerns, or (ii) the activities or views of individuals but not one specific identifiable group | e.g. the consequences of panic or misdirection during a natural disaster (fire, flood) would be felt by individuals - (not a specific group). |
| * relates to a need or desire experienced by individuals – not specific to a particular group | e.g. the consequences of a hoax that a company was giving out 500,000 iPhones were felt by many individuals - not specific to any particular group. See: Entry (33) |
| **Consequences affect a specific group or set of groups, businesses or organisations** | * relates to (i) an issue, incident or event that concerns, or (ii) the activities or views of a specific demographic or social group | e.g. the consequences of a false claim of negative behaviour by a particular demographic group would be felt by that group. I have not identified an example in the first entries studied |
| * relates to the activities or views a specific business or organisation | e.g. the consequences of activists setting up a fake version of an Eritrean media’s website were felt by that media. See: (Entry 108) |
| * relates to a specific social, socio-economic or political issue circumstance or threat that would affect a specific identifiable group. | e.g. the consequences of a false claim that loan defaulters in Kenya faced imminent arrest were felt specifically by loan defaulters in Kenya. See: (Entry 8) |
| * relates to a particular medical disease or condition that would affect only people with or vulnerable to that particular disease or condition | e.g. the consequences of a false claim of a new, cheap cure for HIV/AIDS were felt by people living with the condition and those close to them. See: (Entry 123) |
| **Consequences affect society as a whole** | * relates to an issue that could shape the operation of a public policy affecting society as a whole | e.g. the consequences of a broad economic or social policy based on a false claim. I have not identified this sort of observed consequences. |
| * relates to an issue which could shape the outcome of an election or elections | I have not identified an example of this sort of observed consequences in the first entries studied. |
| * relates to an issue which concerns society as a whole e.g. gender abuse, climate change etc. | I have not identified an example of this sort of observed consequences in the first entries studied. See also discussion in the study. |

1. Severity of observed consequences

Other things being equal, the health consequences for someone who takes the wrong medication for a mild disease will be less severe than if the medication is for a serious condition. To properly assess the consequences for society of different types of misinformation, I have identified the severity of consequences caused by exposure to different types of misinformation – from potentially beneficial to severe. I have made these assessments based on the following criteria.

|  |  |  |
| --- | --- | --- |
| **Severity of consequences** | **Criteria** | **Examples** |
| **Observed beneficial consequences**  i.e. brings benefit for individuals, groups or society overall – without causing substantive negative effects for others | * The false claim causes or contributes to behaviour that has **lasting direct or indirect positive consequences for the health or well-being of the individual or society overall**.\* * The beneficial consequence is **sufficient to outweigh any negative results** for the individual, individuals, or society, including the meta-misinformation effect on trust.   **\*i.e.** **it promotes behaviour with positive results or discourages behaviour with negative results** for the individual or society overall. | In the study, I explore the argument that false information can be beneficial to society overall – e.g. used in a conflict to ‘win’ and thus bring the conflict to an end.  Using these criteria, I have **not observed proven beneficial consequences among the entries studied so far**.  See the discussion of this issue in the study. |
| **Mild negative consequences** | * The false claim causes or contributes to a **substantive but temporary negative emotional effect**[[95]](#footnote-95) (e.g. fear or distress). | e.g. the stress caused to people living with HIV/AIDS when told their government would end free ARV treatment was substantive but, as it was quickly disproven, was thus ‘mild’. See: (Entry 139) |
| * The false claim causes or contributes to a **modest financial or material loss** with no long-term financial or long-term effects for the individual or organisation | e.g. a loss suffered by individuals who paid a small fee to apply for a fake job are generally limited. See: (Entry 98)  \* Note: I assess the severity of financial loss not in absolute terms but relative to the individuals’ financial status |
| **Moderate negative consequences** | * The false claim causes or contributes to a **substantive, chronic (but not permanent) negative emotional effect** – e.g. stress, fear.   Note: I identify stress or fear that lasts for more than four weeks but is not permanent as **chronic**. See:  https://pmc.ncbi.nlm.nih.gov/ articles/PMC2746940/  Studies show chronic stress or fear can have substantive negative consequences for mental and physical health. | e.g. the mental distress caused by a false claim about an anti-FGM activist was chronic but not permanent. I thus identified it as a moderate effect. See: (Entry 26). |
| * The false claim causes or contributes to physical violence or assault, **the physical and psychological effects of which are temporary** – not permanent. | e.g. a man mistakenly identified as ex-Ivorian president Laurent Gbagbo suffered a physical assault but not long-lasting injuries. See: (Entry 90) |
| * The false claim causes or contributes to **serious but recoverable financial loss or destruction of property for individual or organisation** | e.g. a church was attacked leading to “wanton destruction of properties and injuries to many” after rumours a missing child had been found buried there. The effects were short-lived. See: (Entry 121) |
| * The false claim causes or contributes to **serious harm to an individual’s career but the harm is reparable** | e.g. false claims about an anti-FGM activist in Kenya affected her ability to work in this field. I identified this as a moderate effect. See: (Entry 26). |
| * The false claim promotes the use of an **ineffective medication for a mild medical condition – and the harm caused is not permanent** | e.g. a false claim that pineapple leaves cure nosebleeds, diabetes and low immune system. I identified the potential reliable on an ineffective treatment as a moderate effect – compared to some treatments. See: https://africacheck.org/fact-checks/meta-programme-fact-checks/pineapple-leaves-not-wonder-cure |
| **Severe negative consequences** | * The false claim causes or contributes to a **substantive, long-term or permanent negative emotional effect** – e.g. distress, fear &/or other negative psychological effect – **lasting years not weeks or months** | Misinformation, repeated over a long period, has the potential to cause long-term or permanent negative emotional effects. I have not identified any such consequences among the entries studied so far. See discussion in the study. |
| * The false claim causes or contributes to **physical violence or assault, the physical effects of which are long-term or permanent** | e.g. two health researchers were killed and one permanently maimed in Ethiopia after a false rumour spread they were harming children. See: (Entry 241) |
| * The false claim causes or contributes to a combination of verbal and physical abuse – **the combination causes physical and/or psychological effects of which are long-term** | e.g. two women in Togo suffered effects likely to be long-term when they were verbally and physically assaulted as ‘bird witches’. See: (Entry 132) |
| * The false claim causes or contributes to **serious financial loss for individual or organisation or the destruction of property and the loss is not recoverable** | Misinformation used for financial hoaxes has the potential to cause long-term financial harm to individuals and organisations. I have not identified any such examples among entries studied so far. |
| * The false claim causes or contributes to **substantive permanent harm to an individual’s career** | Misinformation has the capacity to end a career permanently. I have not identified any such examples among the entries studied so far. |
| * The false claim distorts health behaviour with **effects on health that are severe** (i.e. life- changing or life-limiting) and/or **long-term or permanent** | Misinformation has the capacity to cause permanent, severe effects to health. I have not identified any such examples among the entries studied so far. |
| * The false claim **distorts public policy, directly through influencing policymakers, or indirectly through distorting the outcome of an election**. [[96]](#footnote-96) | e.g. the governments in Congo & DR Congo allowed sale of an ineffective medication as an AIDS vaccine. See: (Entry 123) |

1. Duration of observed consequences

In order to make the above assessments, I also identify the consequences in the database as being either (i) short-term or reversible or (ii) long-term or irreversible. I identify duration as being a function of:

1. **the duration for which the audience will perceive the claim to be credible**

This is itself a function of (i) the topic, (ii) the initially perceived credibility of the false information and (iii) the perceived credibility of corrective information

1. **the time-limited nature or reversibility of the consequence**

This is a function of (i) the nature of the consequence (ii) the context. Panic will ease when a predicted calamity fails to occur, however, a person killed in vigilante violence caused by panic cannot be brought back to life.

Once the database is complete, I will ask colleagues who review the database to assess whether these findings are reasonable and useful for the study.

|  |  |  |
| --- | --- | --- |
| **Duration of consequence** | **Criteria** | **Examples** |
| **Short-term &/or reversible** | * A predicted or reported event is proven to have neither occurred nor be about to occur and no irreversible action has been taken | e.g. the panic caused when parents believed foreign nationals were kidnapping children outside schools in South Africa was short-lived. See: (Entry 54) |
| * Measures can be taken to counter-effect or reverse the consequences caused | I have not identified any such examples among the entries studied so far. |
| * Credible corrective information is presented, and repeated where necessary, correcting the false understanding caused. | Many studies (See [Porter & Wood, 2021](doi:10.1073/pnas.2104235118), and others)  show that credible corrective information will correct false understanding, particularly if repeated. |
| **Long-term &/or irreversible** | * The false claim is naturally difficult to disprove, for example requiring access to and trust in official statistics to show it is false. | See examples of false claims made about the Biafran separatist cause. The sources of corrective information are not trusted. |
| * The consequences caused are themselves inherently long-term or irreversible | e.g. a government policy with long-term implications (exacerbates the process of climate change; a long-term political realignment) |

1. Whether claim had substantive potential to cause (further) specific consequences – & what field

Based on close examination of evidence in the database, and research of the effects that different types of information may have in given circumstances, I argue in the study it is possible to identify false claims as having a substantive potential (i.e. a real or material chance) to cause or contribute to specific substantive consequences for individuals or society. While it may never be possible to accurately predict the probability that particular types of misinformation and circumstance will cause or contribute to particular outcomes, I argue in the study that it is possible to identify the main factors that affect both the probability or potential for and nature of any consequences that may or do occur.

Based this examination, I assess the potential of the entries in the database to cause or have caused or contributed to substantive consequences, in terms of:

Whether the claim is substantively false, and its capacity to change understanding, attitude or behaviour based on information that is false;

whether sufficient people consider it to be true to have the particular effect;

1. the audience’s capacity and motivation to act on the information, and its power to affect others if it does so

In terms of the audience, I distinguish between the potential for misinformation to cause consequences when perceived as credible by (i) members of the public and (ii) people in positions of authority who thus have the capacity to act in ways that affect society.

Note that I refer in the study to ‘**substantive**’ (i.e. real or material) potential to cause or contribute to consequences, neither suggesting such consequences will necessarily follow nor that there is no possibility that some sort of consequence. Quite how plausible is the potential, is of course complex, as I seek to indicate below and in more depth in the study.   
I use the term ‘**specific’** to make clear I am referring to consequences particular to the claim itself (or in combination with earlier editions and/or variants of the claim) and not to any form of secondary effects – which would be beyond the scope of this study. I use the term ‘**substantive**’ to indicate material, meaningful consequences.

I explain in the table detailed below how I take these factors into account and distinguish between those claims that the model suggests had no plausible potential to cause or contribute to consequences then or later in 10 broad fields of effects identified earlier, and those that did. The table is, of course, not exhaustive. It reflects the consequences identified from the sample in this study, and in other, earlier research. Where I did not identify any such potential consequences after reviewing all entries I will note as such.

1. **Potential consequences for democracy governance** 
   1. Exposure of **public** to **political misinformation during election campaign** has  
       potential to **distort voting preferences**
   2. **Repeated exposure of public** to **political, social misinformation** has potential to **affect political, social attitudes**
   3. Exposure of **public** to **political misinformation during election campaign** has  
       potential to **affect voter motivation & thus election participation**
   4. Exposure of **public** to **misinfo re-risks of participation in election** has potential to  
       **affect voter motivation & thus election participation**
   5. Exposure of **public** to **misinfo re-election rules/procedures** has potential to **distort  
       participation in election**
   6. Exposure of **public** to **misinfo re-likely election outcome, fairness** has potential to  
       **affect election participation, acceptance**
   7. Exposure of **public** to **misinfo re-risks of participation in protests** has potential to  
       **affect protestor’s motivation & thus participation in protests**
   8. Exposure of **policymakers** to **misinfo re-a general policy issue** has potential to **affect  
       policy on the issue directly** by misleading policymakers on a reality
   9. Exposure of **public** to **misinfo re-a general policy issue** has potential to **distort that  
       policy indirectly by creating or undermining public pressure** on an issue
   10. Exposure of **public** to **misinfo re-a particular politician or politicians** has potential to  
        **affect trust in that politician or those politicians**
   11. **Repeated exposure** of **public** to **misinfo re-fairness &/or competence of  
        policymakers** has potential to **affect trust, engagement with policymakers**
   12. **Exposure** of **policymakers to misinfo re-rules, laws & standards** has potential to  
        **distort application of those rules, laws & standards**
   13. **Exposure** of **public to misinfo re-rules, laws & standards** has potential to  
        **distort adherence to those rules, laws & standards**
   14. **Repeated exposure** of **public** to **misinfo re-media, public information** has potential  
        to **affect public trust in, use of info**
2. **Potential consequences for international relations, cross-border conflicts** 
   1. Exposure of **policymakers** to **misinfo re-another country** has potential to **distort  
       inter-governmental relations directly**
   2. Exposure of **public** to **misinfo re-another country** has potential to **distort inter-  
       governmental relations via public pressure**
   3. Exposure of **business, public**, to **misinfo re-country** has potential to **affect relations  
       at business, individual level**
   4. Exposure of **political, military leaders** to misinfo re-a country's actions, capabilities  
       has **potential to cause, distort cross-border conflicts by misleading them re reality**
   5. Exposure of **public** to misinfo re-a country's actions, capabilities has **potential to  
       cause or distort cross-border conflicts by distorting public support** for action
3. **Potential consequences for individual & public health** 
   1. Exposure of **public** to **health misinfo** has potential to **distort health behaviour in a  
       way that causes direct harm to health** of self or others
   2. Exposure of **public** to **health misinfo** has potential to **distort health behaviour in a  
       way that causes indirect harm to health** of self or others
   3. Exposure of **public** to **misinfo re-natural disasters, physical risks** has potential to  
       **cause physical risks to individuals, others**
   4. Exposure of **public** to **misinfo re-sexual health** has potential to cause, contribute to  
       **sexual behaviour with negative effects**
   5. Exposure of **politicians, health policymakers** to **health misinfo** has potential to  
       **distort health policies directly** by misleading them about a reality
   6. Exposure of **public** to **health misinfo** has potential to **distort health policies  
       indirectly by creating or undermining public support for policy**
   7. Repeated exposure of **public** to misinformation **re-actions of health workers** has  
       potential to **affect trust, engagement with health services**
4. **Potential consequences for individuals’ mental health** 
   1. Exposure of **public** to misinfo re-an individual or individuals' character has potential  
       to **cause distress to object/s of misinfo**
   2. Exposure of **public** to misinfo re-a threat or risk has potential to cause, contribute to  
       **public fear, panic or stress**
   3. Exposure of **mental health policymakers** to misinfo **re-mental health** has potential  
       to **distort mental health policies directly** by misleading them about a reality
   4. Exposure of **public** to **misinfo re-mental health** has potential to **distort mental  
       health policies indirectly** by affecting public support for policy
5. **Potential consequences for discrimination & abuse**
   1. **Repeated exposure** of **public** to misinfo re-characteristics of demographic group has  
       potential to **contribute to stereotypes & thus to abuse, discrimination**
   2. Exposure of **public** to misinformation re-character of individual or a demographic  
       group has **potential to be used to enact abuse or discrimination**
   3. Exposure of **policymakers** to **misinfo re-characteristics of a demographic group** has  
       potential to **directly affect policies in way that discriminates against them**
   4. Exposure of **public** to **misinfo re-characteristics of demographic group** has potential  
       to **affect policies in way that discriminates** by affecting public support for policy
6. **Potential consequences for vigilantism, civil unrest, insurgencies** 
   1. Exposure of **public** to misinfo re-individual/s' character or actions has **potential to  
       cause/contribute to vigilantism**
   2. Exposure of **public** to misinfo re-character of political, religious, ethnic groups has  
       **potential to cause, contribute to violence between groups**
   3. Exposure of **policymakers, insurgency leaders** to misinfo **re-actions, capabilities of  
       other side** has potential to **cause, distort course of insurgency by affecting decisions**
   4. Exposure of **public** to **misinfo re-actions, capabilities of parties to an insurgency** has  
       potential to **cause, distort course of insurgency by affecting public support**
7. **Potential consequences for the justice system** 
   1. Exposure of **policymakers in justice system** to **misinfo re-particular court case** has  
       potential **to distort the case**
   2. **Repeated exposure** of **public** to **misinfo re-police, judiciary or prison system** has  
       potential to **affect trust in, engagement with these systems**
   3. Exposure of **policymakers** to **misinfo re-justice system** has potential to **distort  
       justice policies directly** by misleading about a reality
   4. Exposure of public to misinfo re-justice system has potential to distort justice policies  
       indirectly by creating or undermining public support for a policy
8. **Potential consequences for business, economic activity**
   1. Exposure of **public** to **misinfo re-actions, effects of a business or business sector** has  
       potential to **distort operations of that business or sector**
   2. Exposure of **political, economic policymakers** to **misinfo re-economy** has potential  
       to **distort economic policy** by misleading them about a reality
   3. Exposure of **public** to misinfo re-economy has potential to distort macro-economic  
       activity by misleading them about a reality
9. **Potential consequences for the environment**
   1. Exposure of **policymakers** to **misinfo on climate change** has potential to **distort   
       climate policies by misleading policymakers** about a reality
   2. Exposure of **public** to **misinfo on climate change** has potential to **distort  
       climate policies** by creating or undermining public pressure on issue
   3. Exposure of **policymakers** to **misinfo on wildlife, biodiversity** by **misleading  
       policymakers** on a reality
   4. Exposure of **public** to **misinfo on wildlife, biodiversity** by **creating or   
       undermining public pressure** on issue
10. **Potential consequences: financial loss, digital harms for individuals, organisations** 
    1. Exposure of **public** to **hoaxes, clickbait** has potential to cause **financial loss, ID loss  
        &/or digital harms**
    2. Exposure of **public** to **misinfo re-character, actions, rights of individual** has potential  
        to **affect their professional standing, earnings, career**

I set out below the criteria applied for these findings, in the 10 broad fields and 50 sub-fields – listing examples for each case. Note: the findings listed here as showing a “potential” effect do not necessarily indicate a substantive potential, but just that ‘some’ potential exists. For a fuller discussion see (i) database findings (ii) discussion in the book itself.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Type of consequences** | | **Criteria for finding of potential consequences** | **Examples – potential effects** | |
| 1. **MISINFORMATION – POTENTIAL EFFECTS ON DEMOCRACY, GOVERNANCE** | | | | |
| **Potential (i) during election campaigns and (ii) repeatedly - to distort voter preferences** | | | | |
| **1.1.** Whether exposure of the public to **political misinformation**[[97]](#footnote-97) **during a brief election campaign** has substantive potential to **distort** **voting** **preferences**\*  - sufficient to affect an election outcome   \*i.e. which party or candidate to support. | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible[[98]](#footnote-98) by a sufficient number of swing voters to affect the outcome[[99]](#footnote-99) * primarily affective, thus may have potential to shape attitude[[100]](#footnote-100) * completely or mostly false; or element of truth but misleading[[101]](#footnote-101) * on a topic that is both related to a candidate, party or political issue, and new to voters[[102]](#footnote-102)   **The context is**   * an election – providing an opportunity for such an effect to occur * the election is narrow thus a change in a small number of voters may have capacity to affect the outcome   **Those who perceive the claim to be credible**   * include swing or potential swing voters * have capacity and motivation to vote | The evidence from research of the effect of misinformation during election campaigns is mixed.  Studies such as Gunther et al in 2018[[103]](#footnote-103) found misinformation in the US 2016 presidential campaign may have had an effect on voter choices that contributed to the result. Others (Guess et al, 2020,[[104]](#footnote-104) etc) disputed Gunther’s findings and argue the effects in that election were limited or nil.  Based on a review of these and other studies, I conclude misinformation during a campaign *can* change voters preferences in certain circumstances but such effects are (i) rare and (ii) occur primarily when related to issues new to voters – where voters attitudes are not already fixed and elections are narrow.  Based on the criteria shown (left), I did not identify any examples of potential for this effect among the first entries studied. I will update if I do so.  See the discussion in the study. | |
| **1.2.** Whether **repeated** exposure of the public to **political misinformation** in the form of **long-term false narratives** has substantive potential to **shape or distort political, social attitudes** sufficient to affect outcome of an election or have wider effects | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived credible over time by a sufficient number of voters, individuals to affect an election outcome or have wider effects * primarily affective or cognitive, thus may have potential to shape attitude, understanding[[105]](#footnote-105) * completely or mostly false; or element of truth but misleading[[106]](#footnote-106) * on a new or long-standing topic related to political, social or cultural issues.[[107]](#footnote-107)   **The context is**   * vulnerable to influence e.g. a narrow election, or a period of social tension   **Those who perceive the claim credible**   * take an interest in political or social issues * have capacity to vote in the election, &/or to act on their views in other ways that affect society | Media’s long-term influence on political and social attitudes has been a matter of academic debate for decades – back to the 1920s.  While many view media influence as exaggerated, studies do indicate repeated, exposure over time to partisan information can help to shape attitudes.  For example, Foos & Bischof (2021) found a 30-year boycott of anti-EU newspaper The Sun in Liverpool had had a clear effect on the way the city voted in the Brexit referendum. This study show repetition as key to effects on political understanding, attitudes.  Among the first entries in this study, I identified several with this potential effect.  For example, I argue in the study that a **false claim** that exaggerated public support for Kenyan political candidates[[108]](#footnote-108) **had potential**, if repeated over time, **to boost the candidates’ appeal**. See: (Entry 5)  I also argue that a **false claim** about land ownership, had **potential to reinforce a long-standing false narrative** about racial disparities in South Africa. See: (Entry 16)  See the fuller discussion of effects in the study. | |
| **Potential to distort political participation of false information on (i) political issues, (ii) risks of participation (iii) rules & procedures (iv) expected election outcome & fairness** | | | | |
| **1.3.** Whether exposure of the public to **political misinformation**, over the short or long-term, has substantive potential to **distort** **voter motivation\* and thus election participation** – sufficient to affect the outcome of an election – i.e. persuading a voter not to change who they vote for - but persuading them to turn out and vote | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of potential but undecided voters, uncertain whether to vote, to affect an election outcome[[109]](#footnote-109) * affective, behavioural or cognitive – thus may have potential to shape attitude, behaviour, understanding [[110]](#footnote-110) * completely or mostly false; or element of truth but misleading * on a political topic where the issue is either new to most voters and could motivate an undecided voter[[111]](#footnote-111) or is long-standing but the claim repeated sufficiently to motivate unmotivated voters by force of repetition[[112]](#footnote-112)   **The context is**   * an election – providing an opportunity for such an effect to occur * the election is narrow thus a change in the rate of participation by even a small number of voters has capacity to affect the outcome   **Those who perceive the claim credible**   * include supporters in principle of a candidate or party who need motivation in order to vote * have the capacity to vote if motivated to do so by the misinformation (or would stay away if demotivated) | Political operators have long argued it is easier to affect an election by motivating or demotivating active or latent supporters to vote than by persuading voters to switch sides.  Political communication researcher Martin Moore in his 2018 study *Democracy Hacked* (pp 60-61) argues the goal of operations such as Cambridge Analytica was ‘not to change the *way* people voted but to motivate their supporters and suppress their opponents’. For example, the Vote Leave campaign in the UK used false claims on potential Turkish membership of EU to motivate potential supporters to vote – rather than try to persuade opponents to switch (Moore pp 127-128).  I have not identified a political claim with the potential to motivate or demotivate voters among the first entries in this study.  See the discussion of effects in the study. | |
| **1.4.** Whether **exposure** of the public, over the short or long-term,to **misinformation about the risks of participating in an election** has substantive potential to **distort voter motivation & thus participationin an election** sufficient to affect the election outcome | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of voters to affect an election outcome[[113]](#footnote-113) * primarily affective &/or behavioural[[114]](#footnote-114) * completely or mostly false; or element of truth but misleading[[115]](#footnote-115) * related to the risks for those who participate in an election; either a specific new claim or a repeated false narrative[[116]](#footnote-116)   **The context is**   * an election – providing an opportunity for such an effect to occur * there is a history of such violence happening * the election is narrow, thus an effect on participation may affect the outcome.   **Those who perceive the claim credible**   * are uncertain supporters of a candidate or party and need motivation in order to vote * have the capacity to vote if motivated to do so; or to stay away if demotivated | Studies from across Africa (See Bratton et al, 2013. Norris, 2014. Burchard, 2015), Asia and elsewhere show that the real or perceived risk of violence around polling stations deters voters from participating in the polls.    Among the first entries in this study, I identified several examples of this potential effect.  For example, I argue in the study that a **false claim** that supporters of one of the candidates in elections in Nigeria’s Kogi State in 2019 had been shown shooting in the air **had potential to suppress participation by supporters of his opponent**. See: (Entry 91).  See the discussion of effects in the study. | |
| **1.5.** Whether exposure of the public, over the short or long-term, to **misinformation about** **election rules and process** has substantive potential **to distort participation** – sufficient to affect an election’s outcome | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of voters to affect an election outcome[[117]](#footnote-117) * primarily behavioural or cognitive[[118]](#footnote-118) – affecting ability to vote successfully * completely or mostly false; or element of truth but misleading * related to voter registration, voting rights or processes – either a new or repeated claim on a practical issue in the running of the election   **The context is**   * an election – providing an opportunity for such an effect to occur * the election is narrow, thus an effect on participation may affect the outcome.   **Those who perceive the claim credible**   * are targeted in such a way that their exclusion affects the outcome. * have the capacity to vote if they know how to do so correctly. | I have not identified well-researched studies of the specific issue which anecdotal evidence from election monitors indicates may occur.  Among the first entries in this study, I identified several examples of this potential effect.  For example, I argue in the study that a **false claim** that violence had caused election policymakers to cancel voting in a particular local government area during state-elections in Nigeria **had potential to reduce voter participation** – by creating a false understanding that voting was cancelled in the district. See: (Entry 250)  See the discussion of effects in the study. | |
| **1.6.** Whether exposureof the public, over the short or long-term, to **misinformation about the likely outcome &/or fairness of an election** has substantive potential to **distort participation in the election** sufficient to affect an election’s outcome | | I argue the likelihood & nature of any consequences depend on whether:   * perceived to be credible by a sufficient number of voters to affect an election outcome * affective, behavioural or cognitive[[119]](#footnote-119) * completely or mostly false; or element of truth but misleading * related to the likely outcome or fairness of an election   **The context is**   * an election – providing an opportunity for such an effect to occur * there is a history of elections being foregone due to rigging or the particular political situation * the election is narrow, thus an effect on participation may affect the outcome.   **Those who perceive the claim credible**   * include supporters of a candidate or party who need motivation in order to vote * have the capacity to vote if motivated to do so | Studies from around the world have long shown that the perception that an election is a foregone conclusion - for any reason - reduces participation.  For example, during the US presidential election of 1980, reporting on the day that the election was a landslide (for Ronald Reagan) was found to have reduced turnout in those states where voting was still taking place (Jackson, 1983).  In Africa, more recent research found that suggestions that an election outcome was rigged, and the results thus foregone, also reduced voter participation (Shenga & Pereira, 2019).  Among the first entries in this study I identified a number of examples where false or unproven claims of rigging had potential to have this effect.  For example, I argue in the study that **false claims** of a plot to rig elections in Nigeria’s Kogi State had **potential to reduce voter participation**. See: (Entry 91). See discussion of this example and effect in the study. | |
| **1.7.** Whether exposure of the public, over the short or long-term, to **misinformation about the risks of involvement in political or social protests** has substantive potential to **distort participation in protests** – with wider effects. | | I argue the likelihood & nature of consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to distort participation in protests[[120]](#footnote-120) * primarily affective &/or behavioural[[121]](#footnote-121) * completely or mostly false; or element of truth but misleading * related to potential risks of participating in protests; either a new topic or a repeated claim   **The context is**   * protests occur or are planned – providing an opportunity for such an effect to occur * a past history of risks for those participating in protests   **Those who perceive the claim credible**   * support a protest but may be demotivated by fear of risks * have the capacity to protest if they do not fear risks | I have not identified well-researched studies of the effect which anecdotal evidence suggests exists.  Among the first entries in this study, I identified a number of examples of false claims that may have had this effect.  For e.g. I argue that reports in Zimbabwean media and online that **exaggerated the threat** of violence for those attending anti-government protests **had potential to affect participation** See: https://www.bbc.co.uk/ programmes/p07lsh11  See a discussion of this example and this effect in the study. | |
| **Potential to distort public policy by (i) direct, (ii) indirect influence on policymakers** | | | | |
| **1.8.** Whether exposure of people in authority[[122]](#footnote-122), over the short or long-term to **political, social, economic, scientific or other misinformation**[[123]](#footnote-123) has substantive potential to **distort policies directly by misleading people in authority about a reality –** with wider effects for society | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people in authority to distort policies[[124]](#footnote-124) * affective, behavioural or cognitive[[125]](#footnote-125) * completely or mostly false; or element of truth but misleading * related to areas open to action or policy that affect wider society; either a new topic or a repeated claim   **The context is**   * a situation in which policy decisions have to be made [[126]](#footnote-126) * a culture of not challenging policy decisions or actions by those in authority   **Those who perceive the claim credible**   * have capacity to set policy or take action that affects wider society.[[127]](#footnote-127) | Many studies show the effect that misinformation can have on policy directly when false information is judged credible by policymakers and others in positions of authority.  One example is the influence of false information shaping HIV/AIDS policy in South Africa 2000-2006 (See: Chigwedere et al, 2008). There are many other examples in many countries and fields from economic to defence policy.  Among the first entries in the study, I identified one observed to have had this effect: a **false claim** about an HIV/AIDS treatment **which helped convince policymakers in two countries to allow its sale**. See: (Entry 123)  In addition, I identified several entries that had potential to have this effect. For example, I argue in the study that **a false claim** that pregnant women in South Africa drink to harm their unborn babies and become eligible for benefits had **potential to shape policy on benefits**. See: (Entry 171)  See a discussion of this example and this issue in the study. | |
| **1.9.** Whether exposure of the public, over the short or long-term, to **political, social, economic, scientific or other misinformation** has substantive potential to  **distort public policy indirectly by creating or undermining political pressure on people in authority to alter policy** with wider effects for society. | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to alter policy * affective, behavioural or cognitive[[128]](#footnote-128) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim on which those in authority have capacity to set or alter policy; are susceptible to pressure   **The context is**   * those in authority are vulnerable to public pressure on the issue   **Those who perceive the claim credible**   * take an interest in political or social issues * have capacity to put pressure on those in authority to introduce or change a policy in a way that affects society | I have not identified well-researched studies of the effect which anecdotal evidence suggests exists.  Among the first entries in this study, I identified a number that had potential to cause this effect.  For example, I argue in the study a **false claim** about land ownership **had potential to reinforce false beliefs about racial disparities** in South Africa and thus **affect the policies of politicians seeking support from those who agreed**. See: (Entry 16)  See a discussion of this example and this effect in the study. | |
| **Potential to distort trust in, or engagement, with (i) politicians, (ii) government, policymakers (iii) media & information – with wider effects** | | | | |
| **1.10.** Whether the exposure of the public, over the short or long-term, to **misinformation about the actions or character of a particular politician or politicians** has substantive potential to **distort trust in a particular politician, politicians in general** and/or **engagement in politics** with wider effects for society. | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people for this lack of trust to affect society[[129]](#footnote-129) * primarily affective or cognitive[[130]](#footnote-130) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim about the actions or character of a particular politician or politicians in general   **The context is**   * vulnerable to loss of public trust e.g. during an election or when the public need to heed advice during a health crisis   **Those who perceive the claim credible**   * have a capacity to act on the basis of their trust or lack of trust | | | A study of UK public attitudes in 2019 found that the perception that politicians on all sides spread falsehoods distorted trust and reduced the intention of one in five voters to vote (Britain Thinks & Full Fact, 2020).  Among the first entries in this study, while I identified claims that had the potential to distort trust in politicians, I found these all lacked a context in which they would cause substantive wider effects.  See a discussion of these examples and this effect in the study. |
| **1.11.** Whether the **repeated** exposure of the public, over short or long-term,to **misinformation** **about the actions or character of those in gov’t, other policymakers** [[131]](#footnote-131) has substantive potential to **distort trust in the fairness &/or competence of government, civil policymakers** – with wider effects for society. | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people for this lack of trust to affect society * primarily affective or cognitive[[132]](#footnote-132) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim about the actions or character of those in government, others in authority   **The context is**   * vulnerable to loss of public trust e.g. during an election or when the public need to heed advice during a health crisis   **Those who perceive the claim credible**   * have a capacity to act on the basis of their trust or lack of trust | | | A study conducted in four countries in 2020 found information that distorted trust in information from government and civil policymakers reduced public adherence with health guidelines (Hameleers et al, 2020). This was consistent with other studies of effects of changes in trust in the fairness and competence on public behaviour.  Among the first entries in this study, I identified several claims that had the potential to have this sort of wider effects. For example, I argue in the study that **a false claim** that ‘government incompetence in controlling immigration’ led to a popular building being burned down, **had potential to cause vigilante violence** – those who believed the claim seeing government as incompetent to respond. See: (Entry 32)  See a discussion of this example and this effect in the study. |
| **1.12.** Whether the exposure of policymakers over the short or long-term,to **misinformation related to rules, laws & standards** has potential to **distort the application of those rules, laws & standards.** | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a individuals in positions of authority over application of the rules, laws, or standards * primarily behavioural or cognitive[[133]](#footnote-133) * completely or mostly false; or element of truth but misleading * related to the nature or effects of rules, laws or standards or penalties for non-compliance   **The context is**   * the laws, rules or standards apply or would be thought to apply to a substantial section of the population   **Those who perceive the claim credible**   * have a capacity and motivation to act on the basis of their false understanding, in applying these laws, rules or standards | | | I have not identified well-researched studies of this effect.  Among the first entries in this study, I identified a number that had potential to cause this effect. For example, I argue in the study that a **false claim** that the Nigerian military had ordered civilians to follow a strict code of behaviour as part of a new security operation – **had a potential to distort application of the law** – with potential for abuse of civilians found in contravention. See: (Entry 81).  See a discussion of this example and this effect in the study. |
| **1.13.** Whether the exposure of the public, over the short or long-term,to **misinformation related to rules, laws & standards** has potential to **distort the public’s adherence to rules, laws & standards** | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by the public said to be subject to the rules, laws, or standards * primarily behavioural or cognitive[[134]](#footnote-134) * completely or mostly false; or element of truth but misleading * related to the nature or effects of rules, laws or standards or penalties for non-compliance   **The context is**   * the laws, rules or standards apply or would be thought to apply to a substantial section of the population   **Those who perceive the claim credible**   * believe they are subject to these supposed laws, rules or standards | | | I have not identified well-researched studies of this effect. I will update this if I do so.  Among the first entries in this study, I identified a number that had potential to cause this effect. For example, I argue in the study that a **false claim** that the Nigerian military had ordered civilians to follow a strict code of behaviour as part of a new security operation – **had a potential to change their behaviour to comply with these supposed laws**. See: (Entry 81).  See a discussion of this example and this effect in the study. |
| **1.14.** Whether the **repeated** exposure of the public to **misinformation about accuracy, honesty and/or fairness of media or other public information** has substantive potentialto **distort** **trust in & use of media, information** – with wider effects for society. | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people for this lack of trust to affect society * **primarily affective or cognitive**[[135]](#footnote-135) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim about the accuracy, honesty and/or fairness of media or other public information   **The context is**   * vulnerable to loss of public trust e.g. during an election or when the public need to heed advice during a health crisis   **Those who perceive the claim credible**   * have a capacity to act on the basis of their trust or lack of trust in public information (e.g. in a public health crisis) | | | The study noted above, (Hameleers et al, 2020), found that a perception that misinformation was pervasive distorted trust in information and media resulting in reduced adherence with health guidelines in four countries.  Such a perception, true or false, may be exaggerated by false claims of bias in media and information where it does not exist.  Among the first entries in this study, I identified several with potential to cause this effect.  For example, I argue in the study that a **false claim** about media coverage of former SA President Zuma **had potential to confirm a perception of media bias** among those who judged it credible. See: (Entry 19). See a discussion of this example and this effect in the study. |
| **MISINFORMATION - POTENTIAL EFFECTS ON INT’L RELATIONS, CROSS-BORDER CONFLICTS** | | | | |
| **Potential to distort int’l relations between (i) governments, (ii) individuals, businesses** | | | | |
| **2.1.** Whether the exposure of the policymakers over the short or long-term to **misinformation about the actions or character of another country’s government or population** has substantive potential to **distort inter-governmental relations directly by misleading policymakers about a reality** – with wider effects for society. | | I argue the likelihood & nature of consequences depend on whether:   **The false claim is**   * perceived to be credible by a sufficient number in a position to affect policy and actions[[136]](#footnote-136) * affective, behavioural, or cognitive[[137]](#footnote-137) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated false claim about the actions or character of another country’s gov’t or population, or relations between the countries   **The context is**   * one government (or both) is able to act in ways that affect the other if it chooses   **Those in authority who perceive the claim credible**   * have a capacity to affect inter-gov’t relations | I have not identified well-researched studies of the effects of misinformation directly on policymakers’ understanding of other countries actions.  Among the first entries in this study, I identified several false claims with potential to affect or have affected inter-governmental relations.  For example, after reports from South Africa claimed, falsely, that Nigerians had killed in a wave of violence in South Africa[[138]](#footnote-138) false claims quickly emerged that other governments were reacting strongly to the events and led to calls for the Nigerian government to boycott a meeting in South Africa. See (Entry 57 & 63)  I argue in the study that these claimshad thus had the **potential to have affected inter-governmental relations both directly and indirectly**.  See a discussion of this example and this effect in the study. | |
| **2.2.** Whether the exposure of the public over the short or long-term to **misinformation about the actions or character of another country’s government or population** has substantive potential to **distort inter-governmental relations directly by creating or undermining public pressure for actions that would affect inter-governmental relations** | | I argue the likelihood & nature of consequences depend on whether:   **The false claim is**   * perceived to be credible by a sufficient number of people to pressure the policymakers to change policy and actions[[139]](#footnote-139) * affective, behavioural, or cognitive[[140]](#footnote-140) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated false claim about the actions or character of another country’s gov’t or population, or relations between the countries   **The context is**   * one government (or both) is able to act in ways that affect the other if it chooses * the public is able to affect international relations directly, e.g. via a direct vote on foreign policy (See Brexit) * the government is vulnerable to public pressure (e.g. due to election cycle) if public pressure is applied   **Those in authority who feel need to respond to public pressure**   * have a capacity to affect inter-gov’t relations | Evidence from studies of the UK vote on Brexit, indicate that both short-term exposure to false or misleading information online (See Marshall, H. Drieschova, A. 2020) and long-term exposure to partisan media (Foos & Bischof, 2021) can build public pressure for actions which then affect intergovernmental relations.  Among the first entries in this study, I identified several false claims with potential to affect or have affected inter-governmental relations.  For example, after reports from South Africa claimed, falsely, that Nigerians had killed in a wave of violence in South Africa[[141]](#footnote-141) **false claims** quickly emerged that other governments were reacting strongly to the events (See: Entry 57 and 63), and led to calls for the Nigerian government to boycott a meeting in South Africa. I argue in the study that these claimshad had the **potential to have affected inter-governmental relations both directly and indirectly**. | |
| **2.3.** Whether the exposure of (i) business leaders and (ii) the public, over the short or long-term, to **misinformation about the actions or character of another country’s government or population**, has substantive potential to **distort international relations at the level of business,** **individuals** – with wider effects for society | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of businesses &/or the public to have a substantive effect[[142]](#footnote-142) * primarily affective, or behavioural[[143]](#footnote-143) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated false claim about the actions or character of another country’s government or population, or relations between the countries   **The context is**   * there are sufficient relations at the individual level (e.g. travel, cultural) for change to have effect * there are sufficient relations at the business level for change to have effect   **Those who perceive the claim credible**   * have a capacity to change their behaviour in ways that affect relations at the business or individual level | I have not identified well-researched studies of the effect of misinformation on individual and business relations which anecdotal evidence suggests exists. For example, reports that exaggerated the risks of Ebola to travellers to Africa in 2014 appeared to cause a fall in business and personal travel to the continent – on a false basis.  Among the first entries in this study, I identified one that had observed effect: a **misleading claim** that a chain of shops in Nigeria was South African owned **led the shops to be attacked by protestors – thus affecting the businesses attacked**. See:  <https://www.icirnigeria.org/fact-check-are-shoprite-supermarkets-in-nigeria-franchise-outlets/>  I also identified several examples of such potential effects.  For example, I argue in the study that a **false claim** that a Rwandan company was able to listen in to phone conversations in DR Congo had **potential to affect its relationship with its customers**. See: (Entry 138) | |
| **Potential to cause or distort course of wars through (i) policymakers, (ii) public opinion** | | | | |
| **2.4.** Whether the exposure of political, military policymakers, over the short or long-term, to **misinformation about the actions, intentions or capabilities of another country** has substantive potential to **cause or distort course of a cross-border conflict** by **misleading** **policymakers about a reality** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient senior people in political or military policymakers to affect political/miliary policy[[144]](#footnote-144) * affective, behavioural or cognitive[[145]](#footnote-145) * completely or mostly false; or element of truth but misleading * related to a new topic or a repeated claim on the history, politics or resources of a country or the actions, intentions or capabilities of or domestic or foreign support for its government or military   **The context is**   * political, military pressure on the policymakers * perception government has capacity to act successfully   **Those who perceive the claim credible**   * have both the capacity and intent to act militarily against another country | Numerous histories of conflicts from ancient times to WWII (Levine, 2012), Vietnam (McNamara, 1995) and Bosnia (Malcolm, 1994) – have identified the role that false information can play, alongside other factors, in causing or distorting the conduct of wars by shaping or distorting the understanding of political and military leaders about actions and capabilities.  Among the first entries in this study, I identified one false claim with a potential to affect a conflict in this way.  I argue in the study, for example, that a **false claim** that UN troops had collaborated with Uganda-based rebels in a massacre in Eastern DR Congo caused or contributed to subsequent attacks on a UN base in which nine protestors were killed, built pressure for the UN to withdraw, & had **potential to alter the course of the conflict** See: (Entry 160).  See the discussion of effects in the study. | |
| **2.5.** Whether the exposure of the public,over the short or long-term, to **misinformation related to the actions, intentions or capabilities of another country** has substantive potential to  **cause or distort course of cross-border conflicts** by distorting **public support for or against one side or course of action.** | | I argue the likelihood & nature of any consequences depend on whether: **The false claim is**   * perceived to be credible by a sufficient number of people to cause the effect * affective, behavioural, or cognitive[[146]](#footnote-146) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim on the history, politics or resources of a country or the actions, intentions or capabilities of its government or military   **The context is**   * political, military pressure on the policymakers * a perception the government has capacity to act * popular support (e.g. participation, logistical support) is necessary for tactical or strategic reasons   **Those who perceive the claim credible**   * have the capacity to support or oppose action in a way that affects events | Numerous studies (Malcolm, 1994. Baum & Groeling, 2010 & others) have examined the role false information can play contributing to or shaping the conduct of conflicts by shaping public support for or against one side or course of action.  Among the first entries in this study, I identify the claim mentioned above about UN troops in DR Congo as having had potential to alter the course of a cross-border conflict by distorting support for one of the parties.  I did not identify further claims with a potential to affect the course of a conflict in this way – but will update if I do.  See the discussion of effects in the study. | |
| **MISINFORMATION - POTENTIAL EFFECTS ON PUBLIC HEALTH** | | | | |
| **Potential to distort public health (i) directly, (ii) indirectly or (iii) in disasters, risk behaviour** | | | | |
| **3.1.** Whether the exposure of the public, over the short or long-term, to **health-related misinformation**[[147]](#footnote-147) has substantive potential to **distort health behaviour in a way that** **causes direct harm to health of self or others** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to cause the particular health effect directly[[148]](#footnote-148) * affective, behavioural or cognitive[[149]](#footnote-149) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim on (i) the causes, effects, prevalence of medical conditions; (ii) the effects medical treatments.   **The context is**   * the condition concerned is prevalent * the causes and effects of the condition &/or treatment are not widely known, accepted   **Those who perceive the claim credible**   * have limited medical knowledge, &/or are sceptical about effective treatments * have the capacity to act in a way that causes harm to themselves or others | Many studies of the negative effect of misinformation on public health behaviour have been published (See Larson, 2018. Cunliffe-Jones 2020. Motta & Stecula, 2021, others).  In addition to the entries in this study I found to have caused direct harm to health, mentioned earlier, I have identified numerous claims as having had potential to distort health behaviour in a way that would cause direct harm to the health of individuals or society.  For example, posts online in Cameroon **promoted a diabetes treatment** that would (i) not protect against diabetes, (ii) included seeds that can be toxic thus **had potential to cause indirect & direct harm to health**. See: (Entry 153)  Similarly, medical experts argue excessive use of cassava, had **potential to fuel cancer cells** due to the sugar it contains. [See link](https://africacheck.org/fact-checks/spotchecks/no-cassava-doesnt-cure-cancer-and-could-make-it-worse) (Entry 310). See the discussion of effects in the study. | |
| **3.2.** Whether the exposure of the public, over the short or long-term, to **health-related misinformation** has potential to **distort health behaviour in a way that** **causes indirect**[[150]](#footnote-150) **harm to selves or others** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to cause the particular health effect directly * affective, behavioural, or cognitive[[151]](#footnote-151) * completely or mostly false; or element of truth but misleading * related to either a new topic or a repeated claim on (i) the effects of unproven or non-treatments when an effective treatment is available (ii) the behaviour of healthcare professionals who provide an effective treatment.   **The context is**   * an effective treatment for the condition exists * the extent and nature of the threat is not widely known   **Those who perceive the claim credible**   * have limited medical knowledge &/or are sceptical about effective treatments * have the capacity to act on the information in a way that causes a harmful effect to themselves or others | Studies have shown health misinformation to distort health behaviour in a way that causes harm by reducing take-up of effective treatment. [see Thielman et al 2014].  Among the first entries in this study, I identified one shown to have caused indirect effects.  After seeing a false claim that cassava cures prostate cancer, a number of those being treated for the condition may have stopped taking actual cancer medication – an indirect effect of the false claim. See: <https://africacheck.org/fact-checks/spotchecks/no-cassava-doesnt-cure-cancer-and-could-make-it-worse>.  I have identified many other examples. For example, **a false claim** that Islamic jihadists were posing as medics and injecting individuals with HIV had the **potential to cause indirect harm deterring** individuals from taking up a safe, effective treatment. See: (Entry 84). I also argue a **false claim** that manioc juice and a West African drink, *bissap,* were an HIV cure had **potential to stop people using effective treatment**. (See: Entry 145). See the discussion of effects in the study. | |
| **3.3.** Whether the exposure of the public, over the short-term, to **misinformation on risks & behaviour in natural disasters & at other times** has substantive potential to **distort behaviour in a way that increases other physical risks to self or others** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to cause the effect to individuals or groups * primarily affective or behavioural[[152]](#footnote-152) * completely or mostly false; or element of truth but misleading * related to (i) risks of, and behaviour to follow in, a new and current natural disaster or physical threat (e.g. fire, flood, security incident) (ii) the physical risks posed by particular behaviour   **The context is**   * a natural disaster or other threat is occurring * the less is widely known about the threat, the more likely the effect   **Those who perceive the claim credible**   * have limited knowledge of the circumstances * have the capacity to act on the information | Studies have shown that misinformation about the practical risks the public face during a natural disaster or security incident can cause or contribute to physical harm. (Jin et al, 2014. Torpan et al, 2021)  Studies also show risks can be increased by false information that promotes risks behaviour.  I argue in the study that a **video manipulated to appear to show a soldier catch a bullet in his mouth** had the potential to inspire others to try to copy his trick – **risking injury or death.** See: (Entry 36).  See the discussion of effects in the study. | |
| **Potential to cause or contribute to negative sexual or reproductive health behaviour** | | | | |
| **3.4.** Whether the exposure of the public, over the short or long-term, to **misinformation related to sexual or reproductive health** has substantive potential to cause or contribute to **sexual or reproductive health behaviour with negative effects for physical or mental health** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to cause the effect[[153]](#footnote-153) * affective, behavioural, or cognitive[[154]](#footnote-154) * completely or mostly false; element of truth but misleading * related to a new topic or repeated false claim on (i) the effects &/or causes of sexual ill-health; (ii) effects of different treatments; (iii) reproductive anatomy, science; (iv) effects of different reproductive health behaviour e.g. birth practices.   **The context is**   * the condition which the false information concerns is prevalent * the causes and effects of the condition &/or treatment are not widely known, accepted   **Those who perceive the claim credible**   * have limited knowledge of sexual and reproductive health * have the capacity to act on the information (e.g. work as traditional birth attendants) | Studies show the potential for misinformation to cause or contribute to sexual or reproductive health behaviour with negative effects for physical or mental health.  This ranges from the claim that sex with a virgin is a cure for HIV/AIDS, which circulated in some southern African countries in the early years of the AIDS pandemic (Meel, 2003) to false claims about the effects of practices in maternal health which circulated for decades (Amobi, 2020) .  Among the first entries in this study, I identified no example that showed the potential to cause this effect.  See the discussion of these effects in the study. | |
| **Potential to distort public health policy (i) directly (ii) indirectly** | | | | |
| **3.5.** Whether the exposure of people in authority, over the short or long-term, to **health misinformation** has substantive potential to **distort health policies directly by misleading policymakers about a reality** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people in positions of authority on health to cause a particular effect[[155]](#footnote-155) * affective, behavioural, or cognitive[[156]](#footnote-156) * completely or mostly false; or element of truth but misleading * related to new topics or repeated claims about health and medicine that are open to policy action affecting individuals or society   **The context is**   * a situation in which health policy decisions have to be made[[157]](#footnote-157) * a culture of not challenging policy decisions or actions by those in authority   **Those who perceive the claim credible**   * have limited knowledge of health matters * have capacity to set health policy or take action that affects the health of wider society on the basis of the information.[[158]](#footnote-158) | Many studies and reports have shown misinformation to have the potential to affect public health policy directly.  The influence of health misinformation in shaping HIV/AIDS policy in South Africa 2000-2006 (Chigwedere et al, 2008) and Madagascar’s response to the Covid pandemic are two examples (Scudellari, 2021).  In this study, I argue that a false claim that a medication was an effective AIDS vaccine, approved by a prestigious French institute, can be shown to have **altered public policy,** helping persuade policymakers in Congo & DR Congo to allow its sale. See: (Entry 123)  I identified no other example among the first entries that I reviewed as having shown the **potential** to cause this effect but I will update if I do.  See the fuller discussion of effects in the study. | |
| **3.6.** Whether the exposure of the public, over the short or long-term, to **health misinformation** has substantive potential to **distort health policies**  **indirectly by creating or undermining pressure** **on policymakers to make or change health policy.** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to pressure health policymakers on health policy[[159]](#footnote-159) * affective, behavioural or cognitive[[160]](#footnote-160) * completely or mostly false; or element of truth but misleading * related to a new topic or a repeated claim on which health policymakers have capacity to set or alter policy if put under pressure   **The context is**   * those in authority are vulnerable to public pressure (e.g. for electoral reasons)   **Those who perceive the claim credible**   * take an interest in health issues * have capacity to put pressure on health policymakers to introduce or change a policy in a way that affects society | Many studies have identified the effect of misinformation on public attitudes to health issues and how this shaped policy (See Ayodele, 2007 re-polio vaccines in Nigeria).  Among the first entries in this study, I identified no example that showed the potential to cause this effect.  See the fuller discussion of effects in the study. | |
| **3.7** Whether **repeated exposure** of the public, to **misinformation re-the actions of health workers** has potential to **affect trust, engagement with health services** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to have an effect on public trust in and engagement with health services * affective, behavioural or cognitive * completely or mostly false; or element of truth but misleading * the claim &/or similar claims are repeated over time   **The context is**   * low levels of trust in health workers, and/or other sectors in society * the low level of trust is misplaced   **Those who perceive the claim credible**   * require to engage with and trust in health workers due to public health circumstances | Many studies have found that trust in health workers is low – in some countries more than others. Studies of public responses to public health campaigns in situations such as the Ebola crisis in West Africa (2014) or the Covid pandemic show this mistrust has potential to negatively impact public health and can lead to attacks on health workers. | |
| **MISINFORMATION – POTENTIAL EFFECTS ON MENTAL HEALTH** | | | | |
| **Potential to cause or contribute to (I) stress, fear for individual (ii) public stress or panic** | | | | |
| **4.1.** Whether the exposure of the public, over the short or long-term, to **misinformation about the actions or character or individuals or groups** has substantive potential to **cause or contribute to chronic distress for the objects of misinformation** | | I argue the likelihood & nature of any consequences depend on whether: **The false claim is**   * perceived as being seen as credible by a sufficient number of people to distress the subject of the false information[[161]](#footnote-161) * affective[[162]](#footnote-162) * completely or mostly false; or element of truth but misleading * a new or repeated claim about the actions or character of an individual or individuals, in a way that would affect their reputation if believed   **The context is**   * the false claim published or transmitted to an audience the subject sees as important   **The subject of the misinformation**   * is vulnerable to perceived harm to their reputation * suffers from other pressures which exacerbate the stress | Studies show false accusations as having potential to cause psychological distress to the falsely accused. (Brooks & Greenberg, 2020, and more).  Studies identify distress that is experienced for one month or more as chronic (Shalev, 2010 & others) with potential to cause long-term harm to physical, mental health.  Among the first entries in this study, I identified several examples as causing or having potential to cause this effect.  For example, I found that a false claim that hurt the reputation of an anti-FGM activist harmed her mental health. See: (Entry 26). And I argue a **false claim** which caused two women in Togo to be assaulted thereby **had the potential to cause lasting fear and distress**. See: (Entry 132). See the fuller discussion of effects in the study. | |
| **4.2.** Whether the exposure of the public, over the short or long-term, to **misinformation about a supposed practical threat or risk** has substantive potential to **cause or contribute to fear, panic or distress** for a group, society | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a group or across society as a whole * primarily affective, or behavioural[[163]](#footnote-163) * completely or mostly false; or element of truth but misleading * related new topic or repeated claim about a perceived threat to the public or those they care for   **The context is**   * the threat is plausible, e.g. a claim that a forest fire has broken out in a region where such fires occur   **Those who perceive the claim credible**   * is vulnerable to stress, fear * suffers from other pressures which exacerbate the stress. | Reports and research cited in the study show misinformation has the potential to cause fear, stress and panic to the public.  Studies identify distress or fear experienced for >one month as chronic (Shalev, 2010) with potential to cause long-term harm to physical, mental health.  Among the first entries in this study, I identified several where public panic or stress occurred.  For example, a false claim that children were being kidnapped from schools caused stress, fear and panic among parents. See: (Entry 54)  I did not, however, identify examples where a potential effect might plausibly have occurred but did not. See the discussion of effects in the study. | |
| **Potential to distort public policy on mental health (i) directly, (ii) indirectly** | | | | |
| **4.3.** Whether the exposure of people in authority, over the short or long-term, to **misinformation** **related to mental health** has substantive potential to **distort public policy on mental health directly by misleading policymakers about a reality** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people in positions of authority on mental health to affect policy[[164]](#footnote-164) * primarily cognitive[[165]](#footnote-165) * completely or mostly false; or element of truth but misleading * related to new topics or repeated claims about areas of mental health open to policy action, e.g. support services or treatments   **The context is**   * policymakers recognise mental health as a public health issue * a culture of not challenging policy decisions or actions by those in authority   **Those who perceive the claim credible**   * have capacity to set mental health policy or take action on the basis of the information.[[166]](#footnote-166) | Many studies and reports have shown misinformation has the potential to distort general health policy – as set out above.  I have not identified studies of its effects on policy on mental health.  Among the first entries in this study, I identified one example that might have had this effect on mental health policy.  In the study I identify a false claim, made by a nominally credible source, about Nigeria’s suicide rate which I argue **had potential to affect public policy on mental health services**. See: (Entry 173)  See the discussion of effects in the study. | |
| **4.4.** Whether the exposure of the public, over the short or long-term, to **misinformation** **related to mental health** has substantive potential to **distort public policy on mental health indirectly by creating or undermining pressure** **on policymakers to make or change policy.** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to put pressure on those in authority[[167]](#footnote-167) * primarily affective or cognitive[[168]](#footnote-168) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about the nature, prevalence and/or treatment of mental health on which those in authority on health have capacity to set or alter policy   **The context is**   * those in authority are vulnerable to public pressure on the issue   **Those who perceive the claim credible**   * have an interest in mental health as an issue of public health * have capacity to put pressure on those in authority on mental health issues to introduce or change a policy | Many studies have identified the effect of misinformation on public attitudes to health issues & how this shaped policy. [See Ayodele, 2007, on misinformation which shaped public opinion in Nigeria leading politicians to ban polio vaccination there at that time].  I have not identified studies of this effect in the area of mental health specifically.  Among the first entries in this study, I identified no example that showed the potential to cause this effect. See the discussion in the study. | |
| **MISINFORMATION – POTENTIAL EFFECTS ON PREVALENCE OF DISCRIMINATION & ABUSE** | | | | |
| **Potential to cause or contribute to false understanding, stereotype of a demographic group – that may contribute to discrimination & abuse** | | | | |
| **5.1.** Whether the **repeated** exposure of the public, over the long-term, to **misinformation about the actions or character of a demographic group** has substantive potential to **cause or contribute to false understanding, stereotypes that may contribute to**  **abuse, discrimination, against individuals or groups** | | I argue the likelihood & nature of any consequences depend on whether: **The false claim is**   * perceived to be credible by a sufficient number of people, to have a wider social effect[[169]](#footnote-169) * primarily affective or cognitive[[170]](#footnote-170) * completely or mostly false; or element of truth but misleading * on a topic related to the actions or character of a particular demographic group   **The context is**   * the false claim is made and regularly repeated from multiple sources over time * the claim is circulated at a time of heightened attention to identity-related issues or tensions in the country   **Those who perceive the claim credible**   * have existing negative attitudes towards the group in question * have a capacity to act to abuse or discriminate against individuals or the group more widely | As set out above in relation to political and social misinformation, I argue in the study that prolonged exposure to certain types of misinformation can change or shape attitudes to particular demographic groups in ways that can have wider effects.  Among the first entries in this study, I identified numerous false claims with the potential, if repeated over time, to cause or contribute to false understandings or stereotypes that could cause or contribute to discrimination or abuse.  For example, I argue in the study that a **false claim** that foreign nationals were kidnapping children in South Africa caused false understanding that **had potential to contribute to abuse of foreign nationals**. See: (Entry 54). And I argue a **false claim** that two women in Togo deserved abuse as they were ‘bird witches’ **had the potential to cause or contribute to abuse** of other women there and in other countries. See: (Entry 132)  See the discussion of effects in the study. | |
| **5.2.** Whether the exposure of the public, over the short-term, to  **misinformation about the actions or character of a demographic group** has substantive potential to be used to enact **abuse or**  **discrimination of individuals or groups** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to enact abuse or discrimination[[171]](#footnote-171) * primarily affective or behavioural[[172]](#footnote-172) * completely or mostly false; or element of truth but misleading * related to a new or often repeated claim about the actions or character of a particular demographic group   **The context is**   * the false claim aligns with existing negative attitudes toward the group * a time of heightened attention to identity-related issues or tensions in the country more widely   **Those who perceive the claim credible**   * have existing negative attitudes towards the group in question * have a capacity to abuse or discriminate against individuals or groups | In addition to misinformation that shapes attitudes long-term, numerous reports have identified occasions in which false information was used to enact abuse or discrimination.  Among the first entries reviewed in this study, I identified numerous examples where this occurred or had potential to occur. In the study, I argue, for example, that **a false claim** that informal waste collectors were banned from working in a Johannesburg suburb **had potential to be used to abuse** those who contravened the supposed ban. See (Entry 3).  And I argue that a **false claim** that Islamic jihadists were posing as vaccination teams and injecting members of the public with HIV **had potential to cause members of the public to abuse vaccination teams**, and may have done so, though this is not proven. See: (Entry 84). See the discussion of effects in the study. | |
| **Potential to distort public policies in a way that affects discrimination** | | | | |
| **5.3.** Whether the **repeated** exposure of people in authority, to  **misinformation about the actions or character of a demographic group** has substantive potential to **directly** **distort public policy in a way that discriminates against the group concerned.** | | I argue the likelihood & nature of any consequences depend on whether:    **The false claim is**   * perceived to be credible by a sufficient number of people in positions of authority on policy areas in which discrimination can occur (e.g. social policy, justice policy, education etc.)[[173]](#footnote-173) * affective, behavioural or cognitive[[174]](#footnote-174) * completely or mostly false; or element of truth but misleading * related to a new topic, or a repeated claim about the actions or character of a particular demographic group   **The context is**   * heightened attention to identity- issues or tensions in the country more widely * a culture of not challenging policy decisions or actions by those in authority   **Those who perceive the claim credible**   * have capacity to set public policies that discriminate against a particular group | There are numerous examples of people in authority around the world who enacted discriminatory policies on the basis of a false understanding of the group, among other factors.  Among the first entries in this study, I identified several examples showing the potential to shape or contribute to discriminatory attitudes in this way.  For example, I argue in the study that a false claim by a South African minister that exaggerated the number of ‘fraudulent’ marriages of foreign nationals **had potential to lead to discriminatory policies** **against foreign nationals**. See: (Entry 177). And I argue a false claim that hundreds of knives confiscated in Nigeria had been intended for killing Christians had **potential to contribute to policymakers introducing discriminatory policies against Muslims**. See: (Entry 209).  See the discussion of effects in the study. | |
| **5.4.** Whether the **repeated** exposure of the public, to **misinformation about the actions or character of a demographic group** has substantive potential to **indirectly distort policies by creating or undermining pressure** **on policymakers so that policy discriminates against the group concerned.** | | I argue the likelihood & nature of consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to put pressure on policymakers to alter policy * affective, behavioural, or cognitive[[175]](#footnote-175) * completely or mostly false; or element of truth but misleading * relates to a new topic or repeated false claim about a particular demographic group’s actions or character   **The context is**   * a time of heightened attention to identity-related issues or tensions in the country more widely * those in authority are vulnerable to public pressure on the issue (e.g. due to electoral cycle)   **Those who perceive the claim credible**   * Have capacity and motivation to put pressure on those in authority on this issue | There are also numerous examples through history of people in authority who acted following pressure from the public based at least in part on a false understanding about a particular demographic group.  Among the first entries in this study, I identified several examples showing the potential for this occur.  For example, I argue in the study that the **false claim** made by a minister that exaggerated the number of ‘fraudulent’ marriages made by foreign nationals **also had potential to increase public pressure on the policymakers to introduce policies discriminating against foreign nationals** in the way the marriage system operates. See: (Entry 177)  I argue the **false claim** that hundreds of knives confiscated by police in Kano State, Nigeria, had been intended for use killing Christians had the potential to put public pressure on policymakers to act in a way that would discriminate against Muslims. See: (Entry 209)  See the discussion in the study. | |
| **MISINFORMATION - POTENTIAL EFFECTS ON VIGILANTISM, CIVIL UNREST, INSURGENCIES** | | | | |
| **Potential to cause or contribute to vigilantism** | | | | |
| **6.1.** Whether the exposure of the public, over the short or long-term, to **misinformation about the actions or character of an individual or individuals**  has substantive potential to **cause or contribute to** **vigilante violence against the individual or individuals** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to enact vigilante violence * primarily affective, or behavioural[[176]](#footnote-176) * completely or mostly false; or element of truth but misleading * related to a new topic or a repeated claim about the actions or character of a particular individual or individuals – particularly where they are represented either as a threat or as culpable for wrongdoing   **The context is**   * low level of trust in policymakers to protect the public from the threat or provide justice * a culture approving of vigilante responses to threats[[177]](#footnote-177)   **Those who perceive the claim credible**   * approve of vigilante responses to threats * have existing negative attitudes towards the individual or type of individuals in question | Numerous studies have found that misinformation has caused or contributed to vigilante violence against an individual or small group of individuals around the world.  In 2018, for example, images on WhatsApp suggesting strangers were kidnapping children in India caused or contributed to a series of vigilante killings (Banaji et al, 2019, and others).  And false claims made in offline community networks that medical researchers were harming children in Ethiopia caused two researchers to be killed and a third member of the team maimed. (Nur, 2019)  Among the first entries in this study, I identified several examples that caused or had potential to cause vigilante violence against individuals.  For example, I argue in the study that a **false claim** that the body of a missing child had been found under a church **caused a violent attack on the church and injuries ‘to many.’** See: (Entry 121).  See the discussion of effects in the study. | |
| **Potential to cause or contribute to violence between political, religious or ethnic groups** | | | | |
| **6.2.** Whether the exposure of the public, over the short or long-term, to **misinformation about the actions or character of particular political, religious or ethnic groups** has substantive potential to **cause or contribute to violence between political, religious or ethnic groups**. | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to enact political, religious or ethnic violence * affective, behavioural or cognitive[[178]](#footnote-178) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claims about the actions or character of a particular political, religious or ethnic group   **The context is**   * low level of trust in policymakers to protect the public * a culture approving of violent responses to perceived threats or wrongdoing by other groups   **Those who perceive the claim credible**   * approve of violent responses to perceived threats or wrongdoing * have existing negative attitudes towards the group in question | Numerous studies and reports from around the world have shown that misinformation has caused or contributed to political, religious and ethnic violence around the world – as part of conflicts and in civil unrest (See: Adegoke, 2018. Kaye, 2019 and more)  Among the first entries in this study, I identified several that had potential to contribute to political, ethnic or religious violence either as part of organised conflicts or in civil unrest.  I argue in the study, for example, that a false claim that UN troops collaborated with Uganda-based rebels in a massacre of more than 100 people in Eastern DR Congo **had had potential to cause political violence and may have done so**, causing or contributing to an attack on UN troops. See: Entry 160). And, in another case, I argue that a false claim that hundreds of knives confiscated by police in Kano State, Nigeria, had been intended for use killing Christians **had had potential to cause violence against Muslims**. See: (Entry 209)  See the discussion of effects in the study. | |
| **Potential to distort course of insurgencies** | | | | |
| **6.3.** Whether the exposure of political, military policymakers, or leaders of an insurgency over the short or long-term, to **misinformation about** **history, politics, or** **the actions, intentions or capabilities of the other side** has substantive potential to **cause or distort the course of an insurgency by misleading them about a reality** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of political or military or insurgency leaders to cause or distort the course of an insurgency[[179]](#footnote-179) * affective, behavioural or cognitive[[180]](#footnote-180) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about the history, politics, actions, intentions or capabilities of &/or domestic or foreign support for the government or insurgency   **The context is**   * the policymakers, insurgency leaders under pressure to act * policymakers or insurgency leaders feel they have a good prospect of success   **Those who perceive the claim credible**   * have the capacity to act militarily vs. the other side | Numerous studies have examined the role of false information in causing or distorting the conduct of wars (See McNamara, 1995. Baum & Groeling, 2010. Levine, 2012. Freedman, 2022)  I have not identified similar studies in relation to insurgencies.  Among the first entries in this study, I did not identify a false claim with a clear potential to alter the course of an insurgency in this way.  See the discussion of effects in the study. | |
| **6.4.** Whether the exposure of the public over the short or long-term, to **misinformation about history, politics, or the actions, intentions or capabilities of the one of the parties to or against an insurgency** has substantive potential to **cause or distort the course of an insurgency by distorting public support for one side or the other.** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by a sufficient number of people to affect the course of the insurgency * affective, behavioural or cognitive[[181]](#footnote-181) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claims about the history, politics, resources, actions, intentions or capabilities of the other side.   **The context is**   * political, military pressure on the policymakers and/or insurgency leaders * a perception the government or insurgency has capacity to act successfully * popular support (e.g. participation in fighting, and/or logistical support) is necessary for practical or strategic reasons   **Those who perceive the claim credible**   * have the capacity to support or oppose action in a way that affects the course of events | Numerous studies have examined the role of false information in causing or distorting the conduct of wars.  (See McNamara, 1995. Baum & Groeling, 2010. Levine, 2012. Freedman, 2022)  I have not identified similar studies in relation to insurgencies.  Among the first entries in this study, I identified several as having the potential to distort public support to this effect.  For example, I argue that a running series of false claims that foreign powers support Biafran independence **had potential to strengthen support for a Biafran insurgenc**y. See: (Entry 66) as one example among many.  I also argue a false claim by Cameroonian insurgents **had the potential – as part of a wider false narrative – to influence the course of negotiations** taking place at the time. See: (Entry 163)  See the discussion of effects in the study. | |
| **MISINFORMATION - POTENTIAL EFFECTS ON JUSTICE SYSTEM** | | | | |
| **Potential to distort (i) course or outcome of a court case, (ii) trust in police, justice** | | | | |
| **7.1.** Whether the exposure of people in positions of authority in the justice system,[[182]](#footnote-182) over the short or long-term, to **misinformation about the facts of a particular court case** has substantive potential to **distort the course of the case** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people in positions to influence the outcome of the court case.[[183]](#footnote-183) * affective, behavioural or cognitive[[184]](#footnote-184) * completely or mostly false; or element of truth but misleading * related to the facts of the particular case - and/or similar cases in a way that might prejudice approach to this case.   **The context is**   * those involved do not have a culture of questioning such claims * those involved are subject to outside pressure to prejudice outcome   **Those who perceive the claim sufficiently credible**   * have capacity to influence the course &/or outcome of the case | Numerous studies have also been published of the effect of public misinformation in court cases (Frenda et al, 2011. Dando, 2020 etc).  Among the first entries in this study, I identified one example as having the potential to affect the outcome of a particular court case.  I argue in the study that a false claim that a woman had withdrawn an accusation of rape **had** **had the potential to distort the outcome of the case,** building pressure on her to withdraw and providing justification for policymakers seeking to delay it for political and other reasons. See: (Entry 20)  See the discussion of effects in the study. | |
| **7.2.** Whether **repeated** exposure of the public, over the long-term, to **misinformation about the police, judicial or penal system** has substantive potential to **distort trust in,**  **&/or engagement with, the police or judiciary** – **with potential wider effects for society** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people for distortion of trust to have wider effects for society * affective, behavioural, or cognitive[[185]](#footnote-185) * completely or mostly false; or element of truth but misleading * related to a repeated claim or new topic about how police, courts, prisons function that might affect trust   **The context**   * requires trust in police, courts or prisons to ensure the public engages with justice system, avoid vigilantism   **Those who perceive the claim credible**   * have a reason to engage with the police or justice system | I have not identified well-researched studies of the effects of misinformation on trust in and/or engagement with the police and justice system. I will update if I do.  Among the first entries in this study, I have not identified any entries as having this effect. I will update if I do.  See the discussion of effects in the study. | |
| **Potential to distort public policies related to justice system** | | | | |
| **7.3.** Whether the exposure of people in authority in the justice system, over the short or long-term, to **misinformation about the police, judicial or penal system** has substantive potential to **distort public policies on the justice system** – **with potential wider effects for society** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people in positions of authority in the justice system to alter justice policy[[186]](#footnote-186) * affective, behavioural or cognitive[[187]](#footnote-187) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claims about areas of policing or the justice system open to policy action   **The context is**   * policymakers are under pressure over functioning of the justice system * a culture of not challenging policy decisions or actions by those in authority,   **Those who perceive the claim credible**   * have capacity to set policy for police or the justice system.[[188]](#footnote-188) | I have not identified well-researched studies of the effects of misinformation on policy within the police and justice system.  Among the first entries in this study, I identified one as having the potential to affect policing policy.  I argue in the study that a false claim by the Police Minister in the South African parliament that the policymakers needed to increase the police budget to meet a supposed UN standard on police staffing levels, **had had the potential to influence public funding of the police**. See: (Entry 180).  See the discussion of effects in the study. | |
| **7.4.** Whether the exposure of the public, over the short or long-term, to **misinformation about the police, judicial or penal system** has substantive potential to **distort public policy indirectly by creating or undermining pressure** **on policymakers to make or change policy.** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people to pressure policymakers to alter policy * affective, behavioural, or cognitive[[189]](#footnote-189) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about which justice authorities have capacity to set or alter policy   **The context is**   * those in authority are vulnerable to public pressure on the issue   **Those who perceive the claim credible**   * have capacity to put pressure on those in authority on policing or the justice system | I have not identified well-researched studies of the effects of misinformation on public attitudes in a way that would affect policy in the police and justice system.  Among the first entries in this study, I have not identified an example as having the potential to have this effect.  See the discussion of effects in the study. | |
| **MISINFORMATION – POTENTIAL EFFECTS ON BUSINESS, ECONOMIC ACTIVITY** | | | | |
| **Potential to distort (i) business operations, (ii) macro-economic activity** | | | | |
| **8.1.** Whether the exposure of the public, over the short or long-term, to **misinformation about the operations of a business or business sector** has substantive potential to **affect those operations.** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient actual or potential customers of the business or business sector to affect its operations * affective, behavioural, or cognitive[[190]](#footnote-190) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about actions, prospects, ownership structure or effects of a business, or business sector, and its operations   **The context is**   * the business or sector is vulnerable to public and commercial pressure[[191]](#footnote-191)   **Those who perceive the claim credible**   * have a capacity affect the business e.g. boycotting or using its services.[[192]](#footnote-192) | Studies show the potential of misinformation to affect business operations, both negatively (Atkinson, 2019. Katwala, 2019 & more) and positively – short term at least (Hoggan et al, 2009. Tesler, 2018. Supran & Oreskes 2021).  Among the first entries in this study, I have identified a number of false claims that caused or had potential to cause effects for businesses or business sectors. For example, I argue in the study that a false claim that a chain of shops in Nigeria was South African owned **caused the shops to be attacked and damaged by protestors**, in reprisal for supposed events in South Africa. See:  <https://www.icirnigeria.org/fact-check-are-shoprite-supermarkets-in-nigeria-franchise-outlets/>.  I also argue that a false claim that an airline employee shown mishandling luggage worked for South African Airways **had potential to affect its business**. See: (Entry 113). See the discussion of effects in the study. | |
| **Potential to distort macro-economic activity** | | | | |
| **8.2.** Whether the exposure of political, economic policymakers, over the short or long-term, to **misinformation about the state of the economy** has substantive potential to **distort macro-economic policy, by misleading them about a reality** | | I argue the likelihood & nature of consequences depend on whether:   **The false claim is**   * perceived to be credible by sufficient people in authority to alter economic policy[[193]](#footnote-193) * primarily affective or cognitive[[194]](#footnote-194) * completely or mostly false; element of truth but misleading * on a topic related to the state of or prospects for the economy, of the effects of different economic policies   **The context is**   * the economy is vulnerable to changes in or failures of economic policy   **Those who perceive the claim credible**   * have a capacity to alter macro-economic policy based on a false understanding. | Numerous studies of economic policy and major economic events - from market bubbles to financial collapses – show clear evidence of the effects of false information.  Among the first entries in this study, I have identified a number of false claims that had potential to affect macro-economic policy by persuading policymakers of a false claim. For example, I argue in the study that a false claim by Ghana’s president that the country’s deficit had halved as a percentage of GDP since he took power was misleading, and **had the potential to distort economic policy** by making lawmakers believe the deficit was down. See: (Entry 296)  See the discussion of effects in the study. | |
| **8.3.** Whether the exposure of the public, over the short or long-term, to **misinformation about the state of the economy** has substantive potential to **distort macro-economic activity, by misleading them about a reality** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people to distort macro-economic activity * affective, behavioural or cognitive[[195]](#footnote-195) * completely or mostly false; or element of truth but misleading * related to a new topic or a repeated claim about the state of or prospects for the economy, or the effects of different economic policies   **The context is**   * the economy is particularly vulnerable to changes in public sentiment   **Those who perceive the claim credible**   * have a capacity to change their economic activity based on their understanding or sentiment.[[196]](#footnote-196) | Studies of major economic events - from the Californian gold rush to latter day financial collapses – show that exposing the public to false information can distort macro-economic activity.  Among the first entries in this study, I did not identify a false claim that had potential to directly affect broad macro-economic activity specifically. I will update if I do so.  See the discussion of effects in the study. | |
| **MISINFORMATION – POTENTIAL EFFECTS ON THE ENVIRONMENT** | | | | |
| **Potential to distort policy on (i) climate change (ii) wildlife, biodiversity, & affect environment** | | | | |
| **9.1.** Whether the **repeated** exposure of policymakers &/or the public, to **misinformation about climate change** has substantive potential to **distort (i) public policy and (ii) personal behaviour in a way that has effect on the climate** | | I argue the likelihood & nature of consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people in authority to alter public policy affecting climate change * affective, behavioural or cognitive[[197]](#footnote-197) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about the reality, causes, effects & costs of climate change &/or costs and effects of responses to it   **The context is**   * the vulnerability of the climate to policy or personal action or inaction   **Those who perceive the claim credible**   * have a capacity to take action in ways that may affect the climate long term | Numerous studies have explored the effects of misinformation about climate change (Tesler, 2018. Supran & Oreskes 2021. Webb, 2021. WEF, 2020) – on public policy, and personal behaviour, misleading them about a reality.  Among the first entries in the study, I identified one example of a false claim that had potential to affect public policy on climate change by affecting the opinion of policymakers.  For example, I argue that a false claim that Victoria Falls was running dry due to climate change which convinced the Zambian President thus **had potential to reinforce policy positions**. See: (Entry 111).  See the discussion of effects in the study. | |
| **9.2.** Whether the **repeated** exposure of the public, to **misinformation about climate change** has substantive potential to **distort (i) public policy and (ii) personal behaviour in a way that has effect on the climate** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people to affect either pressure politicians to change public policy &/or individual action on a scale affecting climate change * affective, behavioural or cognitive[[198]](#footnote-198) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about the reality, causes, effects & costs of climate change &/or costs and effects of responses to it   **The context is**   * the vulnerability of the climate to policy or personal action or inaction * the vulnerability of politicians to public pressure on the issue if the public is misinformed   **Those who perceive the claim credible**   * have a capacity to pressure politicians on the issue * have a capacity to take action in ways that may affect the climate long term | Numerous studies have explored the effects of misinformation about climate change (Tesler, 2018. Supran & Oreskes 2021. Webb, 2021. WEF, 2020) – on public policy, and personal behaviour, misleading them about a reality.  Among the first entries in the study, I have not identified an example that had potential to affect public opinion in a way that would affect public policy on climate change or personal actions. | |
| **9.3.** Whether the **repeated** exposure of policymakers to **misinformation about wildlife, biodiversity** has substantive potential to **distort public policy on protection of the environment** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people in authority to affect policy * affective, behavioural or cognitive[[199]](#footnote-199) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claims about the reality, causes, effects & costs of loss of biodiversity and wildlife & costs and effects of responses to it   **The context is**   * the vulnerability of the ecosystem to action, inaction   **Those who perceive the claim credible**   * have a capacity to take action or alter policy on wildlife and biodiversity | I have not identified any well-researched broader studies of the wider effect of misinformation on policies towards biodiversity and wildlife.  Among the first entries, I have not identified any false claims that had potential to have this effect.  See the discussion of effects in the study. | |
| **9.4.** Whether the **repeated** exposure of the public, to **misinformation about wildlife, biodiversity** has substantive potential to **distort (i) public policy by creating or undermining public pressure on the issue (ii) personal behaviour on protection of the environment** | | I argue the likelihood & nature of any consequences depend on: **Whether the false claim is**   * perceived to be credible by sufficient people to affect policy or activity * affective, behavioural or cognitive[[200]](#footnote-200) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claims about the reality, causes, effects & costs of loss of biodiversity and wildlife & costs and effects of responses to it   **The context is**   * the vulnerability of the ecosystem to action, inaction * the vulnerability of politicians to public pressure   **Those who perceive the claim credible**   * have a capacity to take action or build pressure on policymakers to alter policy on wildlife and biodiversity | In 2012, Africa Check identified false information about the ‘medicinal’ effects of rhino horn as fuelling demand for the animals – thus providing evidence that false information can directly affect wildlife populations. [See link](https://africacheck.org/fact-checks/reports/medical-claims-rhino-horn-youre-better-aspirin-or-biting-your-nails).  I have not identified any well-researched broader studies of the wider effect of misinformation on policies towards biodiversity and wildlife.  Among the first entries, I have not identified any false claims that had potential to have this effect.  See the discussion of effects in the study. | |
| **MISINFORMATION – POTENTIAL FINANCIAL LOSS, DIGITAL HARMS FOR INDIVIDUALS, ORGS** | | | | |
| **10.1.** Whether the exposure of the public, over the short-term, to misinformation in in the form of **hoaxes and clickbait** has substantive potentialto **cause financial loss, loss of ID data, and digital harms such as exposure to computer viruses** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * **hoax** – perceived to be credible by people vulnerable to the inducements offered (work, travel etc).[[201]](#footnote-201) * **clickbait** – perceived as engaging or entertaining – not necessarily credible * primarily affective or behavioural[[202]](#footnote-202) * **hoax / clickbait**: completely or mostly false; or bearing an element of truth but misleading * related to topic of need or interest for the audience   **The context is**   * the greater the social need (e.g. for work), the greater the likelihood of effect   **Those who perceive the claim credible**   * the greater the need (e.g. for work), the more likely the effect * the more susceptible the individuals are to misinformation, the more likely the effect | A limited number of studies (Cable et al, 2020 and others) have examined the effects of misinformation causing financial or ID loss or digital harms.  Among the first entries in this study, I identified many examples that caused or had potential to have these effects.  For example, I show in the study that a clickbait report claiming that a famous singer had been found dead in a public wastebin contained a computer virus and thus **caused harm to those who responded**. [See link](https://africacheck.org/fbcheck/kenyan-gospel-music-star-ringtone-apoko-is-alive-and-well/) (Entry 175)  And a false claim that the company Airtel was giving away 500,000 iPhones **had potential to cause individuals financial loss**. See: (Entry 33)  See the discussion of effects in the study. | |
| **10.2.** Whether the exposure of the public, over the short-term, to **misinformation about the character, actions or rights of an individual** has substantive potentialto **affect the social standing, earnings and/or career of the individual concerned** | | I argue the likelihood & nature of any consequences depend on whether:  **The false claim is**   * perceived to be credible by sufficient people to affect the social standing &/or career of the subject of the claim [[203]](#footnote-203) * primarily affective or behavioural[[204]](#footnote-204) * completely or mostly false; or element of truth but misleading * related to a new topic or repeated claim about the actions or character of an individual or individuals, in a way that would affect their reputation   **The context is**   * the individual concerned has a social standing &/or career or earnings that would be affected by a negative public understanding   **The subject(s) of the misinformation**   * may lose social standing, earnings and/or their career if the false claim is believed | I have not identified well-researched studies of the effects of misinformation on the social standing &/or career of subjects of misinformation. Among the first entries in this study, I identified one as having had the potential to have this effect - briefly.  The **false claim** made that a district of Johannesburg had banned informal waste collectors from working in the area **had potential to (briefly) restrict their earnings**.  See the discussion of effects in the study. | |

1. If potential consequences – direct, conditional, or cumulative?

To understand the prospects for consequences to occur and the effectiveness of counter misinformation approaches proposed, I set out whether such consequences would be: (i) direct, (ii) conditional, (iii) cumulative or (iv) a combination of these, based on these criteria.

|  |  |  |
| --- | --- | --- |
| **Type of potential consequences** | **Criteria** | **Examples** |
| **Potential direct consequences** i.e.   * Immediate effect * dependent on neither particular conditions nor cumulation to cause the consequence | * **primarily affective** information that **may affect the mood of the subject themselves – directly on its publication** | For example, I argue in the study that a false claim that a well-known Kenyan singer had been ‘found dead in a waste bin **had potential to cause the man distress**. See: (Entry 176) |
| * **primarily affective** or **behavioural** information about **an identity issue** (national, ethnic, religious, gender or other demographic group.) that is **not dependent** on particular conditions or cumulative effects | I also argue that a **false claim** that the government in South Africa wanted to give high school certificates to high school ‘dropouts’ had potential to be used to enact abuse of, or discrimination against black youths. See: (Entry 27) |
| * **affective**, **cognitive or** **behavioural** information about the **prevalence, spread, effects or treatment of or for a health condition** which **could affect the general population** | I argue that a **false claim** that doctors harvest the organs of organ donors - against their will, while they are alive - **had potential to reduce organ donation** among the general population – **harming public health**. See: (Entry 167)  Similarly a **false claim** that aubergine leaves are a cure for breast cancer **had potential to reduce take-up of effective medication**. See: (Entry 162) |
| * **primarily affective**, **or** **behavioural** information about **events or circumstances inducing stress for public** | For example,fabricated news of a calamity or disaster that would affect a general population. I have not identified any examples in the first entries examined but will update if I do. |
| * **Primarily affective, or behavioural** information  about an **opportunity or challenge, or clickbait report appealing to emotions** | In the study, I have identified numerous examples of fake job adverts and clickbait that potential to cause financial or other harms – depending on how individuals responded. |
| **Potential consequences - subject to conditions**  i.e. only occurs   * to particular people * in particular circumstances * particular beliefs | * **affective, behavioural or cognitive** information whose effects are **limited to a particular set of people** | e.g. I argue a **false claim** that a drink made from manioc and bissap juice was a cure for HIV had **potential to affect the health of people living with HIV** – i.e. consequences for those with the particular condition. See: (Entry 145) |
| * **affective, behavioural** or **cognitive** information whose effects are **limited to a particular set of circumstances** | For example, **false claims** that violence was taking place around polling stations in a Nigerian state election – may have had **potential to cause consequences in the particular circumstances of an upcoming election**. See: (Entry 91) |
| * **affective, behavioural or cognitive** information whose **consequences require adherence to a particular belief or beliefs** in order to occur | For example, I argue in the study that **false claims** made in Togo that two women were ‘bird witches’, were seen online and had potential to inspire assaults on other women, **conditional on the individuals’ pre-existing belief in the supernatural**. See: (Entry 133) |
| **Potential consequences – by cumulative effect** i.e. only occurs   * through change or reinforcing of existing views or shaping of new views over time | * **primarily affective or cognitive** information on a topic on which audiences have **no pre-set view, but may form a view over time** if subject to a false claim or claims repeatedly | For example, studies show that **false claims** that vaccines are dangerous to health **have potential – if repeated – to shape the attitudes and behaviour of people who had no previous view on vaccines.** See: (Entry 186) |
| * **primarily affective or cognitive** information on a topic on which audiences have a **pre-existing view which may be subject to reinforcement, or change,** if subject to repeated false claims | For example, studies suggest that **repeated exposure to false claims** that exaggerate the threat posed by Muslims to Christians in Nigeria **have potential to shape attitudes in ways that could cause, contribute to discrimination, abuse, violence.**  See: (Entry 174) |

1. If potential consequences – what scale

To properly assess the nature of potential consequences, and thus understand the proportionality and effectiveness of any responses, it is also necessary to consider the scale, severity and duration they have or may have. As with observed consequences, I have identified the scale of potential consequences as affecting:

1. a specific individual or individuals
2. non-specific individuals (i.e. not specific to a particular person, type or group)
3. a specific group or set of groups, a business or organisation
4. society as a whole

Where the potential consequences depend on undetermined contextual factors, I recognise it is not possible to identify who would be affected. I also note that save where the finding indicates effects for a single individual, these categories indicate less **the number** of those affected and more **the focus**. Misinformation that affects the behaviour of individuals across society may affect more people than if it were to affect the behaviour of a specific group. To assess this I have used the following criteria.

|  |  |  |
| --- | --- | --- |
| **Who might be affected** | **Criteria** | **Examples** |
| **Potential consequences limited to a specific individual or individuals** | * relates to the **health, actions or character** of **a specific individual or individuals** – in a way **that harms their reputation, mental wellbeing**   *\*If the individual holds a position of authority (e.g. local politician) what affects them may have wider effects* | For example, I argue a **false claim** that a well-known Cameroonian singer had died had **potential to cause distress to her and her family personally** – but no substantive potential effects for others. See <https://congocheck.net/faux-la-chanteuse-congolaise-mbilia-bel-nest-pas-morte/>. |
| **Potential consequences affect individuals in general – not a specific set** | * relates to (i) **an issue, incident or event that concerns,** or (ii) **behaviour that could be taken up by individuals** – not a specific identifiable group   For example, **misdirection during a natural disaster** (fire, flood), or **promoting risky behaviour to the public** - not a specific group. | For example, I argue a **false claim** a South Sudanese soldier caught a bullet in his mouth had **potential to cause copycat behaviour among individuals** – not a specific group. See (Entry 36) |
| * relates to **a need or desire experienced by individuals** – not specific to a particular group | For example, the consequences of a hoax give away of 500,000 iPhones affected individuals – desire for a free phone not being specific to any particular group. See: (Entry 33) |
| **Potential consequences would be limited to a specific group or set of groups, businesses or organisations** | * relates to the **actions or behaviour** of a **specific demographic or social group or groups** | For example, I argue in the study that a **false claim** that ‘foreign nationals’ burned down a popular Johannesburg building **had potential to cause violence, discrimination, &/or abuse against foreign nationals**  See: (Entry 32). |
| * relates to the **actions or behaviour** a **specific business or organisation** | For example, a **false claim** a telecoms company was giving away free data **had potential to negatively affect the reputation of that company – when revealed to be a scam**. See: (Entry 75) |
| * relates to an **issue, circumstance or threat** whose effects would be limited to a **specific identifiable group**. | For example, the consequences of a **false claim that loan defaulters** in Kenya faced imminent arrest were felt **specifically by loan defaulters** in Kenya. See: (Entry 8)  Misinformation that promotes mistrust of wider society among a particular demographic group would be another example. |
| * relates to a **particular medical disease or condition** and the consequences affect only **people with or vulnerable to that particular disease or condition** | e.g. a **false claim** that a particular fruit was a cure for diabetes had potential to cause **consequences for people living with diabetes**; not more widely.  See: (Entry 153) |
| **Potential consequences would be felt by society as a whole** | * relates to **an issue which could shape the outcome of an election or elections** | I argue that **misinformation that shapes the outcome of an election**, has the potential consequences for **all of society within the area covered by the election &** potentially beyond. See: (Entry 5) |
| * relates to **an issue that could shape public policy affecting society as a whole** | Misinformation that affects **public policy related to society as a whole** (macro-economy, broad public health matters, broad issues of justice etc) has **potential consequences for society as a whole**. See: (Entry 173) |
| * relates to **an issue the consequences of which affect society as a whole** (e.g. the state of economy, public health, levels of gender violence etc.) | For example, I argue in the study that a **false claim** that had potential to increase gender violence in Nigeria thus had potential consequence for society as a whole. See: (Entry 166) |
| * relates to an **issue that affects levels of trust or polarisation across society as a whole** | See discussion in the study of meta-misinformation effect, and broad societal effects on trust and polarisation. |

1. If potential consequences – what severity

Other things being equal, if, due to false information, an individual takes an ineffective medication for a mild disease, the consequences they may face will be less severe than if the condition concerned is serious. To properly assess the consequences for society of different types of misinformation, I have identified the severity of potential consequences faced. I start with the possibility that some forms of mis/disinformation may in given circumstances have beneficial effects for individuals or society. Where effects are negative, I assess in terms of both the extent of the effect and its permanence or duration. Where the severity of potential consequences depends on multiple undetermined contextual factors, I recognise in the database it is not possible to identify the potential severity of the effects.

|  |  |  |
| --- | --- | --- |
| **Severity of potential consequence** | **Criteria** | **Examples** |
| **Potential beneficial consequences**  i.e. brings benefit for individuals, groups or society overall – without causing negative effects for others | * The false claim (i) **causes or contributes to behaviour with substantive positive effects** or(ii) **discourages behaviour with substantive negative effects** for individuals, groups or society. * The false claim **does not cause a negative effect** for other individuals or groups, including the meta-misinformation effect on trust. | In the study, I explore the argument that false information can be beneficial to society overall. For example, a **false claim** that Victoria Falls was running dry due to climate change **appeared to reinforce the (accurate) view of the Zambian President that climate change represents a danger to his country**. See: (Entry 111) It can be argued this was a potentially beneficial effect. See the discussion of in the study. |
| **Potential consequences would be mildly negative.** i.e.   * A limited emotional,   physical or other effect (e.g. minor financial loss)   * Temporary, or   reparable | * The false claim causes or contributes to a **limited and temporary negative emotional** (e.g. fear or distress) **physical** (e.g. mild illness) or **other effect** (e.g. minor financial loss). | For example, I argue, the emotional effect for people living with HIV/AIDS by a **false claim** their government would end free ARV treatment was limited to stress, and, as the claim was quickly disproven, **temporary**. See: (Entry 138)   * A **false claim** causes or contributes an individual or individuals suffering a mild illness, **the physical and psychological effects of which are also temporary**.   See the discussion of this issue in the study. |
| **Potential consequences would be moderately negative** i.e.   * A chronic, (not permanent) emotional, physical or other effect * Effects more severe (e.g. physical violence) but temporary, reparable | * The false claim causes or contributes to a **medium-term or chronic but not permanent negative emotional** (e.g. stress, fear), or **physical** (e.g. physical violence) **effects**. | For example, the primary consequence of a **false claim** about an anti-FGM activist was **chronic but not permanent emotional stress.** I argue in the study this was a moderate effect. See: (Entry 26).  In another example, a man **mistakenly identified** as ex-Ivorian president Laurent Gbagbo suffered **a physical assault**. Like the psychological effects – his injuries were **not long-lasting**. See: (Entry 90). |
| * The false claim promotes the use of an **ineffective response to a mild medical condition – and the harm caused is not permanent** | For another example, a **false claim** that pineapple leaves cure nosebleeds and boosts the immune system had potential to cause **overreliance on ineffective medication for a very mild condition and a weak immune system**. See: https://africacheck.org/fact-checks/meta-programme-fact-checks/pineapple-leaves-not-wonder-cure |
| * The false claim causes or contributes to **serious but reparable financial loss or destruction of property for individual or organisation** | For example, after a false claim, a church in Nigeria was attacked leading to “wanton destruction of properties and injuries to many”. The effects were serious but short-lived. See: (Entry 121) |
| * The false claim causes or contributes to **serious negative effect to an individual’s career but the harm is temporary, reparable** | e.g. **false claims** about the personal life of a Ugandan musician **affected his ability to work - temporarily.** I identified this as a moderate effect. |
| **Potential consequences would be severely negative**  i.e.  **negative** i.e.   * Effects moderate but long-lasting or permanent * Effects temporary but severe | * The false claim causes or contributes to a **long-term or permanent negative emotional (**e.g. distress, fear) or **physical effect** – **lasting years not weeks or months** | For example, I argue in the study that a **false claim** that pricking the ear lobes of someone suffering a stroke would cure the condition – had **potential to cause or contribute to their death – thus severe effect**. See: (Entry 30)  A **false claim** that three health researchers in Ethiopia were harming children caused **two to be killed and one permanently maimed**. Again – severe effect. See: (Entry 274) |
| The **false claim** causes or contributes to a combination of verbal and physical abuse the **effects of which are, or are likely to be, long-term.** | For example, the two women in Togo suffered effects likely to be long-term when they were falsely identified as ‘bird witches’ and verbally & physically assaulted. See: (Entry 133) |
| The false claim causes or contributes to **serious financial loss for individual or organisation or the destruction of property and the loss is not recoverable** | For example, if misinformation were to harm a business to the extent it had to lay off staff or close – I would identify this as severe. |
| * The false claim **distorts public policy,\* directly through influencing policymakers, or indirectly through public opinion, or distorting the outcome of an election**.   \*Even if the direct negative effects for individuals or society are mild or moderate, I identify the distortion of public policy as a severe thing in itself – as undermining the due process of governance. | For example, a **false claim** that the governments of DR Congo, Malawi and Rwanda were boycotting a summit in South Africa had **potential to cause Nigeria to do the same – distorting its foreign policy**. See: (Entry 57) |

1. If potential consequences – what duration

In order to make the above assessments, I have also sought to identify the potential consequences of entries in the database as being either (i) short-term or reversible or (ii) long-term or irreversible. Based on a close examination of the database, I identified duration as being a function of:

1. **the duration for which the audience will perceive the claim to be credible**

I identify this as depending on (i) the topic, (ii) the initially perceived credibility of the false information and (iii) the perceived credibility of the corrective information

1. **the time-limited nature or reversibility of the consequence**

I identify duration as depending also on the time-limited nature and/or reversibility of the consequence. Panic will ease when a predicted calamity fails to occur, however, a person killed in vigilante violence caused by panic cannot be brought back to life.

Where the potential consequences depend on undetermined contextual factors, I recognise it is not possible to identify who would be affected.

|  |  |  |
| --- | --- | --- |
| **Duration of potential consequence** | **Criteria** | **Examples** |
| **Potential consequences would be short-term &/or reversible** | * A **predicted or reported event** is observed to have **neither occurred nor is about to occur** and **no irreversible action has been taken** | e.g. **the panic** caused when parents believed foreign nationals were kidnapping children from schools in South Africa was **naturally short-lived** – when the parents found no such event had occurred. See: (Entry 54) |
| * The effects are **naturally short-term**, and/or **measures can be taken to counter-effect** or **reverse the consequences caused.** | For example, **injuries heal**; a **damaged property is rebuilt**; **a law** **passed** – all due to misinformation – **is repealed** **and its effects reversed**. |
| * **Credible corrective information** is presented in such a way to **correct the false understanding caused**. | Studies show that **credible corrective information may correct false understanding** - if presented in a timely way and repeated. (See Porter & Wood 2021 etc) |
| **Potential consequences would be long-term &/or irreversible** | * The consequences are inherently **long-term or irreversible processes** | For example, the consequences **exacerbate the process of climate change** or launch a **long-term political realignment**. |
| * The false claim is **naturally difficult to disprove**, for example requires access to and trust in official statistics or information. | For example, many of those who believe false claims about vaccine effects do not trust the sources used to correct false claims. |

1. The IFCN is the primary international standards-setting body for the field of fact-checking, set up in 2015 at the Poynter Institute. As of late 2024, 175 fact-checking organisations around the world were signed up to its Code of Principles: <https://www.ifcncodeofprinciples.poynter.org/>. For full disclosure, the author was from 2016-2024 an advisor to the IFCN, co-author of the original version of the IFCN Code in 2016 and led the first review of the Code in 2019-2020. [↑](#footnote-ref-1)
2. See BBC Editorial guidelines: <https://www.bbc.com/editorialguidelines/guidelines/editorial-standards>. [↑](#footnote-ref-2)
3. This draws on the definition used in the IFCN Code of Principles adopted by 175 fact-checking organisations. <https://ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles>. [↑](#footnote-ref-3)
4. See examples of the definitions used by (i) IFCN-accredited fact-checking organisations such as **Africa Check** and **Politifact** and (ii) misinformation researcher Claire Wardle. [↑](#footnote-ref-4)
5. Ways in which images may be mislabelled, misidentified or misattributed include: (i) genuine still or moving visual images paired with misleading text captions, (ii) genuine still or moving images cropped or presented in a sequence or setting that reframes their meaning [↑](#footnote-ref-5)
6. **See:** Vraga, EK, Bode, L. (2020) ‘Defining Misinformation and Understanding its Bounded Nature: Using Expertise and Evidence for Describing Misinformation’ Political Communication Vol 37. [↑](#footnote-ref-6)
7. Egelhofer, JL, Lecheler, S. (2019) ‘Fake news as a two-dimensional phenomenon: A framework and research agenda’. Annals of the International Communication Association, 43(2), 97–116. <https://doi.org/10.1080/23808985.2019.1602782> [↑](#footnote-ref-7)
8. Bontcheva, K, Posetti, J (Eds). (2020) ‘Balancing Act: Countering Digital Disinformation while respecting Freedom of Expression’. UNESCO. https://www.broadbandcommission.org/Documents/working-groups/FoE\_Disinfo\_Report.pdf [↑](#footnote-ref-8)
9. See: Lewandowsky, S, Ullrich, KH, Ecker, CM, Seifert, NS, and Cook, J. (2012) ‘Misinformation and Its Correction: Continued Influence and Successful Debiasing’. Association for Psychological Science. [https://journals.sagepub.com/doi/10.1177/1529100612451018](https://journals.sagepub.com/doi/10.1177/1529100612451018D). See also: Duffy, B. (2018) *The perils of perception. Why we’re wrong about nearly everything*. Atlantic Books. [↑](#footnote-ref-9)
10. E.g., a false claim that the Botswanan president’s helicopter had crashed. [See:](https://factcheck.afp.com/botswanas-president-wasnt-helicopter-crash-pilots-did-perform-precautionary-landing) (Entry 22) [↑](#footnote-ref-10)
11. E.g., a false claim related to official responses to a ferry accident in Kenya. See: (Entry 80) [↑](#footnote-ref-11)
12. E.g., a hoax about medical needs following accidents (and health issues). [See:](https://factcheck.afp.com/beware-fake-medical-appeals-misuse-sick-childrens-photos) (Entry 97) [↑](#footnote-ref-12)
13. E.g., a false claim related to the South African airline SAA’s operations. See: (Entry 113) [↑](#footnote-ref-13)
14. E.g., a claim re cost of living. See: (Entry 126). (**See**: claims about poverty/wealth in development) [↑](#footnote-ref-14)
15. E.g., a false claim a famous Cameroonian footballer had died. See: (Entry 127) [↑](#footnote-ref-15)
16. E.g., a false claim foreign nationals started a fire at a historic SA building. See: (Entry 32) [↑](#footnote-ref-16)
17. E.g., a false claim on crime rates in an SA city. See: (Entry 69) [↑](#footnote-ref-17)
18. E.g., a false claim related to a rape case due to go to trial in Nigeria. See: (Entry 20) [↑](#footnote-ref-18)
19. E.g., a false claim related to a bill on hate speech in Nigeria. See: (Entry 101) [↑](#footnote-ref-19)
20. E.g., a false claim about appropriate number of police in South Africa. See: (Entry 166) [↑](#footnote-ref-20)
21. E.g., a false claim an image showed conditions in prisons in Senegal. See: (Entry 136) [↑](#footnote-ref-21)
22. E.g., a false claim a missing child’s body had been found, man had confessed. See: (Entry 121) [↑](#footnote-ref-22)
23. E.g., a false claim that Lagos had a (proven) population of 21 million. See: (Entry 193) [↑](#footnote-ref-23)
24. E.g., a false claim that only 8% of Congolese had access to electricity. See: (Entry 282) [↑](#footnote-ref-24)
25. E.g., a false claim that Nigeria had achieved food security. See: (Entry 249) [↑](#footnote-ref-25)
26. E.g., a false claim SA schools teach a crude form of sex education. See: (Entry 102) [↑](#footnote-ref-26)
27. E.g., a false claim about the number of students at university in SA in 1994. See: (Entry 241) [↑](#footnote-ref-27)
28. E.g., a false claim about conditions in schools in Kenya and other countries. See: (Entry 82) [↑](#footnote-ref-28)
29. E.g., a false claim of plans to give exam certificates to high school dropouts. See: (Entry 27) [↑](#footnote-ref-29)
30. E.g., a false claim that climate change caused reduction in size of Lake Chad. See: (Entry 73) [↑](#footnote-ref-30)
31. E.g., a false claim about the production of solar panels in Africa. Link not used [↑](#footnote-ref-31)
32. E.g., a false claim about the protection of a type of zebra foal. See: (Entry 72) [↑](#footnote-ref-32)
33. E.g., a false claim about the volume of waste produced in Nigeria per year. See: (Entry 192) [↑](#footnote-ref-33)
34. E.g., a false claim a Nigerian company was giving away 500,000 cell phones. See: (Entry 33) [↑](#footnote-ref-34)
35. E.g., a false claim that about drinking behaviour by South African women. See: (Entry 171) [↑](#footnote-ref-35)
36. E.g., a false claim the Nigerian military ordered women to ‘dress responsibly. See: (Entry 81) [↑](#footnote-ref-36)
37. I have not identified this topic in the first 120 entries reviewed. I will update if I do [↑](#footnote-ref-37)
38. E.g., an unproven claim there are 2 million incidents of rape in Nigeria a year. See: (Entry 251) [↑](#footnote-ref-38)
39. E.g., a false claim that polygamy is required by law in Eritrea. See: (Entry 124) [↑](#footnote-ref-39)
40. E.g., a false claim that a minister in Cameroon admitted carrying out atrocities. [See:](https://factuel.afp.com/non-le-ministre-camerounais-atanga-nji-na-pas-demande-pardon-pour-les-atrocites-quil-aurait-commises) (Entry 140) [↑](#footnote-ref-40)
41. E.g., a false claim that former Nigerian presidents were arrested for corruption. See: (Entry 38) [↑](#footnote-ref-41)
42. E.g., a false claim about protests related to the high salaries of MPs. See: (Entry 12) [↑](#footnote-ref-42)
43. E.g., a series of false claims about reproductive health in Nigeria. Broken link so not included [↑](#footnote-ref-43)
44. E.g., an unproven claim that 8 in 10 Senegalese suffer poor mental health. See: <https://africacheck.org/fr/fact-checks/articles/pas-de-preuves-que-huit-senegalais-sur-dix-souffrent-de-troubles-mentaux> [↑](#footnote-ref-44)
45. E.g., a false claim a photo showed conditions in Cameroon maternity hospital. See: (Entry 133) [↑](#footnote-ref-45)
46. E.g., a false claim that an effective vaccine against AIDS had been created. See: (Entry 123) [↑](#footnote-ref-46)
47. E.g., a false claim that cases of Ebola had been identified in Kinshasa. See: (Entry 131) [↑](#footnote-ref-47)
48. E.g., a false claim health researchers in Ethiopia were abusing children. See: (Entry 241) [↑](#footnote-ref-48)
49. E.g., a false claim about the effects of cell phones on individuals’ health. See: (Entry 170) [↑](#footnote-ref-49)
50. E.g., a false claim that named firms were hiring staff. See: (Entry 98) [↑](#footnote-ref-50)
51. E.g., a false claim that media suppressed news of African Ebola vaccine. See: <https://congocheck.net/faux-muyembe-na-pas-invente-un-vaccin-contre-ebola-plutot-une-molecule-therapeutique/> [↑](#footnote-ref-51)
52. E.g., an imposter claim to be a particular news organisation in Eritrea. See: (Entry 108) [↑](#footnote-ref-52)
53. E.g., a false claim about fake marriages for immigration purposes in SA. See: (Entry 177) [↑](#footnote-ref-53)
54. E.g., a false claim a South African crowd beat a Nigerian to death. See: (Entry 51) [↑](#footnote-ref-54)
55. E.g., a false claim about land ownership by race in South Africa by race. See: (Entry16) [↑](#footnote-ref-55)
56. E.g., a false claim two women in Togo were ‘bird witches’ with special powers. See: (Entry 132) [↑](#footnote-ref-56)
57. E.g., a false claim a statue was a tribute to those killed in the slave trade. See: (Entry 115) [↑](#footnote-ref-57)
58. E.g., a false claim a giant rock is floating in mid-air as proof of God’s power. See: (Entry 24) [↑](#footnote-ref-58)
59. E.g., a false claim a South Sudanese man caught a bullet in his teeth. See: (Entry 36) [↑](#footnote-ref-59)
60. E.g., a false claim that US restaurant offers cannibalism on menu. See: (Entry 99) [↑](#footnote-ref-60)
61. E.g., a false claim that Rwanda controls phone lines in DRC.See: (Entry 138) [↑](#footnote-ref-61)
62. E.g., a false claim that DR Congo government figures have Rwandan ID cards. See: <https://congocheck.net/desintox-ces-fausses-cartes-didentite-attribuees-a-des-personnalites-congolaises/> [↑](#footnote-ref-62)
63. E.g., a false claim that a photo showed UN troops fighting DRC civilians. See: (Entry 160) [↑](#footnote-ref-63)
64. E.g., a false claim that Ghana had one of world’s fastest growth rates. [↑](#footnote-ref-64)
65. E.g., a false claim the Cameroon will start charging public to access ARVs. See: (Entry 139) [↑](#footnote-ref-65)
66. E.g., a false claim the DR Congo army had found a major rebel arms dump. See: (Entry 151) [↑](#footnote-ref-66)
67. E.g., a false claim a Zimbabwean embassy was in arrears on rent, evicted. See: (Entry 77) [↑](#footnote-ref-67)
68. E.g., a false claim that video of shooting showed threat to voters in Nigeria. See: (Entry 91) [↑](#footnote-ref-68)
69. E.g., a false claim US had banned Nigerian politicians for rigging election. See: (Entry 34) [↑](#footnote-ref-69)
70. E.g., a false claim about ANC leader’s view of Zimbabwe’s economic policy. See: (Entry 39) [↑](#footnote-ref-70)
71. E.g., a false claim photo showed public support for Kenyan politicians. See: (Entry 184) [↑](#footnote-ref-71)
72. E.g., a false claim about ministers to be named in new Nigerian government. See: (Entry 29) [↑](#footnote-ref-72)
73. E.g., a false claim that identified a Zambian man as a Kenyan attacked in SA. See: (Entry 58) [↑](#footnote-ref-73)
74. E.g., a false claim that an anti-FGM campaigner had undergone FGM. See: (Entry 26) [↑](#footnote-ref-74)
75. E.g., a false claim about birth of children outside marriage in Nigeria. See: (Entry 172) [↑](#footnote-ref-75)
76. E.g., a false claim about causes of death in childbirth. [↑](#footnote-ref-76)
77. E.g., a false claim that a zookeeper had made a female orangutan pregnant. See: (Entry 25) [↑](#footnote-ref-77)
78. E.g., a false claim France was providing military support for Boko Haram. See: (Entry 120) [↑](#footnote-ref-78)
79. E.g., a false claim that US was sending 3,500 troops to Cameroon. See: (Entry 117) [↑](#footnote-ref-79)
80. E.g., a false claim exaggerating number of Christians killed in Nigeria. See: (Entry 96) [↑](#footnote-ref-80)
81. See: Lewandowsky, S, Ullrich, KH, Ecker, CM, Seifert, NS, and Cook, J. (2012) ‘Misinformation and Its Correction: Continued Influence and Successful Debiasing’. Association for Psychological Science. <https://journals.sagepub.com/doi/10.1177/1529100612451018>. See also: Duffy, B. (2018) *The perils of perception. Why we’re wrong about nearly everything*. Atlantic Books. [↑](#footnote-ref-81)
82. For example, a false claim that the US politician Hillary Clinton, when US Secretary of State, [sold 20% of the country’s uranium stocks to Russia](https://africacheck.org/fact-checks/meta-programme-fact-checks/unravelling-claim-hillary-clinton-sold-20-us-uranium-russia) circulated on Facebook in South Africa. While the climate might have interested South Africans who followed US politics, most who saw and believed the false claim would have had no meaningful means to act on their false understanding. [↑](#footnote-ref-82)
83. In 2019, the police minister in South Africa made a false claim about police numbers in a parliamentary debate. This claim had substantial potential to influence spending on the police service, as discussed in more detail in the study. [↑](#footnote-ref-83)
84. See this interview of a former head of information security at Facebook, discussing the challenge of attributing the source of false online information if those who create it sought to keep their role covert. <https://firstdraftnews.org/articles/alex-stamos-interview-disinformation-campaigns/> [↑](#footnote-ref-84)
85. See [False reports claim US missionary underwent FGM in Kenya | Fact Check (afp.com)](https://factcheck.afp.com/false-reports-claim-us-missionary-underwent-fgm-kenya) (Entry 25) [↑](#footnote-ref-85)
86. See [Kenyan gospel music star Ringtone Apoko is alive and well - Africa Check](https://africacheck.org/fact-checks/meta-programme-fact-checks/kenyan-gospel-music-star-ringtone-apoko-alive-and-well) (Entry 175) [↑](#footnote-ref-86)
87. Evidence of perceived credibility identified in the study includes responses from notable identifiable individuals and organisations (media, politicians and other) who promote and comment on the claim.

    In a much-cited 2013 paper, researchers also found that social media responses (e.g. ‘likes’, emojis) accurately reflect readers’ personality traits, political views and responses to the material and more.   
    ‘Kosinski, M, Stillwell, Graepel, T. (2013). ‘Private Traits and Attributes are Predictable from Digital Records of Human Behavior’. Proceedings of the National Academy of Sciences. 110-15. 5802-5. [↑](#footnote-ref-87)
88. The assessment of what constitutes a sufficient number varies context to context. See notes in database. [↑](#footnote-ref-88)
89. Klapper, J. 1960. *The effects of mass communication*. Free Press. [↑](#footnote-ref-89)
90. Note I refer to the observed effect of the **claim** or variants of the claim – not necessarily the specific example of it that was identified by the fact-checker. Many claims appear in different channels and settings and it is not always possible to identify which version caused the effect seen. [↑](#footnote-ref-90)
91. See: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7606075/>   
    <https://www.mayoclinic.org/healthy-lifestyle/stress-management/in-depth/stress/art-20046037> [↑](#footnote-ref-91)
92. For example, a false claim that people anyone waking in the night should sit upright but not move further for 3.5 minutes before getting up – to avoid the risk of a stroke. (See: Entry 76) [↑](#footnote-ref-92)
93. In a 2020 report, I described an example. The President of the Nigerian Ophthalmologists’ Association identified the case – which he said occurred often– in which an individual believed a false claim that bathing his eyes in diluted battery acid would cure his conjunctivitis. The treatment was, of course, not a true ‘cure’ and left the man partially blind. [↑](#footnote-ref-93)
94. **Note** If that individual holds a position of authority (e.g. a national or local politician, or business leader) what affects them may have wider consequences [↑](#footnote-ref-94)
95. Note: Other things being equal, negative emotions have more severe effect on mental wellbeing if chronic than if short-term. I thus identify temporary effects as ‘mild’ compared with chronic effects, which I identify as ‘moderate’ or ‘severe’. [↑](#footnote-ref-95)
96. Even if the negative effects for individuals or society are mild or moderate, I would identify the distortion of public policy as a severe thing in itself – as undermining the due process of governance. [↑](#footnote-ref-96)
97. For the purposes of the database, I use the term ‘political misinformation’ to refer to information related to a political candidate, party or issue. I refer to misinformation related to the risks and processes of voting, and the likely outcome of elections in separate fields. See discussion in the study [↑](#footnote-ref-97)
98. As noted in Section 20 I have identified six factors as shaping perception of credibility. These are: (i)repetition of a claim; (ii) perceived plausibility of the claim and its presentation; (iii) ease of processing the claim; (iv) the claim’s affective impact and alignment with audience’s pre-existing views; (v) the perceived credibility of the source; & (vi) the durability, or resistance to correction of the claim. I set out in Section 20 how I identify claims perceived as credible using these and other criteria. [↑](#footnote-ref-98)
99. Both the number of swing voters &/or voters more susceptible to false information (typically older, less educated) reached, and the closer the election, help shape the likelihood of substantive effects. [↑](#footnote-ref-99)
100. Many years of research show that, other things being equal, affective information is more likely to shape or change an individual’s attitude to a political issue than cognitive information, particularly if delivered over a short period. (See Westen, D. *The Political Brain. The role of emotion in deciding the fate of the nation.* Public Affairs. 2008, pp 25-45 and other evidence – in the study). Further, while behavioural misinformation may have an effect on participation it is less likely to affect attitudes. [↑](#footnote-ref-100)
101. I argue that to cause a substantive change in voter preferences, on the basis of false information, a claim needs to be wholly or mostly false or misleading – not narrowly false.. [↑](#footnote-ref-101)
102. We know from much research set out in the study (Lewandowsky et al, 2012 ‘Misinformation and Its Correction: Continued Influence and Successful Debiasing’. Association for Psychological Science. <https://journals.sagepub.com/doi/10.1177/1529100612451018> and others.) that, once core political attitudes are formed, they are often slow to change. However, if the claim relates to a political issue, figure or party little known to most voters, it may shape attitudes over a short campaign. [↑](#footnote-ref-102)
103. Gunther, Richard et al. (2018) ‘Fake News May Have Contributed to Trump’s 2016 Victory’. Ohio State University. <https://www.documentcloud.org/documents/4429952-Fake-News-May-Have-Contributed-to-Trump-s-2016.html> [↑](#footnote-ref-103)
104. Guess, AM, Lockett D, Lyons, B, Montgomery JM, Nyhan, B, Reifler, J. (2020). ‘“Fake news” may have limited effects beyond increasing beliefs in false claims’ <https://misinforeview.hks.harvard.edu/article/fake-news-limited-effects-on-political-participation/> [↑](#footnote-ref-104)
105. E.g., information that (i) mispresents the character of a candidate or party or of events – in a way that affects public attitudes to individuals or issues, or (ii) misrepresents the effects of policies or the state of events in a way that creates a false understanding over time. [↑](#footnote-ref-105)
106. I argue again that to effect a substantive change in public attitudes, a false claim or narrative needs to be substantively, not narrowly false. [↑](#footnote-ref-106)
107. The exact topic of the claim and the broader false narrative of which it is a part will shape the likelihood and nature of any effects that occur. We know from many studies (Lewandowsky et al, 2012 and more) that core political and social attitudes, or ‘worldview’, are typically slow to form and slow to change. We know from other research (Allport & Lepkin, 1945. Swire-Thomson et al, 2017, Fazio et al, 2020. Lacassagne et al, 2022) that repetition increases the effect of even the most implausible claims. Claims are thus more likely to be perceived as credible by those who do not already believe them if frequently repeated. [↑](#footnote-ref-107)
108. Numerous companies have been set up around the world creating fake online accounts posing as supporters on the understanding that voters are more likely to support candidate seen as likely winner. (See Moore, M. *Democracy Hacked*. 2018. p 101) [↑](#footnote-ref-108)
109. Political operators and researchers have long argued it is easier to motivate (or demotivate) political supporters to vote than it is to convert a voter from one party to another [↑](#footnote-ref-109)
110. For example (i) a false claim on a political topic which provokes fear among uncommitted voters of the outcome if they do not vote for a particular party; (ii) a false claim on a political topic that includes a prompt urging individuals to vote or not vote; (iii) a false claim on a political topic that creates a false understanding about the outcome of voting for a particular party. [↑](#footnote-ref-110)
111. Voters attitudes thus more susceptible to being shaped by new information [↑](#footnote-ref-111)
112. A claim on a long-standing is more likely to have effect on attitudes if within a longer false narrative [↑](#footnote-ref-112)
113. A false claim on the risk of election violence, for example, is particularly likely to be perceived as credible if the country has a history of such violence. [↑](#footnote-ref-113)
114. E.g. (i) false claims of election-related violence that provoke fear of participating in an election; (ii) false information which urges people to avoid going to the polling station. [On this issue, cognitive information is less likely to have an effect]. [↑](#footnote-ref-114)
115. Any effect of exposure to information that was broadly true – stating correctly for example that fighting took place, but misstating the precise numbers injured – would be based more on accurate information, not what was narrowly false. [↑](#footnote-ref-115)
116. Studies suggest claims made about practical topics, such has whether or not violence has taken place or is expected may be more easily and quickly accepted as true than claims on political or social topics on which individuals have formed a fixed view. [↑](#footnote-ref-116)
117. The less the audience knows about election rules and processes the more vulnerable it is to misinformation. Typically, older and less educated audiences are more susceptible to misinformation. [↑](#footnote-ref-117)
118. For example providing (i) false information on voter registration or the voting process, (ii) false information on the location of polling stations – and not by affecting attitudes to candidates or parties. [↑](#footnote-ref-118)
119. For example, false information that (i) arouses anger at supposed rigging of an election, (ii) advises individuals not to vote because the election is foregone, (iii) provides a false understanding of polling that indicates the result is foregone. [↑](#footnote-ref-119)
120. A false claim on the risks of participating in protests particularly likely to be perceived as credible if the country has a recent history of arrests, punishment or violent responses. [↑](#footnote-ref-120)
121. E.g. (i) false claims that violence is expected at an upcoming protest creates fear which deters people from protesting, (ii) false information which urges people to avoid (or attend) a protest due to the risk/chance of violence. [On this issue, cognitive information is less likely to have an effect]. [↑](#footnote-ref-121)
122. In a review of Joseph Klapper’s 1960 work *The Effects of Mass Communication*, Professor of Government Colin Seymour-Ure noted, in 1996 (Times Higher Educational Supplement) that studies of media effects typically addressed “effects upon the attitudes of mass audiences” but media also has “effects on parties, parliaments, interest groups and prime ministers”. In the database I assess the potential for misinformation to affect the attitude and understanding not only of the public but also of those in positions of authority, able to shape policy or make decisions affecting individuals and wider society, from politicians in ministries to judges in courtrooms. [↑](#footnote-ref-122)
123. I refer to a broad range of fields of misinformation – from politics to science – over which people in positions of authority may have power to shape policy decisions. See discussion in the study [↑](#footnote-ref-123)
124. The number of people sufficient to effect a policy decision varies from organisation to organisation and context to context – and I do not seek to define it here. Other things being equal, the bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. [↑](#footnote-ref-124)
125. For e.g., (i) a false claim that appeals to the emotions of a minister & thus overcomes the rational analysis [see discussion in the study of the emotional appeal of Prof. Peter Duesberg’s arguments about HIV/AIDS]; (ii) a false claim that promotes a particular policy behaviour – such as calling for the government to introduce a boycott of meetings with another country [see discussion of this effect in relation to Nigeria/South Africa in2019]; (iii) a false claim that creates a false understanding of another country’s state of military preparedness or of the effects of a particular economic policy. [↑](#footnote-ref-125)
126. It is likely to have more effect on society if a minister has a false understanding of vaccines when vaccines are needed as a response to a public health crisis, than when no such crisis exists [↑](#footnote-ref-126)
127. For example, if a health minister believes false information about climate change, she may have no capacity to put her false belief into policy, whereas a minister for the energy sector or economy might. [↑](#footnote-ref-127)
128. For e.g., (i) a false claim that appeals to the emotions of the public – (see for example claims about the supposed killing of Nigerians in South Africa that led to calls for a boycott of a meeting; (ii) a false claim that promotes behaviour such as a vaccine boycott – which then pressures politicians to introduce one as policy (iii) a false claim that creates a false understanding of military preparedness which pressures politicians and the military to a particular course of action. [↑](#footnote-ref-128)
129. For example, sufficient to affect the outcome of an election or responses to calls for the public to follow particular guidance during a public health crisis. [↑](#footnote-ref-129)
130. E.g., information that (i) mispresents the character of a candidate or party or of events – in a way that affects public attitudes to individuals or issues, or (ii) misrepresents the effects of policies or the state of events in a way that creates a false understanding. [↑](#footnote-ref-130)
131. Scientists, medical experts, civil authorities etc. [↑](#footnote-ref-131)
132. E.g., information that (i) creates anger or outrage at a supposed unfairness or (ii) misrepresents the effects of authorities’ policies in a way that creates a false understanding [↑](#footnote-ref-132)
133. E.g., information that (i) creates a false understanding about the terms of particular laws, rules or standards or (ii) creates a false understanding about the penalties for non-compliance with them [↑](#footnote-ref-133)
134. E.g., information that (i) creates a false understanding about the terms of particular laws, rules or standards or (ii) creates a false understanding about the penalties for non-compliance with them [↑](#footnote-ref-134)
135. E.g., information that (i) creates anger or outrage at supposed media bias or (ii) misrepresents the coverage of the media in a way that creates a false understanding [↑](#footnote-ref-135)
136. The number of people in authority sufficient to make a policy decision varies from organisation to organisation. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. The less well-informed the government is about the other country the more likely false information is to be perceived credible. [↑](#footnote-ref-136)
137. E.g. the false claim (i) is emotionally charged, affecting public attitudes to the other country, (ii) explicitly urges particular behaviour e.g. a boycott of relations, (iii) provides false information, creating a false understanding about the status or actions of the other country [↑](#footnote-ref-137)
138. Evidence from the anti-xenophobia group Xenowatch found 18 people had been killed and 22 maimed or injured. Those killed included South Africans, Zimbabweans and Congolese and others whose nationality was unidentified. No Nigerians were known to have died. [↑](#footnote-ref-138)
139. The number of people in authority sufficient to make a policy decision varies from organisation to organisation. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. The less well-informed the government is about the other country the more likely false information is to be perceived credible. [↑](#footnote-ref-139)
140. E.g. the false claim (i) is emotionally charged, affecting public attitudes to the other country, (ii) explicitly urges particular behaviour e.g. a boycott of relations, (iii) provides false information, creating a false understanding about the status or actions of the other country [↑](#footnote-ref-140)
141. Evidence from the anti-xenophobia group Xenowatch found 18 people had been killed and 22 maimed or injured. Those killed included South Africans, Zimbabweans and Congolese and others whose nationality was unidentified. No Nigerians were known to have died. [↑](#footnote-ref-141)
142. What would constitute a sufficient number to have effect varies country to country and context to context. The bar for information to be judged credible by those in business may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate [↑](#footnote-ref-142)
143. E.g. the false claim (i) is emotionally charged, creating fear of travel to the other, (ii) explicitly urges particular behaviour e.g. a business boycott. [↑](#footnote-ref-143)
144. The bar for information to be judged credible by those in government or the military may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. **Note (i)** Where government or military have low levels of trust in &/or poor relations with the other country, they may be unable to accurately verify claims through official channels. **(ii)** If the information comes through their own or allied official channels, false information may be regarded as credible and not checked as rigorously as required.. [↑](#footnote-ref-144)
145. E.g. false information (i) that creates fear of attack – precipitating action. [*The ‘Gulf of Tonkin incident’ of August 1964 is one of the best documented such incidents*] (ii) that explicitly promotes a behavioural response to a particular event or circumstance (iii) on the status or intentions of another country or its military that leads to a response . [↑](#footnote-ref-145)
146. E.g. false information that (i) creates fear of attack or outrage at the supposed actions of the other country. (ii) relates to the behaviour of another country or its military or provokes popular support for action. (iii) misinforms he public on the status or intentions of another country or its military. [↑](#footnote-ref-146)
147. I use the term ‘health misinformation’ here broadly, referring both to false or misleading information about the causes, effects and prevalence of medical conditions, and the effectiveness or ineffectiveness of medical treatments and to the behaviour of health practitioners which might affect attitudes to health services. See discussion in the study [↑](#footnote-ref-147)
148. The perceived credibility of health information is particularly linked with the perceived credibility of the source. For some, this means research in peer-reviewed medical journals; for others anecdotal claims by friends and family, or the claims of fellow believers of health-based conspiracy theories. The number of people sufficient to cause a particular health effect depends on a number of factors including the nature of the condition concerned; fewer people needed spread an easily communicable disease than one which is not, for example. [↑](#footnote-ref-148)
149. E.g. false information that (i) creates misplaced fear, or unwarranted complacency, that directly reduces use of a beneficial treatment; (ii) promotes behaviour directly harmful to health. In a 2020 report, I described an example in which a man suffering conjunctivitis believed a false claim, spread in his community, that bathing eyes in diluted battery acid would cure the condition. The treatment left the man partially blind. (Cunliffe-Jones, 2020) (iii) misinforms the individual about the prevalence of a health condition, leading them to expose themselves to the disease. [↑](#footnote-ref-149)
150. I distinguish between ‘direct’ harm to health, where the behaviour itself causes harm to health, and ‘indirect’ harm, where the behaviour is not harmful *in itself* but causes individuals to, for instance, reduce use of a beneficial medication. [↑](#footnote-ref-150)
151. E.g. false information that (i) creates misplaced hope in an ineffective treatment, and thereby reduces use of a beneficial medication; (ii) promotes behaviour that is indirectly harmful to health e.g. use of cars rather than cycling (iii) misinforms the individual about the behaviour of health practitioners, reducing take-up of a beneficial treatment. [↑](#footnote-ref-151)
152. E.g. (i) false information that dismisses well-placed fears about the threat posed by a forest fire or other sort of natural disaster (ii) false information on the direction in which people should travel to flee a fire or other natural disaster. [↑](#footnote-ref-152)
153. The perceived credibility of health information is particularly linked with the perceived credibility of the source – as set out in reference 130. [↑](#footnote-ref-153)
154. False information about sexual or reproductive health may (i) create a false sense of complacency about the threat posed by a medical condition [*e.g. the false claims in 2020 that Covid was not affecting certain ethnic groups*] (ii) promote harmful behaviour [*e.g. the claim noted above that sex with a virgin is a cure for HIV*] (iii) provide false understanding of reproductive health [*e.g. promoting unsafe birth processes*]. [↑](#footnote-ref-154)
155. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. [↑](#footnote-ref-155)
156. E.g. The false information (i) appeals to the leader on an emotional level [*It is argued President Mbeki was attracted in 2000 to the argument HIV did not cause AIDS for emotional reasons. See Specter 2007*]; (ii) promotes a behaviour that has effect on a disease’s means of transmission; (iii) understates the prevalence of a new and worrying condition – thereby distorting the priority given to it by policymakers. [↑](#footnote-ref-156)
157. It is likely to have more effect on society if a minister has a false understanding of vaccines when vaccines are needed as a response to a public health crisis, than when no such crisis exists [↑](#footnote-ref-157)
158. For example, if the education minister believes false information about a health issue, she may have no capacity to put her false belief into policy, whereas the health minister may do so. [↑](#footnote-ref-158)
159. The perceived credibility of health information is particularly linked with the perceived credibility of the source. For some, this means information in peer-reviewed medical journals; for others anecdotal claims by friends and family [↑](#footnote-ref-159)
160. E.g. False information that (i) appeals to the public on an emotional level – for example unproven claims vaccines are dangerous and should be banned (Ayodele, 2007); (ii) seeks pressure on authorities to approve sale of a health treatment (as happened with a supposed AIDS treatment approved in DR Congo); (iii) creates a false understanding among the public of the prevalence of a health condition – and calls for policymakers act. [↑](#footnote-ref-160)
161. For many subjects of personal misinformation, what causes distress is the perception that the false information is or may be seen as credible by those who see it. [↑](#footnote-ref-161)
162. E.g. false information that shapes the attitudes of individuals about the actions or character of the subject [↑](#footnote-ref-162)
163. E.g. false information that (i) causes fear or distress for a group (e.g. false claims people in default on loans in Kenya are facing arrest) (ii) provokes panicked behaviour (e.g. false claims about kidnapping prompting parents of Johannesburg children to rush to pick them up from school). [↑](#footnote-ref-163)
164. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. [↑](#footnote-ref-164)
165. E.g. false information that promotes a false understanding of the prevalence of mental health issues or the effectiveness of treatments. [↑](#footnote-ref-165)
166. For example, if the foreign minister believes false information about mental health, she may have no capacity to put her false belief into policy, whereas the health minister may do so. [↑](#footnote-ref-166)
167. The perceived credibility of health information is particularly linked with the perceived credibility of the source. For some, this means information in peer-reviewed medical journals; for others anecdotal claims by friends and family – or in the Nigerian case, local religious leaders. [↑](#footnote-ref-167)
168. E.g. false information that (i) creates misplaced scorn of mental health as a public health concern, (ii) creates a false understanding of the prevalence of the problem or effects of different treatments [↑](#footnote-ref-168)
169. The information is more likely to shape or change discriminatory attitudes if the claim and/or close variants of it are repeated frequently over time [↑](#footnote-ref-169)
170. E.g. (i) the false claim creates or exacerbates fear of or contempt for the group in question, (ii) explicitly promotes discriminatory or abusive action towards the group on a false premise (iii) provides false understanding about the actions or character of the group e.g. the number arriving in the country or involved in crime. [↑](#footnote-ref-170)
171. The information has potential to directly, immediately cause discriminatory or abusive behaviour. It is more likely to do this if it aligns or is congruent with a broad false narrative about the group concerned that is accepted by the audience. [↑](#footnote-ref-171)
172. E.g. (i) the false claim creates or exacerbates fear of or contempt for the group in question, (ii) explicitly prompts discriminatory or abusive action towards the group on a false premise [↑](#footnote-ref-172)
173. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. [↑](#footnote-ref-173)
174. E.g. the false claim (i) creates or exacerbates contempt for the group in question, (ii) explicitly promotes on a false premise discriminatory or abusive action towards the group (iii) produces a false understanding about the actions or character of the group e.g. involvement in crime. [↑](#footnote-ref-174)
175. E.g. the false claim (i) creates or exacerbates contempt for the group in question, (ii) explicitly promotes on a false premise discriminatory or abusive action towards the group (iii) produces a false understanding about the actions or character of the group e.g. involvement in crime. [↑](#footnote-ref-175)
176. E.g. the false claim (i) creates or exacerbates a sense of fear of the individual or individuals in question, on a false basis; &/or dismisses the prospect that policymakers will protect those who are fearful; (ii) explicitly promotes vigilante action against the individual or group concerned [↑](#footnote-ref-176)
177. See Banaji et al, 2019 – on the culture of vigilantism [↑](#footnote-ref-177)
178. E.g. the false claim (i) creates or exacerbates a sense of fear of or anger toward the group in question, on a false basis; (ii) explicitly promotes violent action against the group concerned, on a false basis; (iii) creates a false understanding of the number, actions or behaviour of the group. [↑](#footnote-ref-178)
179. As McNamara, Freedman and others show, even a very small number of senior leaders may – if misinformed – take decisions on the course of conflicts that they may not have taken if provided with accurate information. The 1964 ‘Gulf of Tonkin Incident’ is an example. [↑](#footnote-ref-179)
180. E.g. (i) false information that creates fear of or outrage at an attack – precipitating action. (ii) false information that overstates the prospects for success of a course of action. (iii) false information on the intentions or capabilities of the other side in the insurgency. [↑](#footnote-ref-180)
181. E.g. (i) false information that creates fear of attack or outrage at the supposed actions of one of the parties. (ii) false information relates to the actions of the other party or promoting particular behaviour against them. (iii) false information on the intentions or capabilities of the other side. [↑](#footnote-ref-181)
182. Depending on the judicial system concerned, this could mean a justice minister, prosecutor, judge or juror. [↑](#footnote-ref-182)
183. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. If those involved are subject to public or other pressure they may not need to be convinced themselves but simply need an argument to justify their decision. [↑](#footnote-ref-183)
184. E.g. the false claim (i) creates a strong emotional reaction prejudicing the approach to the case (ii) promotes, on a false basis, the argument the case should be dismissed (iii) creates a false understanding about the actions or character of one or both of the parties. [↑](#footnote-ref-184)
185. E.g. a false claim that (i) arouses a misplaced or exaggerated fear of (or complacency about the results of) engaging with the police or courts; (ii) exhorts individuals to engage or not with the police or courts; (iii) creates a false understanding of the police or courts resulting in a change in engagement. [↑](#footnote-ref-185)
186. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate. [↑](#footnote-ref-186)
187. E.g. a false claim that (i) relates to a particular case and affects attitude of authority to justice system; (ii) exhorts authority on a false basis to alter justice policy; (iii) creates a false understanding of how the system operates [↑](#footnote-ref-187)
188. For example, if the health believes false information about the justice system she may have no capacity to put her false belief into policy, whereas the justice minister may do so. [↑](#footnote-ref-188)
189. E.g. a false claim that (i) relates to a particular case and affects individuals’ attitude to justice system; (ii) exhorts individuals to pressure authority on a false basis to alter justice policy; (iii) creates a false understanding among individuals of how the justice system operates [↑](#footnote-ref-189)
190. E.g. (i) a false claim creates fear that a company’s product is dangerous and dissuades customers from use; (ii) a false claim about the company’s activities promotes a boycott, (iii) a false claim about the company’s financial structure persuades investors to divest. [↑](#footnote-ref-190)
191. Other things being equal, a business with limited financial reserves is likely to be more vulnerable to reputational misinformation than one with strong reserves, while a business with a monopoly for an essential service will be less vulnerable than one which has to compete for custom. [↑](#footnote-ref-191)
192. The fact an individual sees and perceives the misinformation as credible may have little effect on the business if they are not an actual or potential client of its services. [↑](#footnote-ref-192)
193. The bar for information to be judged credible by those in authority may be higher than for the general public if the process for deciding actions or adopting policies involves effective debate [↑](#footnote-ref-193)
194. For example, information that (i) affects policymakers’ attitude to the effects of economic policy actions (ii) misinforms policymakers on the state of the country’s finances or economic performance. [↑](#footnote-ref-194)
195. E.g., false information that (i) creates a false optimism about economic prospects may affect economic activity, (ii) promotes particular economic behaviour, (iii) distorts understanding of the prospects for national finances may encourage or discourage economic activity [↑](#footnote-ref-195)
196. Other things being equal, those on a high income have greater potential than those on a lower income to change their economic activity – cutting back or increasing spending – based on information received. [↑](#footnote-ref-196)
197. E.g., false information that (i) arouses misplaced fear of the health effects of wind turbines; (ii) promotes, on a false basis, behaviours that foster climate change (iii) misinforms policy makers about the effects of policy actions against climate change [↑](#footnote-ref-197)
198. E.g., false information that (i) arouses misplaced fear of the health effects of wind turbines; (ii) promotes, on a false basis, behaviours that foster climate change (iii) misinforms policy makers about the effects of policy actions against climate change [↑](#footnote-ref-198)
199. E.g., information that (i) arouses misplaced empathy for a particular species, over the interests of biodiversity (ii) promotes behaviour by the public that has an effect on biodiversity or wildlife (iii) misinforms policy makers about the effects of policy actions [↑](#footnote-ref-199)
200. E.g., information that (i) arouses misplaced empathy for a particular species, over the interests of biodiversity (ii) promotes behaviour by the public that has an effect on biodiversity or wildlife (iii) misinforms policy makers about the effects of policy actions [↑](#footnote-ref-200)
201. Unlike most examples of mis/disinformation, repetition of hoaxes and clickbait may reduce the likelihood of effect [↑](#footnote-ref-201)
202. E.g. false information that (i) creates hope a need (work opportunity) will be met/amuses or enrages if clickbait; (ii) promotes behaviour e.g. applying for the work opportunity; clicking the link [↑](#footnote-ref-202)
203. The publication alone of the accusation causes distress, regardless of perceived credibility. [↑](#footnote-ref-203)
204. E.g. false information that (i) provokes contempt towards the subject of the claim – on a false basis; (ii) promotes a boycott of the subject of the claim. [↑](#footnote-ref-204)