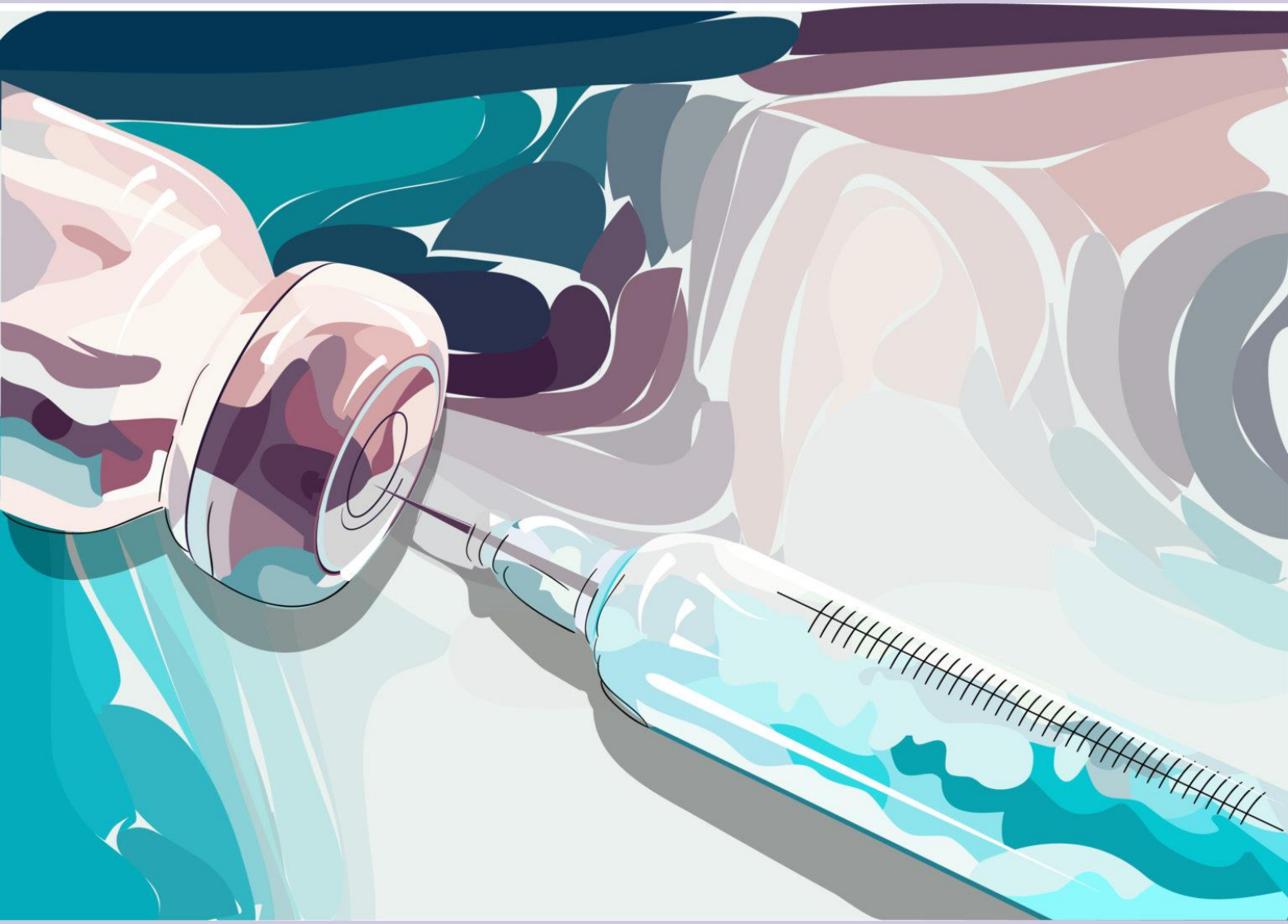
IPPO The International Public Policy Observatory

Debunk, Inform, Avoid?

Debunking vaccine-related misinformation

A RAPID EVIDENCE REVIEW



Ian Shemilt, Gareth Hollands, Claire Stansfield & James Thomas

December 2022

UKRI Economic and Social Research Council

Debunk, Inform, Avoid?

Debunking vaccine-related misinformation: a rapid evidence review

Ian Shemilt, Gareth Hollands, Claire Stansfield & James Thomas

Author affiliations

- Ian Shemilt*, EPPI Centre, UCL Social Research Institute
- Gareth Hollands*, EPPI Centre, UCL Social Research Institute
- Claire Stansfield, EPPI Centre, UCL Social Research Institute
- James Thomas, EPPI Centre, UCL Social Research Institute

* *Joint lead authors.*

Funding & funder involvement

This review was undertaken by the EPPI Centre as a partner of the International Public Policy Observatory (IPPO) (https://covidandsociety.com) which is ESRC-funded.

Conflicts of interest

There were no conflicts of interest in the writing of this report.

Contributions

The opinions expressed in this publication are solely those of the authors and do not necessarily reflect those of the EPPI Centre or the funders.

This report should be cited as: Shemilt I, Hollands GJ, Stansfield C, Thomas J (2022). *Debunk, Inform, Avoid? Debunking vaccine-related misinformation: a rapid evidence review.* London: EPPI Centre, UCL Social Research Institute, UCL Institute of Education, University College London.

Acknowledgements

We would like to thank unnamed stakeholders for their participation in a series of informal 1-to-1 discussions that were instrumental in helping us to identify, formulate and refine the research question addressed in this rapid evidence review. We would also like to thank Sarah O'Meara (International Public Policy Observatory) for leading on stakeholder engagement for this project in collaboration with the authors.

ISBN: 978-1-911605-40-9

Key Messages

- Current research evidence suggests that communications aimed at debunking vaccine-related misinformation should be used judiciously and sparingly.

  This proposal is broadly consistent with existing policies and practices of many national and local public health policymakers and communicators in the UK, which emphasise responding to false or misleading claims about vaccines by communicating accurate information and facts, and reserving debunking strategies for potential use in a more limited set of circumstances.

- Debunking messages may be more successful when they can be formulated, tailored and targeted to specific subgroups of people, and specific misinformation claims and narratives, based on insights derived from:

  - public health surveillance data analysis;
  - (social) media analytics;
  - best practice toolkits and guidance;
  - professional expertise, experience and judgement; and
  - analysis and understanding of why different people hold and maintain specific vaccine-related misinformation beliefs.

- While there is some evidence that debunking messages can reduce vaccine-related misinformation beliefs or improve vaccine attitudes, there is no clear evidence that any such changes in beliefs or attitudes are accompanied by positive or negative changes in people's vaccination intentions; and there is almost no current research evidence concerning the effects of debunking on vaccine uptake.

- Drawing more specific practical insights from current research evidence about when – under what circumstances and for whom – communications aimed at debunking vaccine-related misinformation are more or less likely to be effective is hampered by key differences between research studies, including in the research questions they address, and in the outcomes and differential effects they assess, using various different measures.

- The current evidence base is limited by the lack of studies designed to assess the differential effects of 'real-world' messages developed to debunk evolving vaccine-related misinformation narratives on beliefs, intentions and uptake behaviour over time in a large, diverse UK population sample. If stakeholders judge there is a need to develop further research that more directly addresses uncertainties about if and when to debunk vaccine-related misinformation in a current UK context, then scoping and development of such research could be informed by insights from this review.

Executive Summary

Background

Misinformation about COVID-19 has proliferated in digital and physical environments throughout the pandemic. Exposure to false or misleading claims about COVID-19 and other vaccines has been associated with serious and widespread harms to people's physical and mental health, and rates of severe illness and death from COVID-19 are higher among unvaccinated people. Developing effective ways to counter vaccine-related misinformation and its spread has therefore become a key challenge for policymakers, public health practitioners, and communications professionals.

Debunking can be defined as any form of direct communication aimed at correcting, refuting or rebutting an established misinformation claim or narrative. When debunking reduces people's beliefs in vaccine-related misinformation, changes attitudes to vaccination, reduces vaccine hesitancy and/or increases intentions to be vaccinated, it could help to increase uptake of vaccines. However, when debunking unintentionally strengthens people's beliefs in vaccine-related misinformation, increases vaccine hesitancy and/or reduces intentions to be vaccinated, it could instead discourage uptake.

Based on our initial engagement with UK stakeholders involved in tackling misinformation, we conducted a rapid evidence review that aimed to address the following question:

For whom and/or ***under what circumstances*** is **debunking** misinformation about vaccines likely to be more effective than either providing **accurate information only**, or **not responding**, for:

- Reducing people's vaccine-related ***misinformation beliefs***;
- Changing people's ***attitudes to vaccines***;
- Reducing people's ***vaccine hesitancy or resistance***, or increasing their ***intentions to be vaccinated***; and/or
- Increasing people's ***uptake of vaccines***?

In addition, what are the ***key gaps and uncertainties*** in this body of research evidence?

Methods

We searched multiple sources, including our map of research on COVID-19 misinformation, the *OpenAlex* dataset, and four electronic databases, up to early September 2022. Retrieved records were screened against prespecified eligibility criteria. Only studies with a parallel group, randomised controlled trial design were eligible for consideration. Data were extracted from each included study on its key characteristics, principal findings, implications for practice and policy, and implications for research. Extracted data were tabulated, analysed and summarised, and we have also created an open-access web database of included studies and their key characteristics.

Results

Characteristics of Included Studies

47 randomised controlled trials are included in the review, published since 2010. Most of these studies involved people from the United States, while only one involved people from the UK. Most studies involved young adults, adults and/or older adults, while a few specifically focused on parents.

Misinformation Types

Included studies focused either on misinformation about vaccines or viruses in general, or about specific vaccines or viruses, including measles, mumps and rubella, COVID-19, influenza, human papilloma virus (HPV), and Zika. Most studies investigated debunking of false or misleading claims relating to the safety, efficacy or necessity of vaccines. Fewer investigated debunking of conspiracy theories, claims concerning political or economic motives of people or organisations involved in vaccine development or claims about the vaccine development process.

Interventions

All included studies investigated messages that included a debunking component, with fewer investigating messages that also incorporated explicit repetition of the false or misleading information or myth being debunked, and/or additional accurate information about the vaccine or virus. Corrections used in debunking messages were mostly fact-based, with fewer logic-based, source-based, or fact-check label only corrections. Media channels included social media platforms, websites or blogs, newspapers, radio, and online video. However, in many cases, participants were exposed to debunking messages that were communicated either via printed materials, or via an online survey, specifically developed for the study.

Outcomes and differential effects

Most included studies assessed the effects of debunking messages on participants' vaccine-related misinformation beliefs. Many also assessed effects on participants' attitudes to vaccination and/or their vaccine hesitancy, resistance, or intentions. Only two studies assessed effects on participants' vaccine uptake. Most studies incorporated assessments of 'for whom' debunking messages were most likely to be effective, with fewer investigating 'under what circumstances' they were most likely to be effective.

Main Findings of Included Studies

Overall, the findings of included studies are equivocal about the effects of debunking messages on people's *vaccine-related misinformation beliefs and/or vaccine attitudes*. Some studies found overall beneficial effects of debunking messages on beliefs or attitudes. However, others found adverse effects of debunking on these outcomes, no effects on these outcomes, or mixed and/or differential effects.

The findings of some included studies suggested debunking messages might be *less* likely to have beneficial effects on beliefs or attitudes – and *more* likely to have adverse effects on these outcomes – among people who more *strongly* hold misinformation beliefs, or who are *more* resistant to vaccination. Conversely, these interventions might be *more* likely to have beneficial effects on beliefs or attitudes – and *less* likely to have adverse effects on these outcomes – among people who more *weakly* hold (or do not hold) vaccine-related (and other) misinformation beliefs, or who are *less* resistant to vaccination.

Whilst these findings confirm that debunking messages *can* reduce vaccine-related misinformation beliefs and improve vaccine attitudes among at least some people, most studies that identified a beneficial impact of debunking on such beliefs or attitudes did not find any corollary effects on people's *intentions to be vaccinated*. More generally, the effects of debunking on *vaccine hesitancy, resistance and/or intentions* were inconsistent between included studies.

Overall, there is no clear evidence of any differences in patterns of results between (i) studies that compared debunking with 'information only', (ii) those that compared debunking with 'not responding' (control), and

(iii) those that incorporated both types of comparison: evidence for the effects of debunking appeared similarly equivocal among all three subgroups of studies.

Implications for Practice and Policy

Communications aimed at debunking vaccine-related misinformation clearly *can* have a beneficial influence on vaccine-related misinformation beliefs and/or attitudes, among at least some people. However, such changes are not necessarily accompanied by positive changes in people's intentions to be vaccinated. There is also currently high uncertainty about whether debunking messages encourage uptake of vaccines.

Communications aimed at debunking vaccine-related misinformation should therefore be used judiciously and sparingly, as part of broader communications strategies and vaccine promotion campaigns that also target the wider array of psychological antecedents and other determinants of vaccine hesitancy, acceptance, and uptake.

Debunking messages are also unlikely to influence people's vaccine-related beliefs, attitudes or intentions in the same way. Communications of this kind may therefore be more successful when they can be formulated, tailored and targeted to specific subgroups of people, and specific misinformation claims or narratives, based on insights derived from:

- public health surveillance data analysis;
- (social) media analytics and active monitoring of health communications in media channels;
- best practice toolkits and guidance;
- professional expertise, experience and judgement; and
- analysis and understanding of why different people hold and maintain specific vaccine-related misinformation beliefs.

Policymakers and communicators should continue with the practice of engaging trusted, influential people ('key messengers') to act as co-producers and senders of debunking messages and other communications aimed at countering vaccine-related misinformation to promote vaccine uptake. More generally, policymakers and communicators should also continue to draw on insights from behavioural science when formulating debunking messages and other communications-based responses.

Implications for Research

A key gap in the current research evidence base is the lack of any studies designed to assess the differential effects of 'real-world' messages developed to debunk dynamic and evolving vaccine-related misinformation narratives on misinformation beliefs, and on vaccine intentions and uptake, over time in a large, diverse UK population sample. If stakeholders judge there is a need to develop further research that more directly addresses uncertainties about if and when to debunk vaccine-related misinformation in a current UK context, then scoping and development of such research could be informed by insights from this review.

Our review findings suggest that further efforts could be most appropriately applied to formulating a call for proposals for a randomised study, embedded in sequential waves of a longitudinal UK population survey, designed to address the most important aspects of this complex question, and with the potential to directly inform implementation in the UK context.

1. Background

1.1 COVID-19 and vaccine-related misinformation

From its outset the COVID-19 pandemic has fuelled unprecedented and sustained levels of global demand for, and supply of, related information in digital and physical environments, as governments, scientists and people everywhere rapidly sought to better understand how to protect their own and others' health from the virus (Pian 2021). This situation led the World Health Organization to declare a parallel 'infodemic', warning that too much information about COVID-19, including false or misleading information, would be likely to influence people's behaviour in ways that could severely undermine public health mitigation strategies (WHO 2020).

Misinformation can be defined as information that has features of being false, misleading or unsubstantiated, relative to current expert opinion and evidence (Vraga 2020). However, expert opinion and evidence evolve over time – especially in the context of a new virus – and misinformation and information are also complex phenomena, which are not binary, mutually exclusive or fixed. For example, misinformation claims often refer to at least some factual information aligned with current expert opinion (Peng 2021), and trusted information sources can unintentionally become vectors for spreading misinformation (e.g. if content becomes out of date compared with current evidence). Public use of the terms 'misinformation' and 'information' is therefore inevitably contested, and such debates have social, political, behavioural and other dimensions. Closely related phenomena include disinformation, malinformation, conspiracy theories, myths, fake news and pseudoscience (Tandoc 2017, Wardle 2017).

A range of different types of COVID-19-related misinformation claims became established (or re-established) in digital and physical environments during the early phase of the pandemic (Brennen 2020, Gabarron 2021, van der Linden 2020) including claims relating to the origins of the virus, its spread through communities, actions, policies or responses of public authorities, and preventive measures or potential treatments. Later, once news of the development of COVID-19 vaccines became widespread, vaccine-related misinformation claims also proliferated, including claims relating to the safety, efficacy or necessity of the vaccines, political or economic motives to develop them, and various conspiracy theories (Skafle 2022, Smith 2020).

Exposure to COVID-19-related misinformation has been associated with various harms to people's physical and mental health (do Nascimento 2022, Bierwiaczonek 2022, Rocha 2021, van Mulukom 2022). Exposure to vaccine-related misinformation has also been associated with serious adverse health, public health and social impacts in the context of COVID-19 and other viruses (Lieneck 2022, Ripp 2022). Higher rates of severe illness and death from COVID-19 have also been observed among unvaccinated people (ONS 2022). However, the reasons behind uptake of, or hesitancy towards, COVID-19 and other vaccines are complex (Abba-Aji 2022, Andreas 2022, Crawshaw 2021, Kafadar 2022, Kamal 2022, Smith 2022). Developing effective ways to counter vaccine-related misinformation and its spread has therefore become a key challenge and priority for policymakers, public health practitioners, and communications professionals involved in COVID-19 and other vaccination promotion campaigns.

1.2 Interventions for tackling COVID-19 and vaccine-related misinformation

Various kinds of interventions have been proposed to help tackle misinformation, related phenomena, and their spread, including in the context of COVID-19 (Bontcheva 2020, Janmohamed 2021). These range from measures focused on upstream prevention, such as education to develop media or health literacy skills

(Edwards 2021), to more downstream strategies attempting to alter cognitive and behavioural responses to misinformation, or limit its spread, more directly. The latter include communications-based responses aimed at 'inoculating' people against potential misinformation *before* specific false or misleading claims or narratives become established (pre-bunking) (van der Linden 2020), or aimed at debunking, or fact-checking, such claims or narratives *after* they have become established (Chan 2017, Walter 2018, Walter 2020, Walter 2021); technical and/or algorithmic solutions (Gradoń 2021), including the automated (or semi-automated) identification, removal or blocking of misinformation content on social and other digital media platforms (Krasodomski-Jones 2019, Marsden 2020); and regulatory or legal measures, such as increased regulation of social media platforms and content or criminal sanctions against 'senders' (sources) of harmful misinformation (Helm 2021).

Countermeasures to misinformation can be placed on a continuum from less to more intrusive, in terms of their potential to impinge on freedoms of speech and expression (Helm 2021). For example, interventions focused on improving people's media, digital or health literacy skills (education and training) – which already feature in the UK Government's draft Online Safety Bill (UK Government 2022) – and communications-based responses are less restrictive of individual liberties or autonomy, compared with content removal or blocking, and regulatory or legal measures. If less restrictive interventions are also effective in achieving their intended outcomes, they may be more attractive and feasible for policymakers to implement (Giubilini 2019).

1.3 Debunking COVID-19 and other Vaccine-Related Misinformation

Debunking can be defined as any form of direct communication aimed at correcting, refuting or rebutting an established misinformation claim or narrative (Lee 2021). Exposure to both misinformation claims, and communications aimed at correcting them, is a commonly reported experience in digital and physical environments (Bode 2021b). Debunking communications invariably integrate messaging intended to correct one or more misinformation narratives after they have already become established, with or without explicitly repeating the myth and/or incorporating messaging intended to convey other accurate and/or vaccine promoting information (Lewandowsky 2012, Roozenbeek 2022a).

When debunking reduces people's beliefs in vaccine-related misinformation, changes attitudes to vaccination, reduces vaccine hesitancy and/ or increases intentions to be vaccinated, it could help to increase their uptake of vaccines (Caulfield 2020). However, when debunking unintentionally strengthens people's beliefs in vaccine-related misinformation, increases vaccine hesitancy and/or reduces intentions to be vaccinated, it could instead discourage vaccine uptake.

During the summer of 2022, we convened a series of informal 1-to-1 discussions with selected stakeholders, who had been engaged in tackling health-related and other misinformation during the COVID-19 pandemic within the scopes of their roles in UK national or local government (including public health), or third sector organisations. The aims of this initial engagement were to find out what countermeasures stakeholders had deployed to tackle COVID-19 and other misinformation, and to progressively iterate between (a) understanding demand among stakeholders for research evidence to inform their policies and practice, (b) preliminary scoping to investigate the availability and characteristics of relevant bodies of research evidence that could be summarised to help address those policy and practice questions, and (c) formulating the questions and eligibility criteria for this rapid evidence review.

In these discussions, we heard from multiple stakeholders that countermeasures deployed in the UK at national and local levels during the COVID-19 pandemic had primarily focused on communications-based

responses to health-related misinformation narratives that had already become established, with vaccine-related misinformation becoming a key target for intervention. Communications professionals have been central to the coordination and delivery of such responses, typically working in coordination with data analysts, public health specialists, and 'key messengers', including elected officials and influential members of the public and communities. At a local level, communications-based responses have been highly tailored to address the specific needs and concerns of local populations and communities, informed by public health surveillance data, media monitoring and analytics, and 'on the ground' intelligence. These tailored communications have also used images of local people, and local branding or messaging.

We also heard from stakeholders that communications-based responses to COVID-19 vaccine-related misinformation had typically been deployed as part of broader communications campaigns aimed at reducing COVID-19 vaccine hesitancy and promoting vaccine uptake, which had invariably emphasised providing accurate information with positive framing. National government guidance for local and national health communicators during the COVID-19 vaccination programme had also focused on providing accurate information, and generally advised against debunking – primarily due to concerns about the risks of amplifying misinformation claims. However, multiple stakeholders expressed a common interest in understanding what is known from current research evidence about when – under what circumstances and/or for whom – it might be better to go further, by directly debunking vaccine-related misinformation claims, as well as when it might be better not to respond to such claims.

2. Review Question

This rapid evidence review aimed to address the following question:

For whom and/or ***under what circumstances*** is **debunking** misinformation about vaccines likely to be more effective than either providing **accurate information only**, or **not responding**, for:

- Reducing people's vaccine-related ***misinformation beliefs***;
- Changing people's ***attitudes to vaccines***;
- Reducing people's ***vaccine hesitancy or resistance***, or increasing their ***intentions to be vaccinated***; and/or
- Increasing people's ***uptake of vaccines***?

In addition, what are the ***key gaps and uncertainties*** in current research evidence?

3. Methods

Further details of the methods used in this rapid evidence review are available on request. Brief details of methods and procedures are provided below, with links to supplementary files.

3.1 Study identification

We identified eligible primary studies by searching multiple sources, including our map of research evidence on misinformation about COVID-19 (Shemilt 2022), the OpenAlex dataset (Chawla 2022), and four electronic databases covering health, social sciences and psychology research, from 2010 up to early September 2022. Further details of database searches (including the OpenAlex search) can be viewed and downloaded here.

Retrieved records were screened against prespecified eligibility criteria in two stages. First, for title-abstract screening, bibliographic records were imported into EPPI-Reviewer software (Thomas 2022), deduplicated and dual-screened, using active learning (O'Mara-Eves 2015) to continually re-prioritise the list of records yet to be screened. Second, for full text screening, published articles reporting potentially eligible studies were retrieved and dual-screened, with conflicts resolved by discussion and consensus.

3.2 Data extraction and analysis

A pilot data extraction and coding tool was developed and applied by two researchers to the same five included study reports. After discussion and refinement, the tool was used to extract data from each included study on its characteristics, principal findings, implications for practice or policy, and implications for research.

Each study was coded by one researcher, with key uncertainties resolved by discussion between two researchers. Extracted data were tabulated, analysed and summarised, and we have also created an open-access web database of included studies and their key characteristics (see also '4.1. Characteristics of included studies').

4. Results

4.1 Characteristics of Included Studies

We identified 47 studies eligible for inclusion in this review published since 2010 (Amazeen 2022, Bode 2015 'Study 1', Burger 2022, Carey 2020 'Yellow fever experiment', Chockalingam 2021, Clayton 2021 'General parent study', Clayton 2021 'Non-compliance study', Dixon 2013a, Dixon 2013b, Dixon 2015, Featherstone 2020, Gawronski 2022 'Experiment 1' Gawronski 2022 'Experiment 2', Haglin 2017, Helfers 2022, Horne 2015, Jolley 2014 'Study 2', Jolley 2017 'Study 1', Jolley 2017 'Study 2', Kim 2021, Lee 2022, Lunz Trujillo 2021 'Large survey experiment', Lyons 2019, Nyhan 2014, Nyhan 2015, Pluviano 2017, Pluviano 2019, Pluviano 2022 'Experiment 1', Pluviano 2022 'Experiment 2', Porter 2022, Reavis 2017 'Study 1', Schmid 2019 'Experiment 1', Schmid 2019 'Experiment 2', Schmid 2019 'Experiment 3', Schmid 2019 'Experiment 4', Schmid 2019 'Experiment 6', Steffens 2021, Stojanov 2015, Sullivan 2019 'Experiment', Sullivan 2019 'Replication attempt', Sun 2022, Vraga 2019, Vraga 2020 'Study 2', Xiao 2021, Yousuf 2021, Yue Dai 2022, Zhang 2021).

These 47 included studies (38 articles) were identified from 8,030 records retrieved by our database searches. A PRISMA Flow Diagram (Page 2021) can be viewed and downloaded here. We have also created an open-access web database to enable interactive exploration of these 47 included studies by their key characteristics, including the populations, types of misinformation, debunking messages (interventions), comparators, outcomes and differential effects they investigate. Selected key characteristics of included studies are summarised below[1].

Study designs

Only studies with a parallel group, randomised controlled trial design were eligible for consideration in this review. Almost all included studies (46 of 47) are individual randomised controlled trials (i.e. individual participants were randomly assigned to intervention or comparator groups). The other study (1 of 47) is a

[1] Please note that totals in this section do not necessarily add to 47 studies, because (i) multiple characteristics can apply to each study and (ii) some characteristics are not reported for some studies.

cluster randomised controlled trial (i.e. clusters of participants – in this case 'households – were randomly assigned to intervention or comparator groups). Three included studies were also conducted within the framework of a population survey (Clayton 2021 'General parent study', Clayton 2021 'Non-compliance study', Nyhan 2015).

Populations/ Participants

One included study involved participants from the United Kingdom (Pluviano 2017) and this study also included Italian people. Most of the studies involved people from the United States (34 studies), while around a quarter involved people from other European countries (Germany, Italy, Macedonia, or The Netherlands) (11 studies). Most studies involved young adults (40 studies), adults (38 studies) and/or older adults (32 studies), while 6 studies specifically focused on parents and guardians.

Misinformation Types

Included studies focused either on misinformation (and/or closely related phenomena) about vaccines or viruses in general (22 studies), or misinformation about specific vaccines or viruses, including: measles, mumps and rubella (MMR) (10 studies); COVID-19 (7 studies); influenza (7 studies); human papilloma virus (HPV) (4 studies); or Zika (4 studies). Most included studies investigated debunking of false or misleading claims relating to the safety, efficacy and/or necessity of vaccines (40 studies), with fewer investigating debunking of conspiracy theories (7 studies), claims concerning the political or economic motives of people or organisations involved in vaccine development (7 studies), or about the vaccine development process itself (e.g. concerning the speed of development of new vaccines, or provision of/ access to them) (4 studies) (Smith 2020).

Interventions

Included studies investigated the effects of different debunking messages with a wide variety of characteristics. One key dimension of variation concerns the composition of the debunking message(s) used (Lewandowsky 2020). All investigated messages that included a debunking component (i.e. messaging designed to correct, refute or rebut false or misleading claims about vaccines) (40 studies reported details) with fewer investigating communications that also incorporated explicit repetition of the false or misleading information, or myth, being debunked (17 studies) and/or additional accurate information (not pertaining to the false or misleading information, or myth, being debunked) about the vaccine or virus concerned (11 studies). The number of debunking message variations evaluated within each study ranges between 1 and 5, with a median of 1 (see also 'Differential effects' below).

Correction types used in debunking messages can be classified as 'fact-based', 'source-based', 'logic-based' (Kim 2021), or 'fact-check label only'. Most included studies investigated debunking using fact-based corrections (42 studies), which aim to counter misinformation by explaining relevant facts that conflict with false or misleading claims. Several studies (also) investigated debunking using logic-based corrections (10 studies), which use rhetorical approaches to highlight the logical flaws in misinformation claims. Few studies (also) investigated debunking using source-based corrections (2 studies), which highlight the motives or disreputable traits of misinformation sources; or using fact-check labels only (2 studies).

Included studies investigated debunking messages sent using a variety of media channels, including: social media platforms (Facebook, Twitter or Instagram) (12 studies); websites or blogs (11 studies); newspapers or magazines (4 studies); radio (1 study); or online video (1 study). Other included studies were conducted via online survey platforms (e.g. Amazon Mechanical Turk) (18 studies), or using printed materials specifically

prepared for the study (4 studies). Correspondingly, most studies investigated debunking using a text-only modality (45 studies), several studies investigated use of both text and images (usually via social media) (11 studies), and single studies investigated audio (1 study) or video (1 study) modalities.

Most included studies investigated debunking messages attributed to 'expert' sources (39 studies), including national government or health agencies (17 studies), scientists or researchers, or generic experts, fact-check organisations, or health care organisations/ professionals. Fewer studies investigated debunking messages attributed to 'non-expert' sources messengers (14 studies), typically generic non-experts (9 studies) or media organisations (4 studies). One included study investigated debunking messages attributed to national or local government elected officials or political leader(s), while none investigated debunking by faith organisations/ leaders, nor cultural or sporting organisations.

Comparisons and comparators

All included studies incorporated at least one comparison between (a) 'debunking' and 'information only' (9 studies), and/or (b) 'debunking' and (c) 'control' ('not responding') (45 studies). However, these 47 studies were originally designed to address a variety of specific research questions, which only partially coincide with those we aimed to investigate in this rapid evidence review (see '2. Research questions'). For example, one included study aimed to assess the impacts of "a video containing evidence-based content, social norms modelling and debunking strategies… compared to a video containing information and social norms without debunking strategies" on vaccine-related misinformation beliefs and attitudes to vaccines among older adults in The Netherlands (Yousef 2021). The number of eligible comparisons within each included study (with or without reporting usable data or results) ranged between 1 and 8, with a median of 2.

Another key dimension of variation between included studies concerns whether or not participants assigned to 'debunking', 'information-only', and/or 'control' arms of studies were also separately exposed (typically immediately beforehand) to stimulus materials containing vaccine-related misinformation claims/ narratives. For example, in one included study, some participants were exposed to a Facebook post claiming that COVID-19 vaccines were developed so quickly that they could not be safe (Amazeen 2022). This review includes studies in which all eligible intervention and comparator groups were exposed to separate stimulus materials of this kind (e.g. Jolley 2017 'Study 2'), as well as those in which none of the eligible intervention and comparator groups were exposed to such materials (e.g. Lunz Trujillo 2021). It also includes studies in which participants in the intervention group(s), but not in comparator group(s), were exposed to separate stimulus materials of this kind (e.g. Sullivan 2019 'Experiment'). However, we excluded studies in which participants in the comparator group(s), but not in the intervention group(s), were exposed to separate stimulus materials of this kind, since any such comparisons are likely to exaggerate any beneficial effects of debunking messages on target outcomes, relative to comparator groups.

Outcomes

Many included studies assessed the effects of debunking messages on participants' vaccine-related misinformation beliefs (35 studies). Many (also) assessed the effects of debunking on participants' attitudes to vaccination (21 studies) and/or their vaccine hesitancy, resistance or intentions (31 studies). Two included studies (also) assessed the effects of debunking on vaccine uptake.

In general, there are key differences (i.e. heterogeneity) between included studies in approaches to measuring outcomes for three of these four outcome types (i.e. beliefs, attitudes, and hesitancy/ resistance/ intentions). Approaches varied in terms of specific features of outcome measurement techniques or

instruments, metrics and data formats for analysis, and methods of aggregation (how group data are summarised). However, timepoints for outcome measurement are largely consistent among included studies: outcomes were typically measured immediately after exposing participants to either debunking or comparator messages; and longer-term follow-up was exceptional. Also, classifying specific outcome measures used among included studies into these types was sometimes complicated by included studies having measured (and reported) composite outcomes that combine multiple response items spanning more than one outcome domain.

Differential effects

Included studies investigated a wide variety of potential differential effects of debunking relevant to our review questions concerning for whom, and/or under what circumstances, debunking vaccine-related misinformation is likely to be more effective than providing information only, or not responding.

For whom?

Most included studies incorporated some kind of assessment of whether the outcomes (main effects) of 'debunking' (versus 'information only' and/ or 'control') were modified by characteristics of study participants ('for whom') (38 studies). Figure 1 illustrates the wide variety of participant characteristics investigated as potential effect modifiers (differential effects) across these included studies.

Overall, included studies assessed cognitive variables as potential effect modifiers (e.g. vaccine-related misinformation beliefs, perceptions of risk of vaccination, and/or other perceptions or psychological antecedents – measured at the beginning of the studies), more often than they assessed demographic and/or socio-economic variables (e.g. gender, race, ethnicity, culture or language, education level).

Under what circumstances?

Many included studies incorporated some kind of assessment of whether the outcomes (main effects) of 'debunking' were modified by the type(s) of misinformation (5 studies) and/or characteristics of the debunking messages (22 studies) investigated.

With regards to types of misinformation, 5 included studies assessed differential effects by vaccines or viruses (3 studies: Pluviano 2022 'Experiment 1', Pluviano 2022 'Experiment 2', Zhang 2021), misinformation claim types (2 studies: Lyons 2019, Porter 2022) and/or senders or sources of misinformation (1 study: Porter 2022).

Figure 2 illustrates the variety of characteristics of debunking messages investigated as potential effect modifiers (differential effects) across 22 included studies.

Figure 1. Included studies incorporating analysis of 'for whom' debunking is likely to be effective (N)

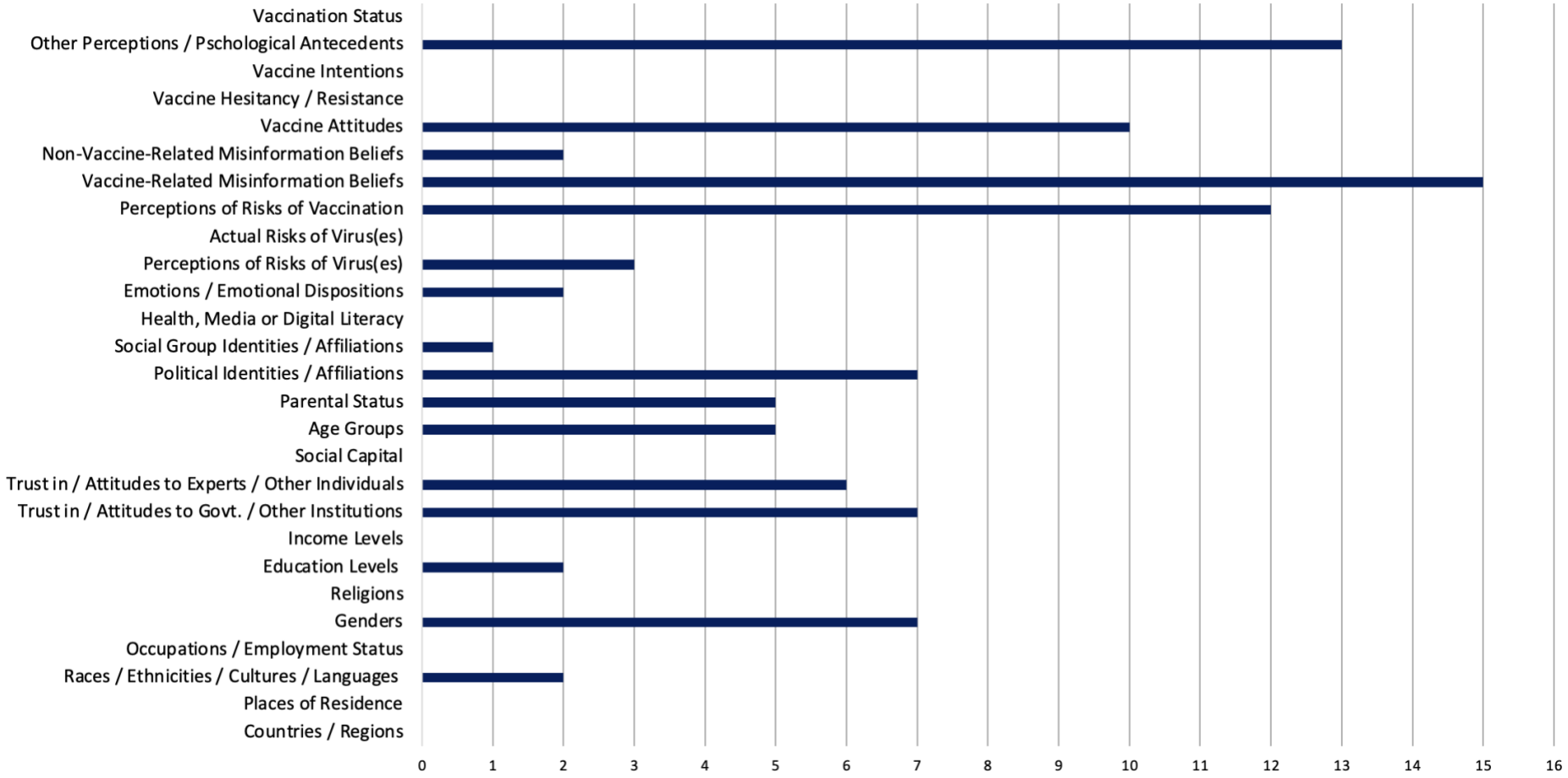
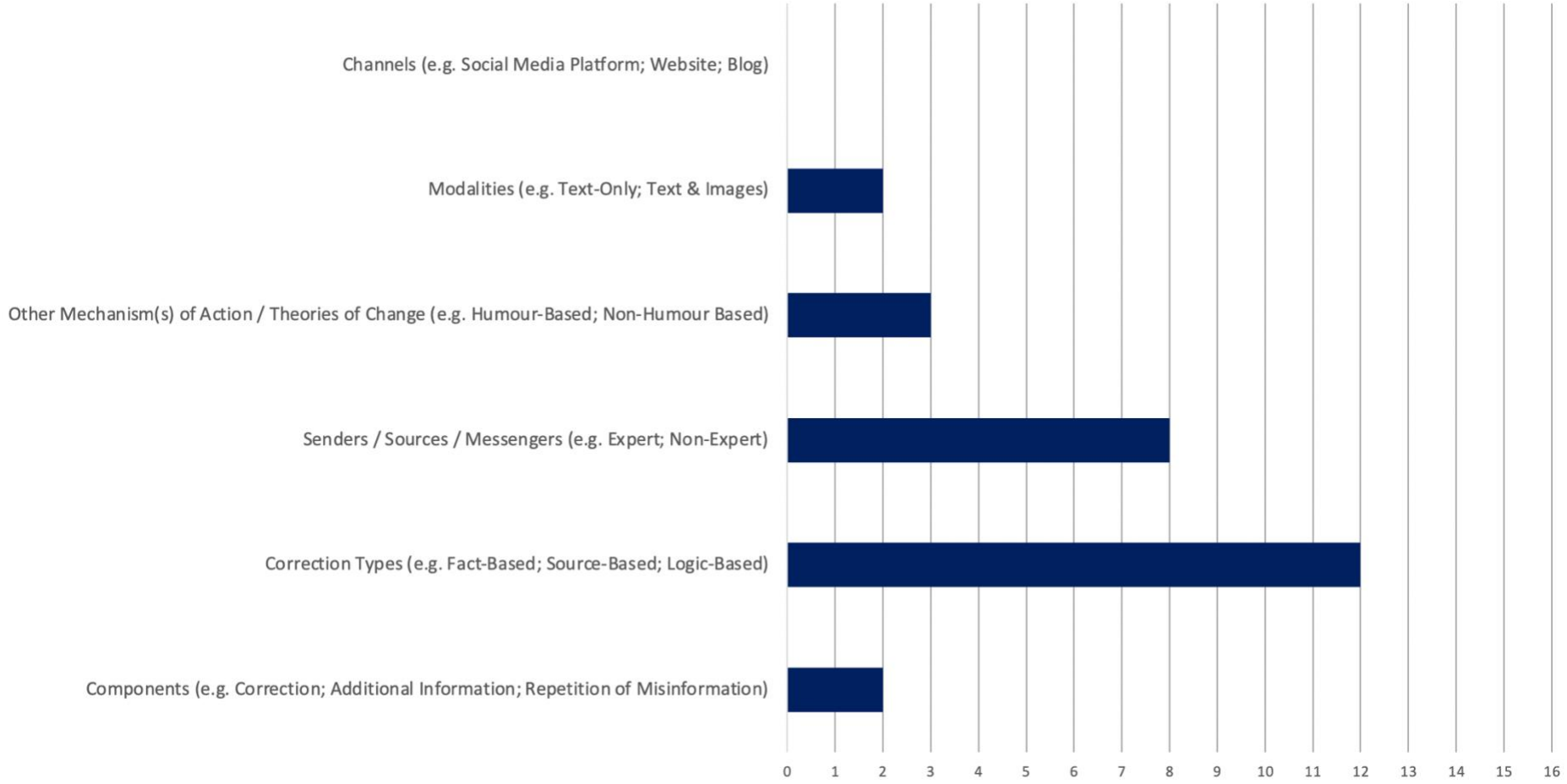


Figure 2. Included studies incorporating analysis of 'under what circumstances' debunking is likely to be effective: debunking message characteristics (N)



4.2 Summary of Main Findings of Included Studies

Main findings

The findings of 47 included randomised studies are equivocal about the effects of debunking messages – compared with 'information-only' messages and/or 'not responding' (control) – on people's ***vaccine-related misinformation beliefs*** and/or ***vaccine attitudes***.

Some individual studies found overall beneficial effects of debunking messages on such beliefs and/or attitudes (e.g. Amazeen 2022, Haglin 2017, Reavis 2017, Yousuf 2021). However, others found adverse effects of debunking on beliefs and/or attitudes (e.g. Pluviano 2019, Pluviano 2022 'Experiment 1', Pluviano 2022 'Experiment 2'); others found no effects on these outcomes (Bode 2015, Chockalingam 2021, Jolley 2017 'Study 1', Steffens 2021); and others found mixed and/or differential effects on these outcomes (Carey 2020, Gawronski 2022 'Experiment 1', Stojanov 2015, Sullivan 2019 'Experiment'). Two included studies of simple fact-check labels aimed at correcting vaccine-related misinformation on Twitter found beneficial (Lee 2022) and mixed (Zhang 2021) effects on misinformation beliefs and/or attitudes. With two exceptions (Nyhan 2015, Pluviano 2019), included studies measured these outcomes immediately after exposure to debunking messages (or comparators) only, with no further follow-up outcome measurement.

For differential effects, some included studies (e.g. Helfers 2022, Lee 2022, Lunz Trujillo 2021 'Large survey experiment') found that debunking messages might be *less* likely to have beneficial effects on beliefs and/or attitudes – and/or *more* likely to have adverse effects on these outcomes – among people who more *strongly* hold vaccine-related (and other) misinformation beliefs, and/or who are *more* resistant to vaccination. Conversely, debunking messages might be *more* likely to have beneficial effects on beliefs and/or attitudes – and/or *less* likely to have adverse effects on these outcomes – among people who more *weakly* hold (or do not hold) vaccine-related (and other) misinformation beliefs, and/or who are *less* resistant to vaccination. For this reason, some authors of included studies have proposed that the effectiveness of debunking messages is likely to be enhanced when they can be used in rapid response to vaccine-related misinformation claims, before people's beliefs in those claims can potentially become strengthened through repeated exposure (e.g. Bode 2015, Pluviano 2022 'Experiment 1', Pluviano 2022 'Experiment 2').

Whilst these findings confirm that debunking messages *can* reduce vaccine-related misinformation beliefs and/or improve vaccine attitudes among at least some people, most included studies that identified a beneficial impact of debunking on vaccine-related misinformation beliefs and/or attitudes did not find any corollary effects on people's ***intentions to be vaccinated*** (e.g. Helfers 2022, Jolley 2014 'Study 2', Nyhan 2014). More generally, the effects of debunking on ***vaccine hesitancy, resistance and/or intentions*** were inconsistent between included studies. For *vaccine uptake*, the only two included studies that measured the effects of debunking messages on uptake did not find any effect, compared with an 'information only' and a 'control condition' (Clayton 2021 'General parent study', Clayton 2021 'Non-compliance study').

Overall, there is no clear evidence of any differences in patterns of results between (i) those included studies that compared debunking with 'information only', (ii) those that compared debunking with 'not responding' (control), and (iii) those that incorporated both of these types of comparison: evidence for the effects of debunking appeared similarly equivocal among all three subgroups of studies. Furthermore, we did not identify any outcomes or differential effects that were clearly patterned between studies with different characteristics. In particular, whether participants assigned to debunking and/or comparator groups were exposed to a separate 'misinformation stimulus' (see '4.1. Characteristics of Included Studies', 'Comparisons and comparators') did not appear to have any discernible influence on patterns of results between studies.

However, assessments of this kind were hampered by key differences between these 47 randomised studies, including the research questions they address, and the outcomes and differential effects they assess (using various different measures – see '4.1. Characteristics of Included Studies').

Several authors of included studies highlighted that debunking messages are likely to be more effective if they can be co-produced by and/or piloted with people in target audiences, and then refined to incorporate the feedback (e.g. Clayton 2021 'General parent study', Clayton 2021 'Non-compliance study', Nyhan 2014, Yousef 2021). However, authors of some included studies (also) cautioned against placing undue emphasis on debunking, and/or highlighted the need to incorporate debunking into broader communications-based (and other) strategies aimed at promoting vaccine uptake (e.g. Helfers 2022, Nyhan 2015, Xiao 2021).

How these findings relate to the UK context

External validity examines the extent to which study findings can be generalised to other contexts. Ecological validity examines, more specifically, the extent to which study findings can be generalised to 'real-life' or 'real-world' settings. Assessing ecological validity entails consideration of how closely study participants, settings, procedures, stimulus materials and outcome measures reflect the behaviour(s) and/or context(s) we are aiming to generalise to. In this review, we are particularly interested in the (main and differential) effects of debunking vaccine-related misinformation on vaccine uptake (target behaviour) among people in the United Kingdom (target context and population).

We have not formally assessed the ecological validity of these 47 randomised controlled trials or other aspects of their applicability to the UK context (Guyatt 2011). However as only included UK participants (Pluviano 2017), and also that (with two exceptions: Clayton 2021 'General parent study', Clayton 2021 'Non-compliance study') these studies did not assess the impacts of debunking on vaccine uptake. There is also considerable variation between these studies in the extent to which their settings (e.g. online survey platforms), procedures and stimulus materials are likely to reflect the ways people in the UK experience exposure to vaccine-related misinformation and its correction in 'real world' digital and physical environments. We cannot assess the extent to which the wide variety of debunking messages (interventions) tested among included studies may differ from those used in practice by communicators across the UK.

5. Implications for Practice and Policy

The findings of this rapid review echo and reaffirm the prime challenge faced when deciding, if at all, to debunk vaccine-related (and other) misinformation: debunking messages are unlikely to influence people's beliefs, attitudes or intentions in the same way.

While debunking messages clearly *can* have a beneficial influence on vaccine-related misinformation beliefs and/or attitudes, among at least some people, such changes are not necessarily accompanied by positive changes in their intentions to be vaccinated. Also, the current lack of evidence from randomised studies concerning the effects of debunking messages on vaccine uptake means that there is currently high uncertainty about potential of communications aimed at debunking vaccine-related misinformation to help improve vaccination uptake rates.

Communications aimed at debunking vaccine-related misinformation should therefore be used judiciously and sparingly, as part of broader communications strategies and vaccine promotion campaigns that also go beyond refuting false or misleading claims (Bode 2021a, Sylvia Chou 2020) to target the broader sets of psychological antecedents and other determinants of vaccine hesitancy, acceptance and uptake (Betsch 2018, Crawshaw 2021, Freeman 2020, Kamal 2021, Lindholt 2021). This proposal is broadly consistent with existing policies and practices of many national and local public health policymakers and communicators in the UK, which emphasise responding to false narratives by communicating accurate information and facts, and reserving debunking strategies for potential use in a more limited set of circumstances (or not at all).

The evidence we have reviewed suggests that 'debunking' might be *less* likely to have beneficial effects on beliefs and/or attitudes – and/or *more* likely to have adverse effects on these outcomes – among people who more *strongly* hold vaccine-related (and other) misinformation beliefs and/or who are *more* resistant to vaccination. Although this finding is tentative and not necessarily generalisable to UK populations (and/or population subgroups), it resonates with well-rehearsed concerns about potential 'backfire effects' of debunking, which remain uncertain (Ecker 2017, Haglin 2017, Featherstone 2020, Swire-Thompson 2020). This is where beliefs in misinformation might unintentionally be strengthened by exposure to messages aimed at correcting or refuting false or misleading claims, among at least some people.

Conversely, debunking might also be *more* likely to have beneficial effects on these outcomes among people who more *weakly* hold (or do not hold) vaccine-related (and other) misinformation beliefs and/or who are *less* resistant to vaccination. If so, this could be important, because our engagement with UK policymakers and communicators suggested that vaccine hesitant people – more than resistant people – were the main target for communications-based countermeasures to vaccine-related misinformation during the COVID-19 vaccination programme. This targeting strategy was based on the premise that hesitant people would be more likely to be persuaded by such countermeasures to take up a COVID-19 vaccine than those who are resistant; and it aligns with the proposal that, rather than aiming to prevent potential 'backfire effects', communications aimed at correcting false or misleading claims need to be targeted to those people most likely to be positively influenced by them (Nyhan 2021).

Likewise, the findings of this review suggest that debunking messages may be more successful when they can be formulated, tailored and/or targeted to specific subgroups of people (Broniatowski 2021) and/or specific misinformation claims/ narratives, based on insights derived from:

- public health surveillance data analysis;
- (social) media analytics and active monitoring of health communications/ media channels;
- best practice toolkits and guidance (ECDPC 2021, GCS 2021, GCS-BST 2022, Lewandowsky 2020, Lewandowsky 2021, LGA 2021, UNICEF 2021);
- professional expertise, experience and judgement; and
- analysis and understanding of why different people hold and maintain specific vaccine-related misinformation beliefs (Pennycook 2021).

However, the evidence in this review does not provide any useful insights to elucidate how the effects of debunking (compared with 'information-only' and/or 'not responding') might differ between people belonging to different demographic and/or socio-economic UK population subgroups. Included studies (conducted mainly in the US or other European countries) more often focused on exploring the extent to

which people's psychological predispositions might influence the effectiveness of communications aiming to debunk vaccine-related misinformation.

Unlike demographic/ socio-economic data (e.g. current levels of vaccine uptake among different ethnicity groups), data on people's psychological predispositions are not generally available in public health surveillance systems to help inform the targeting of specific debunking messages to specific people. If policymakers and communicators had access to data enabling predictions of the likelihood that people with certain demographic characteristics tend to hold certain psychological traits (and vice-versa), this would in principle raise the possibility of using psychographic microtargeting (Gaysynsky 2022, Matz 2020), and related 'audience segmentation' techniques (Dhawan 2023, Evans 2019, Huckvale 2019), to reach specific population subgroups via social media platforms and other digital media channels. Debunking messages could also be tailored to address specific psychological dispositions, by filtering who views the messages using publicly available demographic and other data (Lunz Trujillo 2021).

In principle, microtargeting (and related techniques) may seem to offer an appealing technical response to directly counter the problem of algorithms driving misinformation on social media and other digital platforms, which are in turn shaped by platforms' data on preferences, cognitive styles and consumption patterns. However, the ethics of using these kinds of techniques (Guttman 2017, Kozyreva 2020, Martinez-Martin 2018, Matz 2020), people's resistance to being targeted by them, and current shifts in political, policy and industry contexts (e.g. removal of draft legislation from the UK Government's draft Online Safety Bill that would have placed duties on platforms to tackle legal but harmful content; recent changes in Twitter's policies on COVID-19 and other misinformation) are likely to curtail their use in practice.

The evidence in this review also suggests policymakers and communicators should continue with (or consider adopting) the practice of engaging trusted and/or influential people ('key messengers') to act as co-producers and senders of debunking messages (and other communications) aimed at promoting vaccine uptake. More generally, it suggests that policymakers and communicators should continue to draw on insights from behavioural science when formulating debunking messages aimed at promoting vaccine uptake (GCS-BST 2020).

Because the large majority of included studies measured outcomes immediately following exposure to debunking messages, it is not possible to draw any inferences from this body of evidence regarding the persistence, or longevity, of any beneficial or adverse effects. For the same reason, this body of evidence cannot speak to the issue of potential 'continued influence effects' of misinformation (Ecker 2010), whereby false or misleading claims that may be repeated as part of communications aimed at debunking them, may continue to influence some people's cognitions after exposure, even when debunking may have immediately reduced the strength of their misinformation beliefs. Similarly, the evidence in this review does not provide any insights into the effects of repeated exposure to debunking messages on target outcomes.

Drawing more specific practical insights from current research evidence about when, under what circumstances and for whom communications aimed at debunking vaccine-related misinformation are likely to be more or less effective, is hampered by key differences between included studies on various dimensions.

6. Implications for Research

The extent to which the evidence in this review can be used to inform the design and targeting of communications-based responses to vaccine-related misinformation in digital and physical environments is curtailed by some key research gaps that could be addressed by researchers when designing future randomised controlled trials on this topic in the UK and elsewhere.

First, there is limited evidence for the effects of debunking messages on people's vaccination uptake. This reflects longstanding concerns across the behavioural sciences about insufficient research attention to real-world contexts and actual changes in targeted behaviours (Clarke 2021, Maner 2016). To address this important gap, it may be feasible (with informed consent from participants) to link future studies to objective data sources, such as vaccination registers, or electronic health records, to either replace or supplement self-reported outcome data that could by collected by longer-term follow-up of participants. Also, similar issues elsewhere have been addressed through collaborative and innovative research, for example in generating large-scale and real-world interventional evidence concerning the use of face masks to reduce transmission of COVID-19 (Abaluck 2021), and (especially relevant to the focus of the current review) testing inoculation techniques against online manipulation on YouTube (Roozenbeek 2022b).

Second, while we conclude that debunking messages are *likely* to be most appropriate when they can be formulated, tailored and/or targeted to specific subgroups of people and/or specific misinformation narratives, this is principally derived from observational rather than experimental evidence. In other words, because debunking is observed to work more effectively for some population subgroups than others, it is plausible that formulating messages with such subgroups in mind is more effective. However, to know this with more certainty would require it to be tested directly, to examine if – in any given population subgroup – tailored messaging is more effective than generic messaging. Our review did not identify any studies designed to assess this, and this gap could be addressed in future studies.

Third, the current lack of standardisation across study designs, materials, and measures, make it is difficult to compare across, and draw conclusions from, this body of evidence from randomised studies. Research on debunking and other communications-based countermeasures aimed at tackling vaccine-related misinformation would benefit from development and use of an agreed standard core set of outcomes that should be measured and reported in future randomised controlled trials (Kirkham 2016), to help support future meta-analyses (and other syntheses) of cumulative evidence for intervention effectiveness. We have also highlighted the lack of consistency between studies in the degree to which participants were exposed to misinformation claims or narratives prior to their correction. While it seems improbable that such misinformation could be standardised across studies, the situation may be exacerbated by studies attempting to mimic real-world exposure to misinformation using a variety of artificial online surveys. Were future studies instead able to more closely mimic, or actually take place within, actual real-world contexts (such as social media websites) this would likely limit the scope of possible interventions and exposures, and so could plausibly reduce such heterogeneity.

Finally, the most important gap from a United Kingdom perspective is the lack of evidence from randomised studies of the effectiveness of debunking messages among UK populations. This severely limits the applicability of this body of evidence to UK policy and practice. In principle, this key gap could be addressed by conducting one or more new primary research studies involving UK participants. We therefore encourage preliminary conversations between UK stakeholders (policymakers, communicators, members of the public,

industry/ media, researchers, research funders) aimed at deciding whether reducing uncertainty around the question of when, if at all, to debunk vaccine-related misinformation, is judged sufficiently important to warrant investment in new UK-based primary research.

If there is sufficient demand among UK stakeholders for new primary research on this topic, then further engagement would be needed to specify the key design features of a new primary study, to make its findings as useful as possible to inform the judicious use of debunking by UK policymakers and communicators to help counter vaccine-related misinformation and promote vaccine uptake. The latter conversations can usefully be informed by the findings of this rapid review, which indicate that – along with design features suggested above in this section – any further UK-based randomised studies of the effectiveness and differential effects of debunking messages should also:

- Evaluate debunking messages (and comparators) in a real-world digital environment and context that closely maps on to the intended UK implementation context;
- Be embedded in sequential waves of a longitudinal representative UK population survey, to (a) better reflect the dynamic and evolving nature of misinformation narratives, (b) enable follow-up outcome measurement to investigate the persistence of any effects of debunking, and (c) provide a flexible model for future research that can be readily adapted and rapidly mobilised in response to future viral outbreaks and vaccination programmes;
- Include assessments of: (a) differential effects between key UK population subgroups, (b) the feasibility of tailoring and targeting debunking messages to UK population subgroups using available public health surveillance data, and (b) the value of tailoring debunking relative to a generic approach.

References

Included studies (n=47)

Amazeen 2022

Amazeen, M. A., Krishna, A., & Eschmann, R. (2022). Cutting the Bunk: Comparing the Solo and Aggregate Effects of Prebunking and Debunking Covid-19 Vaccine Misinformation. *Science Communication*, *44*(4), 387–417. https://doi.org/10.1177/10755470221111558

Bode 2015 'Study 1'

Bode, L., & Vraga, E. K. (2015). In Related News, That was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media. *Journal of Communication*, *65*(4), 619-638. https://doi.org/10.1111/jcom.12166

Burger 2022

Burger, M. N., Mayer, M., & Steimanis, I. (2022). Repeated information of benefits reduces COVID-19 vaccination hesitancy: Experimental evidence from Germany. *PLOS ONE*, *17*(6), e0270666. https://doi.org/10.1371/journal.pone.0270666

Carey 2020 'Yellow fever experiment'

Carey, J. M., Chi, V., Flynn, D. J., Nyhan, B., & Zeitzoff, T. (2020). The effects of corrective information about disease epidemics and outbreaks: Evidence from Zika and yellow fever in Brazil. *Science Advances*, *6*(5), eaaw7449. https://doi.org/10.1126/sciadv.aaw7449

Chockalingam 2021

Chockalingam, V., Wu, V., Berlinski, N., Chandra, Z., Hu, A., Jones, E., Kramer, J., Li Xiaoqiu, S., Monfre, T., Ng Yong, S., Sach, M., Smith-Lopez, M., Solomon, S., Sosanya, A., & Nyhan, B. (2021). The limited effects of partisan and consensus messaging in correcting science misperceptions. *Research & Politics*, *8*(2). https://doi.org/10.1177/20531680211014980

Clayton 2021 'General parent study'

Clayton, K., Finley, C., Flynn, D. J., Graves, M., & Nyhan, B. (2021). Evaluating the effects of vaccine messaging on immunization intentions and behavior: Evidence from two randomized controlled trials in Vermont. *Vaccine*, *39*(40), 5909-5917. https://doi.org/10.1016/j.vaccine.2021.08.047

Clayton 2021 'Non-compliance study'

Clayton, K., Finley, C., Flynn, D. J., Graves, M., & Nyhan, B. (2021). Evaluating the effects of vaccine messaging on immunization intentions and behavior: Evidence from two randomized controlled trials in Vermont. *Vaccine*, *39*(40), 5909-5917. https://doi.org/10.1016/j.vaccine.2021.08.047

Dixon 2013a

Dixon, G., & Clarke, C. E. (2013a). The effect of falsely balanced reporting of the autism-vaccine controversy on vaccine safety perceptions and behavioral intentions. *Health Education Research*, *28*(2), 352-359. https://doi.org/10.1093/her/cys110

Dixon 2013b
Dixon, G., & Clarke, C. E. (2013b). Heightening Uncertainty Around Certain Science. *Science Communication*, *35*(3), 358-382. https://doi.org/10.1177/1075547012458290

Dixon 2015
Dixon, G. N., McKeever, B. W., Holton, A. E., Clarke, C., & Eosco, G. (2015). The Power of a Picture: Overcoming Scientific Misinformation by Communicating Weight-of-Evidence Information with Visual Exemplars. *Journal of Communication*, *65*(4), 639-659. https://doi.org/10.1111/jcom.12159

Featherstone 2020
Featherstone, J. D., & Zhang, J. (2020). Feeling angry: the effects of vaccine misinformation and refutational messages on negative emotions and vaccination attitude. *Journal of Health Communication*, *25*(9), 692-702. https://dx.doi.org/10.1080/10810730.2020.1838671

Gawronski 2022 'Experiment 1'
Gawronski, B., Brannon, S. M., & Ng, N. L. (2022). Debunking Misinformation about a Causal Link between Vaccines and Autism: Two Preregistered Tests of Dual-Process versus Single-Process Predictions (With Conflicting Results). *Social Cognition*. https://doi.org/10.1521/soco.2022.40.6.580

Gawronski 2022 'Experiment 2'
Gawronski, B., Brannon, S. M., & Ng, N. L. (2022). Debunking Misinformation about a Causal Link between Vaccines and Autism: Two Preregistered Tests of Dual-Process versus Single-Process Predictions (With Conflicting Results). *Social Cognition*. https://doi.org/10.1521/soco.2022.40.6.580

Haglin 2017
Haglin, K. (2017). The limitations of the backfire effect. *Research & Politics*, *4*(3). https://doi.org/10.1177/2053168017716547

Helfers 2022
Helfers, A., & Ebersbach, M. (2022). The differential effects of a governmental debunking campaign concerning COVID-19 vaccination misinformation. *Journal of Communication in Healthcare*, 1-9. https://doi.org/10.1080/17538068.2022.2047497

Horne 2015
Horne, Z., Powell, D., Hummel, J. E., & Holyoak, K. J. (2015). Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences*, *112*(33), 10321-10324. https://doi.org/10.1073/pnas.1504019112

Jolley 2014 'Study 2'
Jolley, D., & Douglas, K. M. (2014). The Effects of Anti-Vaccine Conspiracy Theories on Vaccination Intentions. *PLOS ONE*, *9*(2), e89177. https://doi.org/10.1371/journal.pone.0089177

Jolley 2017 'Study 1'
Jolley, D., & Douglas, K. M. (2017). Prevention is better than cure: Addressing anti-vaccine conspiracy theories. *Journal of Applied Social Psychology*, *47*(8), 459-469. https://doi.org/10.1111/jasp.12453

Jolley 2017 'Study 2'
Jolley, D., & Douglas, K. M. (2017). Prevention is better than cure: Addressing anti-vaccine conspiracy theories. *Journal of Applied Social Psychology*, *47*(8), 459-469. https://doi.org/10.1111/jasp.12453

Kim 2021
Kim, S. C., Vraga, E. K., & Cook, J. (2021). An Eye Tracking Approach to Understanding Misinformation and Correction Strategies on Social Media: The Mediating Role of Attention and Credibility to Reduce HPV Vaccine Misperceptions. *Health Communication*, *36*(13), 1687-1696. https://dx.doi.org/10.1080/10410236.2020.1787933

Lee 2022
Lee, J., Kim, J. W., & Lee, H. Y. (2022). Unlocking Conspiracy Belief Systems: How Fact-Checking Label on Twitter Counters Conspiratorial MMR Vaccine Misinformation. *Health Communication*, 1-13. https://dx.doi.org/10.1080/10410236.2022.2031452

Lunz Trujillo 2021 'Large survey experiment'
Lunz Trujillo, K., Motta, M., Callaghan, T., & Sylvester, S. (2021). Correcting Misperceptions about the MMR Vaccine: Using Psychological Risk Factors to Inform Targeted Communication Strategies. *Political Research Quarterly*, *74*(2), 464-478. https://doi.org/10.1177/1065912920907695

Lyons 2019
Lyons, B., Merola, V., & Reifler, J. (2019). Not Just Asking Questions: Effects of Implicit and Explicit Conspiracy Information About Vaccines and Genetic Modification. *Health Communication*, *34*(14), 1741-1750. https://dx.doi.org/10.1080/10410236.2018.1530526

Nyhan 2014
Nyhan, B., Reifler, J., Richey, S., & Freed, G. L. (2014). Effective Messages in Vaccine Promotion: A Randomized Trial. *Pediatrics*, *133*(4), e835-e842. https://doi.org/10.1542/peds.2013-2365

Nyhan 2015
Nyhan, B., & Reifler, J. (2015). Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine*, *33*(3), 459-464. https://doi.org/10.1016/j.vaccine.2014.11.017

Pluviano 2017
Pluviano, S., Watt, C., & Della Sala, S. (2017). Misinformation lingers in memory: Failure of three pro-vaccination strategies. *PLOS ONE*, *12*(7), e0181640. https://doi.org/10.1371/journal.pone.0181640

Pluviano 2019
Pluviano, S., Watt, C., Ragazzini, G., & Della Sala, S. (2019). Parents' beliefs in misinformation about vaccines are strengthened by pro-vaccine campaigns. *Cognitive Processing*, *20*(3), 325-331. https://doi.org/10.1007/s10339-019-00919-w

Pluviano 2022 'Experiment 1'

Pluviano, S., Watt, C., Pompeia, S., Ekuni, R., & Della Sala, S. (2022). Forming and updating vaccination beliefs: does the continued effect of misinformation depend on what we think we know? *Cognitive Processing*, *23*(3), 367-378. https://dx.doi.org/10.1007/s10339-022-01093-2

Pluviano 2022 'Experiment 2'

Pluviano, S., Watt, C., Pompeia, S., Ekuni, R., & Della Sala, S. (2022). Forming and updating vaccination beliefs: does the continued effect of misinformation depend on what we think we know? *Cognitive Processing*, *23*(3), 367-378. https://dx.doi.org/10.1007/s10339-022-01093-2

Porter 2022

Porter, E., Velez, Y., & Wood, T. J. (2022). Factual Corrections Eliminate False Beliefs About COVID-19 Vaccines. *Public Opinion Quarterly*, *86*(3), 762–773. https://doi.org/10.1093/poq/nfac034

Reavis 2017 'Study 1'

Reavis, R. D., Ebbs, J. B., Onunkwo, A. K., & Sage, L. M. (2017). A self-affirmation exercise does not improve intentions to vaccinate among parents with negative vaccine attitudes (and may decrease intentions to vaccinate). *PLOS ONE*, *12*(7), e0181368. https://doi.org/10.1371/journal.pone.0181368

Schmid 2019 'Experiment 1'

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour*, *3*(9), 931-939. https://doi.org/10.1038/s41562-019-0632-4

Schmid 2019 'Experiment 2'

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour*, *3*(9), 931-939. https://doi.org/10.1038/s41562-019-0632-4

Schmid 2019 'Experiment 3'

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour*, *3*(9), 931-939. https://doi.org/10.1038/s41562-019-0632-4

Schmid 2019 'Experiment 4'

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour*, *3*(9), 931-939. https://doi.org/10.1038/s41562-019-0632-4

Schmid 2019 'Experiment 6'

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour*, *3*(9), 931-939. https://doi.org/10.1038/s41562-019-0632-4

Steffens 2021
Steffens, M. S., Dunn, A. G., Marques, M. D., Danchin, M., Witteman, H. O., & Leask, J. (2021). Addressing Myths and Vaccine Hesitancy: A Randomized Trial. *Pediatrics*, *148*(5), e2020049304. https://doi.org/10.1542/peds.2020-049304

Stojanov 2015
Stojanov, A. (2015). Reducing conspiracy theory beliefs. *Psihologija*, *48*(3), 251-266. https://doi.org/10.2298/psi1503251s

Sullivan 2019 'Experiment'
Sullivan, M. C. (2019). Leveraging Library Trust to Combat Misinformation on Social Media. *Library & Information Science Research*, *41*(1), 2-10. https://doi.org/10.1016/j.lisr.2019.02.004

Sullivan 2019 'Replication attempt'
Sullivan, M. C. (2019). Leveraging Library Trust to Combat Misinformation on Social Media. *Library & Information Science Research*, *41*(1), 2-10. https://doi.org/10.1016/j.lisr.2019.02.004

Sun 2022
Sun, Y., & Lu, F. (2022). How Misinformation and Rebuttals in Online Comments Affect People's Intention to Receive COVID-19 Vaccines: The Roles of Psychological Reactance and Misperceptions. *Journalism & Mass Communication Quarterly*. https://doi.org/10.1177/10776990221084606

Vraga 2019
Vraga, E. K., Kim, S. C., & Cook, J. (2019). Testing Logic-based and Humor-based Corrections for Science, Health, and Political Misinformation on Social Media. *Journal of Broadcasting & Electronic Media*, *63*(3), 393-414. https://doi.org/10.1080/08838151.2019.1653102

Vraga 2020 'Study 2'
Vraga, E. K., Bode, L., & Tully, M. (2020). Creating News Literacy Messages to Enhance Expert Corrections of Misinformation on Twitter. *Communication Research*, *49*(2), 245-267. https://doi.org/10.1177/0093650219898094

Xiao 2021
Xiao, X., & Su, Y. (2021). Integrating Reasoned Action Approach and Message Sidedness in the Era of Misinformation: The Case of HPV Vaccination Promotion. *Journal of Health Communication*, *26*(6), 371-380. https://dx.doi.org/10.1080/10810730.2021.1950873

Yousuf 2021
Yousuf, H., van der Linden, S., Bredius, L., Ted van Essen, G. A., Sweep, G., Preminger, Z., van Gorp, E., Scherder, E., Narula, J., & Hofstra, L. (2021). A media intervention applying debunking versus non-debunking content to combat vaccine misinformation in elderly in the Netherlands: A digital randomised trial. *EClinicalMedicine*, *35*, 100881. https://doi.org/10.1016/j.eclinm.2021.100881

Yue Dai 2022

Yue Dai, N., Wufan, J., Lunrui, F., Mengru, S., & Crystal Jiang, L. (2022). The Effects of Self-generated and Other-generated eWOM in Inoculating against Misinformation. *Telematics And Informatics*, *71*, 101835. https://doi.org/10.1016/j.tele.2022.101835

Zhang 2021

Zhang, J., Featherstone, J. D., Calabrese, C., & Wojcieszak, M. (2021). Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Preventive Medicine*, *145*, 106408. https://doi.org/10.1016/j.ypmed.2020.106408

Excluded studies (n=98)

A list of references of the 98 studies (articles) excluded from this review at the full text screening stage can be downloaded here.

Additional references

Abaluck, J., Kwong, L.H., Styczynski, A., Haque, A., Kabir, M.A., Bates-Jefferys, E., Crawford, E., Benjamin-Chung, J., Raihan, S., Rahman, S., Benhachmi, S., Bintee, N.Z., Winch, P.J., Hossain, M., Reza, H.M., Jaber, A.A., Momen, S.G., Rahman, A., Banti, F.L., Huq, T.S., Luby, S.P., & Mobarak, A.M. (2022). Impact of community masking on COVID-19: A cluster-randomized trial in Bangladesh. *Science*, *14* 375(6577), eabi9069. https://doi.org/10.1126/science.abi9069

Abba-Aji, M., Stuckler, D., Galea, S., & McKee, M. (2022). Ethnic/racial minorities' and migrants' access to COVID-19 vaccines: A systematic review of barriers and facilitators. *Journal of Migration and Health*, *5*, 100086. https://doi.org/10.1016/j.jmh.2022.100086

Andreas, M., Iannizzi, C., Bohndorf, E., Monsef, I., Piechotta, V., Meerpohl, J.J., & Skoetz, N. (2022). Interventions to increase COVID-19 vaccine uptake: a scoping review. *Cochrane Database of Systematic Reviews*, *8*, CD015270. https://doi.org/10.1002/14651858.CD015270

Betsch, C., Böhm, R., & Chapman, G.B. (2015). Using Behavioral Insights to Increase Vaccination Policy Effectiveness. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 61–73. DOI: https://doi.org/10.1177/2372732215600716

Betsch, C., Schmid P., Heinemeier, D., Korn, L., Holtmann, C., & Böhm, R. (2018). Beyond confidence: Development of a measure assessing the 5C psychological antecedents of vaccination. *PLoS One*, *13*(12), e0208601. https://doi.org/10.1371/journal.pone.0208601

Bierwiaczonek, K., Gundersen, A.B., & Kunst, J.R. (2022) The role of conspiracy beliefs for COVID-19 health responses: A meta-analysis. *Current Opinion in Psychology*, *46*, 101346. https://doi.org/10.1016/j.copsyc.2022.101346

Bode, L., & Vraga, E. (2021a). The Swiss cheese model for mitigating online misinformation. *Bulletin of the Atomic Scientists*, *77*(3), 129-133. https://doi.org/10.1080/00963402.2021.1912170

Bode, L., & Vraga, E. (2021b). Correction Experiences on Social Media During COVID-19. *Social Media and Society*, *7*(2). https://doi.org/10.1177/20563051211008829

Bontcheva, K., Posetti, J., Teyssou, D., Meyer, T., Gregory, S., Hanot, C., & Maynard, D. (2020). *Balancing Act: Countering Digital Disinformation while respecting Freedom of Expression*. Geneva: UNESCO, Broadband Commission for Sustainable Development. Retrieved December 05, 2022, from https://en.unesco.org/publications/balanceact

Brennen, J.S., Simon, F.M., Howard, P.N., & Nielsen, R.K. (2020). *Types, Sources, and Claims of COVID-19 Misinformation*. Oxford: Reuters Institute. Retrieved December 05, 2022, from https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation

Broniatowski, D.A., Dredze, M., & Ayers, J.W. (2021). "First Do No Harm": Effective Communication About COVID-19 Vaccines. *American Journal of Public Health*, *111*(6), 1055-1057. https://doi.org/10.2105/AJPH.2021.306288

Chan, M.S., Jones, C.R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation. *Psychological Science*, *28*(11), 1531–1546. https://doi.org/10.1177/0956797617714579

Chawla, D.S. (2022). Massive open index of scholarly papers launches. (News]). *Nature*, 24th January. https://doi.org/10.1038/d41586-022-00138-y

Clarke, N., Pechey, E., Kosīte, D., König, L., Mantzari, E., Blackwell, A.K.M., Marteau, T.M., & Hollands, G.J. (2021). Impact of health warning labels on selection and consumption of food and alcohol products: systematic review with meta-analysis. *Health Psychology Review*, *15*(3), 430–453. https://doi.org/10.1080/17437199.2020.1780147

Crawshaw, J., Konnyu, K., Castillo, G., van Allen, Z., Smith, M., Trehan, N., Gauvin, F-P., Grimshaw, J.M., & Presseau, J. (2021). *Factors affecting COVID-19 vaccination acceptance and uptake among the general public: a living behavioural science evidence synthesis (v5, Aug 31st, 2021)*. COVID-19 Evidence Network to Support Decision Making in Canada. Retrieved December 05, 2022, from https://www.mcmasterforum.org/docs/default-source/product-documents/living-evidence-syntheses/covid-19-living-evidence-synthesis-4.5---factors-affecting-covid-19-vaccination-acceptance-and-uptake-among-the-general-public.pdf

Caulfield, T. (2020). Does Debunking Work? Correcting COVID-19 Misinformation on Social Media. In: Flood, C.M., Macdonnell, V., Philpott, J., Thériault, S., & Venkatapuram, S. *Vulnerable: The Law, Policy and Ethics of COVID-19*. Ottawa: University of Ottawa Press. https://www.ualberta.ca/law/media-library/faculty-research/hli/media/images/caulfield-debunking-works-vulnerable-caulfield.pdf

Dhawan, S., & Hegelich, S. From Outside In: Profiling, Persuasion and Political Opinion in the Age of Big Data. In: Montag ,C., Baumeister, H. (eds). (2023). *Digital Phenotyping and Mobile Sensing. Studies in Neuroscience, Psychology and Behavioral Economics*. Springer, Cham. https://doi.org/10.1007/978-3-030-98546-2_10

do Nascimento, I.J.B., Pizarro, A.B., Almeida, J.M., Azzopardi-Muscat, N., Gonçalves, M.A., Björklunde, M., & Novillo-Ortiz, D. (2022). Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization*, *100*, 544–561. https://doi.org/10.2471/BLT.21.287654

Ecker, U.K.H., Lewandowsky, S., & Tang, D.T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory and Cognition*, *38*(8), 1087–2000. DOI: https://doi.org/10.3758/MC.38.8.1087

Ecker, U.K.H., Hogan, J.L., & Lewandowsky, S. (2017). Reminders and Repetition of Misinformation: Helping or Hindering Its Retraction? *Journal of Applied Research in Memory and Cognition*, *6*(2): 185–192. https://doi.org/10.1016/j.jarmac.2017.01.014

Edwards, L., Stoilova, M., Anstead, N., Fry, A., El-Halaby, G., & Smith, M. (2021) *Rapid Evidence Assessment on Online Misinformation and Media Literacy: Final Report for Ofcom*. London: London School of Economics. 2021. Retrieved December 05, 2022, from www.ofcom.org.uk

European Centre for Disease Prevention and Control (ECDC). (2021). *Countering online vaccine misinformation in the EU/EEA*. Stockholm: ECDC. https://doi.org/10.2900/329304

Evans, W.D., Thomas, C.N., Favatas, D., Smyser, J., & Briggs, J. (2019). Digital Segmentation of Priority Populations in Public Health. *Health Education and Behavior*, *46*(2S), 81–89. https://doi.org/10.1177/1090198119871246

Freeman, D., Loe, B.S., Chadwick, A., Vaccari, C., Waite, F., Rosebrock, L., Jenner, L., Petit, A., Lewandowsky, S., Vanderslott, S., Innocenti, S., Larkin, M., Giubilini, A., Yu, L.M., McShane, H., Pollard, A.J., & Lambe, S. (2020). COVID-19 vaccine hesitancy in the UK: the Oxford coronavirus explanations, attitudes, and narratives survey (Oceans) II. *Psychological Medicine*, *12*, 1–15. https://doi.org/10.1017/s0033291720005188

Gabarron, E., Oyeyemib, S.O., & Wynn, R. (2021). COVID-19-related misinformation on social media: a systematic review. *Bulletin of the World Health Organization*, *99*, 455–463A. https://doi.org/10.2471/BLT.20.276782

Gaysynsky, A., Heley, K., & Chou, W-Y.S. (2022). An Overview of Innovative Approaches to Support Timely and Agile Health Communication Research and Practice. *International Journal of Environmental Research and Public Health*, *19*(22), 15073. https://doi.org/10.3390/ijerph192215073

Giubilini, A. (2019). Vaccination Policies and the Principle of Least Restrictive Alternative: An Intervention Ladder. In: Giubilini, A. *The Ethics of Vaccination. Palgrave Studies in Ethics and Public Policy*. Palgrave Pivot, Cham. https://doi.org/10.1007/978-3-030-02068-2_3

Gradoń, K.T., Hołyst, J.A., Moy, W.R., Sienkiewicz, J., & Suchecki, K. (2021). Countering misinformation: A multidisciplinary approach. *Big Data & Society*, *8*(1). https://doi.org/10.1177/20539517211013848

Guttman, N. (2017). Ethical Issues in Health Promotion and Communication Interventions. *Oxford Research Encyclopaedia of Communication.* Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228613.013.118

Guyatt, G.H., Oxman, A.D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Falck-Ytter, Y., Jaeschke, R., Vist, G., Akl, E.A., Post, P.N., Norris, S., Meerpohl, J., Shukla, V.K., Nasser, M., & Schünemann, H.J. on behalf of the GRADE Working Group. (2011). GRADE guidelines: 8. Rating the quality of evidence--indirectness. *Journal of Clinical Epidemiology*, *64*(12), 1303-1310. https://doi.org/10.1016/j.jclinepi.2011.04.014

Helm, R.K., & Nasu, H. (2021). Regulatory Responses to 'Fake News' and Freedom of Expression: Normative and Empirical Evaluation. *Human Rights Law Review*, *21*(2), 302–332. https://doi.org/10.1093/hrlr/ngaa060

Huckvale, K., Venkatesh, S., & Christensen, H. (2019). Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *npj Digital Medicine*, *2*, 88. https://doi.org/10.1038/s41746-019-0166-1

Janmohamed, K., Walter, N., Nyhan, K., Khoshnood, K., Tucker, J.D., Sangngam, N., Altice, F.L., Ding, Q., Wong, A., Schwitzky, Z.M., Bauch, C.T., De Choudhury, M., Papakyriakopoulos, O., & Kumar, N. (2021). Interventions to Mitigate COVID-19 Misinformation: A Systematic Review and Meta-Analysis. Journal of *Health Communication*, *26*(12), 846–857. https://doi.org/10.1080/10810730.2021.2021460

Kafadar, A.H., Tekeli, G.G., Jones, K.A., Stephan, B., & Dening, T. (2022). Determinants for COVID-19 vaccine hesitancy in the general population: a systematic review of reviews. *Zeitschrift für Gesundheitswissenschaften*, *19*, 1–17. https://doi.org/10.1007/s10389-022-01753-9

Kamal, A., Hodson, A., & Pearce, J.M. (2021). A Rapid Systematic Review of Factors Influencing COVID-19 Vaccination Uptake in Minority Ethnic Groups in the UK. *Vaccines*, 9(10), 1121. https://doi.org/10.3390/vaccines9101121

Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E., Moher, D., Schmitt, J., Tugwell, P., Tunis, S., & Williamson, P.R. (2016). Core Outcome Set-STAndards for Reporting: The COS-STAR Statement. *PLoS Medicine*, *13*(10), e1002148. https://doi.org/10.1371/journal.pmed.1002148

Kim, S. C., Vraga, E. K., & Cook, J. (2021). An Eye Tracking Approach to Understanding Misinformation and Correction Strategies on Social Media: The Mediating Role of Attention and Credibility to Reduce HPV Vaccine Misperceptions. *Health Communication*, *36*(13), 1687–1696. https://dx.doi.org/10.1080/10410236.2020.1787933

Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools. *Psychological Science in the Public Interest*, *21*(3), 103–156. https://dx.doi.org/10.1177/1529100620946707

Krasodomski-Jones, A., Jones, E., Judson, E., & Miller, C. (2019) *Warring Songs: Information Operations in the Digital Age*. Demos. Retrieved December 05, 2022, from https://demos.co.uk/project/warring-songs-information-operations-in-the-digital-age

Lee, E-J., & Shin, S.Y. (2021). Debunking misinformation. In: Tumber, H., Waisbord, S. (eds). *The Routledge Companion to Media Disinformation and Populism*. London: Routledge. https://doi.org/10.4324/9781003004431

Lewandowsky, S., Ecker, U.K., Seifert, C.M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Lewandowsky, S., Cook, J., Ecker, U.K.H., Albarracín, D., Amazeen, M.A., Kendeou, P., Lombardi, D., Newman, E.J., Pennycook, G., Porter, E., Rand, D.G., Rapp, D.N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C.M., Sinatra, G.M., Swire-Thompson, B., van der Linden, S., Vraga, E.K., Wood, T.J., & Zaragoza, M.S. (2020). *The Debunking Handbook 2020*. https://doi.org/17910/b7.1182

Lewandowsky, S., Cook, J., Schmid, P., Holford, D.L., Finn, A., Leask, J., Thomson, A., Lombardi, D., Al-Rawi, A.K., Amazeen, M.A., Anderson, E.C., Armaos, K.D., Betsch, C., Bruns, H.H.B., Ecker, U.K.H., Gavaruzzi, T., Hahn, U., Herzog, S., Juanchich, M., Kendeou, P., Newman, E.J., Pennycook, G., Rapp, D.N., Sah, S., Sinatra, G.M., Tapper, K., & Vraga, E.K. *The COVID-19 Vaccine Communication Handbook. A practical guide for improving vaccine communication and fighting misinformation*. 2021. Retrieved December 05, 2022, from https://sks.to/c19vax

Lieneck, C., Heinemann, K., Patel, J., Huynh, H., Leafblad, A., Moreno, E., & Wingfield, C. (2022). Facilitators and Barriers of COVID-19 Vaccine Promotion on Social Media in the United States: A Systematic Review. *Healthcare*, *10*(2), 321. https://doi.org/10.3390/healthcare10020321

Lindholt, M.F., Jørgensen, F., Bor, A., & Petersen, M.B. (2021). Public acceptance of COVID-19 vaccines: cross-national evidence on levels and individual-level predictors using observational data. *BMJ Open*, *11*, e048172. https://doi.org/10.1136/ bmjopen-2020-048172

Lunz Trujillo, K., Motta, M., Callaghan, T., & Sylvester, S. (2021). Correcting Misperceptions about the MMR Vaccine: Using Psychological Risk Factors to Inform Targeted Communication Strategies. *Political Research Quarterly*, *74*(2), 464–478. https://doi.org/10.1177/1065912920907695

Maner, J.K. (2016). Into the wild: Field research can increase both replicability and real-world impact. *Journal of Experimental Social Psychology*, *66*, 100–106. https://doi.org/10.1016/j.jesp.2015.09.018

Marsden, C., Meyer, T., & Brown, I. (2020). Platform values and democratic elections: How can the law regulate digital disinformation? *Computer Law & Security Review*, 36, 105373. https://doi.org/10.1016/j.clsr.2019.105373

Martinez-Martin, N., Insel, T.R., Dagum, P., Greely, H.T., & Cho, M.K. (2018). Data mining for health: staking out the ethical territory of digital phenotyping. *npj Digital* Medicine, *1*, 68. https://doi.org/10.1038/s41746-018-0075-8

Matz, S.C., Appel, R.E., & Kosinski, M. (2019). Privacy in the age of psychological targeting. *Current Opinion in Psychology*, *31*, 116–121. https://doi.org/10.1016/j.copsyc.2019.08.010

Nyhan, B. (2021). Why the backfire effect does not explain the durability of political misperceptions. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(15), e1912440117. https://doi.org/10.1073/pnas.1912440117

O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., & Ananiadou, S. (2015). Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews 4*, 5. https://doi.org/10.1186/2046-4053-4-5

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch V.A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *British Medical* Journal, 372, n71. https://doi.org/10.1136/bmj.n71

Peng, W., Lim, S., & Meng, J. (2022). Persuasive strategies in online health misinformation: a systematic review, Information. *Communication & Society*. https://doi.org/10.1080/1369118X.2022.2085615

Pennycook, G., & Rand, D.G. (2021). The Psychology of Fake News. Trends in Cognitive Sciences, 25(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Pian, W., Chi, J., & Ma, F. (2021). The causes, impacts and countermeasures of COVID-19 "Infodemic": A systematic review using narrative synthesis. *Information Processing & Management, 58*(6), 102713. https://doi.org/10.1016/j.ipm.2021.102713

Ripp, T., & Röer, J.P. (2022). Systematic review on the association of COVID-19-related conspiracy belief with infection-preventive behavior and vaccination willingness. *BMC Psychology*, *10*, 66. https://doi.org/10.1186/s40359-022-00771-2

Rocha, Y.M., de Moura, G.A., Desidério, G.A., de Oliveira, C.H., Lourenço, F.D., & de Figueiredo, N.L.D. (2021). The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review. *Journal of Public Health*, *9*, 1–10. https://doi.org/10.1007/s10389-021-01658-z

Roozenbeek, J., & van der Linden, S. (2022a). How to Combat Health Misinformation: A Psychological Approach. *American Journal of Health Promotion*, *36*(3), 569–575. https://doi.org/10.1177/08901171211070958

Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022b). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, 26, 8(34), eabo6254. https://doi.org/10.1126/sciadv.abo6254

Shemilt, I., Hollands, G.J., Vigurs, C., Stansfield, C., & Thomas, J. (2022). *Misinformation in COVID-19: a living map of research evidence*. (Web Database). London: EPPI Centre, UCL Social Research Institute, University College London. https://eppi.ioe.ac.uk/eppi-vis/login/open?webdbid=232

Skafle, I., Nordahl-Hansen, A., Quintana, D.S., Wynn, R., & Gabarron, E. (2022). Misinformation About COVID-19 Vaccines on Social Media: Rapid Review. *Journal of Medical Internet Research*, *24*(8), e37367. https://doi.org/10.2196/37367

Smith, L.E., Amlôt, R., Weinman, J., Yiend, J., & Rubin, G.J. (2017). A systematic review of factors affecting vaccine uptake in young children. *Vaccine*, *35*(45), 6059–6069. https://doi.org/10.1016/j.vaccine.2017.09.046

Smith, R., Cubbon, S., & Wardle, C. (2020). *Under the surface: Covid-19 vaccine narratives, misinformation and data deficits on social media*. First Draft. Retrieved December 05, 2022, from https://firstdraftnews.org/vaccinenarratives-full-report-november-2020

Swire-Thompson, B., DeGutis, J., & Lazer, D. (2020). Searching for the backfire effect: Measurement and design considerations. *Journal of Applied Research in Memory and Cognition*, *9*(3), 286–299. https://doi.org/10.1016/j.jarmac.2020.06.006

Sylvia Chou, W.Y., Gaysynsky, A., & Cappella, J.N. (2020). Where We Go From Here: Health Misinformation on Social Media. *American Journal of Public Health*, *110*(S3), S273-S275. https://doi.org/10.2105/AJPH.2020.305905

Tandoc, E.C., Zheng, W.L., & Ling, R. (2017). Defining "Fake News". *Digital Journalism, 6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Thomas, J., Graziosi, S., Brunton, J., Ghouze, Z., O'Driscoll, P., & Bond, M. (2022). *EPPI-Reviewer: advanced software for systematic reviews, maps and other evidence synthesis*. (Software). London: EPPI Centre. https://eppi.ioe.ac.uk/CMS/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4

UK Government. (2022). *Online Safety Bill (Draft. 121 2022-23 – as amended in Public Bill Committee)*. London: House of Commons. Retrieved December 05, 2022, from https://bills.parliament.uk/bills/3137

UK Government Communication Service Behavioural Science Team (Cabinet Office) (GCS-BST). (2022). *The Wall of Beliefs: A toolkit for understanding false beliefs and developing effective counter-disinformation strategies.* London: UK Government Communication Service. Retrieved December 05, 2022, from https://gcs.civilservice.gov.uk/publications/the-wall-of-beliefs

UK Government Communication Service (GCS). (2021). *RESIST 2 Counter Disinformation Toolkit.* London: UK Government Communication Service. Retrieved December 05, 2022, from https://gcs.civilservice.gov.uk/publications/resist-2-counter-disinformation-toolkit

UK Local Government Association (LGA). (2021). *Applying behavioural insights to improve COVID vaccination uptake: a guide for councils*. UK Local Government Association. Retrieved December 05, 2022, from https://www.local.gov.uk/publications/applying-behavioural-insights-improve-covid-vaccination-uptake-guide-councils

UK Office of National Statistics (ONS). (2022). *Deaths involving COVID-19 by vaccination*

status, England: deaths occurring between 1 January 2021 and 31 May 2022. London: Office of National Statistics. Retrieved December 05, 2022, from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19byvaccinationstatusengland/deathsoccurringbetween1january2021and31may2022

UNICEF. *Countering Online Misinformation Resource Pack*. UNICEF Regional Office for Europe and Central Asia (CD4 Section). Retrieved December 05, 2022, from https://www.unicef.org/eca/media/13636/file

van der Linden, S., Roozenbeek, J., & Compton, J. (2020). Inoculating Against Fake News About COVID-19. *Frontiers in Psychology*, 11, 566790. https://doi.org/10.3389/fpsyg.2020.566790

van Mulukom, V., Pummerer, L.J., Alper, S., Bai, H., Čavojová, V., Farias, J., Kay, C.S., Lazarevic, L.B., Lobato, E.J.C., Marinthe, G., Banai, I.P., Šrol. J., & Žeželj, I. (2022). Antecedents and consequences of COVID-19 conspiracy beliefs: A systematic review. *Social Science & Medicine*, *301*, 114912. https://doi.org/10.1016/j.socscimed.2022.114912

Vraga, E.K., & Bode, L. (2020). Defining Misinformation and Understanding its Bounded Nature: Using Expertise and Evidence for Describing Misinformation. *Political Communication, 37*(1), 136–144. https://doi.org/10.1080/10584609.2020.1716500

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policy making*. Strasbourg: Council of Europe. Retrieved December 05, 2022, from https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html

Walter, N., & Murphy, S.T. (2018). How to unring the bell: A meta- analytic approach to correction of misinformation, *Communication Monographs*, *85*(3), 423–441. https://doi.org/10.1080/03637751.2018.1467564

Walter, N., Cohen, J., Holbert, R.L., & Morag, Y. (2020). Fact-Checking: A Meta-Analysis of What Works and for Whom. *Political Communication*, *37*(3), 350–375. https://doi.org/10.1080/10584609.2019.1668894

Walter, N., Brooks, J.J., Saucier, C.J., & Suresh, S. (2021). Evaluating the Impact of Attempts to Correct Health Misinformation on Social Media: A Meta-Analysis. *Health Communication*, *36*(13), 1776–1784. https://doi.org/10.1080/10410236.2020.1794553

World Health Organisation. (2020). *Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation. Joint statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC (23rd September 2020).* Geneva: World Health Organisation. Retrieved December 05, 2022, from https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation

International Public Policy Observatory (IPPO)
IPPO is an ESRC funded initiative to provide decision-makers in government at all levels with access to the best available global evidence on the social impacts of the COVID-19 pandemic, and the effectiveness of policy responses. IPPO is a collaboration between the Department of Science, Technology, Engineering and Public Policy (STEaPP) and the EPPI Centre at UCL; Cardiff University; Queen's University Belfast; the University of Auckland and the University of Oxford, together with think tanks including the International Network for Government Science (INGSA) and academic news publisher The Conversation.

EPPI Centre
Founded in 1996, the EPPI Centre is a specialist centre in the UCL Social Research Institute. It develops methods: (i) for the systematic reviewing and synthesis of research evidence; and (ii) for the study of the use research. As well as being directly involved in the academic study and the practice of research synthesis and research use, the centre provides accredited and short course training programmes in research synthesis and social policy and research.

UCL Social Research Institute (SRI)
The SRI (formerly the Department of Social Science) is one of the leading centres in the UK for multidisciplinary teaching and research in the social sciences. With more than 180 academic, research and professional staff, it works to advance knowledge and to inform policy in areas including gender, families, education, employment, migration, inequalities, health and child/adult wellbeing.

.....

First produced in 2022 by:

International Public Policy Observatory
Department of Science, Technology, Engineering and Public Policy (STEaPP)
Shropshire House (4th Floor)
11-20 Capper Street London

WC1E 6JA

https://covidandsociety.com

Twitter: @covidandsociety

ISBN: 978-1-911605-40-9