



**IPIE**  
International Panel on the  
Information Environment

# **Countermeasures for Mitigating Digital Misinformation**

A Systematic Review

*Synthesis Report 2023.1*



**IPIE**  
International Panel on the  
Information Environment

# **Countermeasures for Mitigating Digital Misinformation**

A Systematic Review

## SYNOPSIS

This *Synthesis Report* (2023.1) provides a formal systematic review of scientific literature on countermeasures for mitigating digital misinformation. 588 peer-reviewed global publications from many disciplines were the focus of this study in order to highlight the most effective countermeasures for mitigating potential effects of misinformation, disinformation, and a range of related phenomena.

According to the report's selected publications, the four most often endorsed countermeasures are corrective information materials, information and media literacy content, content moderation, and content labeling. More than 10% of the analyzed publications validated these countermeasures. Research reveals several patterns in the investigation of countermeasures for combating misinformation across disciplines. Social sciences emerge as a leading area of scholarship in exploring various strategies, with one exception: content moderation, which is more actively tested in publications from the physical sciences. Simultaneously, experiment-based methodologies highlight content labeling and content reporting as the most effective countermeasures. There is no substantial geographic variation in what researchers are finding.

Five important limitations in current research were identified:

1. Few publications test specific countermeasures with real-world data;
2. Some of the solutions offered in the literature are too broad to be tested;
3. Methods that are more likely to bring critical perspectives, such as interviews, focus groups, and discourse analysis, are used less often than quantitative methods;
4. Some countermeasures are understudied in particular disciplines. For example, redirection, or information and media literacy are understudied in the health and physical sciences;
5. The literature in English that is analyzed pays insufficient attention to the problem beyond a few Western countries.

Researchers from many disciplines agree there is a need for more focus, greater awareness of present and past scholarship, more experimental and quasi-experimental methods, and the development of research in global contexts. The connection between misinformation and real-life consequences is difficult to determine, and research on the effects of exposure to misinformation requires higher-quality data than technology firms currently provide. Moreover, studies on the efficacy of proposed countermeasures, especially on combinations of interventions by creators and consumers of online misinformation, require further investigation.

# CONTENTS

<b>SYNOPSIS .....</b>	<b>3</b>
<b>SECTION 1. INTRODUCTION .....</b>	<b>5</b>
OBJECTIVES AND RESEARCH QUESTIONS.....	6
METRICS AND PROTOCOLS.....	7
<b>SECTION 2. METHODS.....</b>	<b>11</b>
SCREENING: CODING ELIGIBILITY AND AVAILABILITY .....	13
INCLUSION: SYNTHESIS AND ANALYSIS.....	14
LIMITATIONS.....	15
<b>SECTION 3. STATE OF THE RESEARCH.....</b>	<b>17</b>
CURRENT KNOWLEDGE.....	17
GAPS IN SCHOLARSHIP .....	26
<b>SECTION 4. CONCLUSION .....</b>	<b>28</b>
<b>REFERENCES .....</b>	<b>30</b>
<b>SUPPLEMENTARY INFORMATION .....</b>	<b>34</b>
<b>ACKNOWLEDGMENTS .....</b>	<b>38</b>
CONTRIBUTORS .....	38
FUNDERS .....	38
DECLARATION OF INTERESTS .....	38
PREFERRED CITATION .....	38
COPYRIGHT INFORMATION .....	39
<b>ABOUT THE IPIE .....</b>	<b>40</b>

## SECTION 1. INTRODUCTION

The quality of information that individuals encounter online remains a source of critical contemporary concern. It is now well documented that digital platforms facilitate the spread of misinformation [1], [2]. This content is often identified using terms such as “misleading information,” “disinformation,” “fake news,” “rumors,” or “computational propaganda” [3]–[5]. In this *Synthesis Report* (2023.1), these phenomena will be referred to with the umbrella term “misinformation.” To combat the spread of digital misinformation, a great deal of work must be done to understand its implications, how to address them, and how to design the most effective countermeasures.

The rapidly expanding debates offer multiple solutions to the problem of misinformation across the domains of health [6], politics [7], and news media [8]. However, those studies that have a broader scope mainly focus on separate contexts rather than trying to generalize about design principles that could work across platforms, contexts, or types of content. Broad, systematic reviews are needed to map divergent solutions mitigating the spread of misinformation that are most likely to be transportable across platforms and countries, so that solutions can be implemented as soon as possible. This report offers a comprehensive cross-platform interdisciplinary systematic review of digital misinformation literature derived from the analysis of 4,798 peer-reviewed publications, 588 of which satisfy the publication inclusion criteria designed following the recommendations of the rigorous Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2009 [9].

## **Objectives and Research Questions**

There is no universally agreed-upon definition of misinformation. As a result, this investigation does not distinguish between various kinds of incorrect, false, or misleading information, or the intentions behind its spread. The literature often does not conclude intentions and, thus, does not distinguish whether a piece of information is spread with an intent to deceive, or not [10]. Such content can include rumors and conspiracies, inaccurate content or false beliefs [11], [12]. This report examines literature that focuses on “misinformation” and synonymous concepts such as “propaganda,” “disinformation,” “misleading information,” or “fake news” [1], [3], [4].

These multiple forms of misinformation are investigated by asking the following questions:

- How has research published in English-language, peer-reviewed journals covering countermeasures mitigating misinformation on digital platforms developed over time and across disciplines?
- Which countermeasures that can be implemented on digital platforms are most likely to be effective in mitigating the impact of misinformation, according to the latest scientific research?

This report systematically evaluates 588 peer-reviewed publications that discuss countermeasures across disciplines in order to synthesize existing knowledge and identify limitations and blind spots in the literature. While these studies propose some countermeasures for mitigating misinformation, less than one-fifth of them report the effectiveness of a countermeasure they propose.

This systematic approach has several advantages compared to recent reviews of misinformation [13], [14]. It includes the broadest range of publications possible, then analyzes each of them to determine the effects of countermeasures across

national and platform contexts. This results in one of the largest scholarly article datasets on digital misinformation. Broad search terms are used to select publications, to avoid focusing on a single aspect of misinformation such as “conspiracy theories” [14], a specific area of application such as marketing [15], or a specific platform.

### **Metrics and Protocols**

Scientists have produced many reviews that cover the causes and consequences of misinformation and have explained why misinformation is problematic, however, relatively few of those reviews follow the PRISMA recommendations [9]. PRISMA guidelines constitute an industry standard for systematic reviews and are recommended by the communities of scholars focused on designing protocols that are aimed at enabling high-quality information to be gathered from research, such as Cochrane. And even fewer—if any—studies are broad enough to capture evidence across disciplines, platforms, and contexts.

The studies that have approached the analysis of digital misinformation systematically have often focused on its taxonomies [15]–[17], factors promoting the diffusion of misinformation [18], or the analysis of the consequences of diffusion [6], [19]. Since this domain of research lacks consistent definitions, researchers have been reviewing and synthesizing the definition of misinformation, especially the various content types that propagate it [16]. Some systematic review studies have examined the impact of misinformation [6], [19] and have unpacked its causes [18]–[21] in order to find workable countermeasures [14]–[17]. These systematic reviews have found that social network analysis and content analysis are the most common approaches used to examine misinformation [17], [23]. In health misinformation research, experiments [23] that primarily focus on the accuracy of content and sentiment analysis are also common methods [17]. One systematic review argues that a good deal of the science of misinformation is atheoretical [24]. This review

relies on one of the largest databases of peer-reviewed literature that focuses on misinformation on digital platforms to analyze the evidence about the most effective platform countermeasures mitigating misinformation systematically.

One of the most common variables used to evaluate the success of a countermeasure is a platform users' perception of the information encountered. As one misinformation research review notes, individuals' judgments about the accuracy and credibility of misinformation have become "a central concern for both theory and practice" [25]. The same review suggests that the literature presents two major approaches to countering misinformation: first, information and media literacy interventions designed to equip users with a tool to combat misinformation; second, adding labels to content to trigger increased scrutiny by information consumers. Another review states that information and media literacy interventions help users to reduce misinformation sharing [13]. A report that summarizes countermeasures to combat COVID-19 misinformation suggests that providing consumers with accurate information about best practices in healthcare is an effective approach to reducing the believability of misinformation [21].

Studies that summarize the effects of helping individuals to judge accuracy and credibility of information find that not all remedies work well. For example, the effectiveness of content labeling with the aim of reducing the likelihood of users believing misinformation varied across different types of information and interventions [13], [25]. In a randomized field experiment, the specific countermeasure of exposing users to news credibility labels did not significantly decrease the consumption of low-quality news [26]. While some individual studies show the potential of content labels to reduce belief in misinformation [27], [28], other publications show a very limited effect of applicability for this type of intervention [29], [30]. These limited effects are often associated with specific categories of users, such as those supporting a particular political party, or with

certain types of labels, such as humorous content. Moreover, algorithmic detection techniques for identifying—necessary for labelling at scale—misinformation were found to be inefficient due to the deficiency of datasets [20].

Finally, many scholars suggest that the scarcity of research on platform interventions hinders the development and evaluation of countermeasures [13]. Furthermore, since the connection between misinformation and real-life consequences appears inconsistent and fragile [31], research on the effects of exposure to misinformation requires higher-quality data than technology firms currently provide [32]. Table 1 summarizes eleven common countermeasures that are proposed in this literature that can be implemented by digital platforms. This systematic review addresses the knowledge gap by identifying and analyzing two interventions that are likely to be effective in mitigating the impact of digital misinformation: content labeling and corrective information.

**Table 1. Strategies for Improving the Global Information Environment.**

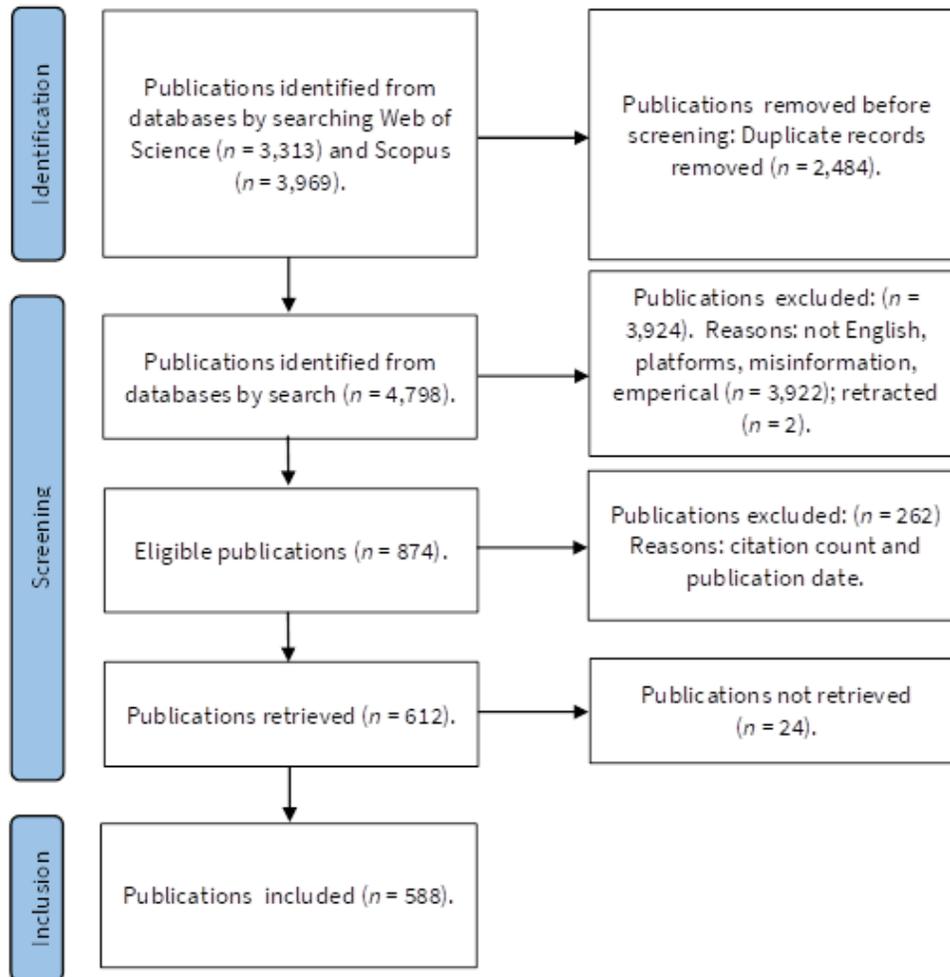
Countermeasures	Examples
Advertisement policy–Modifying the advertisement policy of the platform, which often adds a user-facing component to the advertising mechanisms.	Facebook requires the “Paid for by” label or introduces an information button for advertisements.
Content labeling–Labeling posts, accounts, and stories with tags about fact-checking, funding, or advertising, or any other forms of tagging or flagging, including providing further context without the user having to click through to receive the additional information.	A platform adds a “disputed” label to a user post, or a platform labels posts by state media with a “warning” sign.
Content or account moderation–Taking down or marking content; using human or algorithmic moderation to suspend and block accounts.	YouTube downranks content, or Twitter reduces interactions with accounts that users do not follow.
Content reporting–Changing how users report potential misinformation on a platform.	TikTok introduces a “misinformation” option in the content reporting options.
Content user sharing–Targeting the distribution of misleading content by users.	WhatsApp limits opportunities to forward a message, or Pinterest prevents pinning or saving posts.
Corrective information–An organization, platform, or individual provides accurate information without regard to whether users have preconceptions about it.	Governments or private enterprises publicly debunk a rumor on social media in a separate, unlinked piece of content, or user-generated content debunks a conspiracy.
Disclosure–Informing a user that they have come in contact, shared, or interacted with misinformation.	Reddit tells users they have interacted with misinformation.
Information & media literacy–Educating users to identify misinformation by giving them tips or suggestions or by training them.	Facebook offers Tips to Spot False News, including “be skeptical of headlines,” “look closely at the URL,” and “investigate the source.”
Redirection–Redirecting users to additional information, accounts, or posts, usually by taking users to different content or by overlaying accurate information and alerts.	Instagram shows content from local health authorities when users search for COVID-19 information, or Facebook and Twitter introduce election hubs before the election period.
Security or verification–Increasing or decreasing the security or verification requirements on a platform.	Twitter’s protection program for political officials.
Self-fact-checking–Providing users with an opportunity to fact-check information for themselves.	A platform offers users an opportunity to interact with fact-checkers to verify the information they consume using private messages.

**Source:** IPIE based on data collected; some examples are paraphrased from [12].

## SECTION 2. METHODS

This report is a comprehensive interdisciplinary systematic review that examined evidence in scientific literature, published in English, on countermeasures that combat misinformation and that can be implemented by digital platforms. It included articles from peer-reviewed scientific journals and covered the theme of misinformation on platforms. Figure 1 presents the flow of publications through the stages of systematic review.

**Figure 1. Sampling Stages of Systematic Review.**



Note: Flow of publications is presented based on a standard design suggested by the PRISMA recommendations [8].

Source: IPIE calculations based on data collected.

### **Identification: Search Strategies and Eligibility Criteria**

Two academic databases, Web of Science and Scopus, were used in this analysis because they offer a valid instrument for evaluating scholarly contributions in social science [33] and have been used in past systematic reviews and meta-analyses that asked similar questions [14], [25]. In addition, they incorporate key publications in social sciences, humanities, health, and other areas that publish relevant work, including journals that publish qualitative research. These databases also provide an interface that allows the extraction of large amounts of bibliographic information that were needed for this large-scope comprehensive study.

With a focus on studies of misinformation, the presence of one of the synonyms for “misinformation” in an article’s abstract, title, or keywords allowed for the identification of relevant articles (see Table S2 in Supplementary Information). This was done by including publications which used any of the umbrella terms of “misinformation” as discussed in the Introduction: “misinformation”, “misleading information”, “disinformation”, “propaganda”, “fake news” or “rumors” to describe the main research questions of the publication. Only those peer-reviewed studies that were published between 1st of November 2006 (after Facebook, the first modern social media platform, was made publicly available) and 31st of December 2022, were included.

Eligibility criteria (see details in Table S2):

1. English—Study available in English;
2. Misinformation—Study discusses any aspects of misinformation, disinformation, fake news, propaganda, or similar concepts;
3. Digital platforms—Mention of any digital platform as an object of study; a digital platform was defined as a website that is dependent on user-generated content and facilitates interactive and networked communication [34].

Studies offering solutions exclusively in relation to misinformation spreading on non-social media websites, such as those of mainstream media or political blogs were not included;

4. Empirical—Study is based on empirical evidence—information acquired by observation or experimentation that is analyzed in a scientific publication. This includes qualitative and quantitative research designs.

Several search terms combinations were tested, but the one chosen ensured the highest recall. In addition, the researchers deliberately refrained from including specific prominent platform names, including Twitter and YouTube (although did test this approach), as this would have skewed the search results by systematically overlooking lesser-known platforms. Gray literature, such as non-peer-reviewed conference abstracts or presentations, was not included because articles that have been through peer-review are more likely to provide rigorous findings, complete methodology information, include more subjects and more refined analysis, and transportable statistics on effects [35]. The final search request consisted of eight terms for “misinformation” and three terms for “social media.” See Table S1 for the full Boolean search string, which used identical terms for the two databases with a few modifications introduced to account for the differences in syntax requirements. The databases were last consulted on 28 February 2023.

### **Screening: Coding Eligibility and Availability**

Once selections from each database were finalized, the results were merged from the databases ( $N_{Scopus} = 3,313$ ,  $N_{WOS} = 3,969$ ). After eliminating duplicates, there were a total of 4,798 publications. A coding scheme was developed to assess the eligibility of collected publications (Table S2). These codes covered eligibility criteria, methods, proposed countermeasures, and their measurements. Five researchers ran a pilot test in pairs (four graduate students and a postdoctoral researcher) to achieve an optimum level of reliability in the proposed coding template. Intercoder reliability

showed high agreement based on a random sample of 163 publications for eligibility criteria (see Table S3 for Krippendorff's  $\alpha$  measures). Disagreements were discussed in the group and resolved [36].

The five coders read the titles, abstracts, and keywords of the publications to affirm the eligibility of a study based on the criteria listed above. Some publications contained all the search terms but did not consider misinformation as an object of study, were not empirical, or did not focus on digital platforms. 874 publications met the eligibility criteria. Due to the size of the resulting publication dataset and resource constraints, all publications were prioritized and reviewed that appeared online before July 2022 and had been cited more than 11 times since their publication. To account for newer studies that may have reached this citation threshold, all articles published between January 2020 and July 2022 were also reviewed. This allowed focus on the prominent and recent scientific publications (the average age of a publication in the final dataset was 2 years 7 months as of March 2023) and on those studies that reached the threshold of 11 citations. This mixed approach to study sampling is an accepted strategy in systematic analyses where quality of assessments takes precedence over extensive literature searches to prevent including trials of low methodological quality [35], [36]. 24 publications could not be retrieved, so they were excluded in line with PRISMA recommendations. This was because either the library of the University of Oxford that was used for access did not have a subscription to the relevant publisher, or a URL provided by a database was not active. This entire process resulted in the identification of 588 empirical publications about misinformation circulating on digital platforms, all of which were reviewed.

### **Inclusion: Synthesis and Analysis**

After coding for eligibility, the coding process switched to three sets of codes that summarized the literature: methods, countermeasures, and their measurement.

Existing typologies [13], [37] were used to develop a pilot coding instrument to classify countermeasures. The pilot coding process validated the instrument and helped to finalize it as a typology to classify publications according to the types of countermeasures they proposed. Following the pilot coding exercise, coding experiences were compared and the final coding template was adopted. Intercoder reliability measures were calculated based on 50 randomly selected publications (see Table S3 for Krippendorff's  $\alpha$  measures and agreement). The publications were randomly divided among the five researchers who coded them manually and independently. Pairs of researchers compared their article coding and resolved any discrepancies through conversations with the lead coder, as well as with the broader research team [36]. The team also separately synthesized evidence about the overall effectiveness of different countermeasures obtained through experiments and quasi-experimental designs. Please see *IPIE Synthesis Report (2023.2)* for the results of this analysis [38]. Risk of bias was assessed and reported for the studies included in *Synthesis Report (2023.2)* using the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) and revised tool for Risk of Bias in randomized trials (RoB 2) protocols. Some database manipulations were performed with the bibliometrix package for R [39].

### **Limitations**

It is possible that some studies were excluded from this systematic review even though an extensive search was undertaken that included two large social science databases and numerous search terms. More recently available databases, such as Microsoft Academic, could provide additional sources for an even more comprehensive analysis. It may have been possible to include additional sources and gray literature by using databases, for example, Google Scholar, to add more literature to reduce the chance of file-drawer bias occurring. However, it was critical that the sampling strategy privilege publications had gone through peer-review. With such a broad search strategy, a meaningful boundary had to be determined in sample

selection. Any update to this work should include publications in other formats, such as peer-reviewed conference reports and books.

Reviews of this type are limited by language, especially as the major databases primarily contain research published in English. This limits the scope of the evidence and skews it away from research published in places where English is not the main language of academic inquiry. Despite the presence of multiple countries in the focus of the reviewed publications, this review remains Western-centric, with scholarship originating in the USA dominating the dataset. Gaps in scholarship coverage are especially evident in relation to the Caribbean, Central and Eastern Europe, Central Asia, Latin America, and West Africa. Unfortunately, attempts to diversify a language scope in similar reviews often brings limited results [40]. Focusing on the literature in other languages is a natural next step in knowledge aggregation. Because the literature is rapidly expanding, it should soon be possible to test a greater variety of countermeasures as more publications provide empirical evidence about their effectiveness.

Similarly, more attention should be paid to the validity of the analyzed literature. The quality of the evidence reviewed varied from publication to publication even when a countermeasure was empirically tested. A discussion of the quality of the methodologies of relevant publications appears in *Synthesis Report (2023.2)*. Hence, the publications analyzed should be validated beyond reliance on the peer-review procedures.

## SECTION 3. STATE OF THE RESEARCH

In this section, the state of research on countermeasures to misinformation is reviewed by focusing on current knowledge and then discussing important gaps in scholarship.

### **Current Knowledge**

The scientific literature identifies several countermeasures to mitigate digital misinformation. Out of the 588 publications analyzed, 359— almost two-thirds of the total—proposed at least one countermeasure. To provide an overview, these solutions were grouped into 11 categories, as shown in Table 2. Four countermeasures were promoted by at least 10% of the publications: corrective information, information and media literacy, content or account moderation, and content labeling.

For instance, the advocates of an information literacy approach recommend that media organizations design and implement health education and user-empowerment campaigns [41]. Reportedly, this would help users to identify, process, share, and amplify high-quality information, thus correcting misinformation on social media. Some authors recommend content or account moderation in the form of algorithmic monitoring of text and images to identify suspicious accounts and content, though this intervention would require the development of such monitoring algorithms [42]. Another publication found that content moderation strategies deployed by the Chinese social media platform Sina Weibo after a crisis event failed to detectably decrease the proliferation of misinformation. However, the authors predict that this effect is masked because of the heavy censorship that Weibo applies through content moderation [43]. The examples of content labeling and corrective information countermeasures are discussed in SR2023.2 [38].

**Table 2. Proposed Countermeasures.**

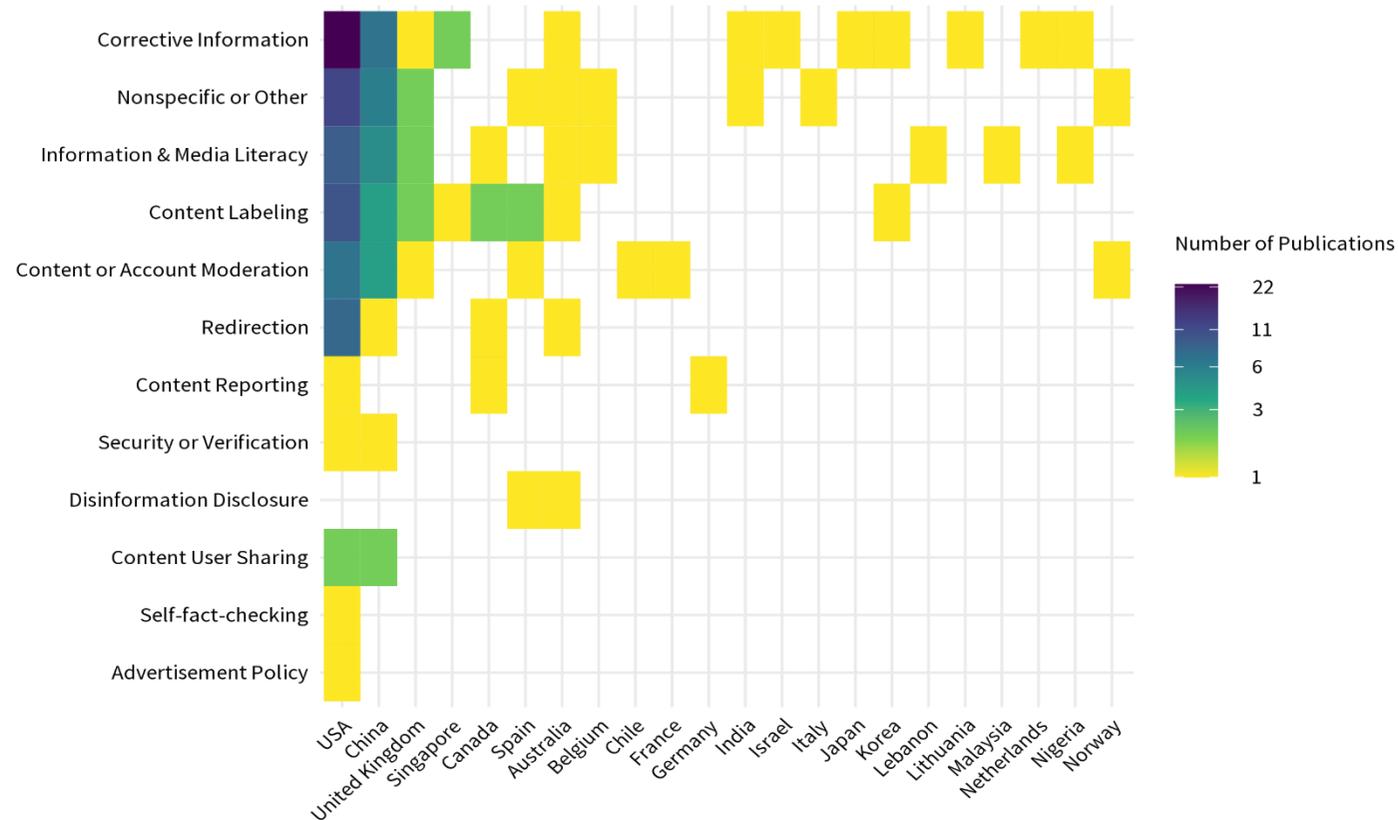
Countermeasures	Number of publications	Number of publications where countermeasures are linked to research	Number of publications where counter-measure effect reported
Corrective information	129	120	41
Information & media literacy	97	87	22
Content or account moderation	67	64	17
Content labeling	51	49	27
Redirection	24	23	12
Self-fact-checking	17	17	1
Content user sharing	16	16	4
Content reporting	6	6	3
Security or verification	5	5	2
Disinformation disclosure	4	4	2
Advertisement policy	3	3	1
Nonspecific or other	157	115	28

**Note:** This figure presents a subsample of the 588 publications—the publications that proposed any countermeasure ( $n = 359$ ). A publication was counted separately for each countermeasure it reported. Not every publication proposed a countermeasure, so some are excluded from this summary table.

**Source:** IPIE calculations based on data collected.

Countermeasures were also compared in relation to specific geographic contexts and found no substantial country-level divide between regions in relation to countermeasures offered and tested by scholars. Figure 2 provides a glimpse of what solutions are offered in which geographic contexts, based on the location of the corresponding author of a publication. Many countermeasures, such as content labeling and content moderation, are discussed around the world.

**Figure 2. Distribution of Countermeasures by Country of Author Affiliation.**



**Note:** This figure presents a subsample of the 588 publications—only those publications are included where the databases provided relevant information about the country affiliation of the first author whose publication proposed a countermeasure and reported on its effect ( $n = 97$ ). For the remaining 262 publications, full author affiliation information was not available. Countries are arranged by the number of studies for each category.

**Source:** IPIE calculations based on data collected.

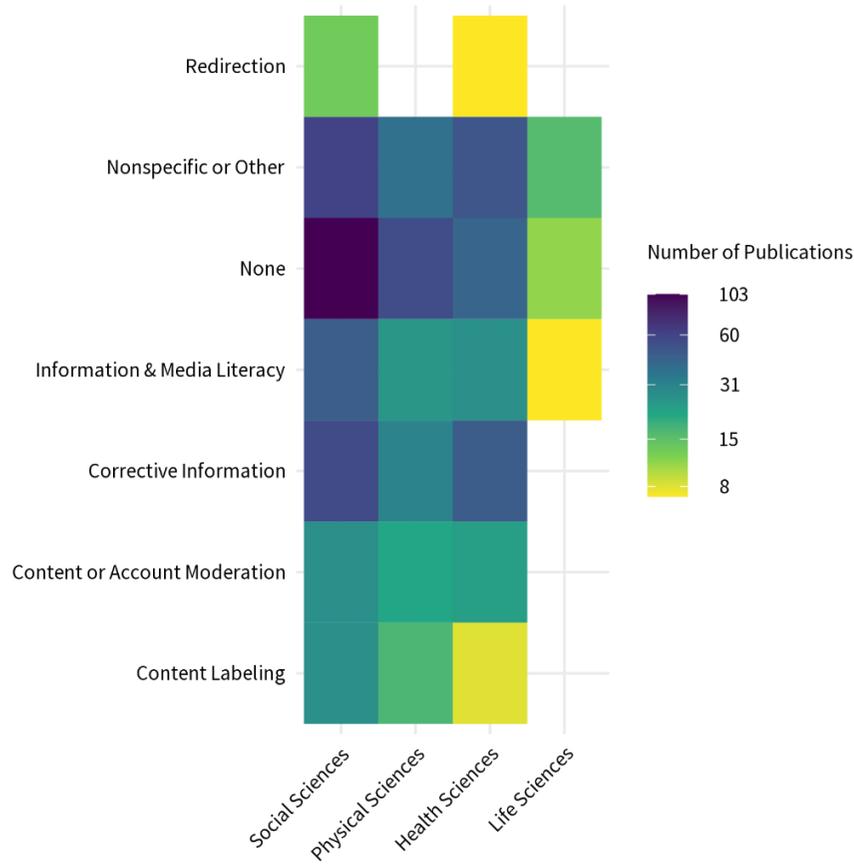
However, research from some countries emphasizes certain solutions more often than others. For example, the content user sharing countermeasure is slightly more often tested by scholars based in China and the USA. Disinformation disclosure is emphasized by researchers located in Spain, while scholars in Singapore pay more attention to content moderation, as the available data shows.

Figures 3 and 4 take a detailed look at what solutions for mitigating digital misinformation are offered across different disciplines. Most of the publications suggesting any solutions focus on three key areas: population health misinformation, with the cluster of COVID-19 and vaccination-related studies dominating this focus; the individual psychological aspects of misinformation, such as perception of information; and the political and social aspects of communication (Figure 4, a). The population health misinformation cluster expanded significantly largely due to academic interest in the COVID-19 pandemic.

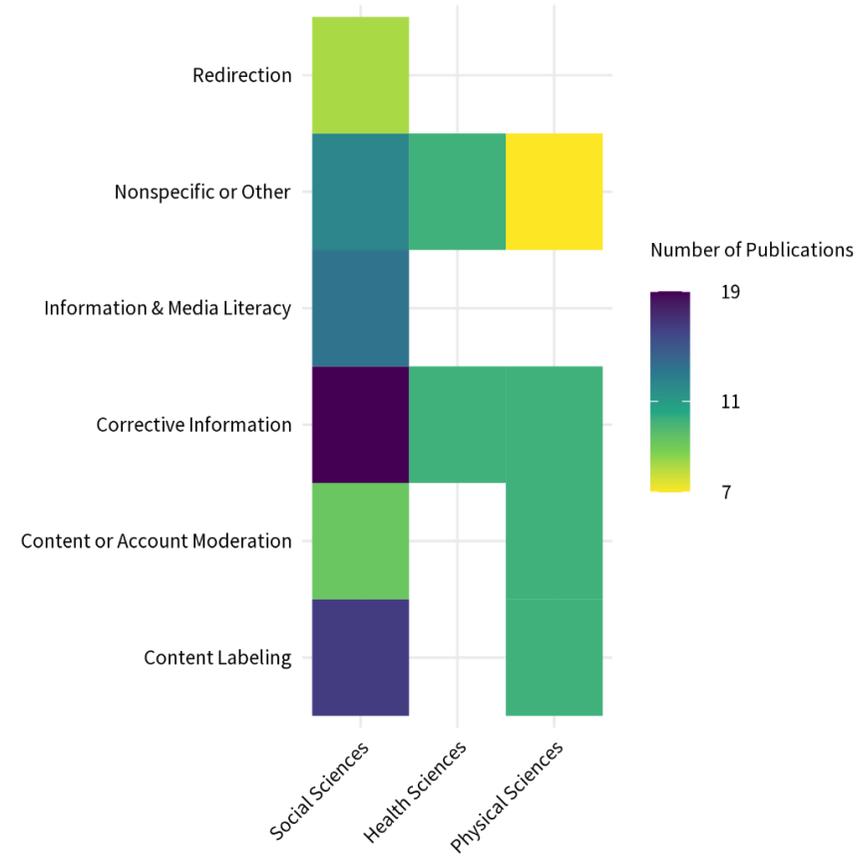
By comparing scholarship across different disciplines, the report finds that social sciences are at the forefront of countermeasure studies, apart from content moderation, a countermeasure more actively tested and offered in publications across physical sciences. At the same time, health science journals almost exclusively focus on corrective information as tested countermeasures (Figure 3, b). Countermeasures suggested in studies on politics-linked topics often favor content labeling, while psychology-linked research favors corrective information materials and information and media literacy content. Content moderation is looked at more carefully by studies focusing on population health, as well as the contexts of the USA, and China (Figure 4, b).

**Figure 3. Distribution of Countermeasures by Journal Field.**

a)



b)

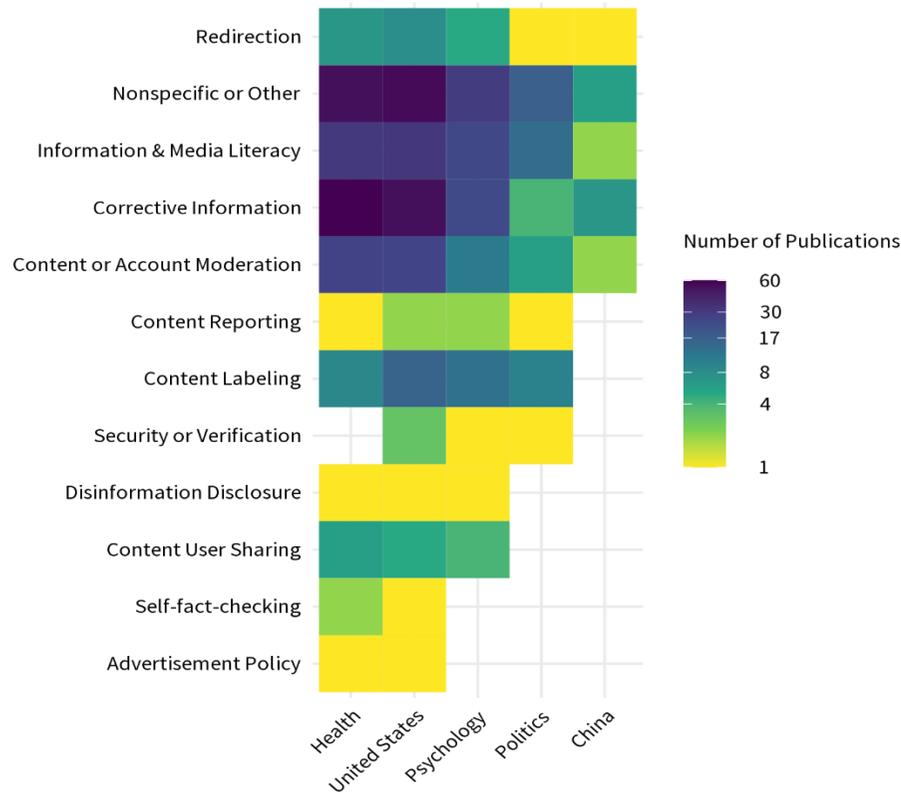


**Note:** This figure presents a subsample of the 588 publications—only those countermeasures that were reported at least seven times in one of the science categories are shown. Information about a journal field is taken from the Scopus database, and some journals had several field tags. In (b), only publications reporting on the effect of a countermeasure are included.

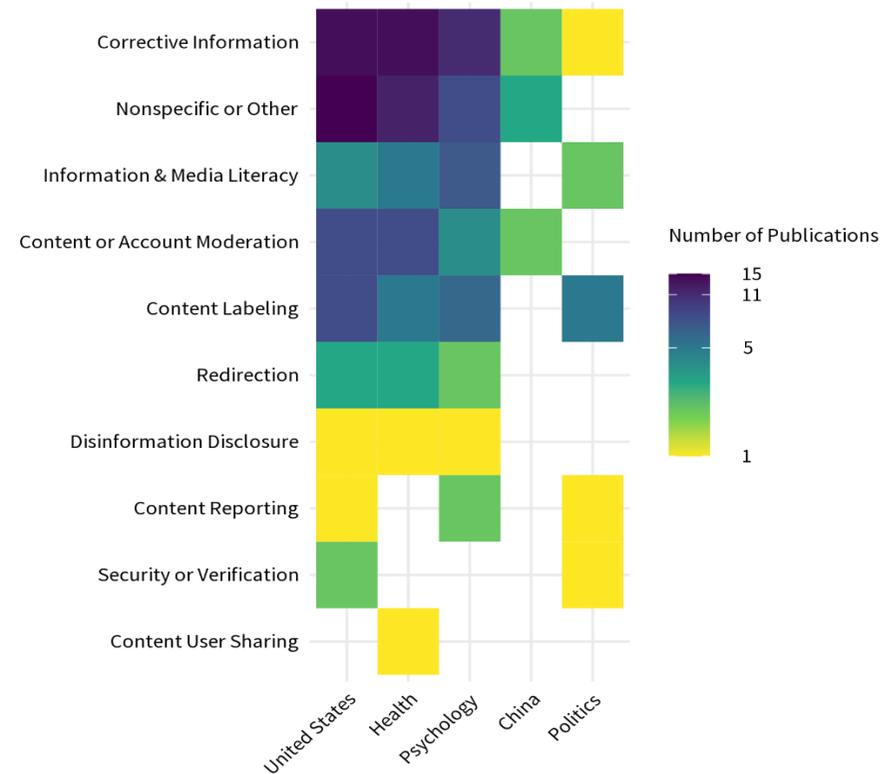
**Source:** IPIE calculations based on data collected.

**Figure 4. Distribution of Countermeasures by a Publication Focus Identified through Keywords.**

a)



b)



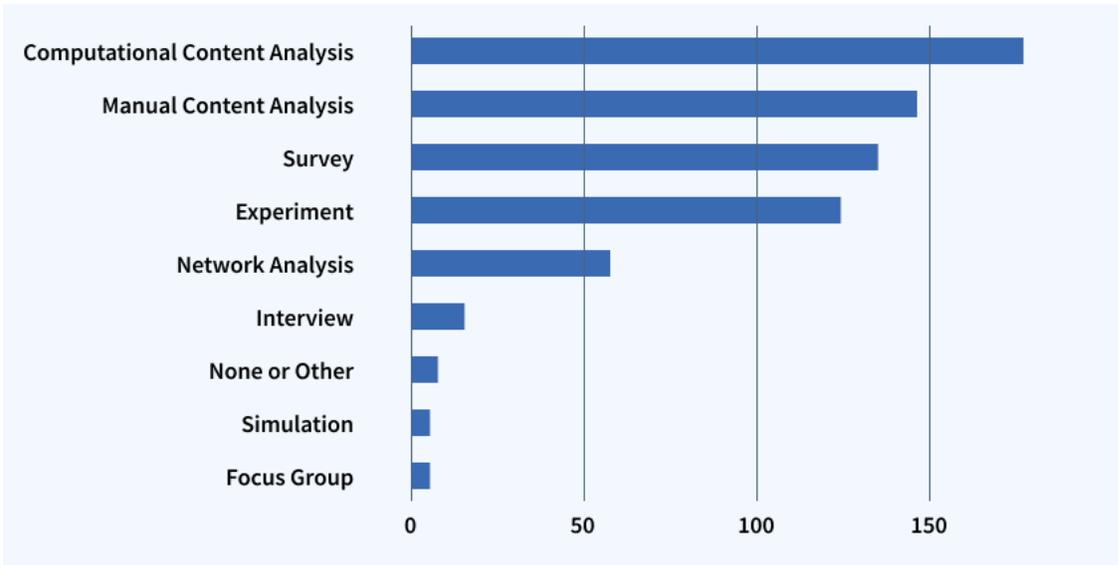
**Note:** These figure present subsamples of the 588 publications—the publication focus was identified by analyzing keywords provided by publication authors, journal editors, or Web of Science or Scopus; includes only those publications for which keyword data was available through Web of Science or Scopus: (a) publications proposing a countermeasure ( $n = 203$ ); (b) publications reporting on an effect of a countermeasure ( $n = 59$ ).

**Source:** IPIE calculations based on data collected.

This study also focused on how contemporary misinformation scholarship uses scientific methods, not least because a particular methodology can influence what exact countermeasures are proposed. The most common methods of data collection and analysis adopted in studies that linked the countermeasures they proposed to the object of study were computational and manual content analysis of social media data, surveys, and experiments (Figure 5). It is noted that previous systematic reviews had not highlighted a prominent role of survey research designs.

Publications relying on content analysis approaches tend to advocate remedies that are based on corrective information materials, while the authors of survey-based studies often promote information & media literacy remedies (Figure 6, a). At the same time, experiment-based methodologies feature content labeling or content reporting especially prominently. There is little alteration in this picture if focus shifts to only those reports that propose a countermeasure linked to the object of the study (Figure 6, b). However, the picture evolves if publications that report on the impact, effect, or consequences of a countermeasure they propose are included. Figure 6 (c) shows how proposals for labeling countermeasures are boosted by an interest in them by studies based on experimental designs, and how information & media literacy lose their appeal across an array of methodologies, with content moderation being largely supported by computational content analysis-based studies. This finding reminds us how important it is to scrutinize a solution from different perspectives and examine both the evidence and the method by which it was derived.

**Figure 5. Methods of Data Collection and Analysis.**

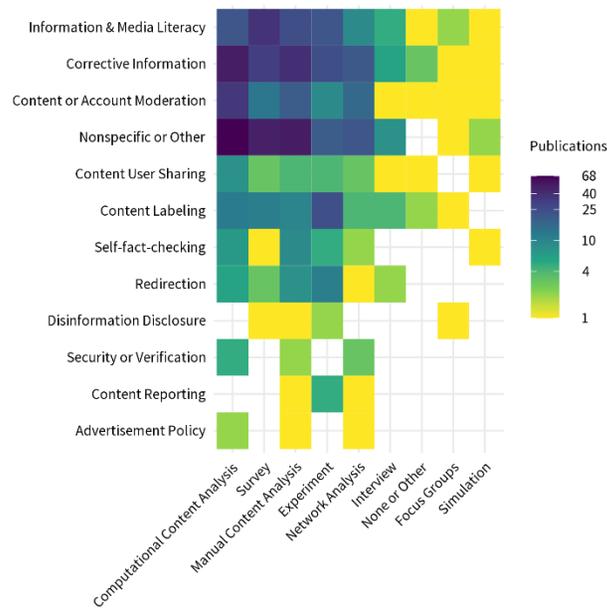


**Note:** This figure presents a subsample of the 588 publications—only those publications that linked reported countermeasures to the object of study ( $n = 307$ ). Some publications used several methods of data collection and analysis.

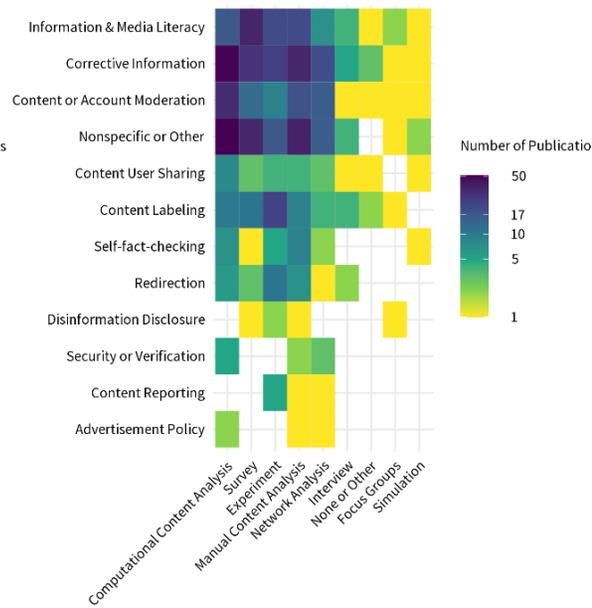
**Source:** IPIE calculations based on data collected.

**Figure 6. Methods of Data Collection and Analysis, by Proposed Countermeasures.**

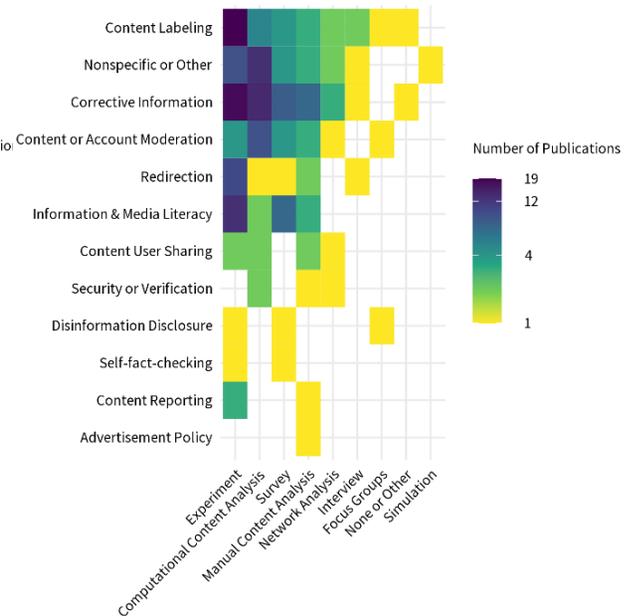
a)



b)



c)



**Note:** This figure shows the distribution of countermeasures over methods of data collection and analysis was counted separately whenever a publication reported several countermeasures or methods. The figure presents a subsample of the 588 publications: (a)  $n = 359$  reports on any countermeasure proposed, (b)  $n = 307$  reports on only those countermeasures that were linked to the object of study, and (c)  $n = 106$  reports on only those countermeasures that reported on the impact, effect, or consequences of a countermeasure.

**Source:** IPIE calculations based on data collected.

### **Gaps in Scholarship**

Four gaps in the scholarship appear through this systematic review. First, relatively few research publications test the specific countermeasures they propose using real-world data. While 52% of publications recommend a countermeasure that is linked to empirical research, only 18% of them report on the effectiveness of a countermeasure. This low percentage might be due to such barriers as ethical considerations, methodological limitations, research funding, or platform data availability. Indeed, it is increasingly hard to collect social media data: some platforms make it very hard to collect appropriate data, while others actively deter it. Identifying the unnecessary barriers to conducting research and working with regulators and platforms to remove those barriers, will greatly advance our knowledge of how to mitigate digital misinformation.

Those publications testing the effects of these mitigating strategies often focus on the link between misinformation and online behavior, such as news sharing and social media engagement. The rigorously tested countermeasures include situations that provide users with labeled content or corrective information.

Second, some of the solutions offered in the literature are too broad to be grouped in any category. Such nonspecific solutions are offered by 44% of the studies. They include “providing relevant information,” building “fake news awareness” [44], and promoting “a culture of humility aiming [to] demolish walls and barriers between tribes” [45]. This shows that much of the research does not offer or evaluate specific countermeasures for redressing digital misinformation.

Third, a gap is found in applying qualitative methodological approaches. Such methods as interviews, focus groups, and discourse analysis are being used less often than quantitative methods. Moreover, discourse analysis was relatively underutilized, thus its use is summarized under the “Other” category in the

analysis. Figure 3 shows that scholarship relying on qualitative methodologies is in the minority in this sample.

Fourth, some countermeasures seem to be understudied in certain disciplinary contexts. For example, information and media literacy, and redirection are widely discussed in humanities and social sciences but are particularly understudied across health sciences and physical sciences (Figure 3). However, once a focus is on keywords rather than a journal field (Figure 4), it is found that health-related and medical topics cover a broad range of countermeasures linked to studies that emerged during the pandemic.

Fifth, as shown in Section 2, the literature in English that is analyzed pays insufficient attention to many regions of the world.

## SECTION 4. CONCLUSION

This *Synthesis Report* (2023.1) provides a formal, systematic review of the literature on misinformation and the countermeasures it proposes to map the state of the field. The researchers systematically examined various types of countermeasures and how these countermeasures are validated in empirical research.

While most studies investigate the incidence of misinformation, many assess the effectiveness of countermeasures. Four countermeasures received the support of multiple studies (Table 3). These are corrective information, information and media literacy content, content moderation, and content labeling. However, this does not mean these countermeasures are the most effective ones. They are the most offered solutions but not necessarily the most tested. Some publications measure the effectiveness of the countermeasures they recommend, but most publications propose countermeasures without providing generalizable evidence for validation. Moreover, it is found that approximately one-fifth of the publications propose only nonspecific countermeasures to address the issue of digital misinformation. Policymakers should be aware of the implications of the diverse designs of the scholarship when looking for empirical evidence of the effectiveness of countermeasures. This research underscores the importance of a more strategic validation approach to better understand the effectiveness of the countermeasures in mitigating misinformation.

**Table 3. Strategies for Improving the Global Information Environment.**

Consensus	Countermeasures	Examples
Endorsed	Content labeling–Labeling posts, accounts, and stories with tags about fact-checking, funding, or advertising, or any other forms of tagging or flagging, including providing further context without the user having to click through to receive the additional information.	A platform adds a “disputed” label to a user post, or a platform labels posts by state media with a “warning” sign.
	Content or account moderation–Taking down or marking content; using human or algorithmic moderation to suspend and block accounts.	YouTube downranks content, or Twitter reduces interactions with accounts that users don’t follow.
	Corrective information–An organization, platform, or individual provides accurate information without regard to whether users have preconceptions about it.	Governments, private enterprises, or users publicly debunk a rumor or conspiracy on social media in a separate, unlinked piece of content.
	Information & media literacy–Educating users to identify misinformation by giving them tips or suggestions or by training them.	Facebook offers Tips to Spot False News, including “be skeptical of headlines,” “look closely at the URL,” and “investigate the source.”
Uncertain	Advertisement policy–Modifying the advertisement policy of the platform, which often adds a user-facing component to the advertising mechanisms.	Facebook requires the “Paid for by” label or introduces an information button for advertisements.
	Content reporting–Changing how users report potential misinformation on a platform.	TikTok introduces a “misinformation” option in the content reporting options.
	Content user sharing–Targeting the distribution of misleading content by users.	WhatsApp limits opportunities to forward a message, or Pinterest prevents pinning or saving posts.
	Disclosure–Informing a user that they have come in contact, shared, or interacted with misinformation.	Reddit tells users they have interacted with misinformation.
	Redirection–Redirecting users to additional information, accounts, or posts, usually by taking users to different content or by overlaying accurate information and alerts.	Instagram shows content from local health authorities when users search for COVID-19 information, or Facebook and Twitter introduce election hubs before the election period.
	Security or verification–Increasing or decreasing the security or verification requirements on a platform.	Twitter’s protection program for political officials.
	Self-fact-checking–Providing users with an opportunity to fact-check information for themselves.	A platform offers users an opportunity to interact with fact-checkers to verify the information they consume using private messages.

**Note:** Gray indicates interventions with uncertain effects; green indicates the strategies endorsed by the authors of studies that have been aggregated in this systematic review.

**Source:** IPIE, [13], [38].

## REFERENCES

- [1] C. Wittenberg and A. J. Berinsky, 'Misinformation and Its Correction', in *Social Media and Democracy: The State of the Field, Prospects for Reform*, Cambridge University Press, 2020, pp. 163–198. Accessed: Jul. 03, 2023. [Online]. Available: <https://www.cambridge.org/core/books/social-media-and-democracy/misinformation-and-its-correction/61FA7FD743784A723BA234533012E810>
- [2] M. Broussard, *More Than a Glitch: Confronting Race, Gender, and Ability Bias in Tech*. Cambridge, Massachusetts: MIT Press, 2023.
- [3] A. Herasimenka, J. Bright, A. Knuutila, and P. N. Howard, 'Misinformation and Professional News on Largely Unmoderated Platforms: The Case of Telegram', *J. Inf. Technol. Polit.*, vol. 20, no. 2, pp. 1–15, 2023, doi: 10.1080/19331681.2022.2076272.
- [4] P. N. Howard, *Lie machines: How to Save Democracy From Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. New Haven: London: Yale University Press, 2020.
- [5] A. Herasimenka, A. George, A. Lawson, A. Pavliuc, X. Wang, and P. N. Howard, 'Responding To Digital Misinformation: A Systematic Review and Meta-Analysis of Effective Countermeasures', presented at the International Communication Association, Toronto, Canada, May 2023.
- [6] Y. M. Rocha, G. A. de Moura, G. A. Desidério, C. H. de Oliveira, F. D. Lourenço, and L. D. de Figueiredo Nicolete, 'The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review', *J. Public Health*, Oct. 2021, doi: 10.1007/s10389-021-01658-z.
- [7] Y. M. Kim *et al.*, 'The Stealth Media? Groups and Targets behind Divisive Issue Campaigns on Facebook', *Polit. Commun.*, vol. 35, no. 4, pp. 515–541, Oct. 2018, doi: 10.1080/10584609.2018.1476425.
- [8] M. Tully, E. K. Vraga, and L. Bode, 'Designing and Testing News Literacy Messages for Social Media', *Mass Commun. Soc.*, vol. 23, no. 1, pp. 22–46, Jan. 2020, doi: 10.1080/15205436.2019.1604970.
- [9] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and The PRISMA Group, 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement', *PLOS Med.*, vol. 6, no. 7, p. e1000097, Jul. 2009, doi: 10.1371/journal.pmed.1000097.
- [10] S. Valenzuela, D. Halpern, and F. Araneda, 'A Downward Spiral? A Panel Study of Misinformation and Media Trust in Chile', *Int. J. Press.*, vol. 27, no. 2, pp. 353–373, Apr. 2022, doi: 10.1177/19401612211025238.
- [11] B. G. Southwell, E. A. Thorson, and L. Sheble, 'Introduction: Misinformation among Mass Audiences as a Focus for Inquiry', in *Misinformation and Mass Audiences*, University of Texas Press, 2018, pp. 1–12. doi: 10.7560/314555-002.

- [12] S. Valenzuela, D. Halpern, J. E. Katz, and J. P. Miranda, 'The Paradox of Participation Versus Misinformation: Social Media, Political Engagement, and the Spread of Misinformation', *Digit. Journal.*, vol. 7, no. 6, pp. 802–823, Jun. 2019, doi: 10.1080/21670811.2019.1623701.
- [13] L. Courchesne, J. Ilhardt, and J. N. Shapiro, 'Review of Social Science Research on the Impact of Countermeasures Against Influence Operations', *Harv. Kennedy Sch. Misinformation Rev.*, Sep. 2021, doi: 10.37016/mr-2020-79.
- [14] D. Mahl, M. S. Schäfer, and J. Zeng, 'Conspiracy Theories in Online Environments: An Interdisciplinary Literature Review and Agenda for Future Research', *New Media Soc.*, p. 14614448221075760, Feb. 2022, doi: 10.1177/14614448221075759.
- [15] G. D. Domenico, J. Sit, A. Ishizaka, and D. Nunan, 'Fake News, Social Media And Marketing: A Systematic Review', *J. Bus. Res.*, vol. 124, pp. 329–341, Jan. 2021, doi: 10.1016/j.jbusres.2020.11.037.
- [16] E. Kapantai, A. Christopoulou, C. Berberidis, and V. Peristeras, 'A Systematic Literature Review On Disinformation: Toward A Unified Taxonomical Framework', *New Media Soc.*, vol. 23, no. 5, pp. 1301–1326, May 2021, doi: 10.1177/1461444820959296.
- [17] V. Suarez-Lledo and J. Alvarez-Galvez, 'Prevalence of Health Misinformation on Social Media: Systematic Review', *J. Med. Internet Res.*, vol. 23, no. 1, p. e17187, Jan. 2021, doi: 10.2196/17187.
- [18] K. M. Douglas *et al.*, 'Understanding Conspiracy Theories', *Polit. Psychol.*, vol. 40, no. S1, pp. 3–35, 2019, doi: 10.1111/pops.12568.
- [19] R. Kabha, A. M. Kamel, M. Elbahi, A. M. D. Hafiz, and W. Dafri, 'Impact Of Fake News And Myths Related To COVID-19', *J. Content Community Commun.*, vol. 12, pp. 270–279, 2020, doi: 10.31620/JCCC.12.20/25.
- [20] I. Hassan, M. N. L. Azmi, and A. M. Abdullahi, 'Evaluating the Spread of Fake News and its Detection. Techniques on Social Networking Sites', *Romanian J. Commun. Public Relat.*, vol. 22, no. 1, pp. 111–125, Apr. 2020, doi: 10.21018/rjcpr.2020.1.289.
- [21] W. Pian, J. Chi, and F. Ma, 'The Causes, Impacts And Countermeasures Of Covid-19 "Infodemic": A Systematic Review Using Narrative Synthesis', *Inf. Process. Manag.*, vol. 58, no. 6, p. 102713, Nov. 2021, doi: 10.1016/j.ipm.2021.102713.
- [22] N. Walter, J. J. Brooks, C. J. Saucier, and S. Suresh, 'Evaluating the Impact of Attempts to Correct Health Misinformation on Social Media: A Meta-Analysis', *Health Commun.*, vol. 36, no. 13, pp. 1776–1784, Nov. 2021, doi: 10.1080/10410236.2020.1794553.
- [23] Y. Wang, M. McKee, A. Torbica, and D. Stuckler, 'Systematic Literature Review on the Spread of Health-related Misinformation on Social Media', *Soc. Sci. Med.*, vol. 240, p. 112552, Nov. 2019, doi: 10.1016/j.socscimed.2019.112552.
- [24] O. Abu Arqoub, A. Abdulateef Elegu, B. Efe Özad, H. Dwikat, and F. Adedamola Oloyede, 'Mapping the Scholarship of Fake News Research: A Systematic Review',

- Journal. Pract.*, vol. 16, no. 1, pp. 56–86, Jan. 2022, doi: 10.1080/17512786.2020.1805791.
- [25] K. Bryanov and V. Vziatysheva, ‘Determinants of Individuals’ Belief in Fake News: A Scoping Review Determinants of Belief in Fake News’, *PLoS ONE*, vol. 16, no. 6, p. e0253717, Jun. 2021, doi: 10.1371/journal.pone.0253717.
- [26] K. Aslett, A. M. Guess, R. Bonneau, J. Nagler, and J. A. Tucker, ‘News Credibility Labels Have Limited Average Effects on News Diet Quality and Fail to Reduce Misperceptions’, *Sci. Adv.*, vol. 8, no. 18, p. eabl3844, May 2022, doi: 10.1126/sciadv.abl3844.
- [27] K. Clayton *et al.*, ‘Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media’, *Polit. Behav.*, vol. 42, no. 4, pp. 1073–1095, Dec. 2020, doi: 10.1007/s11109-019-09533-0.
- [28] U. K. H. Ecker *et al.*, ‘The Psychological Drivers of Misinformation Belief and Its Resistance to Correction’, *Nat. Rev. Psychol.*, no. 1, pp. 13–29, Jan. 2022, doi: 10.1038/s44159-021-00006-y.
- [29] R. K. Garrett and S. Poulsen, ‘Flagging Facebook Falsehoods: Self-Identified Humor Warnings Outperform Fact Checker and Peer Warnings’, *J. Comput.-Mediat. Commun.*, vol. 24, no. 5, pp. 240–258, Sep. 2019, doi: 10.1093/jcmc/zmz012.
- [30] D. S. Morris, J. S. Morris, and P. L. Francia, ‘A Fake News Inoculation? Fact Checkers, Partisan Identification, and the Power of Misinformation’, *Polit. Groups Identities*, vol. 8, no. 5, pp. 986–1005, Oct. 2020, doi: 10.1080/21565503.2020.1803935.
- [31] F. Miró-Llinares and J. C. Aguerri, ‘Misinformation About Fake News: A Systematic Critical Review of Empirical Studies on the Phenomenon and Its Status as a “Threat”’, *Eur. J. Criminol.*, Apr. 2021, doi: 10.1177/1477370821994059.
- [32] J. A. Tucker *et al.*, ‘Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature’, William and Flora Hewlett Foundation, Mar. 2018. Accessed: Jun. 12, 2019. [Online]. Available: <https://dx.doi.org/10.2139/ssrn.3144139>
- [33] M. Norris and C. Oppenheim, ‘Comparing Alternatives to the Web of Science for Coverage of the Social Sciences’ Literature’, *J. Informetr.*, vol. 1, no. 2, pp. 161–169, Apr. 2007, doi: 10.1016/j.joi.2006.12.001.
- [34] A. Bechmann and S. Lomborg, ‘Mapping Actor Roles in Social Media: Different Perspectives on Value Creation in Theories of User Participation’, *New Media Soc.*, vol. 15, no. 5, pp. 765–781, Aug. 2013, doi: 10.1177/1461444812462853.
- [35] M. Egger, P. Juni, B. CJ, F. Holenstein, and J. Sterne, ‘How Important are Comprehensive Literature Searches and the Assessment of Trial Quality in Systematic Reviews? Empirical Study’, *Health Technol. Assess. Winch. Engl.*, vol. 7, pp. 1–76, Feb. 2003, doi: 10.3310/hta7010.
- [36] A. P. Siddaway, A. M. Wood, and L. V. Hedges, ‘How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and

- Meta-Syntheses’, *Annu. Rev. Psychol.*, vol. 70, pp. 747–770, Jan. 2019, doi: 10.1146/annurev-psych-010418-102803.
- [37] K. Yadav, ‘Platform Interventions: How Social Media Counters Influence Operations’, Carnegie Endowment for International Peace. Accessed: Jul. 25, 2022. [Online]. Available: <https://carnegieendowment.org/2021/01/25/platform-interventions-how-social-media-counters-influence-operations-pub-83698>
- [38] International Panel on the Information Environment, *Platform Responses to Misinformation: A Meta-Analysis of Data*, SR2023.2. Zurich, Switzerland: IPIE, 2023. [Online]. Available: [www.IPIE.info](http://www.IPIE.info)
- [39] M. Aria and C. Cuccurullo, ‘bibliometrix: An R-tool for comprehensive science mapping analysis’, *J. Informetr.*, vol. 11, no. 4, pp. 959–975, 2017.
- [40] UNESCO, ‘Evidence on the Gendered Impacts of Extended School Closures: a Systematic Review’, UNESCO, 2022. Accessed: Mar. 17, 2023. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000380935>
- [41] N. A. Atehortua and S. Patino, ‘COVID-19, a Tale of Two Pandemics: Novel Coronavirus and Fake News Messaging’, *Health Promot. Int.*, vol. 36, no. 2, pp. 524–534, Apr. 2021, doi: 10.1093/heapro/daaa140.
- [42] L. H. X. Ng and A. Taeihagh, ‘How does fake news spread? Understanding pathways of disinformation spread through APIs’, *Policy Internet*, vol. 13, no. 4, pp. 560–585, 2021, doi: 10.1002/poi3.268.
- [43] J. Zeng, C. Chan, and K. Fu, ‘How Social Media Construct “Truth” Around Crisis Events: Weibo’s Rumor Management Strategies After the 2015 Tianjin Blasts’, *Policy Internet*, vol. 9, no. 3, pp. 297–320, 2017, doi: 10.1002/poi3.155.
- [44] O. D. Apuke and B. Omar, ‘Fake News and COVID-19: Modelling the Predictors of Fake News Sharing Among Social Media Users’, *Telemat. Inform.*, vol. 56, p. 101475, Jan. 2021, doi: 10.1016/j.tele.2020.101475.
- [45] F. Zollo *et al.*, ‘Debunking in a World of Tribes’, *PLOS ONE*, vol. 12, no. 7, p. e0181821, Jul. 2017, doi: 10.1371/journal.pone.0181821.

## SUPPLEMENTARY INFORMATION

**Table S1. Boolean Search Term in the Web of Science 2021 Syntax.**

---

AB=(disinformation OR misinformation OR propaganda OR “fake news” OR rumo\* OR  
“misleading information” OR “false information” OR “computational propaganda”) OR  
TI=(disinformation OR misinformation OR propaganda OR “fake news” OR rumo\* OR  
“misleading information” OR “false information” OR “computational propaganda”) OR  
KP=(disinformation OR misinformation OR propaganda OR “fake news” OR rumo\* OR  
“misleading information” OR “false information” OR “computational propaganda”)

AND

(AB=(“social media” OR “social networking site” OR “digital platform\*”) OR TI=(“social media”  
OR “social networking site” OR “digital platform\*”) OR KP=(“social media” OR “social networking  
site” OR “digital platform\*”))

AND

(LA==(“ENGLISH”) AND DT==(“ARTICLE” OR “EARLY ACCESS”))

---

Source: IPIE.

**Table S2. Codes Used to Organize and Synthesize Research Evidence.**

Code	Description of the code	Variables
Eligibility		
English	Publication is available in English.	0 = no 1 = yes
Misinformation	Publication discusses any aspects of misinformation, disinformation, fake news, propaganda, rumors, “credibility” of information or digital/automated manipulation as a key object of study. If there are several objects, misinformation should be at least one out of three of them.	0 = no 1 = yes
Digital platforms	Social media platforms (websites that are dependent primarily on user-generated content that facilitates two-way interaction) constitute a significant focus of the research. The study names any social media platform as a major study object or its key background, field, or context. Forums are not considered as social media.	0 = no 1 = yes
Empirical	A study is based on empirical evidence that is analyzed in the publication, which should be clear from Abstract.	0 = no 1 = yes
Full paper coding		
Methods	Methods used to collect and analyze data. Multiple choice.	0 = none 1 = survey 2 = interviews 3 = focus groups 4 = ethnography 5 = experiment 6 = content analysis (manual) 7 = content analysis (automated) involving social media data extraction 8 = network analysis 9 = agent-based modeling, simulation 10 = process tracing and (or) case study 100 = other
Measures proposed	Countermeasures mitigating the impact of misinformation that are related to the functioning of platforms proposed in the study. The proposed measures should mitigate the impact of misinformation that spreads on digital platforms, according to this study. Multiple choice.	NA = nothing proposed 0 = broad 1 = advertisement policy 2 = labeling 3 = content or account moderation 4 = content reporting 5 = content distribution & sharing 6 = corrective information materials

		7 = disinformation disclosure 8 = information literacy 9 = redirection 10 = security or verification 11 = other 12 = fact-checking infrastructure
Measures proposed specify	Shortly describe specific countermeasures proposed to address the problem of misinformation.	[open coding]
Link between proposed countermeasure and study object(s)	Is there a link between the proposed countermeasures and the object of the study? The countermeasures should be linked to and derived from the research presented in the paper.	0 = no 1 = yes
Proposed countermeasures target any human misinformation creators	Do the proposed countermeasures actively target any human creators/sources of misinformation content?	0 = no 1 = yes
Measure effects	Does the study report on the impact/effect/consequences of countermeasures using quantitative measures?	0 = no 1 = yes
Measure effects specify	Please specify, if “yes” for the previous question.	[open coding]

**Notes:** Codes are ordered as they appeared in the codebook provided to coders. For replication purposes, a version with examples is available upon request. Table 1 presents the summary of definitions and examples used to code the Measures proposed variable.

**Source:** IPIE.

**Table S3. Inter-coder Reliability Assessment (Krippendorff's  $\alpha$  and Percentage Agreement).**

Category	$\alpha$	%
Eligibility		
Digital platforms	0.92	98.00
Empirical	0.88	93.90
English	1.00	100.00
Misinformation	0.76	87.80
Methods and countermeasures		
Link between proposed countermeasures and study object(s)	0.72	81.80
Methods	0.74	77.30
Proposed countermeasures target human misinformation creators	0.83	97.70

**Source:** IPIE calculations based on data collected.

## ACKNOWLEDGMENTS

### Contributors

Drafting Authors: Aliaksandr Herasimenka (Consulting Scientist, Belarus/UK), Sebastián Valenzuela (Chair of the Methodology Panel, Chile), Wendy Hui Kyong Chun (Vice Chair of the Methodology Panel, Canada), Young Mie Kim (Vice Chair of the Methodology Panel, South Korea/USA), Philip Howard (IPIE Chair, Canada/UK), Anna George (USA), Adrienn Lawson (Hungary), Alexandra Pavliuc (Canada), and Xianlingchen Wang (China). Independent methodology reviews: Shelley Boulianne (Canada/France), Simon Munzert (Germany), Xiaoli Nan (China/USA), Valentina Proust (Chile/USA), and Alex Wood (UK). Design, Formatting, and Layout: Domenico Di Donna, Raffaele Fulgente. Copyediting: Beverley Sykes. Independent fact-checking: Bill Ktepi. We gratefully acknowledge support from the IPIE Secretariat: Sheldon Himelfarb, Egerton Neto, Donna Seymour, and Alex Young.

### Funders

The International Panel on the Information Environment (IPIE) gratefully acknowledges the support of the Oxford Martin School, Ford Foundation, Conrad N. Hilton Foundation, Children's Investment Fund Foundation, Rockefeller Brothers Fund, Alfred P. Sloan Foundation, and Skoll Foundation. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the IPIE and do not necessarily reflect the views of the funders.

### Declaration of Interests

IPIE reports are developed and reviewed by a global network of research affiliates and consulting scientists who constitute focused Scientific Panels and contributor teams. All contributors and reviewers complete declarations of interests, which are reviewed by the IPIE at the appropriate stages of work.

### Preferred Citation

An IPIE *Summary for Policymakers* provides a high-level precis of the state of knowledge and is written for a broad audience. An IPIE *Synthesis Report* makes use of scientific meta-analysis techniques, systematic review, and other tools for evidence aggregation, knowledge generalization, and scientific consensus building, and is written for an expert audience. An IPIE *Technical Report* addresses particular questions of methodology, or provides a policy analysis on a focused regulatory problem. All reports are available on the IPIE website ([www.IPIE.info](http://www.IPIE.info)). This document should be cited as:

International Panel on the Information Environment, 2023. *Countermeasures for Mitigating Digital Misinformation: A Systematic Review*. SR2023.1. Zurich, Switzerland: IPIE.

### Copyright Information



This work is licensed under an Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).

## ABOUT THE IPIE

The International Panel on the Information Environment (IPIE) is an independent and global science organization committed to providing the most actionable scientific knowledge about threats to the world's information environment. Based in Switzerland, the mission of the IPIE is to provide policymakers, industry, and civil society with independent scientific assessments on the global information environment by organizing, evaluating, and elevating research, with the broad aim of improving the global information environment. Hundreds of researchers from around the world contribute to the IPIE's reports.

For more information, please contact the International Panel on the Information Environment (IPIE), [secretariat@IPIE.info](mailto:secretariat@IPIE.info). Seefeldstrasse 123, P.O. Box, 8034 Zurich, Switzerland.



International Panel on  
the Information  
Environment

Seefeldstrasse 123,  
P.O. Box, 8034 Zurich,  
Switzerland

ISBN 978-1-9168600-0-1



9 781916 860001 >

