



Delft University of Technology

A Comprehensive Review on Countering Rumours in the Age of Online Social Media Platforms

Ebrahimi Fard, Amir; Verma, Trivik

DOI

[10.1007/978-981-16-5268-4_11](https://doi.org/10.1007/978-981-16-5268-4_11)

Publication date

2022

Document Version

Final published version

Published in

Causes and Symptoms of Socio-Cultural Polarization

Citation (APA)

Ebrahimi Fard, A., & Verma, T. (2022). A Comprehensive Review on Countering Rumours in the Age of Online Social Media Platforms. In *Causes and Symptoms of Socio-Cultural Polarization: Role of Information and Communication Technologies* (pp. 253-284). Springer Nature. https://doi.org/10.1007/978-981-16-5268-4_11

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

A Comprehensive Review on Countering Rumours in the Age of Online Social Media Platforms



Amir Ebrahimi Fard and Trivik Verma

Abstract The power of rumour spreading in the age of online social media is intimidating. It can incite to insurrection, denigrate people, and damage financial markets, proving catastrophic for society. Despite widespread scholarly research and practice of developing a constellation of counter-rumour strategies, the massive waves of rumours are still sweeping over individuals, organisations, and societal institutions. To systematically tackle this issue, we present a comprehensive review and an epidemic framework to resolve three challenging aspects of rumour dissemination in online social media. First, we identify and explain the various forms of false and unverified information, relevance, and impact. Second, we address how social media can exacerbate the phenomenon of rumour spreading. Using the framework, the classification of rumour disseminating mechanisms on social media, allows us to develop counter-rumour strategies. Finally, we inspect past strategies employed in addressing rumour dissemination and use the framework to explore parallels between epidemic management and addressing rumour. We identify the highly neglected aspects of the current cumulative rumour response and factors that may be effectively targeted in the future. Our approach might support understanding social media's role in propagating rumours and devising active measures in quelling this epidemic.

Keywords Rumour · Unverified information · Disinformation · Propaganda · Conspiracy theory

1 Introduction

In one of the most famous plays – Henry the Fourth, Part II – Shakespeare writes “rumour is a pipe, blown by surmises, jealousies, conjectures, and of so easy and

A. E. Fard (✉) · T. Verma

Faculty of Technology, Policy and Management, Delft University of Technology, Delft, the Netherlands

e-mail: a.ebrahimifard@tudelft.nl; T.Verma@tudelft.nl

so plain a stop, that the blunt monster with uncounted heads, the still-discordant wavering multitude, can play upon it". His words elegantly express how easy, widespread, and vicious the emergence and circulation of rumours can be. Since he first wrote the play, the phenomenon of rumour spreading is exacerbated and turned into a far-reaching phenomenon to the extent that the World Economic Forum ranked the spread of misinformation as one of the most prominent risks to democracy (Howell et al. 2013).

Although rumour spreading is mostly associated with political contexts due to the excessive use of rumours by political figures to disparage their rivals and critics, this phenomenon's scope is much bigger than politics (Koenig 1985; Allport and Postman 1947; Knopf 1975; DiFonzo 2009). It is a domain-agnostic phenomenon that arises in any circumstance in which meanings are uncertain, questions are unsettled, information is missing, and lines of communications are absent (DiFonzo 2009). From the content perspective, rumours are false or unverified statements about instrumentally important topics. Thus any incident – no matter if it is political or not – could be a rumour-mongering subject. People engage in the rumour process since it attributes a ready-made justification to unexplained events (DiFonzo and Bordia 2007a). It increases the comprehension and understanding of the situation by offering detailed reasoning and clarification. Rumours might also be initiated deliberately as a psychological tool for strategic purposes such as character assassination, influence operations, and financial benefits (DiFonzo and Bordia 2007a, b; Koenig 1985). The dissemination of rumours, whether being intentional or inadvertent, may feed on hate, create fear, and raise false hopes (Knapp 1944). It may tarnish reputation of individuals (Allport and Postman 1947), organisations (Koenig 1985), or even countries (Ellick and Westbrook 2018a), provoke rioting and unrest (Knopf 1975), shake financial markets (Aral 2020), influence decision-making processes (Farrell et al. 2019), and disrupt global aid operations (Vosoughi et al. 2018; DiFonzo and Bordia 2007a).

One of the contributing factors in the escalation of rumour dissemination is polarised beliefs toward rumours (DiFonzo et al. 2013). Polarisation emerges since individuals prefer to interact with those who share similar views and orientations. This phenomenon divides societies into different echo chambers, reinforcing one's belief due to the repeated interactions with like-minded peers (Cota et al. 2019). Within an echo chamber, people are in high consensus and often say and hear similar messages. Research has shown the repetition of a message makes it more believable. Such a setting precipitates in social clusters with heterogeneous views toward a rumour. Such an environment deprives people of opposing views; thus, they increasingly become more confident regarding their echo chamber induced beliefs (Boutyline and Willer 2017; Flaxman et al. 2016).

In this vein, the media's role is crucial as it streamlines communication and increases the rate of reach to the audience (Franks and Attia 2011) (and subsequent exposure to rumours). Thus media could polarise individuals by feeding them with certain content and driving them toward echo chambers. On the other hand, it could expose them to diverse viewpoints, painting a fuller picture on every argument or debate. Research has shown that social media platforms exacerbate

polarisation by insulating people from opposing viewpoints (Cota et al. 2019). The size and diversity of social networks (Christakis and Fowler 2009), and automation mechanisms (Li et al. 2020; Shao et al. 2018; Woolley and Howard 2018) play a central role. Besides, other factors such as a lack of media literacy (Guess et al. 2020), minimal supervision (Lazer et al. 2018), low barrier to entry (Tsfati et al. 2020), and the lack of social media regulation (Marsden et al. 2020) facilitate polarisation which subsequently lead to the creation and circulation of rumours.

The escalation of rumour diffusion through social media platforms may lead to severe consequences that can influence political, economic, and social well-being of our society (Vosoughi et al. 2018). For instance, on April 23rd of 2013, the Twitter account of *Associated Press* released a tweet saying “Breaking: Two explosions in the White House and Barack Obama has been injured.” This tweet went viral by 4000 tweets in less than 5 minutes. This false news spread precipitated a significant drop (to the amount of \$140 billion dollars) in the stock market in a single day. Automated trading algorithms immediately began trading based on the potentials and consequences of the white house explosion and the reported death or injury of the U.S. president (Aral 2020). There are plenty of rumour dissemination cases in other domains such as elections (Aral and Eckles 2019), business issues (Farrell et al. 2019), and healthcare (Li et al. 2020) which lead to severe outcomes, that often have a disproportionate impact on the well-being of multiple communities.

In response to the detrimental effects of rumour propagation, a series of confrontation strategies have been devised (Knopf 1975; Ponting 1973; Allport and Postman 1947; DiFonzo and Bordia 2007a). Although taking the potential danger of rumour spreading into account and countering this phenomenon was a significant action, it was often an intermittent effort with ephemeral impacts. There was no long-term plan behind the confrontation strategies. Whenever a major incident happened or was about to happen, rumours started to thrive and then countering techniques were proposed and practised. While this whack-a-mole approach might have worked previously, it could not keep up with the rumour supply and circulation rate due to the sudden growth of social media in the past decade. On the one hand, various stakeholders such as social media platforms, governments, academia, and media organisations collaboratively developed new solutions. On another, lax in policy efforts at multiple social media platforms allowed for more innovative rumour spreading. Although scholars have proposed and practised a constellation of counter-rumour strategies on different scales, the massive waves of rumours still negatively impact individuals, organisations, and societal institutions (Vosoughi et al. 2018). To counter the alarming trend and repercussions of rumour propagation, we need to revise our former tackling approaches by developing a comprehensive understanding of the problem and past counter-strategies.

This chapter reviews three significant domains in the emerging field of rumour studies: the problem, its implications, and strategies for tackling them. The problem definition is a crucial phase as it entails determining what exactly has to be curbed and controlled in rumour propagation. There are different variations of false and unverified information (e.g., fake news, disinformation, misinformation, conspiracy

theory, etc.) which are recognised by the scholars as similarly harmful phenomena (Difonzo and Bordia 2007b; Woolley and Howard 2018). We study all aspects within the context of social media – its role and relevance in facilitating a spread – to provide a fuller understanding of the phenomenon. Because of the relatively long period of rumour confrontation in the societies, it is indispensable to obtain an overview of the past counter-rumour strategies and clarify the importance of tackling some rumour variations. It would provide information about the strengths and weaknesses of the rumour responses in the past. Scholars and decisionmakers could utilise that information later in the development of confrontation plan against rumour dissemination.

We organise the rest of the chapter as follows. In Sect. 2, we introduce different forms of false and unverified information. Section 3 investigates social media's role and its features and mechanisms in promoting rumours. Section 4 sets the past counter-rumour strategies in a common framework and evaluate them. Finally, Sect. 5 presents our findings and gives some suggestions for future work.

2 Background

There are many concepts in the English language implying false or unverified information. Terms such as misinformation, disinformation, rumour, urban legend, fake-news, propaganda, and conspiracy theory are just a few of these concepts that intermittently appear in the scientific arena. What academia has experienced regarding the conceptualisation of those terms and their conceptual siblings is an epistemic crisis. Although there are plenty of studies exploring various kinds of false and unverified information from different angles, there is considerable disagreement between proposed definitions. They are often conflated or used interchangeably (Freelon and Wells 2020; O'Connor and Weatherall 2019; Molina et al. 2019; Lazer et al. 2018; Vosoughi et al. 2018). The lack of consensus on conceptualisation would create confusion and drains the community's efforts in countering the surge of false information.

Although the genesis of such information might be deliberate or inadvertent and with different purposes, they primarily disseminate based on similar motives and follow the same dynamical process (i.e. creation and dissemination) (DiFonzo 2009; Difonzo and Bordia 2007b; Difonzo 2018). Regardless of the variations, the dynamics of false and unverified information is similar across the board.

2.1 *The Variations of False and Unverified Information*

This section investigates rumour, its variants, and its conceptual siblings to understand the differences among variations of false and unverified information. The notion of rumour refers to declarative statements composed of nouns and verb

statements that purport to inform, explain, predict and thus provide information (DiFonzo 2009; Bordia et al. 2014). It is a collective process that arises in the collaboration of many. It involves a division of labour among participants, each of whom makes a different contribution (Shibutani 1966). Rumour existence is contingent on its circulation (Bordia et al. 2014), and end of the communication activity equals the death of rumour (Shibutani 1966). A rumour spreads if it relates to, affects, or threatens rumour participants in some way. Rumours are unverified in some context, and they are not accompanied by substantial evidence for at least some group of people (Difonzo and Bordia 2007b). Rumours tend to thrive amid situations that are ambiguous, confusing, uncertain and threatening (DiFonzo 2009). Rumours circulate primarily as a sense-making or threat management mechanism. They offer details and reasons as well as meanings, clarifications, and justifications.

Rumours may take different shapes; however, they evolve into a similar dissemination dynamic which mimics rumour spreading after the first generation of the transmission. Propaganda and fake-news are planted into public deliberately to induce psychological threats and take advantage of the people; however, the audience treats them in the same way they treat rumours. People circulate their impressions, interpretations, or reactions among themselves to make sense of those information (Cook 2020; Difonzo and Bordia 2007b; DiFonzo 2009). Similarly, conspiracy theories and pseudoscience might emerge as a coping strategy among a group of people to manage psychological threats in response to uncertain or threatening situations (Difonzo 2018). Legends, myths, and urban legends are also very similar phenomena; however, their life-cycle is much longer than rumours; hence, we consider them a separate phenomenon. Similarly, Gossip is also a distinct phenomenon in local group levels with a slightly different dynamic. Table 1 summarises those variations and subsequent sections explain them in more detail.

2.1.1 Gossip

Gossip is an evaluative social talk about an individual's personal life. It is a group-level phenomenon which glues down groups and adjusts people's relationships. Gossiping can maintain group cohesiveness, establish, change, maintain group norms, group power structure, and group membership. It can also function as an entertainment mechanism (DiFonzo 2009; DiFonzo and Bordia 2007a). Gossiping is also an effective strategy for intragroup competition (DiFonzo 2009; Fine and Rosnow 1978). Nefarious self-serving motives may drive Gossip to slanderous outcomes. They may break groups apart or taint people's reputation. However, there is benevolent Gossip that functions as a warning against the harmful behaviour of particular individuals. Gossip may also regulate individuals' behaviour regarding the context. Gossip works on group dynamics between friends, not those who do not know each other. It is the signal of affiliation, closeness, and camaraderie (DiFonzo 2009; Fine and Rosnow 1978).

Table 1 Comparison between different forms of false and unverified information (Inspired by DiFonzo and Bordia 2007a)

Context	Evidentiary basis	Perceived importance	Content slanderous	Message theme & structure	Function	Function
Rumour , Ambiguity, danger, threat, or change	Low	High	Maybe	The message is a declarative statement, consisting mostly nouns and verbs.	The message is a declarative statement, consisting mostly nouns and verbs.	Maybe
Gossip , Social network formation and maintenance	Maybe	Low	High	The message is evaluative, informal and refers to individuals.	Allowing groups to become more cohesive and to define their membership, norms, and power structure	High
Legend , Socially cohesive contexts	Low	Low	Low	The narratives pertain to issues that are important for successive generations such as birth, marriage, and death. They have story like structures, including setting plot, characters, climax, and denouement.	To entertain and to convey mores, norms, and cultural truths	High
Propaganda , Threatening situation when much is at stake	Low	High	High	The messages are about supporting or challenging particular viewpoints or ideologies mostly in politics and business. The message structure is composed of fabricated materials as well as manipulated and vivid images, symbols, and slogans.	To simultaneously induce psychological threats and to function as a sense-making and threat management mechanism.	Low
Conspiracy theory , Psychological threats	Low	High	High	The messages follow a narrative about covert and hostile activities of secret and powerful groups.	To cope with threats by providing alternative explanations for events and incidents.	Low
Fake-news , Attention-gaining context	Low	High	High	The messages are emotionally charged and written in narrative style; They are not fact-checked and their source is not verified. In order to draw attentions, elements of (i) threat-related information, (ii) sexually related information, or (iii) elements associated to disgust, are incorporated.	To gain financial benefits by drawing eyeballs.	Maybe
Pseudoscience , Ambiguity and threat	Low	High	Maybe	The message pertains to an issue within the domain of science in a broad sense.	Sense-making and threat management by rejecting the empirical studies and/or promoting fabricated claims	Low
Misinformation , Ambiguity and threat	Low	High	Maybe		For sense-making	Maybe

2.1.2 Legend

Legends are narratives about unusual, humorous, or horrible events with moral lessons¹ (DiFonzo and Bordia 2007a). After being recounted for many years, a prior history of distortion and transformation, legends converge to stable forms and become part of people's folklore and verbal heritage. Legends are immortal because they capture the universal aspects of human character. Legends follow a storytelling framework. They have a setting, plot, characters, climax, and denouement. They function as a mechanism for entertainment and propagation of values and mores. Legends also make sense of the world by conforming to answers to the persistent riddles of life. Legends are about subjects considered important for successive generations. If the legends are about primal forces, cosmology, or religious beliefs, then they are called myths (Allport and Postman 1947; DiFonzo and Bordia 2007a).

2.1.3 Propaganda

Propaganda refers to persuasive tactics devised deliberately by governments or corporations to promote or challenge a particular viewpoint by manipulating symbolic representations (Shibutani 1966; O'Connor and Weatherall 2019). It can occur in computational settings over social media, "using algorithms, automation, and human curation to purposefully manage and distribute misleading information over social media networks" (Woolley and Howard 2018). The computational setting in computational propaganda allows automation which brings scalability and anonymity. Many state and non-state actors use computational propaganda to suppress their oppositions, promote their viewpoints, divert or destroy movements, and create fake trends (Woolley and Howard 2018).

Propaganda may take on particular shapes. One of them is innuendo which functions as a character assassination technique to discredit the reputed individuals. For instance, since early times innuendos tarnished U.S. presidential elections by accusing candidates with illicit sexual relations, racism, brutal treatment of wives, drunkenness and the alleged possession of certain blood types (Allport and Postman 1947). It may also serve as a projection technique to accuse another person of the same things that the accuser is guilty of (DiFonzo and Bordia 2007b). One of the most notorious shapes of propaganda rumour is disinformation invented by KGB in 1923\$. It is black propaganda based on forgeries. Disinformation includes forged and fabricated narratives, letters, documents, photographs, reports, and press releases (Martin 1982; Romerstein 2001). One of the kinds of forgery that is getting increasingly popular is audiovisual (A.V.) manipulation. It includes both

¹ The term "legend" refers to both traditional legends (about knighthood, ogres, witches, sleeping princesses, etc.) and modern or contemporary legends (about dating, technology, organ removal, etc.) Modern/Contemporary legends are also called urban legends which is a misnomer because those narratives do not necessarily occur in an urban environment.

the cutting edge AI-reliant technologies of deepfakes and cheap-fakes; conventional audiovisual manipulation techniques such as speeding, slowing, cutting, re-staging or re-contextualising footage (Paris and Donovan 2019).

2.1.4 Conspiracy Theory

Conspiracy theories are unverified pieces of information circulating about events or incidents initiated for deliberate hostile purposes by a coalition of actors operating in secret. A conspiracy theory assumes predesigned patterns govern the universe, and there is no room for randomness and coincidence. That is why conspiracy theories try to randomly connect the dots and find the secret patterns (Van Prooijen 2018). Conspiracy theories may arise in a variety of subject domains such as scientific research (e.g. global warming is a hoax created by China (Hendricks and Vestergaard 2019)), sport (e.g. referee bribing conspiracy theory (Van Prooijen 2018)), or the government (e.g. deep state conspiracy theory (Benkler et al. 2018)). Among the commonly used conspiracy tactics are contradictory explanations, overriding suspicion, nefarious intent, something must be wrong, persecuted victim, immunity to evidence, re-interpreting randomness (Cook 2020).

2.1.5 Fake-News

The notion of fake-news² is defined as “fabricated information that mimics news media contents in form but not in the process and intent”. Fake-news outlets do not follow editorial norms and guidelines (Lazer et al. 2018). Such outlets narrate emotionally charged articles that are devoid of fact-checking and source verification; sometimes, pieces have inconsistencies with the registration date (Molina et al. 2019). Although fake-news reports have mostly arisen in a political context, there are plenty of cases in other domains such as vaccination and stock markets (Lazer et al. 2018).

Since the early days of journalism, fake-news found its way into various news outlets (Uberty 2016). Fake-news articles could draw attention easier than real news as there is no constraint for fabrication. We can be as creative as we want to develop appealing, attention-grabbing and memorable articles (Acerbi 2019). More attention means higher readership, leading to a more significant profit margin for the news outlets (Standage 2017). One of the earliest and most successful fake-news articles was the New York Sun’s “Great Moon Hoax” of 1835, which claimed there was an alien civilisation on the moon. This fabricated story drew much attention to

² In the current political climate, there is a significant disagreement in academia regarding the consumption of the term “fake news” as it became a value-loaded term linked to particular political figures (Vosoughi et al. 2018; Lazer et al. 2018); however, due to the lack of an alternative name and to avoid adding further confusion to the existing fluid terminology, we have elected to retain the term “fake news”.

New York Sun to the extent that its circulation reached from 8000 to 19000 copies, which meant overtaking Times of London as the world's bestselling daily newspaper (Standage 2017; Posetti and Matthews 2018).

2.1.6 Pseudoscience

A statement is pseudoscientific if it satisfies three criteria of (i) scientific domain, (ii) unreliability, and (iii) deviant doctrine³. The measure of scientific domain entails a pseudoscientific statement to be about an issue within the realm of science.⁴ A pseudoscientific statement suffers from a severe lack of reliability and trust. Besides, it has no use for knowledge production nor any practical cases. The deviant doctrine criteria indicate the support of pseudoscientific statement proponents to represent that statement as the most reliable knowledge on the subject matter (Hansson 2017).

Pseudoscience can take two different forms of science denialism, and pseudo-theory promotion (Hansson 2017). Science denialism refers to “the rejection of empirically supported propositions despite scientific consensus and the effort to create the appearance of debate when there is none” (Schmid and Betsch 2019). Some typical examples of denialism are climate change, Holocaust, vaccination, and negative impacts of tobacco (Hansson 2017; Cook 2020). The other category of pseudoscience is pseudo-theory promotion, which refers to the fabrication of a set of claims to advance the pseudoscientist's theory. Sometimes it leads to the rejection of parts of science. Some typical examples of pseudo-theories are astrology, homoeopathy, iridology, Scientology, transcendental meditation, and ancient astronaut theories (Hansson 2017, 2018).

2.1.7 Misinformation

Misinformation is a widespread form of false and unverified information. This concept originates in cognitive psychology and developed by the scholars who were studying misinformation effects on memory formation, visual object classification, children's ability to infer the mental states of others, and performance on multiple-choice tests (Freelon and Wells 2020). The misinformation effect refers to “the distorting effects of misleading post-event information on memory for words, faces, and details of witnessed events” (Frenda et al. 2011). Nevertheless, nowadays, the term has found a much broader yet loose meaning: any kind of deceptive message that might be harmful but spreads inadvertently (Freelon and Wells 2020). If its

³ To view a statement pseudoscientific, all the three conditions require to be confirmed (Hansson 2017). For example, if a commentary satisfies the first two criteria but not the third one, probably it is fraud in science or mistake in science, but not pseudoscience.

⁴ The term “science” implies science in a broad sense which comprises humanities as well.

truthfulness is unverified for an individual and s/he spreads it (without malice), that piece is considered misinformation.

2.2 *Process-Based Perspective*

This section investigates the dynamics of false and unverified information by explaining the emergence and dissemination phases. False and unverified information is initially shared to serve four general purposes of (i) social manipulation, (ii) sense-making and threat management, (iii) social dynamics, and (iv) cultural dynamics. Social manipulation refers to “the purposeful, systematic generation and dissemination of information to produce harmful social, political, and economic outcomes in a target area by affecting beliefs, attitudes, and behaviour” (Mazarr et al. 2019). Planting misleading information into public is a long-standing manipulation strategy when much is at stake (e.g. in wartime, elections, highly competitive markets). False and unverified information can also appear at the time of uncertainty and threat as a coping strategy and “to give meaning to our sensations and to put a context around them, so they gain significance and fit into an understanding that coheres” (DiFonzo 2009). They may also function as a social mechanism to entertain, to supply social information, and to establish, change, or maintain group membership, group power structure, or group norms (DiFonzo and Bordia 2007a). The cultural dynamics is the other purpose of spreading false and unverified information to establish, maintain, or impart cultural mores or values. It also provides answers to the persistent riddles of life (DiFonzo and Bordia 2007a; Allport and Postman 1947).

After the creation of a message on social media – deliberately or inadvertently – the message spreading starts. Humans used to circulate their messages through word of mouth or letter; however, the communication technologies and the emergence of new media facilitated this procedure. Although media is one of the most efficient ways of releasing false and unverified information, not all media outlets are the same. They vary from highly professional ones with codes of conduct, style guides, and journalistic guidelines to those without any policy or standard. Due to the controversial aspects of false or unverified information, they cannot always meet the professional journalism standards and reach credible outlets (Lazer et al. 2018). Sometimes message circulation starts from state-owned media agencies and low credibility news outlets (O’Connor and Weatherall 2019). Stakeholders whose purpose is social manipulation use state-owned media since they operate as the information operation wing of many governments (Ellick and Westbrook 2018b).

Additionally, blogs and low credibility websites may also be the host of non-credible information. They are solely searching for more visits to increase their revenue; therefore, what matters for them is catching eyeballs not complying with journalism norms, and standards (Acerbi 2019). An alternative or complementary channel is online social media platforms. They have a massive user base as well as minimum restrictions for the authorised contents. Thus it can be a perfect

environment for any false or unverified information to begin their circulation no matter its purpose (Christakis and Fowler 2009; Lazer et al. 2018).

After releasing false and unverified information, it would be extremely challenging to control its passage and keep it in check, as it forms spontaneously. Its development depends upon fortuitous events, momentary emotional reactions, and the particular interest of those who make up the public. Sometimes, the discussions might be directed and redirected, especially in social manipulation; however, the extent of interventions effectiveness is circumscribed (Shibutani 1966). When false and unverified information goes public, spontaneous discussions, including different communication postures among the people who come across that information, propels it further (DiFonzo and Bordia 2007a). People often participate in those discussions collaboratively by raising questions, providing information, indicating beliefs and opinions, expressing feelings, or suggesting a course of action no matter what is the type of false and unverified information (Bordia and DiFonzo 2004). Thus, whether it is black propaganda to tarnish a presidential candidate or an honest mistake about confusing a firecracker's sound with an explosion, the same dynamic will happen (the scale and life-time of the process depend on a set of internal and external factors). This dynamic process is precisely similar to what occurs in rumour spreading when people engage in the shared sense-making process through interaction with others.

Although the genesis of false and unverified information might be for different reasons, they evolve into a similar dissemination dynamic which mimics rumour spreading after the first generation of the transmission. Propaganda and fake-news are planted into public deliberately to induce psychological threats and take advantage of the people; however, the audience treats them in the same way they treat rumours. People circulate their impressions, interpretations, or reactions among themselves to make sense of those information (Ruths 2019; Difonzo and Bordia 2007b; DiFonzo 2009). Similarly, conspiracy theories and pseudoscience might emerge as a coping strategy among a group of people to manage psychological threats in response to uncertain or threatening situations (Difonzo 2018). Legends, myths, and urban legends are also very similar phenomena; however, their life-cycle is much longer than rumours; hence, we consider them a separate phenomenon. Similarly, Gossip is also a distinct phenomenon in local group levels with a slightly different dynamic.

2.3 Forms of False and Unverified Information

This section draws attention to the message forms in the variations of false and unverified information. As discussed before, rumours, legends, and Gossip are three broad variations of false and unverified information. The least harmful one is a legend and its siblings, namely urban legend and myth. Although they had been rumour once, they are distinct phenomena with crucial differences after years of transmission. Legends' primary goal is to share values and conform to the vacuums

of riddles of life. Thus, it is implausible that legends lead to harmful consequences. The other one is Gossip which mainly maintains group-level mechanisms such as cohesiveness, power structure, norms, and membership. However, it may transform into slander based on nefarious self-serving motives. Nevertheless, it is implausible that the impact of gossiping transcends the boundaries of social cliques and reaches more extensive social networks.

The other construct is rumour and its offspring. In a broad sense, rumour functions as a sense-making or threat management mechanism. However, depending on the form of rumour, both sense-making and threat management may take different shapes. In propaganda rumour, it mostly serves a pernicious function. Although there are different types of propaganda rumour, their primary intent is malicious. The conspiracy rumours are also relatively harmful and harass their subjected groups by falsely accusing them. Fake-news rumours might become harmful by promoting appealing yet fabricated materials to lure individuals. The pseudoscientific rumour is a toxic phenomenon that attacks the science institution by tarnishing scientists, scientific evidence, and scientific methods. Misinformation does not inflict any harm wittingly; however, as discussed before, it may appear when uninformed individuals are engaged in the rumour process. Besides, even if misinformation rumour does not take the shape of derogatory rumours and spread with a benign yet inadvertent motive, it may lead to harm (DiFonzo 2009).

Thus among the variations of false and unverified information, the rumour family operates on a large scale, even if they start spreading unwittingly. Without malign intent, they may still lead to severe consequences. Therefore, it is crucial to take the variations of rumour into account and develop a solution to curb and control this phenomenon; otherwise, the repercussions would be inevitable. They may influence political, economic, and social well-being (Vosoughi et al. 2018).

3 Rumour Facilitation Features in Online Social Media Platforms

This section discusses how the introduction of new communication technologies could facilitate message transmission and thus rumour dissemination. The development of communication technologies is a complex and continuous process. Here, we slice up this developing process and take snapshots of a few technologies with significant importance in rumour dissemination.

In the pre-printing era, there was no synchronous and mass communication technology, and the possibility of long-distance message transmission was quite limited (Chakravarthi 1992). Within this period, only local and small-scale communication prevailed. Thus, rumours were also often about local issues and remained within the communities. The invention of the printing press was a turning point in communication technologies as it made mass communication possible (Buringh and Van Zanden 2009). This technology increased the chance of exposure to rumours by

replication of the same message and sharing them among many people (Standage 2017; Posetti and Matthews 2018; O'Connor and Weatherall 2019).

Another crucial technology was the telephone which introduced *synchronicity* to the communication process. The phone could accelerate and expand the rumour circulation by offering fast *long-distance* communication. The key distinctive features of the printing press and telephone technologies (i.e., synchronicity and mass communication) were incorporated into radio and television and created synchronous mass communication mediums. It was the first time in history that distant live contact with masses became possible. They could also draw more attention because of *multimedia elements*. Besides, compared to written media, the radio and television comprised a wider audience beyond those who could read. Due to the features mentioned above, radio and television were extensively utilised in rumour spreading (Dowe 1997).

The mass spread of rumours began with the emergence of the World Wide Web (WWW) using *distant synchronous/asynchronous multi-lateral communication* (Lehmann 2019). Through the forums, chat rooms, and other WWW-based applications, people could communicate without knowing each other. Such platforms allow individuals to *hide their identities* or use anonymous, or even fake avatar in their profiles (Christakis and Fowler 2009; Hussain 2019). *Minimal supervision* is another feature of the WWW that fostered rumour spreading. Earlier, only a self-selected group of people were linked to media outlets and had access to communication channels. With the advent of a free venue for people to express their thoughts and opinions (in the forums, chat-rooms, their blog or website), WWW created a global platform for people to communicate no matter who they are and how credible their messages (Lazer et al. 2018; Dowe 1997).

The rise of online social media platforms was the landmark in the history of rumour spreading. This technology supports distant multi-lateral communications with different synchronicity modes. Social media is a complex phenomenon that offers novel features in three layers of social, institutional, and technological. From social perspective, it is an enormous hyper-connected network (Christakis and Fowler 2009) pursuing power-law degree distribution (Barabási et al. 2016). The large scale of social networks vastly increases the number of people who might get exposed to rumours. It also increases the chance of creating communities with similar values, beliefs, and interests (Briones et al. 2011; Schor 1995).

This turns social media to an environment with a high potential for polarisation. It allows people who had not any tribune before, find like-minded peers and freely communicate their controversial thoughts quickly. Besides, within social media, high-degree nodes or hubs that play influencers and opinion leaders' role affect the small-world property in the network, which eventually leads to the virality of information and rumours (Barabási et al. 2016; Vosoughi et al. 2018). Furthermore, when a network's size increases, the chance of diversity among the users improves, expanding the scope of potential rumours (Christakis and Fowler 2009).

From the institutional perspective, social media platforms allow their users to participate in information dissemination while leveraging anonymity. It specifically gives a safety margin to the initiators or moderators of inorganic rumours (Fox

2020; McGonagle 2017) since they can “say whatever they want, whenever they want, and yet be shielded by anonymity” (Christopherson 2007). The platforms also enable the democratisation of the content by allowing individuals to consume, create, and distribute their content without governmental control (Tufekci 2017) and with minimal supervision (Lazer et al. 2018). This means people of different ages, education, and nationality are free to share their thoughts, and discuss their ideas about various topics in politics, sport, or trivial daily-basis incidents, to name but a few. They can produce and share whatever content they want as long as it does not violate the platforms guidelines developed at a minimal level not to curtail freedom of speech (Hussain 2019; Roth and Pickles 2020). This policy would lower the barrier to entry for not only those who have not received any training in journalism but also for the ones who bluntly reject the journalistic norms of objectivity and balance (Lazer et al. 2018). The low barrier to entry also contributes to a more polarised community. It allows those who support controversial ideas to easily find their echo chamber and make it bigger and more isolated.

Besides, any attempt to control such an extravagant system is perceived as the censorship, since social media platforms are considered the manifestation of freedom of speech and whatever that restricts this sphere will be interpreted as a violation of freedom of speech. That is why codes of conduct, style guides, and journalistic guidelines in online social media platforms are kept to the minimum level (Ruths 2019). This is an ideal setting for rumour spreading because the information is not verified before releasing into the sphere of online social media.

From a technological perspective, social media is equipped with mechanisms such as recommendation systems and social bots that can easily be used to facilitate the spread of rumours (Woolley and Howard 2018). The recommendation system is a specific type of algorithm used to enhance the user experience by reducing the information overload (Jannach and Jugovac 2019) and helping users find compelling content in large corpora by personalised suggestions (Berkovsky and Freyne 2015; Wagh and Patil 2012). The other automation mechanism in online social media platforms is social bots. They are computer programs that tend to emulate and alter human behaviour and produce content and interact with other humans (Ferrara et al. 2016). Scholars have shown that recommender systems and social bots exacerbate polarisation in social media. Recommender systems drive social media users toward more radical content by deliberately avoiding to expose them with opposing views. Social bots also amplify the spread of information that is pleasing for some people to trap them in echo chambers by inducing certain beliefs. Social bots are primarily active in political contexts (Dandekar et al. 2013; Stella et al. 2018). For instance, research by Dutch newspaper NRC revealed that of over 900 Russian Twitter accounts, listed by the U.S. investigation into Russian meddling in the 2016 elections, as attempting to influence the Dutch debate on Twitter in 2016 and 2017. By far, these accounts focused on amplifying anti-Islam sentiment, and polarising attitudes towards migration and refugees (Hazenburger et al. 2018).

Both social bots and recommendation systems allegedly play a central role in spreading rumours by amplifying the messages with strategic goals and leading people toward rumour rabbit holes, respectively. Some studies show the significant role of bots in the circulation of rumours. They drive rumours by liking, sharing, and searching for information. Notably, they are responsible for a substantial amount of contents during political events such as the 2016 U.S. Presidential election and 2017 French election, to name but a few (Lazer et al. 2018; Woolley and Howard 2018; Shao et al. 2018). On the other hand, there are studies, for example, the most extensive research on the digital spread of rumours, that show the insignificance of social bots compared to humans in the spread of rumours (Vosoughi et al. 2018).

Similarly, in the case of recommendation systems, some studies show the effectiveness of such systems in the circulation of rumours (Matz et al. 2017). In contrast, some others raise doubt about the effectiveness of recommendation systems in the spread of rumours. Although controversial on the surface, all those paradoxical results might be part of a bigger picture (Ruths 2019) which currently does not exist.

4 The Evaluation of Strategies for Countering Rumours

It was merely a century ago that the first systematic efforts against the mass spread of rumours began (Allport and Postman 1947). Since then a variety of techniques have been exercised by media organisations (Fabry 2017), researched at academic institutions (Van Der Linden et al. 2017; Allport and Postman 1947), and recently implemented by social media platforms (Google: How Google Fights Disinformation 2018; Mosseri 2017) and by governments within digital policy frameworks (DW: German justice minister to set up task force on Internet hate speech 2015).

To develop an effective and comprehensive plan to quell rumours and develop novel techniques, it is crucial to be aware of the existing strategies and potential capabilities. We present a review which explores both the targeted and neglected aspects of addressing rumour diffusion. Figure 1 illustrates the variety of prevalent strategies for countering rumours. Those strategies are classified based on their impacts on the components of the communication process. A communication process comprises three major elements of the sender, channel, and receiver in which senders transmit messages to receivers through communication channels. The first group of strategies takes those who initiate rumours into account and restrain them (in case of deliberate rumour spreading). The second group tends to secure communication channels and minimise rumour emergence and circulation within the channels. The purpose of the third group is to protect those who were targeted by the rumours. In the following, we explain the strategies in more detail.

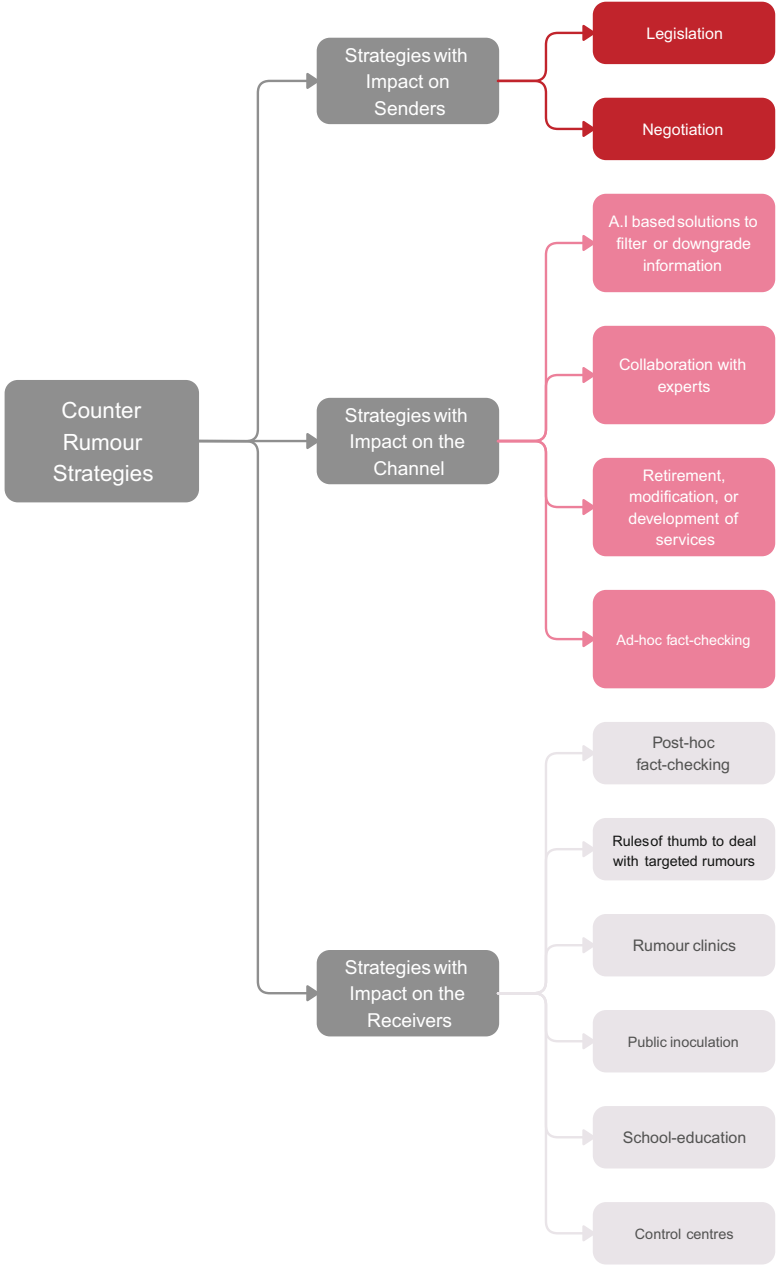


Fig. 1 The classification of counter-rumour strategies

4.1 Strategies with Impact on Senders

In this section, we discuss counter-strategies developed to confront those who deliberately participate in the rumour process. To this end, we propose two main approaches to legislation and political solutions.

4.1.1 Legislation

One of the oldest and most widespread rumour counter-strategies is to take the government's legal action against those who contribute to the rumour spreading. The earliest form of this approach is to protect individuals against baseless defamation and character assassination. Although this is a practical approach to serve this purpose, it could be costly and take a long time. However, on the bright side, legal actions raise awareness by drawing media attention (Farrell et al. 2019). With the growth of online social media platforms, the expansion of legal actions from conventional mediums to online social media gradually started. Germany was among the first countries taking online environments into account (DW: German justice minister to set up task force on Internet hate speech 2015); however, the number of countries with regulations regarding rumour spreading in online environments is increasing (Funke and Flamini 2018).

4.1.2 Negotiation

Social manipulation campaigns are becoming an essential tool for information operations. On the one hand, different governments and organisations use this approach to influence public opinion. On the other hand, platforms and media organisations try to flag such operations and take down the accounts and messages linked to those operations (Fung and Garcia 2019). Reactive measures often seem like a cat and mouse game since new manipulation campaigns keep popping up while platforms adapt to take them down. One of the strategies to confront or limit this spread is negotiating with the offenders and working toward an agreement with them. This tactic is not always a feasible option due to lack of interest from either sides, unknown identity of the offenders, or unrealistic expectations; however, this is a promising approach with meagre cost, and high output (Bodine-Baron et al. 2018).

4.2 Strategies with Impact on the Channel

In this section, we discuss the strategies concerning the protection of communication channels. Media organisations or social media companies propose strategies

that aim to make channels rumour-proof and minimise the likelihood of rumour emergence and circulation.

4.2.1 Solutions Based on Artificial Intelligence

The spread of rumours is a rapid, widespread, and profound phenomenon within social media platforms (Vosoughi et al. 2018). One of the most efficient approaches to confront this rampant diffusion is Artificial Intelligence (A.I.). It is a fast and cheap approach, and it can tackle the spread of rumours at scale, and across languages and time zones. A.I. techniques are useful for two primary purposes of content filtering and downgrading.

Filtering

Among the massive streams of information flowing in social media platforms, there are mischievous pieces that could make their way to the platform. Although compared to the information circulating in social media, the volume of such contents is insignificant. They have to be taken down before making public; otherwise, there might be unexpected ramifications. The tremendous amount of circulating information in social media eliminates the option of manual inspection. Instead, platforms benefit from the power of predictive analytics and machine learning. They often first flag the contents that are deemed problematic with high confidence, then those pieces that have a borderline status are sent for human-based evaluation.

Downgrading

Social media platforms are often using a timeline-based design within which there is running list of posts in users main page showing most recent updates from their social network. This list -at least in default setting- does not show contents in chronological order. Platforms use algorithms to show posts in an order that is more appealing for the user. If rumours can game the ranking algorithm and climb up to the top of the list, they likely see them. This dynamic would increase the visibility of rumours which positively affect their circulation rate. To confront this issue, platforms cannot remove rumour related posts as they do not break any law or violate any term; thus they try to identify such posts and relegate them in the timeline (Lyons 2018; Harvey 2018; Google: How Google Fights Disinformation 2018). This action would subsequently reduce rumours spreading chance.

4.2.2 Collaborative Solutions

To protect communication channels from rumours, the collaboration between experts from media organisations, social media platforms, and academia is essential. They all have complementary expertise that would help reduce the chance of rumour emergence and circulation.

Collaboration with Fact-Checkers

One of the biggest concerns for social media platforms is the verification of dubious information. The verification is a difficult problem to address because of two reasons. First, social media platforms purport they cannot arbitrate information truthfulness as they are technology firm and not a media company. Second, the scale of information circulation in social media is massive and conventional methods cannot keep up with information growth. To tackle these issues, social media platforms started to collaborate with fact-checking institutes considered information verification experts. After a primary AI-based screening or receiving reports from users, they sent inappropriate posts to fact-checkers for further assessment. Although this approach can address the first issue, the second problem is still there. The platforms' response to this challenge is to expand the collaboration with fact-checking institutions and turn it into a crowdsourced activity. To this end, they are planning to give a reviewer privilege to particular users and ask them to evaluate flagged social media posts. With this solution, the capacity of fact-checking significantly increases, thus could solve the data volume issue.

Collaboration Between News Outlets and Social Media Platforms

Media literacy and digital journalism are essential in building resilient communication channels against rumours. Social media, along with major media organisations, can promote quality journalism in the digital era. To this end, major social media platforms and news outlets formed partnerships, initiated training programs, and even developed products and services to empower journalists (Mosseri 2017; Policy team: Update on Twitter's review of the 2016 US election 2018; Google News Initiative).

Collaboration with Academia

Tackling large-scale rumours in social media is a recent phenomenon with plenty of unknown and unexplored aspects. A social media platform is an organisation with limited resources which make having a comprehensive understanding of social media rather impossible. On the other hand, universities and scientific institutions are specialised organisations for research. Social media platforms could benefit from cutting edge knowledge without a huge investment in their research and development to collaborate with academia. Such a collaboration is also beneficial for scientific institutions as they would have access to a suitable environment to test old social theories and hypotheses and propose new ones (Al-Tabbaa and Ankrah 2019).

4.2.3 Design Based Solutions

In this category of solutions, the goal is to permanently retire a service, refine a service with issues, or develop new service.

Service Retirement

Sometimes, a service that was initially developed to add value to a business, gradually turns to a source of headache for that organisation. One of the approaches to tackle this issue is to retire the problematic service. Although this might seem like a naive strategy, sometimes the stake of keeping the service open (for modifications) is too high, and the best thing that we can do is to shut down the service as soon as possible. One example of service retirement is Facebook APIs such as Events API, Search API, and App Insights API after Cambridge Analytica scandal (Archibong 2018).

Service Modification

Another approach to deal with problematic services is the modification strategy. Using this approach, the service provider can preserve the service by addressing its shortcomings. One example of this approach is WhatsApp message forwarding limitation policy issued to reduce the rumour diffusion speed (Kastrenakes 2019).

New Service Development

The design and development of new services are essential for the survival and growth of organisations. In the case of rumour response, social media platforms have built services to mitigate the chance of rumour emergence and transmission. They mostly try to put the information in the right context by providing extra information. One of those services is the context button which adds more context to the Facebook posts (Lyons 2018).

4.2.4 Ad-hoc Factchecking

Media organisations are in the front-line of combating rumours. They are on the first tier of the news supply chain, making them responsible for sharing accurate, impartial, and evidence-based information. The significance of the accurate journalism entailed defining an exclusive role in news outlets to inspect information integrity before making public. Later on, this auditing role got a new title of ad-hoc fact-checking it aims to eliminate errors before a piece goes live (Fabry 2017).

4.3 Strategies with Impact on the Receivers

In this section, we discuss strategies to protect people subjugated to rumours. In the following, we discuss six counter-strategies.

4.3.1 Post-hoc Fact-checking

In the previous section, we explained ad-hoc fact-checking as a preemptive counter rumour strategy. Although it is a highly effective approach to confront rumours, it

is not always feasible to flag and corrects misleading information beforehand. For instance, during live programs (e.g., presidential debates) it does not make sense to use ad-hoc fact-checking. In such situations, another variant of fact-checking called post-hoc fact-checking is used. It identifies and corrects errors after they go public. Post-hoc fact-checking is less effective than ad-hoc fact-checking as in the earlier people are already exposed to rumours. Nevertheless, keeping the number of people exposed to the minimum has a tremendous impact on tackling rumour diffusion (Graves 2016).

4.3.2 Rule of Thumbs to Deal with Targeted Rumours

When a rumour particularly targets an entity (e.g., an individual, an organisation, a country), the entity’s spokesperson should take a stance. Figure 2 displays possible ways of responding to a rumour. Initially, one should decide whether to comment about the rumour or to ignore it. In case of commenting, then there are three avenues they could be followed: (i) Confirmation of the rumour and giving detailed information about it, (ii) Denial of the rumour and giving a rebuttal, or (iii) Withholding to comment about the rumour (DiFonzo et al. 1994).

The other option is to ignore it, which is the weakest quelling strategy. It is doubtful that a rumour dies on its own because something unusual to someone may be deemed as plausible by another person. In other words, even if some people drop the rumour, some others pick it up.

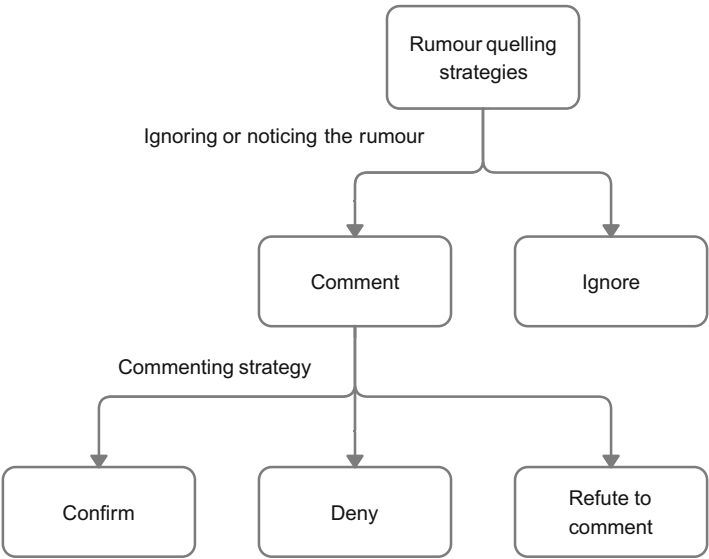


Fig. 2 The classification of counter-rumour strategies. (DiFonzo et al. 1994)

Some rumours are fully or partially truthful. Scholars have shown that confirming the genuine parts of the rumours reduce the chance of rumour transmission. One of the most common strategies to rein in the rumours is denial. There are a few factors that influence denial effectiveness. A denial should be based upon the truth. Denial of a true rumour is a dangerous strategy that may permanently destroy the responsible party's credibility. The denial effectiveness is enhanced when the denial source is trusted and is aligned with the rumour context. It is also crucial to avoid repeating the rumour during the denial. The repetition fosters beliefs to the rumours. The responsible party should also issue the denial as soon as possible rather than after reaching a certain critical mass. Finally, a denial message should deliver a clear, detailed explanation with strong evidence indicating the rumour falsehood and convey to listeners a straightforward course of action about what they should do when they come across the rumour. A no-comment response would work the same as an ignorance strategy, and at worst gives more credence to the rumour. It reinforces the cover-up hypotheses and transfers the message that "we have something to hide" (DiFonzo and Bordia 2007a; DiFonzo 2009; DiFonzo et al. 1994).

4.3.3 Rumour Clinics

The rumour clinic was a workshop composed of experts collecting and analysing major circulating rumours and then training participants about the techniques embedded in those rumours. Although in the beginning, the government was supposed to organise rumour clinics, they stepped back due to the uncertain consequences of sharing and repeating rumours. Despite the initial withdrawal, they could not stand the idea of public rumour revelation, and eventually, in 1943 they could shut down rumour clinics (Allport and Postman 1947; Faye 2007).

4.3.4 Control Centres

In response to widespread rumour circulation during civil right movements in the U.S. in the 1960s, the government created a telephone service called rumour control centre (RCC) to provide reliable information about floating rumours to the general public. Those who called RCCs could ask about the veracity of rumours and stories they heard, and then RCC staffs could verify that information via police and intelligence units (Young et al. 2014). From an organisational perspective, RCCs were small organisations with a few employees and a simple structure. They only require communication hookups with police and fire departments and other city agencies to get the most recent updates regarding the incidents (Ponting 1973; Knopf 1975). From an institutional perspective, they were counted as governmental agencies (Young et al. 2014). RCCs gradually vanished in the 70s due to the funding and legitimacy issues (Ponting 1973; Faye 2007); however, a digital variant of RCCs has recently been introduced by the Federal Emergency Management Agency

(FEMA) website. This web page provides information about running rumours during disasters. Although it is just a static web page, it is continuously updated by monitoring the circulating rumours in the digital sphere.

4.3.5 Education and Media Literacy

Education is supposed to be highly useful to confront rumour dissemination. By providing knowledge on propaganda techniques and critical thinking, it merely gives a powerful and long-lasting shield against all sorts of misleading information. Open Society Institute calls education “the best all-round solution” and mentions “high-quality education and having more and more educated people is a prerequisite for tackling the negative effects of fake news and post-truth” (Lessenski 2018).

This policy aims to improve media literacy and critical thinking among citizens. It would help them to reflect on the information they receive before believing or sharing them. Despite all the great qualities of education policy, too much emphasis on being critical might backfire and come with unintended consequences of undervaluing respectable and high-quality news outlets (Lazer et al. 2018).

4.3.6 Public Inoculation

Despite shutting down rumour clinics in 1943, the idea of creating immunity against rumours did not die and reframed in public health literature as public inoculation against rumour spreading. As in the rumour clinics, the goal was to create awareness on rumours to not being tricked by them; public inoculation also follows the same idea but with a different style. In biology, inoculation or vaccination refers to the process of training the immune system to produce anti-body by exposing it with a weakened virus. A similar philosophy can be fitting for public inoculation. Still, instead of training the immune system here the idea is to train the brain to build up resistance against rumours (Matz et al. 2017; Van Der Linden et al. 2017).

One of the challenging aspects of public inoculation is its implementation. One of the common strategies in this vein is to resort to media, elites, and thought-leaders to echo inoculation messages (Farrell et al. 2019). It has also shown that serious gaming is a practical approach to inoculate people against rumours (Roozenbeek and van der Linden 2019). Despite the success and effectiveness of public inoculation in countering rumours, it is still considered an experimental solution for rumour confrontation in limited contexts (Roozenbeek and van der Linden 2019). The expansion of this strategy would show its effectiveness in other domains and on a larger scale.

4.4 *Evaluation of Strategies*

This section first presents a framework for evaluating the counter-rumour strategies, followed by their assessment.

4.4.1 **An Evaluation Framework Based on the Spread of Infectious Diseases**

The spread of rumours bears many similarities to the evolution and transmission of contagious diseases (Kucharski 2016). Almost half a century ago, Goffman and Newill (Goffman and Newill 1964) directed attention to the analogy between the spread of infectious diseases and disseminating information (Daley and Kendall 1964). They argued that transmission of ideas does not need to be restricted to infectious diseases but is a more general process applied to many contexts. For example, the development of the psychoanalytic movement in the early twentieth century was no less an epidemic than the outbreak of influenza in 1917 and 1918 (Goffman and Newill 1964). This similarity between biological and intellectual epidemics is even caused the same modelling paradigm to be adopted to explain the dynamic of propagation (Daley and Kendall 1964; Moreno et al. 2004).

In epidemiology, control frameworks have proven successful for reining in epidemics (Daley and Gani 2001). It is composed of three mechanisms of (i) education, (ii) immunisation, and (iii) screening and quarantine. The first two are prevention measures that aim to minimise exposure to the disease and give a complete protection to a person against infection. At the same time, the third one has a more interventional nature to reduce the transmission rate.

Education is one of the simplest and cheapest ways to control epidemics by training the public about simple techniques such as wearing masks, washing hands, social distancing, and gargling to reduce the likelihood of exposure to the disease. It is mostly about raising awareness about dos and don'ts regarding a particular condition. For example, in the case of AIDS epidemics, the educational campaigns in February 1987 tried to discourage risk-prone behaviours such as unprotected sex or needle exchange for drug users. The campaign was successful by reducing the spread of the virus in countries where the state organised educational campaigns (Daley and Gani 2001; Leung and Nicoll 2010).

Immunisation through vaccination is one of the most influential and cost-effective strategies to control epidemics (Kato et al. 2011; Daley and Gani 2001; Magnusson 2017). It is referred to as “one of the great public health triumphs of all time” due to achieving landmark gains over a relatively short period. For example, in the case of smallpox, a worldwide vaccination campaign succeeded in eradicating the disease. For instance, the global immunisation program against diphtheria, pertussis, tetanus, poliomyelitis, measles and tuberculosis in 1974, immunised only 5% of the world's children. However, in less than 20 years, more than 90%

of the world's children had received BCG vaccine, and 75–85% had received immunisation against diphtheria, tetanus, pertussis, poliomyelitis and measles (Ruff 1999).

The third approach to control the epidemic is screening and quarantine. It is an interventional approach that plays a role after an epidemic has already started. It is a core public health approach as it can reduce and delay the disease's spread somewhat at the earliest stage. During the epidemics, after screening susceptible individuals, those that pose a risk are quarantined. "Many countries do not attempt these measures because of logistics, and cost-benefit considerations" (Leung and Nicoll 2010; Daley and Gani 2001; Magnusson 2017).

Due to the substantial similarity between the propagation of diseases and the information dissemination from one hand, and a comprehensive framework in disease eradication, this study proposes to adopt the same framework for the rumour confrontation. To this end, the past counter-strategies are set into the epidemic framework to understand which phase of the epidemic control is less emphasised in rumour confrontation.

4.4.2 Effectiveness of Strategies

This section evaluates the counter rumour strategies introduced in the previous section using the epidemic framework. The epidemic framework presents three approaches to control the spread of disease: (i) exposure minimisation, (ii) immunisation or vaccination, and (iii) reducing the transmission rate. In this section, we use the same framework to assess the goal of counter-rumour strategies. As Table 2 displays, rumour counter strategies pursue at least one of the epidemic control approaches.

The role of legislation mitigates the transmission rate and prevents further spread of rumours. It takes punitive measures against those who participate in rumour dissemination. Such measures would make people more careful and cautious about sharing information with their peers. In the political solutions, the goal is to cut the rumour from the source through negotiation by the main sponsors behind the rumours, so it belongs to the strategies with exposure minimisation view.

The A.I. techniques for the filtering and downgrading aim to reduce the visibility of rumours in the online social network by taking down rumour related contents and accounts; thus, they are considered strategies with exposure minimisation approach. If they are used to just flag the misleading contents, they also fall into the third category (reducing the transmission rate). The collaboration between online social media platforms and fact-checking institutes provides information about the truthfulness of the posts circulating on the media. Fact-checking the posts may dissuade people from sharing the news with their network. The collaboration can also improve the accuracy of filtering and downgrading algorithms. The other type of platforms' cooperation is with a media organisation. It would empower professional journalism in the digital era and reduce the likelihood of rumour emergence within news channels. The partnership between platforms and the scientific community would contribute to improving filtering and downgrading techniques

Table 2 Analysis of the quelling strategies against epidemic control framework

		Exposure minimisation	Giving complete protection (vaccination)	Reducing the transmission rate
Rumour counter strategies	Legislation			✓
	Political solution	✓		
	Filtering	✓		✓
	Downgrading	✓		✓
	Collaboration with fact-checking	✓		✓
	Collaboration for media literacy	✓		
	Collaboration in scientific project	✓		
	Service retirement	✓		✓
	Service modification	✓		✓
	New service development	✓		✓
	Ad-hoc fact-checking	✓		
	Post-hoc fact-checking			✓
	Rule of thumbs to deal with targeted rumours			✓
	Rumour clinics		✓	
	Public inoculation		✓	
	Educational policy		✓	
	Control centres			✓

which would eventually minimise the rumour exposure. Within the design-based strategies, the service retirement and the service modifications aim to reduce the rumour appearance and transmission likelihood. The new service development can also serve the same function; however, they are mostly for reducing the rumour transmission among the services that already exist.

The purpose of ad-hoc fact-checking is to prevent mistakes and false information before making public; therefore, it falls into the exposure minimisation category. Post-hoc fact-checking reviews and fixes incorrect information after they go public and decreases the likelihood of rumour transmission. The rule of thumb to deal with targeted rumours is a set of principles that a rumour audience could use to reduce rumour transmission. The rumour clinics and public inoculation try to create immunisation by teaching people how rumours deceive the mind. The purpose of the educational approach is to raise awareness and educate people. It works like a shield for the brain as it aims to train the brain not to be trapped by rumours. Control centres reduce the rumour transmission rate by filling the news channels gap by providing information about the floating rumours.

The introduced counter-measures are exercised intermittently over the past century. Despite tremendous efforts and developing all those strategies, diffusion of rumours not only has not shrunk but also escalated. The quick reactions to the sudden rise of rumours in different periods and the absence of a comprehensive plan are amongst the reasons for the failure of curbing the rumours. Besides, focusing on the rumour exposure minimisation and reducing the rumour transmission rate while neglecting the more effective immunisation approach is another reason for the failure of the current set of strategies against rumour dissemination.

5 Conclusion

The dissemination of large-scale rumour spreading regardless of the purpose could precipitate catastrophic repercussions. This research aims at addressing this challenge by posing three questions: what is a rumour? How does media facilitate its circulation? And how did past counter-rumour strategies perform?

We discuss rumour, Gossip, legend, propaganda, conspiracy theory, fake news, misinformation, and pseudoscience and analyse them from a process-based perspective. We infer that except Gossip and legend – which are fundamentally different regarding the context of emergence, content, and functionality – the rest of false and unverified information variations are other forms of rumour. We also argue legends are improbable to be harmful; gossips might be detrimental, but in small-scales; rumours might become extremely dangerous in large-scale, therefore it is of utmost importance to develop resilience against this phenomenon. The spread of rumours since the pre-printing press till the social media era is taken into account to respond to the second question. The focus was primarily on social media since the spread of rumours scaled-up, diversified, and accelerated during this period. For the third question, we investigated counter-rumour strategies. Based on their impact on different communication components, we classify them into three groups of sender-, channel-, and receiver-related methods. After introducing the strategies, they are assessed based on the proposed epidemic framework. In this framework, strategies are evaluated based on three criteria: exposure dissemination, giving complete protection, and reducing the transmission rate. The analysis shows that the past processes against rumours unevenly cover all aspects of the epidemic framework. Although vaccination is recognised as the most effective approach in controlling the epidemics, there is more emphasis on the other two techniques.

To effectively address the existing gaps in the tackling mass rumour spreading, what is essentially required is a long-term and comprehensive plan covering different variations of rumour spreading and incorporating all aspects of the epidemic framework. This plan mainly focuses on (i) immunisation approach due to its proven effectiveness and (ii) AI-based techniques due to the scale, scope, and speed of rumour spreading in online social media platforms. Besides, we should develop the

new generation of strategies for all variations of false and unverified information. Additionally, we must closely monitor social media mechanisms that contribute to the circulation of rumours on a large-scale, such as recommendation systems and social bots.

References

- Acerbi A (2019) Cognitive attraction and online misinformation. *Palgrave Commun* 5(1):1–7
- Allport GW, Postman L (1947) *The psychology of rumor*, 1st edn. Henry Holt and Company
- Al-Tabbaa O, Ankrah S (2019) Engineered university-industry collaboration: a social capital perspective. *Eur Manag Rev* 16(3):543–565
- Aral S (2020) *The hype machine: how social media disrupts our elections, our economy, and our health-and how we must adapt*. Currency
- Aral S, Eckles D (2019) Protecting elections from social media manipulation. *Science* 365(6456):858–861
- Archibong I (2018) API and other platform product changes. <https://developers.facebook.com/blog/post/2018/04/04/facebook-api-platform-product-changes/>
- Barabási AL et al (2016) *Network science*. Cambridge University Press, Cambridge
- Benkler Y, Faris R, Roberts H (2018) *Network propaganda: manipulation, disinformation, and radicalization in American politics*. Oxford University Press, Oxford
- Berkovsky S, Freyne J (2015) Web personalization and recommender systems. In: *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp 2307–2308. <https://doi.org/10.1145/2783258.2789995>
- Bodine-Baron EA, Helmus TC, Radin A, Treyger E (2018) *Countering Russian social media influence*. RAND Corporation, Santa Monica
- Bordia P, DiFonzo N (2004) Problem solving in social interactions on the internet: rumor as social cognition. *Soc Psychol Q* 67(1):33–49
- Bordia P, Kiazad K, Restubog SLD, DiFonzo N, Stenson N, Tang RL (2014) Rumor as revenge in the workplace. *Group Org Manag* 39(4):363–388
- Boutyline A, Willer R (2017) The social structure of political echo chambers: variation in ideological homophily in online networks. *Polit Psychol* 38(3):551–569
- Briones RL, Kuch B, Liu BF, Jin Y (2011) Keeping up with the digital age: how the American red cross uses social media to build relationships. *Public Relat Rev* 37(1):37–43
- Buringh E, Van Zanden JL (2009) Charting the “rise of the west”: manuscripts and printed books in Europe, a long-term perspective from the sixth through eighteenth centuries. *J Econ Hist* 69(2):409–445
- Chakravarthi P (1992) The history of communications-from cave drawings to mail messages. *IEEE Aerosp Electron Syst Mag* 7(4):30–35
- Christakis NA, Fowler JH (2009) *Connected: the surprising power of our social networks and how they shape our lives*. Little, Brown Spark, New York
- Christopherson KM (2007) The positive and negative implications of anonymity in internet social interactions: “on the internet, nobody knows you’re a dog”. *Comput Hum Behav* 23(6):3038–3056
- Cook J (2020) A history of FLICC: the 5 techniques of science denial. <https://www.skepticalscience.com/history-FLICC-5-techniques-science-denial.html>
- Cota W, Ferreira SC, Pastor-Satorras R, Starnini M (2019) Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Sci* 8(1):35
- Daley DJ, Gani J (2001) *Epidemic modelling: an introduction*. Cambridge University Press, Cambridge

- Daley DJ, Kendall DG (1964) Epidemics and rumours. *Nature* 204(4963):1118. <https://doi.org/10.1038/2041118a0>
- Dandekar P, Goel A, Lee DT (2013) Biased assimilation, homophily, and the dynamics of polarization. *Proc Natl Acad Sci* 110(15):5791–5796
- DiFonzo N (2009) *The watercooler effect: an indispensable guide to understanding and harnessing the power of rumors*. Avery, New York
- DiFonzo N (2018) Conspiracy rumor psychology. Conspiracy theories and the people who believe them
- DiFonzo N, Bordia P (2007a) Rumor psychology: social and organizational approaches. American Psychological Association
- DiFonzo N, Bordia P (2007b) Rumors influence: toward a dynamic social impact theory of rumor. *Frontiers of social psychology*. Psychology Press, New York
- DiFonzo N, Bordia P, Rosnow RL (1994) Reining in rumors. *Organ Dyn* 23(1):47–62. [https://doi.org/10.1016/0090-2616\(94\)90087-6](https://doi.org/10.1016/0090-2616(94)90087-6)
- DiFonzo N, Bourgeois MJ, Suls J, Homan C, Stupak N, Brooks BP, Ross DS, Bordia P (2013) Rumor clustering, consensus, and polarization: dynamic social impact and self-organization of hearsay. *J Exp Soc Psychol* 49(3):378–399
- Dowe T (1997): News you can abuse. <https://www.wired.com/1997/01/netizen-6/>
- DW: German justice minister to set up task force on Internet hate speech (2015). <https://www.dw.com/en/german-justice-minister-to-set-up-task-force-on-internet-hate-speech/a-18714334>
- Ellick AB, Westbrook A (2018a) Scopus vs. Web of Science vs. Google Scholar. <https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html>
- Ellick AB, Westbrook A (2018b) Operation InfeKtion: Russian Disinformation from Cold War to Kanye. <https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html>
- Fabry M (2017) Here's how the first fact-checkers were able to do their jobs before the Internet. <https://time.com/4858683/fact-checking-history/>
- Farrell J, McConnell K, Brulle R (2019) Evidence-based strategies to combat scientific misinformation. *Nat Clim Chang* 9(3):191–195. <https://doi.org/10.1038/s41558-018-0368-6>. <http://www.nature.com/articles/s41558-018-0368-6>
- Faye C (2007) Governing the grapevine: the study of rumor during world war ii. *Hist Psychol* 10(1):1
- Ferrara E, Varol O, Davis C, Menczer F, Flammini A (2016) The rise of social bots. *Commun ACM* 59(7):96–104
- Fine GA, Rosnow RL (1978) Gossip, gossipers, gossiping. *Personal Soc Psychol Bull* 4(1):161–168
- Flaxman S, Goel S, Rao JM (2016) Filter bubbles, echo chambers, and online news consumption. *Public Opin Q* 80(S1):298–320
- Fox J (2020) 'fake news'—the perfect storm: historical perspectives. *Hist Res* 93(259):172–187
- Franks B, Attia S (2011) Rumours and gossip as genres of communication. In: *The social psychology of communication*. Springer, pp 169–186
- Freelon D, Wells C (2020) Disinformation as political communication. *Polit Commun* 37(2):145–156
- Frenda SJ, Nichols RM, Loftus EF (2011) Current issues and advances in misinformation research. *Curr Dir Psychol Sci* 20(1):20–23
- Fung B, Garcia A (2019) Facebook has shut down 5.4 billion fake ac-counts this year. <https://edition.cnn.com/2019/11/13/tech/Facebook-fake-accounts/index.html>
- Funke D, Flammini D (2018) A guide to anti-misinformation actions around the world. <https://www.poynter.org/ifcn/anti-misinformation-actions/>
- Goffman W, Newill VA (1964) Generalization of epidemic theory: an application to the transmission of ideas. *Nature* 204(4955):225–228
- Google News Initiative. <https://newsinitiative.withgoogle.com/>

- Google: How Google Fights Disinformation (2018). https://storage.googleapis.com/gweb-uniblog-publish-prod/documents/How_Google_Fights_Disinformation.pdf
- Graves L (2016) Deciding what's true: the rise of political fact-checking in American journalism. Columbia University Press, New York
- Guess AM, Lerner M, Lyons B, Montgomery JM, Nyhan B, Reifler J, Sircar N (2020) A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proc Natl Acad Sci* 117(27):15536–15545
- Hansson SO (2017) Science denial as a form of pseudoscience. *Stud Hist Philos Sci A* 63:39–47
- Hansson SO (2018) Dealing with climate science denialism: experiences from confrontations with other forms of pseudoscience. *Clim Pol* 18(9):1094–1102
- Harvey D (2018) Serving healthy conversation. https://blog.twitter.com/en_us/topics/product/2018/Serving_Healthy_Conversation.html
- Hazenburg H, Van den Hoven M, Cunningham S, Alfano M, Ashgari H, Mumm EE, Fard AE, Rodriguez ET (2018) Micro-targeting and ICT media in the Dutch parliamentary system: Technological changes in Dutch democracy. Technical report. Delft University of Technology
- Hendricks VF, Vestergaard M (2019) Fact resistance, populism, and conspiracy theories. In: Reality lost. Springer, pp 79–101
- Howell L et al (2013) Digital wildfires in a hyperconnected world. *WEF Rep* 3:15–94(2013)
- Hussain M (2019) An Iranian activist wrote dozens of articles for right-wing outlets. But is he a real person?. <https://theintercept.com/2019/06/09/heshmat-alavi-fake-iran-mek/>
- Jannach D, Jugovac M (2019) Measuring the business value of recommender systems. *ACM Trans Manag Inf Syst (TMIS)* 10(4):1–23
- Kastrenakes J (2019) WhatsApp limits message forwarding in fight against mis-information. <https://www.theverge.com/2019/1/21/18191455/whatsapp-forwarding-limit-five-messages-misinformation-battle>
- Kato F, Tainaka KI, Sone S, Morita S, Iida H, Yoshimura J (2011) Combined effects of prevention and quarantine on a breakout in sir model. *Sci Rep* 1:10
- Knapp RH (1944) A psychology of rumour. *Public Opin Q* 8(1):22–37. <https://doi.org/10.1086/265665>
- Knopf TA (1975) Rumors, race, and riots. *Trans Publ*
- Koenig F (1985) Rumor in the marketplace: the social psychology of commercial hearsay. Auburn House Publishing Company
- Kucharski A (2016) Study epidemiology of fake news. *Nature* 540(7634):525–525
- Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D, Schudson M, Sloman SA, Sunstein CR, Thorson EA, Watts DJ, Zittrain JL (2018) The science of fake news. *Science (New York, NY)* 359(6380):1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lehmann S (2019) Fundamental structures in temporal communication networks. In: Temporal network theory. Springer, pp 25–48
- Lessenski M (2018) Common sense wanted: resilience to 'post-truth' and its predictors in the new media literacy index 2018. Technical report. Open Society Institute. https://osis.bg/wp-content/uploads/2018/04/MediaLiteracyIndex2018_publishENG.pdf
- Leung GM, Nicoll A (2010) Reflections on pandemic (H1N1) 2009 and the international response. *PLoS Med* 7(10):e1000346
- Li HOY, Bailey A, Huynh D, Chan J (2020) Youtube as a source of information on covid-19: a pandemic of misinformation? *BMJ Glob Health* 5(5)
- Lyons T (2018) Hard questions: what's Facebook's strategy for stopping false news? <https://newsroom.fb.com/news/2018/05/hard-questions-false-news/>
- Magnusson R (2017) Advancing the right to health: the vital role of law. World Health Organization, Geneva
- Marsden C, Meyer T, Brown I (2020) Platform values and democratic elections: how can the law regulate digital disinformation? *Comput Law Secur Rev* 36:105373
- Martin LJ (1982) Disinformation: an instrumentality in the propaganda arsenal. *Polit Commun* 2(1):47–64

- Matz SC, Kosinski M, Nave G, Stillwell DJ (2017) Psychological targeting as an effective approach to digital mass persuasion. *Proc Natl Acad Sci* 114(48):12714–12719
- Mazarr MJ, Casey A, Demus A, Harold SW, Matthews LJ, Beauchamp-Mustafaga N, Sladden J (2019) Hostile social manipulation present realities and emerging trends. Technical report. RAND National Defense Research Institute, Santa Monica
- McGonagle T (2017) “fake news” false fears or real concerns? *Neth Q Human Rights* 35(4):203–209
- Molina MD, Sundar SS, Le T, Lee D (2019) “fake news” is not simply false information: a concept explication and taxonomy of online content. *Am Behav Sci*:0002764219878224
- Moreno Y, Nekovee M, Pacheco AF (2004) Dynamics of rumor spreading in complex networks. *Phys Rev E* 69(6):066130
- Mosseri A (2017) Working to stop misinformation and false news. <https://newsroom.fb.com/news/2017/04/working-to-stop-misinformation-and-false-news/>
- O'Connor C, Weatherall JO (2019) The misinformation age: how false beliefs spread. Yale University Press
- Paris B, Donovan J (2019) Deep fakes and cheap fakes. Data & Society, New York
- Policy team: Update on Twitter's review of the 2016 US election (2018). https://blog.twitter.com/en_us/topics/company/2018/2016-election-update.html
- Ponting JR (1973) Rumor control centers. *Am Behav Sci* 16(3):391–401
- Posetti J, Matthews A (2018) A short guide to the history of ‘fake news’ and disinformation. *Int Center J* 7
- Romerstein H (2001) Disinformation as a kgb weapon in the cold war. *J Intel Hist* 1(1):54–67
- Roozenbeek J, van der Linden S (2019) The fake news game: actively inoculating against the risk of misinformation. *J Risk Res* 22(5):570–580
- Roth Y, Pickles N (2020) Updating our approach to misleading information. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html
- Ruff T (1999) Immunisation strategies for viral diseases in developing countries. *Rev Med Virol* 9(2):121–138
- Ruths D (2019) The misinformation machine. *Science* 363(6425):348–348
- Schmid P, Betsch C (2019) Effective strategies for rebutting science denialism in public discussions. *Nat Hum Behav* 3(9):931–939
- Schor EL (1995) Developing communality: family-centered programs to improve children's health and well-being. *Bull N Y Acad Med* 72(2):413
- Shao C, Ciampaglia GL, Varol O, Yang K, Flammini A, Menczer F (2018) The spread of low-credibility content by social bots. *Nat Commun* 9(1):4787. <https://doi.org/10.1038/s41467-018-06930-7>
- Shibutani T (1966) Improvised news. Ardent Media
- Standage T (2017) The true history of fake news. <https://www.1843magazine.com/technology/rewind/the-true-history-of-fake-news>
- Stella M, Ferrara E, De Domenico M (2018) Bots increase exposure to negative and inflammatory content in online social systems. *Proc Natl Acad Sci* 115(49):12435–12440
- Tsfati Y, Boomgaarden H, Strömbäck J, Vliegenthart R, Damstra A, Lindgren E (2020) Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis. *Ann Int Commun Assoc*:1–17
- Tufekci Z (2017) Twitter and tear gas: the power and fragility of networked protest. Yale University Press
- Uberti D (2016) The real history of fake news. https://www.cjr.org/special_report/fake_news_history.php
- Van Der Linden S, Maibach E, Cook J, Leiserowitz A, Lewandowsky S (2017) Inoculating against misinformation. *Science* 358(6367):1141–1142
- Van Prooijen JW (2018) The psychology of conspiracy theories. Routledge, London
- Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359(6380):1146–1151

- Wagh RB, Patil JB (2012) Web personalization and recommender systems: an overview. In: Proceedings of the 18th international conference on management of data, p 114
- Woolley SC, Howard PN (2018) Computational propaganda: political parties, politicians, and political manipulation on social media. Oxford University Press, Oxford
- Young S, Pinkerton A, Dodds K (2014) The word on the street: rumor, "race" and the anticipation of urban unrest. *Polit Geogr* 38:57–67