



UvA-DARE (Digital Academic Repository)

Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination

Hameleers, M.

DOI

[10.1093/ct/qtac021](https://doi.org/10.1093/ct/qtac021)

Publication date

2023

Document Version

Final published version

Published in

Communication Theory

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Hameleers, M. (2023). Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. *Communication Theory*, 33(1), 1-10. <https://doi.org/10.1093/ct/qtac021>

General rights

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

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination

Michael Hameleers *

Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands

*Corresponding author: Michael Hameleers, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Nieuwe Achtergracht 16, 1018 WV Amsterdam, Netherlands. Email: m.hameleers@uva.nl

Abstract

Although disinformation has become a popular concept, we lack an integrative conceptualization that connects the actors, intentions and techniques underlying deceptive information. In this article, we conceptualize disinformation as a context-bound deliberate act for which actors covertly deceive recipients by de-contextualizing, manipulating or fabricating information to maximize utility with the (targeted) outcome of misleading recipients. This conceptualization embeds fragmented accounts of disinformation in a networked and participatory information logic, and offers a comprehensive account of the conditions under which different actors may decide to deceive, how they deceive, and what they aim to achieve by deceiving recipients. Our conceptualization may inform (machine-learning) techniques to detect disinformation and interventions that aim to trigger suspicion by breaking through the truth-default state.

Keywords: deception, disinformation, fake news, misinformation, Truth-Default theory (TDT)

Scholarly attention to disinformation and related concepts has increased rapidly over the past years. Although the field is gravitating toward conceptual consensus on its key constituents (e.g., Chadwick & Staney, 2022; Freelon & Wells, 2020), unclarity on the context, intentions and actor-perspective of disinformation still prevails. Responding to calls in literature to view disinformation as a socio-political order that involves different sources, contextual factors and psychological mechanisms (e.g., Bennett & Livingston, 2018; Wardle & Derakhshan, 2017), this article aims to offer a systematic overview and forward-looking perspective of the context-bound nature of disinformation. To do so, we first of all review the state-of-the-art in disinformation literature and map the overlap and discrepancies between different conceptualizations foregrounded in extant literature. Based on this, we offer a revised conceptualization that integrates the context of disinformation's creation, dissemination and reception.

To offer a short minimal working definition of disinformation, we understand disinformation as the intentional creation and dissemination of false and/or deceptive information (e.g., Bennett & Livingston, 2018; Dan et al., 2021; Freelon & Wells, 2020; Hancock & Bailenson, 2021). As disinformation entails a targeted attempt to deceive and persuade recipients, it is important to focus on intentions. By revealing the covert intentions of disinformation agents, future research may be better able to detect goal-directed attacks on citizens' beliefs. Although extant research has mainly focused on the content and styles of false information (e.g., Damstra et al., 2021), mapping the context of intentional deception can offer a starting point for more refined distinctions between types of harmful content. In addition, a focus on intentions can help to uncover more comprehensive disinformation campaigns that go beyond isolated false messages. Finally, revealing the

intentions of disinformation should contribute to more resilience among recipients because specific knowledge on the context and motivations behind false information may help them to recognize and resist deliberate attacks on their beliefs.

The intentions and the deceptive nature of disinformation are contingent upon the communicator and their role in creating or disseminating disinformation. We therefore need to explicate the motives of sources that create or disseminate disinformation. By tracing disinformation back to the malign actors that orchestrated deceptive campaigns, and by exposing the dissemination context, we can start to unravel the hidden agendas and strategies behind the dissemination of false information, which may include exploiting the dynamics of many-to-many communication in digital information ecologies (e.g., Lukito et al., 2020). Focusing on the origins of disinformation may help to intervene at the roots of the problem, and prevent the further dissemination of misleading information through other communicators in the chain. In addition, mapping the intentions of malign actors and their targeted disinformation campaigns can offer a starting point for legal and policy interventions addressing the causes of deliberately false information.

Mapping the state-of-the-art: existing definitions and conceptualizations

The High Level Expert Group on Fake News and Online Disinformation of the European Commission (2018) has defined disinformation in the following way: "Disinformation (...) includes all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit" (Disinformation, 2018, p. 3). This rather broad definition emphasizes the deceptive and goal-directed nature of disinformation—which sets it apart

from misinformation that also refers to false or misleading information. In other scholarly definitions of disinformation, the notion of deception and the intention to cause harm is also featured (e.g., Chadwick & Stanyer, 2022; Hancock & Bailenson, 2021; Wardle & Derakhshan, 2017). For Wardle and Derakhshan (2017), the intention to harm an individual, group, organization or country is central—which assumes that disinformation is intended to exert a negative influence on the targeted recipients. In exploring the societal impact of deepfakes, Hancock and Bailenson (2021) note that “at the core of deepfakes is deception, which involves intentionally, knowingly, and/or purposely misleading another person.” (Hancock & Bailenson, 2021 p. 149). Disinformation agents may thus intend to create misperceptions, and their efforts are intended to result in a worldview or interpretation among recipients that is based on false beliefs.

In a more contextual definition of a “disinformation order,” Bennett and Livingston (2018) emphasize that the socio-political embeddings and regional contexts of disinformation are crucial to consider. Disinformation—across various contexts—can be seen as a disruptive form of communication that has to be understood in the setting of the media and political landscape it is created and consumed in. Hence, disinformation may thrive in contexts where distrust, populism and corruption are more central, and when the alternative and deceptive narratives forwarded in disinformation offer a credible interpretation of the socio-political world (also see Humprecht et al., 2020 for a discussion on regional resonance). When focusing on the political context of disinformation, it is important to note that although most research has focused on radical right-wing (populist) settings (e.g., Egelhofer & Lecheler, 2019; Marwick & Lewis, 2017), disinformation can also cling on to (radical) left-wing ideologies (e.g., Nikolov et al., 2021) or other issue positions. What matters is that disinformation may strategically exploit socio-political cleavages to create confusion, amplify conflicts and further divide societies.

Beyond contexts of deception, another element considered in disinformation literature is the extent to which deceptive information deviates from objective facts, empirical knowledge, and verified information (e.g., Lewandowsky, 2021; Tandoc et al., 2018; Van der Linden et al., 2021). Deception is arguably more credible when it uses some elements of the truth, and disinformation agents may capitalize on the truth bias by offering concrete linkages to real facts to frame deception. People generally have a hard time detecting deception (Allcott & Gentzkov, 2017; Edgerly et al., 2020; Oh & Park, 2021), and, unless suspicion is triggered, people are intrinsically wired to accept the honesty and accuracy of information, which is explicated in the Truth-Default Theory (TDT) (Levine, 2014). Based on this, disinformation agents may use techniques of deception and dissemination that make inauthentic content seem real (i.e., through deepfake techniques and astroturfing), which avoids the activation of suspected deception among recipients who do not have the resources to systematically scrutinize all arguments. Thus, although the TDT is a reception theory, it may help to explain how targeted disinformation is processed by recipients, and how malign actors may exploit a truth bias in an overloaded information setting by manipulating public opinion.

Based on a synthesis of the disinformation literature, and the overlapping elements in different definitions, we propose the following more inclusive working definition of disinformation: *Disinformation refers to all practices of intentionally*

creating or disseminating deceptive content to cause harm, sow discord, or create financial and/or political gain. Its practices of deception may range from the decontextualization of known facts to the fabrication of alternative narratives. With this definition, we aim to integrate as much of the consensus in extant literature as possible. Based on a closer mapping of the contexts of disinformation—in terms of the actors, platforms, strategies, techniques and envisioned goals—we will now work toward a revised conceptualization that focuses on the intentional dimension of disinformation embedded in contexts of deception.

The networked logic of disinformation’s origins and dissemination

When distinguishing between sources and actors of disinformation, we should acknowledge the hybrid nature and networked reality of disinformation campaigns operating through different actors with different intentions and strategies (Starbird, 2019; Zhang et al., 2021). More specifically, the complete cycle of a disinformation campaign may consist of multiple behaviors of different actors that create, disseminate and consume disinformation via different platforms. Among other actors, authentic social media users, opinion leaders, domestic political actors, foreign political actors, alternative or conventional news media and trolls or bots are all central players in the disinformation order, and the net influence of disinformation campaigns may be the consequence of organized and spontaneous interactions between these actors. Importantly, initiators of disinformation campaigns may strategically exploit this networked logic and anticipate on the involvement of other actors who may not be driven by the intention to deceive recipients—such as opinionated social media users or news media.

To offer an example of this networked and hybrid logic, disinformation campaigns originating from the Russian Internet Research Agency (IRA) may be disseminated by authentic ordinary citizens, which we define as people that interact with disinformation online or offline, but do not have a professional, political or other strategic aim in the creation or dissemination of deceptive information. The same message may also be picked up by domestic right-wing populist leaders, whose anti-establishment and cynical discourse aligns with disinformation campaigns of the IRA (Lukito, 2020). The mainstream media, in turn, may directly quote Tweets by successful populist politicians who engaged with disinformation from the IRA because such conflict-oriented statements resonate with their news values (Lukito et al., 2020). Therefore, we should understand the impact of disinformation through multiple (authentic and inauthentic) actors and sources as a networked or participatory logic (Starbird, 2019) – involving many different players who intentionally and unintentionally amplify the reach of deceptive information.

It is important to stress that—even if they are not necessarily initiators of disinformation themselves—social media platforms and big tech companies (i.e., Microsoft, Amazon, Google, Apple, Meta) play a crucial role as potential gatekeepers and amplifiers of disinformation (e.g., Kim et al., 2018). Embedded in an attention economy, these commercial platforms are endowed with the power to restrict freedom of speech whilst platforming unregulated forms of deceptive information. The decisions and behaviors of these companies are arguably motivated by profit, and are not driven by the

same gatekeeping principles as used by professional journalists. As argued by Kim et al. (2018), the system of digital media that offers a context for disinformation can be referred to as “stealth media” that enable disinformation campaigns through practices such as sponsored content and (micro) targeting. Platforms thus create a fertile opportunity structure for malign actors to disseminate falsehoods and maximize reach and profit.

Intentions motivating disinformation

Politically motivated disinformation: delegitimization or mobilization

Most literature on disinformation has connected its creation and dissemination with specific political agendas—such as the destabilization of foreign states (e.g., Wagnsson & Barzanje, 2019) or the legitimization of radical right-wing political agendas (e.g., Bennett & Livingston, 2018; Marwick & Lewis, 2017). In line with the delegitimizing nature of many disinformation campaigns, malign information can be disseminated to amplify polarized divides in targeted states, sow discord, or cultivate distrust in the established order (e.g., Bennett & Livingston, 2018). Such antagonistic narratives may be used by hostile foreign or domestic political actors to confound and pollute public discourse, impede citizens’ abilities to make well-informed political decisions, and hereby destabilize the political order (e.g., Flore, 2020). Destabilizing or delegitimizing narratives exploit socio-political weaknesses and vulnerabilities, such as cultural, ethnic or economic inequalities, uncertainties or (temporary) crises.

Political disinformation may also be created and disseminated to create a more favorable image of the disseminator—for example by controlling public opinion and pushing pro-state narratives in the international arena (Antoniades et al., 2010). Two prominent examples of such targeted attempts are Russian disinformation campaigns that promoted narratives falsely depicting their role in the MH17 disaster (Rietjens, 2019) or disinformation spread by the Chinese state to discredit protests in Hong Kong (La Cour, 2020). Politically motivated disinformation can be created and disseminated by many different actors—such as radical politicians, foreign states and domestic politicians. Its key aim is to strategically influence democracy or create momentum for specific political movements. By attacking or de-legitimizing opposed ideas or movements versus legitimizing congruent (alternative) ideas and movements, politically motivated disinformation can steer electoral outcomes and manipulate voters. Revealing these dynamics (i.e., through pointing at delegitimizing campaigns attacking established politicians during elections) may help voters to better discern between honest and dishonest persuasive campaigns encountered around key political moments.

Disinformation motivated by ideology

Ideologically-motivated disinformation is different from politically motivated disinformation as it is not directly driven by the aim to affect electoral outcomes, democracy or the political system as such, but rather aims to persuade recipients about ideological values and ideas. To offer an example, information about refugees entering a host country may intentionally be manipulated to exaggerate the threat posed by non-native citizens, hereby reflecting a nativist anti-immigration ideology.

Ideologically-motivated disinformation may deceive out of the conviction that the supported ideology is morally good and superior. Whereas politically motivated disinformation aims to make a political profit (i.e., gaining momentum for a party or issue position) or cause harm (i.e., creating cynicism toward the government), ideologically motivated disinformation aims to persuade recipients about certain ideas, values and/or identities. Although this goal may overlap with political intentions, its targeted consequences are different.

We should emphasize here that although most literature has looked at disinformation that emphasizes a radical right-wing ideological frame (e.g., Bennett & Livingston, 2018; Marwick & Lewis, 2017), disinformation can also be created and disseminated to amplify left-wing ideologies. The nature of disinformation’s ideological bias may largely depend on the (national) context in which it is embedded (Humprecht et al., 2020). Yet, irrespective of the ideological nature of disinformation, deceptive content may occur most at the fringes of the electoral spectrum—where the antagonism between established truth claims and alternative narratives is most pronounced (e.g., Potthast et al., 2018). As indicated by Potthast et al. (2018), 97% of disinformation statements identified by fact-checkers also contains hyper partisan (left- or right-wing) content. The identification of ideological motives may help to reveal which beliefs and identities are targeted by malign actors, and which segments of the audience are potentially reached by content that reassures or attacks their ideological beliefs.

Disinformation driven by financial gains

Disinformation may be disseminated to make a financial profit or to safeguard economic interests (e.g., Mulvey et al., 2015). This is also stressed in the definition of disinformation provided by the High Level Expert Group on Fake News and Online Disinformation of the European Commission, who refer to “causing harm” or “make profit” as the two central intentions underlying disinformation. The creation or dissemination of disinformation to gain financially can first of all be connected to the spread of deceptive narratives by online (alternative or hyper-partisan) information platforms, who may amplify deceptive content as such information sells better in a competitive market. In that sense, more clicks and engagement means more advertising revenues—and therefore an incentive to amplify disinformation that generates more engagement than authentic content. The same motivation may apply to platforms and big tech companies that platform disinformation motivated by the engagement it generates.

Although political disinformation is not likely to be initially created with the intention to make a profit, financial rewards may be a strong motivation to disseminate and offer a platform for disinformation from other sources without critically assessing its veracity—or warning recipients about potentially misleading content. Yet, there are instances where disinformation may also be created to gain profit, for example, in the case of health disinformation where deceptive content is created to induce anxiety which may stimulate purchasing behavior (i.e., medicines or alternative treatments that allegedly alleviate the threat).

Another case of financially-driven deception can be identified in the corporate sphere. In the corporate realm, for example, deception has been associated with fossil fuel companies and associated actors who aim to mislead the public about climate change developments (see Mulvey et al., 2015). Based on a comprehensive analysis of internal documents and

information revealed in lawsuits, [Mulvey et al. \(2015\)](#) conclude that a large number of fossil fuel companies have intentionally disseminated disinformation for multiple decades to sow confusion and obstruct policies to reduce human-induced emissions that cause global warming.

Using information manipulation theory to connect actors and intentions

By connecting the aforementioned actor perspectives and intentions of disinformation, we can tentatively conceptualize the deliberate nature of deception in disinformation as follows: Actors of disinformation deliberately engage in deception techniques to maximize actor-congruent utility or to create an interpretation of reality that is congruent with their perspective in a given deception context. The manipulation of information central to disinformation corresponds to the *covert* presentation of false information described in the Information Manipulation Theory (e.g., [McCornack et al., 2014](#)). According to Information Manipulation Theory, there are four main ways of deception, which include the selective sharing of relevant information, presenting irrelevant information, sharing false information or presenting information in a deliberately ambiguous and vague manner. The guiding principle behind all strategies is that manipulations remain covert, and that deception will succeed if the manipulations remain undetected by recipients.

The conditions motivating the goal-directed use of manipulation differ (see [McCornack et al., 2014](#)), but the reduction of cognitive load (efficiency) for initiators of dissemination and the opportunistic problem-solving nature of deception are important to consider in the realm of disinformation. More specifically, agents of disinformation may manipulate information strategically when deception (instead of telling the truth or offering recipients with accurate and complete information) is the most efficient strategy to solve problems. For example, when foreign states aim to increase cynicism among Western democracies, an efficient way to achieve this goal is to selectively share false or manipulated information that harms and delegitimizes established politicians and present such information as legitimate news.

Against this backdrop, deception in disinformation may be understood as offering a solution to problems in a way to maximize utility using the best fitting available arguments ([McCornack et al., 2014](#)). These best fitting arguments are selectively constructed, fabricated or manipulated in order to fit the aims of the communicator. This understanding of the roots of deception may help to explain under which conditions different actors may create or disseminate disinformation. Here, it has to be stressed that deception does not necessarily, and in most cases does not, include the complete fabrication of storylines: Easily available information that links to external reality is used to alter and doctor information that best fits the context-bound agenda of the disinformation actor.

Extrapolating these premises to the connection of actors and intentions, we can arrive at some theses about *why* different actors would deceive. A politician, for example, may be faced with the complex situation of a discrepancy between reality and their own ideology (i.e., crime rates and immigration are decreasing, whereas right-wing populists want to claim that rising immigration contributes to increasing crimes). In this context, deception can offer a viable way to reduce the

discrepancy and solve the problem in an opportunistic way: New (discrepant) information on reducing crime rates can be left out (violating the quantity assumption of cooperation, as stated in the Information Manipulation Theory), whilst better fitting and cognitively available associations about crime rates may be activated (i.e., making the argument that crimes are becoming increasingly more severe).

Techniques and practices of disinformation

There are many ways by which agents of disinformation may alter or fabricate content to create deceptive narratives: They may either stay close to reality and simply change some interpretations to strategically alter the meaning of information (i.e., a de-contextualization of facts) or create completely fabricated narratives that only reflect reality to a limited extent ([Lewandowsky, 2021](#); [Van der Linden et al., 2021](#)). In this section, we reflect on the most prominent techniques that can be used to create disinformation that mimics or mirrors elements of reality whilst altering the meaning of content in a goal-directed way. Here, we distinguish between techniques of creation and dissemination.

Mimicking news values and routines

One way of creating deceptive narratives is to mimic the routines, presentation styles, frames and values of established news as close as possible—a technique of deception that extant literature has also described as Fake News (e.g., [Allcott & Gentzkow, 2017](#)). Actors of disinformation can, for example, use deceptive (clickbait) headlines to stimulate engagement (e.g., [Wardle, 2017](#)), or use multimodal forms of doctoring (using decontextualized image-text pairs) to signal authenticity and closeness to reality (e.g., [Dan et al., 2021](#)). The degree to which they deviate from facticity can vary from the simple decontextualization of information, the manipulation or doctoring of existing content, to the complete fabrication of deceptive narratives that do no longer refer to facts as they happened. This technique is likely to be used by hyper partisan media or alternative media platforms that oppose the established press and sympathize with extremist voices ([Allcott & Gentzkow, 2017](#); [Heft et al., 2019](#)).

To make references to authentic news formats, deceptive content can be framed using the same styles and routines as journalistic content, for example, by integrating journalistic norms as objectivity, balance, expert-knowledge and verification. Dominant frames in journalistic content may be employed to further simulate authentic coverage (i.e., the human-interest frame or the responsibility attribution frame). Existing formats and presentation styles can be exploited to activate receivers' available schemata and news interpretations, creating an illusion of truth by activating associations related to honesty and authenticity.

Mimicking the vox populi and opinion leaders: bots, trolls and astroturfing

Another prominent technique of disinformation is to simulate the engagement and interactions of the everyday public on social media. The underlying strategy may be the same as inauthentic news: To make deceptive content seem like the type of information people receive on a daily basis from sources they trust and/or feel similar to. One prominent disinformation type in this regard is astroturfing: Coordinated campaigns

that employ real people to present certain beliefs, identities or interpretations imitating grassroots movements that signal widespread social support for these beliefs (Keller et al., 2020; Zhang et al., 2013). Astroturfing is deceptive because the displayed beliefs, identities or interpretations are not really held by the ones communicating them. Moreover, such techniques mislead recipients by creating the false belief that certain interpretations are held by a majority of people.

Using ordinary people as credible cues in disinformation campaigns has been found to be an effective way of mainstreaming disinformation campaigns (e.g., Lukito et al., 2020). Ordinary citizen cues can be both real social media profiles or automated behaviors—also referred to as social bots (e.g., Howard & Kollanyi, 2016). Bots are automated non-human profiles that can disseminate (dis)information autonomously (Zhang et al., 2013). The affordances of AI may be used to make a bot interact in seemingly human manners, and the automated behaviors of bots may be coordinated by humans, which has also been referred to as cyborgs (Zhang et al., 2013). Different from social bots, trolls are actual humans that are hired to disseminate false information and to engage with disinformation to increase the virality and prominence of dishonest information (e.g., Starbird, 2019). A prominent example is the use of troll armies as a coordinated influence campaign by the Russian IRA.

The use of inauthentic references to citizen sources can consist of different techniques employed by the different state and non-state actors described in the preceding sections. The role of inauthentic citizen cues in the disinformation order can best be understood at the intersection of sources and techniques: It reflects inauthentic coordinated behaviors in which techniques are employed to make it seem as if deceptive messages come from a certain source—whereas the actual source is kept hidden from the recipient. Bot or trolling activity may be detected by revealing its temporal and event-driven life cycle. More specifically, such coordinated inauthentic behavior typically involves profiles that only exist around an election or other key event, or correspond to a simultaneous peak in publishing and engagement behavior (Cao et al., 2014). Coordinated inauthentic behavior through bots and trolls may thus be identified by looking at the networked and event-driven nature of social media accounts (also see Keller et al., 2020).

Techniques of multimodal deception: multimodal doctoring, cheapfakes and deepfakes

Although astroturfing refers to deceptive techniques of dissemination used to simulate authenticity and natural interactions on social media, agents of disinformation may also employ techniques to make the false content itself seem real. Here, we argue that the addition of cue-rich media, such as visuals or videos, can enhance the perceived authenticity of deceptive content and herewith increase the likelihood that suspicion is not actively triggered—and thus that the truth-default state of message processing is not overruled (Levine, 2014). In the realm of multimodal disinformation, we specifically discern three techniques: Multimodal doctoring based on visual information, cheapfakes and deepfakes (also see e.g., Dan et al., 2021).

We regard multimodal doctoring as all practices involved in altering, decontextualizing or fabricating visual information to deceive recipients, and to change the meaning and interpretation of content through inauthentic image-text

pairings (Hameleers et al., 2020). It may first of all involve the deceptive pairing of a real image to a real text to deliberately decontextualize visual information and alter its meaning (i.e., showing real footage of a large number of boat refugees in a different context paired with a real story on immigrants entering a country with the intention to overstate the negative impact of an alleged “flood” of refugees). It may also involve the manipulation and fabrication of either the visual (i.e., photoshopping or cropping) or textual component of information with the intention to create a deceptive narrative.

Multimodal doctoring may be an effective technique as the specific qualities of visual information can add persuasive power to narratives (e.g., Geise & Baden, 2014; Powell et al., 2015). Specifically, visual information bears a stronger resemblance to the depicted reality than text—a quality of “indexicality” (Messaris & Abraham, 2001). Indexicality refers to visual information’s ability to offer a more direct picture or index of the real world than textual descriptions. This unique quality should make the activation of suspicion less likely for disinformation that relies on visual information. Moreover, as visual stimuli are more attention-grabbing and emotionally eliciting than texts (Powell et al., 2015), they may be more likely to be heuristically than systematically processed, which also corresponds to a lower likelihood that deception is spotted by recipients.

Cheapfakes or deepfakes profit from the same persuasive advantages. However, they mimic an even closer resemblance to or index of reality by manipulation both the audio and visual component of information, resulting in a dynamic rather than static deceptive narrative. Here, we regard deepfakes as synthetic videos that are made using techniques in deep learning (AI) to make authentic (political) actors express things they never said in real life (Chesney & Citron, 2018; Dobber et al., 2020; Paris & Donovan, 2020; Vaccari & Chadwick, 2020). The creation of such AI-generated synthetic videos is deceptive as the intention is to attack/delegitimize opponents or create support for/mobilize congruent viewpoints (also see Hancock & Bailenson, 2021).

Although deepfakes (still) require substantial resources and effort to create (see Dobber et al., 2020), there are also less resourceful techniques that can be used to manipulate audio-visual content—a technique of audio-visual manipulation referred to as cheapfakes (Paris & Donovan, 2020). Cheapfakes may involve a wide variety of low-effort techniques, such as cropping or editing video fragments (i.e., leaving out disconfirming evidence, placing words out of context, mixing audio with decontextualized footage). It may also involve the use of deceptive subtitles under a video originally in another language.

Irrespective of the techniques employed, the addition of visual elements in deceptive content may be an effective technique as it makes content look more realistic and closer to the external reality it deviates from (Dan et al., 2021). These techniques should consequentially be understood as deceptive acts by which agents of disinformation deliberately exploit the persuasiveness and credibility of authentic communication modes to motivate a truth-default state and to not trigger suspicion.

Disinformation as a context-bound phenomenon: a revised conceptualization

Based on a synthesis of extant literature on disinformation, and the mapping of intentions, source perspectives and techniques, we propose a comprehensive conceptualization of

disinformation that connects actors, intentions and techniques of deception. **Table 1** offers an overview of the different intentional forms of disinformation that can be distinguished, and includes suggestions on how these modes of deception may be detected by looking at the context of disinformation, such as the actors, ideological biases and platform embeddings of deceptive information.

In this revised conceptualization, we stress that the context of deception is crucial for understanding disinformation's intentions and techniques. Here, we understand the general context of disinformation as a problem-solving need for which a given actor (mis)uses information strategically to efficiently solve a given problem (McCornack et al., 2014). As highlighted in the previous sections of this article, the context of disinformation may be identified by revealing and reconstructing the intersections between actor-perspectives, applied techniques and their intended impact on targeted audiences. Concretely, the context of disinformation may contain information about the ideological affinity between senders and recipients of disinformation, the extent to which vulnerable audiences are targeted (also see Humprecht et al., 2020), the accounts used to disseminate disinformation (i.e., their history of malign communication and their interaction with other accounts), or details on the overall campaign of disinformation (i.e., the presence of activity peaks around key events and targeting strategies). These contextual factors together may offer a proxy or probability of intentional deception.

Against this backdrop, we conceptualize disinformation as *a context-bound deliberate act for which actors covertly deceive recipients by de-contextualizing, manipulating or fabricating information to maximize utility with the (targeted) outcome of misleading recipients*. Our revised conceptualization has six central pillars: (1) disinformation is context-bound; (2) it is a deliberate act; (3) it aims to deceive in a

covert manner; (4) it involves different techniques of altering information; (5) it is directed at maximizing personal or organizational profit or gain; and (6) its targeted outcome is to mislead recipients.

The forwarded conceptualization is actor-centered as its core aim is to explicate the motivated and strategical creation and dissemination of disinformation by malign actors. However, disinformation involves a close interaction between senders and receivers. Actors of disinformation, for example, may deliberately target recipients that they perceive to be most susceptible to their message, for example, as they already demonstrate moderate levels of doubt in the established order. By explicating the overlap and discrepancies between the actor-perspective (i.e., what is the intended impact of disinformation?) and recipients of disinformation (i.e., what is the realized impact of disinformation?), we may eventually arrive at new insights into the extent to which different agents of disinformation are successful, and among which segments of the audience their message is most disruptive.

Empirical applications of the revised conceptualization

All six elements of this definition may be reflected in empirical applications. First of all, by analyzing disinformation beyond individual messages and statements, empirical research can reveal the deceptive context and embedding of disinformation. An empirical example is the study of Lukito (2020), who explored the IRA's activity on different media platforms, also taking into account the temporal nature of disinformation campaigns. By reconstructing which platforms and sources are targeted by disinformation actors, and by focusing on the "peaks" in their activity around key events such as elections, empirical research can arrive at a closer assessment of the specific goals that are driving deception.

Table 1. Distinguishing between Intentions of Disinformation Campaigns and Suggestions on How to Detect Them Embedded in Their Communicative Context

Intentions	Detection
De-legitimization (sow discord, increase polarized divides, fuel cynicism, harm foreign countries, mal-information warfare)	<ul style="list-style-type: none"> • Look at broader influence campaigns: Who is harmed for what reason? • Identify which actors aim to sow discord • Reveal political background of source and compare this with targeted recipients • Focus on key moments of influence and peaks in disseminator's behaviors • Reveal (micro-)targeting strategies • Dissect techniques of manipulation
Mobilization (increase support or legitimize for (national) political agendas, consolidate power imbalances/silence the opposition, manage (inter)national image)	<ul style="list-style-type: none"> • Look at broader influence campaigns: Who seeks profit for what reason? • Identify which actors are seeking support • Reveal political background of source and match this with targeted recipients • Focus on key moments of influence and peaks in disseminator's behaviors • Reveal (micro-)targeting strategies • Dissect techniques of manipulation
Ideological motives (create support for alternative ideologies/share personal perspective on truth)	<ul style="list-style-type: none"> • Distill the ideological message from influence campaigns • Analyze the ways in which the ideological narrative is congruent with senders and targeted receivers • Relate construction of truth claims to conventional truth claims
Financial gain (increase advertising revenues, maximize profits, 'greenwashing' corporate image)	<ul style="list-style-type: none"> • Reveal the advertising strategies and sponsoring of influence campaigns • Assess the potential financial profits of deceptive information: What can the actor gain by creating or disseminating falsehoods? • Reveal the profile of the actor disseminating deceptive information and their overall corporate interests

To uncover the covert nature of disinformation campaigns, empirical research may reveal the funding structure of disinformation campaigns, and trace back disinformation messages to the original source that has created it. Agents of disinformation may hide their funding structure and the inauthenticity of disseminators by making deceptive campaigns seem like legitimate news media or interactions between ordinary citizens. A relevant example is the use of astroturfing, typically understood as the employment of paid communicators whose artificial involvement in public opinion may create the impression that an idea or belief is widespread in society (Zhang et al., 2013). As different tools have been developed to empirically identify such artificial involvement (e.g., Metaxas et al., 2011), future disinformation research can employ classification schemes to identify whether actors are covertly using artificial sources.

The covert nature of influence campaigns may also be revealed by exposing the funding and advertising structure of malign information. Ferreira (2021), for example, analyzed how the IRA's political ads exploited Facebook's targeted advertisement infrastructure, herewith revealing how polarizing content (i.e., ads on immigration) were strategically targeted at vulnerable segments of society. Concretely, by exposing the advertising strategies and targeted sub-populations of disinformation campaigns and by comparing the ideological slant of disinformation with the ideological profiles of targeted recipients, disinformation research can potentially reconstruct the motives driving the dissemination of false information.

Another concrete application and suggestion for empirical research is the use of different techniques to detect manipulated content. One empirical application is the use of machine learning to algorithmically detect deepfakes (Groh et al., 2021). At this stage, human coders that have background information on actors shown in the video, and knowledge about the discrepancy between the voiced statements and the actual profile of depicted persons in a deepfake, are found to outperform computer vision methods by a wide margin (Groh et al., 2021). Based on this, the detection of deception in deepfakes may be substantially improved when contextual features (i.e., the discrepancy between the ideological messages communicated and the known profile of the depicted actor) can be factored into models aiming to detect deceptive content.

Another prominent example used in disinformation research and employed by fact-checkers is reverse image search. As reverse image search reveals the context in which an image was used, the comparison between the depiction of the image in the manipulated versus original context may give an indication of how and why multimodal information was taken out of its context. Here, it is important to collect information on the accounts and sources that have shared the decontextualized image or video, as this can reveal the specific intentions behind the alteration and manipulation of materials. A key question in this regard is to assess whether the decontextualized application of original materials fits the ideological or strategic aims of a specific account or source.

Deception may also be detected more accurately by revealing if, and if so how, personal or organizational profit are maximized. For example, an analysis of the context of disinformation in terms of timing, targeting, and ideological slant may offer detailed insights into the gains that were envisioned by disinformation agents. As a concrete example, disinformation campaigns disseminated by troll armies peaking at election times (e.g., Lukito, 2020) and targeting citizens opposing

the status quo (Ferreira, 2021) can be interpreted as goal-directed attacks on established democracies in a context where there is something to gain from manipulation (i.e., changing election outcomes).

Although the examples offered here are not exhaustive, they illustrate how the analysis of the context and (social media) embeddings of disinformation may offer important information on the intentions driving disinformation. Here, an important suggestion is to not regard the aforementioned factors as distinct predictors, but to integrate them into a more comprehensive set of indicators that can reveal intentional deception, and enhance the precision of deception detection research. But which methods may be used to reveal these intentions?

Methodological approaches to focus on the embedding of disinformation

Although extant research identifying disinformation campaigns and statements has mostly been quantitative or automated, more interpretative methods can be used to reveal the embedding of disinformation campaigns. More concretely, digital ethnographies of communities and accounts that are likely to spread false information may help to explore how different communicators deceive, and how they interact with other accounts and platforms. Such interpretative methods allow for a closer analysis of disinformation-in-context, as they can explore the connection between disinformation actors and their deceptive discourse. The qualitative mapping of contextual factors that may predict intentional deception can subsequently be used in automated and quantitative analyses.

Although agents of disinformation may be a difficult to reach population, future research should acknowledge their perspective when aiming to understand the intentional dimension of disinformation. Interviews with creators and disseminators of deceptive content may help to understand the motivations behind disinformation from the perspective of creators themselves. For such research initiatives to be successful, a careful approach to recruit interviewees is important, which is further complicated by the fact that disinformation agents often act in a covert manner. To overcome this issue, former actors involved in disinformation campaigns (i.e., those that were involved in information manipulation in prior elections) may be approached—as they could also feel the urge to reveal the strategies behind malign influence campaigns they no longer support.

As these methods come with important challenges, existing research may also be adjusted to factor in the context of disinformation campaigns. Quantitative content analyses and automated detection tools may include variables on the actor-, platform- and campaign-level of disinformation. To offer a few examples, the context can be taken into account by coding which platforms were used, whether similar malign messages appeared on different platforms, what the ideological bias of the disseminator was, what the similarity of other messages disseminated by the actor was, and how the interaction of the actor's profile with other communicators and profiles looks like. In addition, analyses can reveal the hidden "black box" of campaign structures by looking at the (micro)targeting of the message through the used advertising strategies. Together, the use of such a set of contextual indicators may be used as a proxy mapping the likelihood that false

information was disseminated deliberately, and that specific intentions were driving deceptive communication.

Discussion

In this article, we forwarded a revised conceptualization that regards disinformation as a context-bound deliberate act for which actors covertly deceive recipients by decontextualizing, manipulating or fabricating information to maximize utility with the (targeted) outcome of misleading recipients. As an important contribution to the disinformation literature, we explicitly connect an actor-centered approach to a synthesis of intentions and techniques used to create and disseminate deceptive content. Here, informed by Information Manipulation Theory (McCornack et al., 2014), we postulate that malign actors may decide to deceive as this is a profit-maximizing strategy that is regarded as the most efficient manner to achieve their intended goal. When deceiving recipients, they may exploit the truth bias or truth-default in information processing (Levine, 2014). As recipients may be more likely to regard information as honest than dishonest, deceiving recipients may be an effective strategy for malign actors to reach their goals.

Based on the forwarded framework, we can formulate theses on the specific conditions under which different actors—such as foreign states, domestic political actors or corporations—may intentionally deceive recipients. Specifically, it can be expected that disinformation is created and disseminated to solve problems in an efficient manner, and employed when covert deception is the best-fitting strategy to maximize utility. Different techniques can be used to do so: Leaving out information to strategically guide perceptions of recipients, the complete fabrication of deceptive narratives (lies), or using available associations related to the phenomenon to create a deceptive narrative that fits the political goals of the disinformation actor whilst resonating with the associations that the recipient has available.

This latter technique resonates most with the ways in which disinformation has been described in extant literature: Disinformation agents typically do not create false narratives that are completely out of touch with reality, but alter, doctor or manipulate information to present deceptive narratives that still relate to elements of the truth (e.g., Lewandowsky, 2021; Tandoc et al., 2018; Van der Linden et al., 2021). When manipulating information, agents of disinformation may exploit a truth bias through the various techniques reviewed in this article. They may, for example, rely on inauthentic ordinary citizen cues to mimic the everyday exchanges between social media users, or rely on synthetic videos that make real persons say inauthentic things (deepfakes) to offer a fake index of reality (e.g., Dan et al., 2021). Exploiting truth biases, disinformation can be just as effective as authentic information as long as suspicion of deception is not actively triggered.

Next to improving detection deception research aiming to identify disinformation, the framework forwarded here can also be used to inform effective strategies to counter the negative impact of disinformation. Solutions such as fact-checking (e.g., Nyhan et al., 2019) or media literacy interventions (Vraga & Bode 2020) have been found to be effective under some conditions, but may be resisted by audience segments that are vulnerable to disinformation (e.g., Thorson, 2016). As the proposed conceptualization offers more insights into

the context of disinformation and indicators that increase the likelihood that deception is driven by specific intentions, recipients can be made more aware of signals of information manipulation and deception beyond message elements. More specifically, future empirical research may reveal how targeting strategies are used to deceive, and how malign actors strategically aim to manipulate audiences that are vulnerable to their messages. It may additionally reveal how disinformation messages are part of broader campaigns that peak around key periods of activity. By revealing these “covert” dynamics to citizens, interventions may teach citizens how to detect deception by critically scrutinizing disinformation in context.

Although these solutions are mostly aimed at instilling resilience on the side of recipients, the conceptualization also has implications for interventions targeting producers of disinformation and the platforms used to communicate falsehoods. Concretely, the conceptualization may inform deception detection tools that can arrive at an assessment of likely intentions, which may be used to hold actors of disinformation accountable for the dissemination of harmful content. For example, if microtargeting is misused to deceive vulnerable audience segments, interventions may restrict microtargeting strategies for certain malign purposes, and offer more transparency to users on what targeting strategies influence campaigns are based.

The forwarded conceptualization comes with limitations and challenges for future research. Most importantly, despite the empirical suggestions forwarded here, intentions are extremely difficult to detect empirically—especially when actors of disinformation are not directly involved in research. This implies that we have to identify intentions indirectly, and arrive at a likelihood that certain motivations were driving false information. Although this comes with a degree of uncertainty, revealing the context of disinformation campaigns—for example through mapping micro-targeting strategies, financing structures and peaks in malign accounts’ activity—can help to make better inferences about motivations.

Additionally, just like intentions are hard to identify, not all cases of malign information contain a clear deviation from a detectable and universal truth (see Woolley & Howard, 2018). To offer an example, truthful content may also be used to deceive—a strategy that has been referred to as mal-information (Wardle & Derakhshan, 2017). However, the proposed conceptualization takes into account that disinformation may involve different practices of altering, decontextualizing and fabricating information. The focus on intentions and contexts of deception makes deviations from universal truths less central, as the conceptualization may still “detect” malign information campaigns that decontextualize factually correct information to achieve a certain goal. Yet, the lack of a clear distinction between truths and untruths in many forms of malign information poses a real challenge to all deception detection studies.

Despite these limitations, we hope that the revised conceptualization offered in this article can inform theoretical consensus as well as future research agendas in the realm of deception detection (i.e., automated tools to detect disinformation), the psychology of disinformation, and actor- or network oriented perspectives in disinformation campaigns (i.e., how is disinformation spread in hybrid information ecologies?). At the very least, we hope to offer a conceptual contribution by emphasizing how disinformation needs to be studied as a context-bound phenomenon that involves the

deliberate acts of disinformation agents who employ different techniques with the intention to deceive recipients as efficiently and credibly as possible.

Acknowledgements

Dr Michael Hameleers (PhD, University of Amsterdam) is an Assistant Professor of Political Communication at the Amsterdam School of Communication Research (ASCoR). His research interests include (right-wing) populism, disinformation, and selective exposure.

References

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Antoniades, A., Miskimmon, A., & O'Loughlin, B. (2010). *Great power politics and strategic narratives*. The Centre for Global Political Economy. <https://doi.org/10.4324/9781315770734>
- Bennett, L. W., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Cao, Q., Yang, X., Yu, J., & Palow, C. (2014). Uncovering large groups of active malicious accounts in online social networks. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (pp. 477–488).
- Chadwick, A., & Stanyer, J. (2022). Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: Toward a holistic framework. *Communication Theory*, 32(1), 1–24. <https://doi.org/10.1093/ct/qtab019>
- Chesney, R., & Citron, D. K. (2018). *Deep fakes: A looming challenge for privacy, democracy, and national security* (California Law Review Research Paper No. 692). Retrieved from <https://ssrn.com/abstract=3213954>. <https://doi.org/10.2139/ssrn.32139>
- Dan, V., Paris, B., Donovan, J., Hameleers, M., Roozenbeek, J., van der Linden, S., & von Sikorski, C. (2021). Visual mis- and disinformation, social media, and democracy. *Journalism & Mass Communication Quarterly*, 98(3), 641–664. <https://doi.org/10.1177/10776990211035395>
- Damstra, A., Boomgaarden, H. G., Broda, E., Lindgren, E., Strömbäck, J., Tsfati, Y., & Vliegenthart, R. (2021). What does fake look like? A review of the literature on intentional deception in the news and on social media. *Journalism Studies*, 22(14), 1947–1963. <https://doi.org/10.1080/1461670X.2021.1979423>
- Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. H. (2020). Do (microtargeted) deepfakes have real effects on political attitudes? *The International Journal of Press/Politics*, 26(1), 69–91. <https://doi.org/10.1177/1940161220944364>
- Edgerly, S., Mourão, R. R., Thorson, E., & Tham, S. M. (2020). When do audiences verify? How perceptions about message and source influence audience verification of news headlines. *Journalism & Mass Communication Quarterly*, 97(1), 52–71.
- Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, 43(2), 97–116.
- Ferreira, R. R. (2021). Liquid disinformation tactics: Overcoming social media countermeasures through misleading content. *Journalism Practice*, 1–22. <https://doi.org/10.1080/17512786.2021.1914707>
- Flore, M. (2020). *Understanding citizens' vulnerabilities (II): From disinformation to hostile narratives*. European Commission. <https://doi.org/10.2760/271224>
- Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, 37(2), 145–156. <https://doi.org/10.1080/10584609.2020>
- Geise, S., & Baden, C. (2014). Putting the image back into the frame: Modeling the linkage between visual communication and frame-processing theory. *Communication Theory*, 25(1), 46–69. <https://doi.org/10.1111/comt.12048>
- Groh, M., Epstein, Z., Firestone, C., & Picard, R. (2021). Deepfake detection by human crowds, machines, and machine-informed crowds. *Proceedings of the National Academy of Sciences*, 119(1) <https://doi.org/10.1073/pnas.2110013119>
- Hameleers, M., Powell, T. E., van der Meer, G. L. A., & Bos, L. (2020). A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals disseminated via social media. *Political Communication*, 37(2), 281–301. <https://doi.org/10.1080/10584609.2019.1674979>
- Hancock, J. T., & Bailenson, J. N. (2021). The social impact of deepfakes. *Cyberpsychology, Behavior, and Social Networking*, 24(3), 149–152. <http://doi.org/10.1089/cyber.2021.29208.jth>
- Heft, A., Mayerhöffer, E., Reinhardt, S., & Knüpfer, C. (2019). Beyond Breitbart: Comparing right-wing digital news infrastructures in six Western democracies. *Policy & Internet*, 12(1), 20–45. <https://doi.org/10.1002/poi3.219>
- High Level Expert Group on Fake News and Disinformation (2018). A multi-dimensional approach to disinformation: Report of the independent high level group on fake news and online disinformation. *European Commission*. <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>
- Howard, P., & Kollanyi, B. (2016). Bots, #Strongerin, and #Brexit: Computational propaganda during the UK-EU referendum. *SSRN Electronic Journal*, 1–6. <https://doi.org/10.2139/ssrn.2798311>
- Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, 25(3), 493–516. <https://doi.org/10.1177/1940161219900126>
- Keller, F. B., Schoch, D., Stier, S., & Yang, J.-H. (2020). Political astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication*, 37(2), 256–280. <https://doi.org/10.1080/10584609.2019.1661888>
- Kim, Y.-M., Hsu, J., Neiman, D., Kou, C., Bankston, L., Kim, S.-Y., ... Raskutti, G. (2018). The stealth media? Groups and targets behind divisive issue campaigns on Facebook. *Political Communication*, 35(4), 515–541. <https://doi.org/10.1080/10584609.2018.1476425>
- La Cour, C. (2020). Theorising digital disinformation in international relations. *International Politics*, 57(4), 704–723. <https://doi.org/10.1057/s41311-020-00215-x>
- Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33(4), 378–392. <https://doi.org/10.1177/0261927X14535916>
- Lewandowsky, S. (2021). Conspiracist cognition: chaos, convenience, and cause for concern. *Journal for Cultural Research*, 25(1), 12–35. <https://doi.org/10.1080/14797585.2021.1886423>
- Lukito, J. (2020). Coordinating a multi-platform disinformation campaign: Internet Research Agency activity on three U.S. social media platforms, 2015 to 2017. *Political Communication*, 37(2), 238–255. <https://doi.org/10.1080/10584609.2019.1661889>
- Lukito, J., Suk, J., Zhang, Y., Doroshenko, L., Kim, S. J., Su, M.-H., ... Wells, C. (2020). The wolves in sheep's clothing: How Russia's Internet Research Agency Tweets appeared in U.S. news as vox populi. *The International Journal of Press/Politics*, 25(2), 196–216. <https://doi.org/10.1177/1940161219895215>
- Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online (pp. 1–104). *Data and Society Research Institute*. Retrieved September 18, 2018, from <https://datasociety.net/output/media-manipulation-and-disinfo-online/>
- McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A., & Zhu, X. (2014). Information Manipulation Theory 2 (IMT2): A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology*, 33(4), 348–377.

- Messariss, P., & Abraham, L. (2001). The role of images in framing news stories. In Reese S. D., Gandy O.H. & Grant A. E. (Eds.), *Framing public life* (pp. 215–226). Erlbaum.
- Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011). How (not) to predict elections. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)* (pp. 165–171). <https://doi.org/10.1109/PASSAT/SocialCom.2011.98>
- Mulvey, K., Shulman, S., Anderson, D., Cole, N., Piepenburg, J., & Sideris, J. (2015). *The climate deception dossiers: internal fossil fuel industry memos reveal decades of corporate disinformation*. Retrieved 4 October 2021 from <https://apo.org.au/node/55839>
- Nikolov, D., Flammini, A., & Menczer, F. (2021). Right and left, partisanship predicts (asymmetric) vulnerability to misinformation. *Harvard Kennedy School (HKS) Misinformation Review*, 1(7), 1-13. <https://doi.org/10.37016/mr-2020-55>
- Nyhan, B., Porter, E., Reifler, J., & Wood, T. J. (2019). Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior*, 42(3), 939–960. <https://doi.org/10.1007/s11109-019-09528-x>
- Oh, Y. W., & Park, C. H. (2021). Machine cleaning of online opinion spam: Developing a machine-learning algorithm for detecting deceptive comments. *American Behavioral Scientist*, 65(2), 389–403. <https://doi.org/10.1177/0002764219878238>
- Paris, B., & Donovan, J. (2020). Deepfakes and cheapfakes: The manipulation of audio and visual evidence. *Data & Society Report*. <https://datasociety.net/library/deepfakes-and-cheap-fakes/>
- Pothast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2018). A stylometric inquiry into hyperpartisan and fake news. *Proceedings of the 56Th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. <https://doi.org/10.18653/v1/p18-1022>
- Powell, T. E., Boomgaarden, H. G., De Swert, K., & de Vreese, C. H. (2015). A clearer picture: The contribution of visuals and text to framing effects. *Journal of Communication*, 65(6), 997–1017. <https://doi.org/10.1111/jcom.12184>
- Rietjens, S. (2019). Unraveling disinformation: The case of Malaysia Airlines flight MH17. *The International Journal of Intelligence, Security, and Public Affairs*, 21(3), 195–218. <https://doi.org/10.1080/23800992.2019.1695666>
- Starbird, K. (2019). Disinformation's spread: bots, trolls and all of us. *Nature*, 571(7766), 449.
- Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2018). Defining “Fake News.” *Digital Journalism*, 6(2), 137–153. <https://doi.org/10.1080/21670811.2017.1360143>
- Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, 33(3), 460–480.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- Van der Linden, S., Panagopoulos, C., Azevedo, F., & Jost, J. T. (2021). The paranoid style in American politics revisited: An ideological asymmetry in conspiratorial thinking. *Political Psychology*, 42(1), 23–51. <https://doi.org/10.1111/pops.12681>
- Vraga, E. K., & Bode, L. (2020). Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation. *Political Communication*, 37(1), 136–144. <https://doi.org/10.1080/10584609.2020.1716500>
- Wagnsson, C., & Barzanje, C. (2019). A framework for analysing antagonistic narrative strategies: A Russian tale of Swedish decline. *Media, War and Conflict*, 1–19. <https://doi.org/10.1177/1750635219884343>
- Wardle, C. (2017). Fake news. It's complicated. *First Draft*. Retrieved January 22, 2019 from <https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79>
- Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. *Council of Europe report*. <http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf>
- Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.
- Zhang, J., Carpenter, D., & Ko, M. (2013). Online astroturfing: A theoretical perspective. In *AMCIS 2013 Proceedings*.
- Zhang, Y., Lukito, J., Su, M.-H., Suk, J., Xia, Y., Kim, S. J., Doroshenko, L., & Wells, C. (2021). Assembling the networks and audiences of disinformation: How successful Russian IRA Twitter accounts built their followings, 2015–2017. *Journal of Communication*, 71(2), 305–331. <https://doi.org/10.1093/joc/jqaa042>