3 A threat to democracies?

An overview of theoretical approaches and empirical measurements for studying the effects of disinformation

Wolf J. Schünemann

Disinformation is everywhere. This short and simple insight does not only hold as a concise summary of the widely perceived fundamental problems with (political) information provision in a digitally transformed public sphere. It is also not meant as a fatalistic reaction to the sometimes alarmist warnings that characterize the relevant debate. Disinformation, understood as deliberately spread false information and rumors, has developed into a major concern in modern, highly connected societies. This phenomenon is very visible and can be studied as under a magnifying glass, albeit at a global scale, given the current so-called “infodemic”, i.e. the spread of disinformation regarding COVID-19, its origins, its spread and effects, or suitable treatments (for an overview, cf. Ball and Maxmen 2020). Disinformation even stands out as a widely perceived threat to liberal democracies and their institutions, as these seem particularly vulnerable to manipulative information operations due to their liberal stance toward media freedom and other regime-specific features. Given the unheard-of spread of disinformation – even the emergence of “the disinformation order” (Bennett and Livingston 2018) – induced by the digital transformation, the issue has developed into a widely studied subject in the field of political communication and media studies.

However, disinformation is neither confined to domestic politics nor constrained by disciplinary boundaries in academic discourse. On the contrary, the most intensely debated political information operations in recent years, namely during the Ukrainian crisis in 2014, the British referendum in 2016, and the US presidential election campaign in the same year, were at least partly ascribed to foreign, specifically Russian activity (Maréchal 2017; Pomerantsev 2014). Consequently, the entire problem needs to be viewed in terms of domestic as well as international conflict and has developed into one of the core threats in cyber security discourse and strategy. The current debate on the COVID-19 “infodemic” (Strick et al. 2020), which has seen (dis)information operations seemingly constituting a preferential mode of rivalry between the great powers, is likely to corroborate these trends.

While recent events in international politics have thus raised the awareness of information operations as an essential component of hybrid warfare and produced anxiety among actors in the security and defense sectors (Lanoszka 2019: 228),

there are growing uncertainties about adequate responses, in particular for democratic governments. In international conflict constellations, their “bias against [media] control” (McQuail 2008: 234), applied to the internet, might disadvantage democratic systems compared to autocratic states that protect their so-called information sovereignty by restricting internet freedom (Jamieson 2018: 11; Lanoszka 2019: 228; Omand 2018: 12; Pope 2018: 36) while presumably expending considerable effort on information operations at the domestic and international levels (King et al. 2017). Realistic accounts from international security studies have articulated this asymmetrical vulnerability assessment at the expense of liberal democracies very clearly without deriving any decisive strategy for counteraction (Goldsmith and Russel 2018) or effective deterrence (Goldsmith 2016; Shackelford et al. 2016: 666). In terms of media governance, trends toward more restrictive measures of online control can already be observed even in democratic states (Freedom House 2018).

All in all, disinformation stands out as a major threat to liberal democracies and national security alike. It is likely to appear as one of the crucial ingredients of future securitizing moves by respective political actors across the world. As for cyber security in general, it seems fair to say that threat perceptions in the field are based on a high level of uncertainty. Thus, for disinformation the same critical question needs to be posed as for other cyber threats: Are threat perceptions appropriate or exaggerated? To answer this question, I combine different perspectives: Theoretical reflections mostly developed in the field of cyber security, and empirical insights mainly provided by scholars of political communication and media.

The remainder of the chapter is structured as follows: After giving a basic definition of disinformation in the following section, I present a knowledge-based concept of (dis)information operations, followed by a critical review of threat perceptions as expressed in the core policy documents issued by major democracies and international organizations. Then, I contrast threat perceptions with the evidence as provided by empirical research. In this section, I differentiate three levels of analysis at which effects can be assessed: The micro, meso, and macro levels. Given the lack of knowledge on the actual impact of alleged disinformation and the obvious ambiguities in its interpretation and attribution, I conclude with a call for caution. Increased levels of uncertainty in digital communication must not cause us to stumble into a new phase of international threat politics and the securitization of cyberspace with potentially detrimental effects on liberal democratic values and international peace.

Disinformation as a concept and as a new element of international threat politics

Disinformation is a complex term. After all, in political communication there are probably not many statements made by any strategically motivated actors that would not be depicted as disinformation by any of their adversaries, if only for

the sake of casting doubt on the other party's credibility. The somewhat turbulent career of the compelling term fake news, often used synonymously with disinformation (Allcott and Gentzkow 2017), is illustrative in this regard. Following its widespread use in public discourse in the wake of major political events of the year 2016 (Brexit referendum, US elections), and as one of the main terms used to describe the allegedly emergent age of post-truth (Harsin 2015), it has suffered from its ubiquitous use as a weapon in political conflict. Today, it is almost disqualified for use in academic discourse (Vosoughi et al. 2018: 1146; however, used in Lazer et al. 2018).

In contrast, disinformation as preferred in scholarly discourse is a more technical term that refers to information that is misleading by design (European Commission 2018a: 11). Disinformation is classified as such on the basis of the function ascribed to it. Specifically, this means that disinformation needs to have an actor behind it who intentionally produces and disseminates untrue or incorrect information. Therefore, it can be differentiated from all unintentionally misleading information, so-called misinformation, as well as from other forms of distorted facts that are not meant to mislead but to produce amusement or deeper insights such as jokes or satire (Allcott and Gentzkow 2017). Disinformation can also be related to other terms associated with information operations that are more commonly found in strategic studies, such as propaganda or public diplomacy. Disinformation overlaps with propaganda but is not the same. What both terms have in common is purposeful deception in the form of misleading or extremely one-sided information. There is a difference, however, in that propaganda includes other means of strategic communication that are not falsifying. Moreover, propaganda according to a traditional understanding is bound to government actors in totalitarian or authoritarian regimes and mainly oriented toward the domestic population. At the other end of the spectrum, a more traditional, realistic understanding of public diplomacy also focuses on state-led informational activity toward foreign publics. While newer conceptions of public diplomacy include a broader range of actors and the use of digital networks ("public diplomacy 2.0", Tago 2018: 457), communicating false information does not fall under their definition, as this would conflict with the supposed soft power goals also pursued by democratic regimes. However, in the face of current heightened geopolitical tensions, public diplomacy risks being regarded as a euphemism for international disinformation campaigns, similar to its perception as an "*outgrowth* of propaganda" during the Cold War (Tago 2018).

Given the difficulties that arise when categorical distinctions made in abstract terms are applied to real-world phenomena, it is not surprising that disinformation can also become blurry when it is applied to empirical reality. Aside from the question of what is objectively true and what is false (Vosoughi et al. 2018), it remains difficult to attribute information operations and to find evidence for the intent to mislead. While uncertainty has always been a constitutive element of international relations, given the porosity of national borders especially with regard to information flows in cyberspace and the digitally enhanced means of so-called "hybrid warfare", ambiguity stands out even more clearly as a constant

feature in discussions on disinformation. As in cyber security more generally, ambiguity extends not only to the attacker and their intent but also to the actual quality and effects of an attack and the appropriate reactions to it.

Overall, it seems fair to assume that the issue of disinformation has caused us to enter a new phase of threat politics (Dunn Cavelty 2008) in which we observe a new quality of securitization in cyber security discourse. Even the notion of cyber war, which is compelling in the use of the “told catastrophes” in cyber security and critically discussed in scholarly discourse, has reappeared in a new form, now fed by anxiety about disinformation as a threat to democracy (Jamieson 2018: 7). With information technology and the allegedly threatened public discourse, two issues of high complexity are combined in the most recurrent threat perceptions around disinformation. In public use, this new version of the told catastrophes in the wider cyber security discourse might serve as the grounds for even more urgent reactions than the stories of earlier days. While those have mostly “remained just that – scenarios” (Dunn Cavelty 2008: 3), the new cyber doom scenarios of disinformation are currently happening or have already become reality, e.g. with Donald Trump in office and the United Kingdom having left the EU.

Disinformation and the question of knowledge

As in cyber security more generally and in current debates about disinformation and its alleged threats to democracy more specifically, actors tend to fill the voids of uncertainty with preexistent elements of knowledge that help produce consistent narratives even without actual empirical evidence. Knowledge is thus an important factor for understanding threat perceptions and political responses toward an alleged attacker and the intentions behind an attack. Moreover, knowledge itself can be seen as under attack by disinformation (Farrell and Schneier 2018). When taking a knowledge-oriented perspective on disinformation, it seems important to avoid two flaws that are frequently evident in extant literature. The first is about attribution, the second about the conception of (relevant) knowledge itself.

First, as for cyberattacks, it is often difficult to know the responsible parties and the original source of a piece of malign disinformation. Even if an act can be traced back to a certain source, it might be impossible to find proof of coordinated activity, e.g. by a foreign power. Thus, attribution often relies on *cui bono* assessments: “Both domestic and foreign disinformation aim to disrupt the institutional order, undermine politicians, stir anti-refugee sentiments and create confusion around elections” (Bennett and Livingston 2018: 130). Such forms of “meaning-making” – in public and even academic attribution (Egloff 2020) – seem particularly problematic in the field of disinformation, as its effects are so unclear, probably even for a potential malevolent actor (Thornton and Miron 2019: 263, 269–270). But as *cui bono* does not work for the observer, it maybe does not for the alleged perpetrator either. Moreover, when a *cui bono* logic is applied, the assumed effects of an information operation become part of the criteria for detection and evaluation (see also Jamieson 2018: 144). Thus, the entire threat assessment risks becoming

circular, if the alleged effects are a cornerstone of attribution while the assessment of the effects depends on the attribution.

Second, when assessing the societal effects of disinformation, it seems not very helpful to follow individualistic conceptions of knowledge familiar from electoral studies and major parts of political sociology, where knowledge is understood as personal accumulations of pieces of information an individual is exposed to (Lupia and McCubbins 1998; for a critical discussion see Schünemann 2014, 2018). First, “information by itself usually has no value: it is a raw material that gains value if further processed in specific ways and if meaning and a certain quality are attached to it” (Dunn Cavelty 2008: 15). Knowledge does not belong to the features of an individual (voter) but is produced and processed in discourses. Thus, second, information is “consumed” by individuals only through these collectively built filters of perception. This would necessitate the introduction of more macro-oriented perspectives on political discourse in a target society. From a macro-level knowledge-oriented perspective on disinformation, however, it seems problematic to assume that informational flows are simply discharged into a more or less helpless discourse. In analogy to cyberattacks, it would seem more appropriate to expect successful disinformation campaigns exploiting vulnerabilities already built into the target system (here the discourse), especially if campaigns are steered by a foreign power.

Farrell and Schneier (2018) have introduced a theoretical approach to assessing the vulnerabilities of political systems with respect to information operations from a comparative perspective. Among the various theoretical explorations of the issue, theirs comes closest to the knowledge-oriented macro-perspective, this chapter posits, as it is based on a concept of knowledge orders, namely the distinction between regime-specific stocks of common and contested knowledge, with disinformation campaigns being understood as knowledge attacks. Despite the important re-orientation that the authors present, the underlying conceptual dichotomy of knowledge types seems still too rough so that “common political knowledge” is more or less indistinguishable from other very general categories like trust in institutions or the basic requirement of support for the political system according to functionalist theory, while “contested political knowledge” comprises the broad range of knowledge elements floating around and being processed in public discourse. Consequently, the distinction does not seem truly open to more complex social constructivist conceptions and studies of social knowledge orders. Moreover, the authors do not provide any guidance on how to study the effects of disinformation as knowledge attacks.

Disinformation is all around – the threat perception

Before turning to a discussion of the effects of disinformation, this section sheds light on how the issue has influenced key actors’ foreign and security policies on the international scene. How disinformation has developed into a major concern for modern societies can be easily illustrated by examining recent, relevant policy documents published by national state governments and international

organizations. Thus, first, I present a critical review of core documents. I selected security strategies and reports published by the governments or dedicated committees of the United States, the United Kingdom, Germany, and at EU level. In accordance with my interest in threat politics, I build on a social constructivist meta-theory and the respective ontology as developed by Dunn Cavelty (2008: 30), which views threat frames as the central analytical concept for identifying the most frequently recurrent interpretive schemes or frames that enshrine ideas and beliefs regarding both the problem definition (“diagnostic framing”) and the problem solution (“prognostic framing”). This allows a contrastive picture to be drawn initially that can later be checked against the empirical “evidence” that studies on political communication provide.

Examples from core policy documents

Disinformation has made it into the security strategies of major democracies (Omand 2018: 8). The 2017 “US National Security Strategy” mentions disinformation as one of the primary threats in cyberspace, as “[m]alicious state and non-state actors use cyberattacks for extortion, information warfare, disinformation, and more”. Attacks are seen as able to “undermine faith and confidence in democratic institutions” (White House 2017: 31). Germany’s Cyber Security Strategy also lists disinformation in its threat analysis and explicitly highlights the potential dangers disinformation poses for liberal societies and the democratic order. In the so-called “White book on Security Policy and the Future of the German Army”, the explicit association with “elements of hybrid warfare” (Weißbuch: 37) is illustrative of a prognostic framing, as incidents are interpreted as malicious activity to be countered by security political measures. This notion is even more obvious in the Annual Report 2016–2017 issued by the Intelligence and Security Committee of the British Parliament, which lists the core elements of the foreign disinformation campaigns it discusses (“generally undermining the integrity of the UK’s political system”, “subverting a specific election or referendum”, “poisoning public discourse”) as potential objectives of cyberattacks conducted by hostile foreign actors against Great Britain (UK Parliament 2017: 32–33).

At the level of international organizations, the strategic communications units that have been established by NATO and the EU can be seen as telling examples, as they were founded with the explicit goal to counter Russian disinformation, which societies in the Baltic states and Eastern Europe more generally seem particularly vulnerable to. In the EU, relevant efforts were underlined in 2018, when the EU Commission and the EEAS published their “Action Plan against Disinformation” (European Commission 2018b), proposing intensified activities to counter Russian propaganda in Eastern Europe.

Some reflections on recurrent frames

Diagnostic framing is key to understanding threat politics, as it indicates whether disinformation is an element of political threat perception and vulnerability

assessment. In this sub-section I want to briefly reflect on some of the most frequently recurrent diagnostic frames in the policy documents examined. At first, it seems particularly important to note that disinformation is mostly not perceived as a standalone activity or incident but is instead seen as one element of a more or less coordinated campaign. Disinformation campaigns – at least the ones that cause most anxiety – are assumed to be led by some kind of a strategically behaving actor, a populist movement or party or a foreign state, with current threat perceptions in Europe and the United States mostly relating to Russia.

Another important element of diagnostic framing is the narrative of national elections being targeted and potentially distorted by disinformation campaigns. This narrative adds to the perceived vulnerability of the target as a national society and polity and to the understanding of what might be the intention behind hostile information operations. Having said this, it is above all this narrative that seems constitutive for the interpretation of disinformation as a threat to national security. Moreover, the narrative extends the threat perceptions of disinformation also to other forms of so-called foreign interference such as foreign propaganda activities, tainted leaks, etc. that do not necessarily contain incorrect information (Jamieson 2018: 144). Finally, another element illustrated by the narrative about threats to national elections is its obvious dependency on supposed agency and attribution as highlighted in the theoretical sections above.

The impact question – the supposed effects of disinformation

In order to accurately assess the appropriateness of current threat perceptions, it is important to look at the empirical evidence presented by academic research so far. Different strands of research have reacted to the recently grown interest in disinformation. Political communication in mediatized environments is a particularly complex subject of social science research. Discerning and measuring the effects of any particular kind of information is difficult, especially in situations of intense political communication like an election campaign, where every bit of (dis-)information is thrown into the troubled ocean of public discourse. However, the attempts made so far can very broadly and superficially be grouped into three major categories that are very unequally populated. Firstly, researchers measure exposure to (dis)information and its audience effects. Secondly, the impact of (dis-)information is assessed by studying its effects on mass media communication. Finally, there is a third dimension often neglected or only incidentally touched upon: the dimension of public discourse. As the three strands address layered dimensions of communicative activity, I refer to them as the micro, the meso, and the macro levels. In addition, research and scholarly debate on disinformation and its effects need to take the dynamic evolvments of digitally transformed information markets and public spheres into account that cut across these different levels. Frequently discussed new phenomena such as echo chambers or the automation of political campaigns are said to induce structural changes that might facilitate the spread of disinformation overall (Vosoughi et al. 2018).

The micro level – individual exposure to disinformation and immediate audience effects

Most studies on the impact of (dis-)information are aimed at the level of individual exposure to information. Scholarly research has relied on survey data, web tracking, other kinds of digital trace data, or – though more seldom – experimental designs in order to study how users were exposed to false information and how this might have affected their attitudes and behaviors (Allcott and Gentzkow 2017; Grinberg et al. 2019; Guess et al. 2018; Keersmaecker and Roets 2017; Pennycook et al. 2018). An influential study on the US presidential elections in 2016 (Allcott and Gentzkow 2017) revealed that fake news – categorized as such by using data from fact-checking sites – indeed played a considerable role in the election campaign regarding the sheer volume of content produced, as fake news was able to attract millions of views and clicks. Moreover, a majority of the detected fake news examined in the study were in favor of Donald Trump. Measured effects, based on a market-based model of media consumption, were rather moderate, though. According to the study, the average US adult had only been exposed to 1.14 fake stories that he/she was able to remember (Allcott and Gentzkow 2017; see also Grinberg et al. 2019).

Vosoughi et al. (2018) also relied on assessments by independent fact-checking websites for categorizing news items as true or false, respectively. The study combined this approach with big data analysis, as approximately 126,000 stories posted or shared by about 3 million individual users on Twitter were included as data. Its findings indicate that, while false stories spread just as rapidly as true ones, they reached far more people and were much more likely to become viral than true stories (Vosoughi et al. 2018: 1148), suggesting that they would have greater potential impact.

Guess et al. (2018) combined an online public opinion survey with web traffic data. By measuring the relationship between selective exposure and disinformation, they were able to reveal the fundamental problems of measuring exposure alone, as exposure by itself does not necessarily equate to impact, at least not in the sense of changing actual political attitudes or behaviors. Instead, the authors found that, while a quarter of US Americans visited so-called fake news websites in the run-up to the 2016 election, most of the measured exposure was “attitude-consistent”, suggesting that selective exposure has an attenuating effect on the potential impact of disinformation. With selective exposure, their example makes us aware that broader media logics and discursive predispositions need to be taken into account if a more comprehensive picture is to be obtained.

The meso level – the effects of disinformation on mass media

Digital media have “a complex relationship with traditional media” (Tucker et al. 2018: 4). While they tend to undermine the fundamental structures of traditional mass media (Shirky 2008) on the one hand, they have also “clearly become a tool for traditional media reporting” (Tucker et al. 2018) on the other. Consequently,

it is necessary to understand the interactions between so-called new and traditional media (Marwick and Lewis 2017). Among recently published works, Hall Jamieson published the most comprehensive account of how a disinformation campaign allegedly steered by the Russian government had a decisive impact on the US presidential elections by influencing the mass media. Tellingly titled *Cyberwar: How Russian Hackers and Trolls Helped Elect a President*, her book argues that foreign actors successfully influenced, during critical phases of the campaign, the mass media and its most crucial mechanisms for political communication, such as agenda setting, framing and priming, and thereby might have ultimately turned around the public vote in the most decisive battleground states. Thus, her aim is to show how “[t]he legacy media were complicit in this effort” to undermine the US elections (Jamieson 2018: 38). The empirical chapters of her book offer illustrative findings from qualitative analyses of campaign material, analyses of national telephone surveys (oriented toward the micro level but focused on media effects) and ample anecdotal evidence of alleged Russian information operations during the presidential election campaign. All in all, Jamieson concludes that “Russian-hacked content and disinformation not only infected the news agenda but also tilted the balance of discourse in battleground states against the Democratic Party nominee” (Jamieson 2018: 7).

As Jamieson, however, has to admit at the end of her book, the reflections and reported findings cannot provide evidence-based certainty but can only make the conclusions drawn more or less plausible. The fundamental problem of how to distinguish the effects of deceptive information operations from all the other instances of strategic communication that make up a national election campaign and above all seek to achieve the same outcome (namely the victory of a favored candidate) remains unresolved (Jamieson 2018: 208).

The macro level – the effects of disinformation on political discourse and knowledge

Disinformation campaigns in general and foreign meddling in elections in particular are expected to target public discourse. They constitute attempts to move “the discourse in a particular way” (Morgan 2018: 41). It is these presumed effects at the macro level that particularly feed the perception of disinformation as a threat to national security that can, for instance, be put in the context of “a wider Russian operation to disrupt and agitate Western political discourse” (UK Parliament 2017: 52). Discourse is a concept with diverse meanings. Discourse theories, however, tend to converge in the shared assumption of collective knowledge production and processing through symbolic communication. Unfortunately, but not surprisingly, there is particularly little knowledge on the more substantial effects a disinformation campaign might have on public discourse and knowledge (Farrell and Scheier 2018). As seen above, this discourse is often referred to as being threatened by emerging *counter cultural narratives* emanating from Russian information campaigns or on the domestic level by such narratives spread by new radical right movements (Bennett and Livingston 2018: 128). Narratives – being a core

concept of discursive structuring – are expected to be inserted or critically manipulated in a deliberate maneuver from the outside (Omand 2018: 5–6). However, these assumptions are mostly made without presenting empirical findings on the actual effects and successes of the presumed maneuvers (for a notable exception from data science, see Bessi et al. 2015).

In contrast, Alexander Lanoszka's work (2019) on "Disinformation in international politics" concentrates on the strategic level and argues that it is unlikely that alleged disinformation campaigns have any substantial effect on policy, as public discourse and public opinion would serve as a "second barrier" for disinformation that would need to be overcome. The most important part of his argument is that the acceptance of disinformation by a target audience is dependent on what he calls the "pre-existing ideological commitments and mindsets" of the people that would make them unlikely to believe and to change their minds on the basis of some novel pieces of information: "[s]haping the information that they receive are ideological commitments and partisan identities in addition to their own experiences, rules of thumb, and the pieces of conventional wisdom that they have acquired over their lifetimes" (Lanoszka 2019: 236). Lanoszka applies his conceptions and empirically tests his assumptions using Russian information operations between 2014 and 2017. All in all, he finds his assumptions and argument confirmed.

Lanoszka's findings are consistent with a discourse theoretical perspective on disinformation, as it seems indeed unlikely that information operations, no matter how sophisticated and coordinated they might be, can disrupt or even turn around public discourse as intended. The idea that it is possible to "give history a nudge" through disinformation campaigns, as Sir David Omand (2018: 5) put it, must thus be regarded as highly questionable from a discourse angle. However, measuring the macro-level effects of disinformation is obviously very difficult, as this requires an all-encompassing view, probably a greater historical distance from the events of interest, and profound knowledge of the sociocultural and political configurations of a given society (Lanoszka 2019: 233).

Facilitating factors of disinformation in the digital public sphere

Disinformation is certainly not a new phenomenon. It has been part of politics throughout the ages. As Omand (2018: 7) rightly states, "there is nothing very new about expecting subversive threats to an established state authority" (see also Morgan 2018: 41). As illustrated above, threat perceptions have, however, changed recently under the impression of the structural changes of the digital public sphere. Indeed, with the attenuation of traditional gatekeepers, everybody can publish content and thus spread information potentially at scale (Allcott and Gentzkow 2017; Keersmaecker and Roets 2017) – albeit without any guarantee that they will be listened to (Hindman 2009; Stier et al. 2018). In addition, Facebook and other social media are accused of serving as "gateway[s] to fake news website[s]" (Guess, et al. 2018). Against this backdrop, more broadly discussed phenomena of online communication such as echo chambers and the high

degree of automation through social bots are believed to aggravate polarization (Sunstein 2017) and induce paranoia (Farrell and Schneier 2018: 2). These potentially facilitating factors also need to be checked against the evidence presented by empirical research. What does this mean for the effects of disinformation? While a polarized political system and public might be more vulnerable to the spread of fake news, as hypothesized elsewhere (Zettl 2019), ascribing polarization to digitally enhanced fragmentation (echo chambers) is highly controversial.

As to the *echo chamber*, increasing political polarization is widely perceived as a fertile ground for the spread of disinformation (Jamieson 2018; Lanoszka 2019: 229). Empirical studies have tested the echo chamber hypothesis mainly for the United States, but also for a number of other democracies (Bakshy et al. 2015; Jacobson et al. 2016). The results vary of course, but more recent works have mostly questioned the echo chamber hypothesis (Dubois and Blank 2018; Fletcher and Nielsen 2017; Gentzkow and Shapiro 2011; Guess et al. 2018).

The effects of *automation*, widely observed and discussed in the form of so-called social bots – software-tools that allow for the automated generation of content and/or activity online – are also far from clear. While a number of empirical works have highlighted the potential or factual role social bots have played in political processes, including the spread of (dis-)information as part of information operations based on quantitative measurements (Ferrara 2017; Hegelich and Janetzko 2016; Howard and Kollanyi 2016; Keller and Klinger 2019), their methodology and results have been contested due to the problematic flaws of bot detection (Gallwitz and Kreil 2019) or the hybrid use of automation by real users in actual practice (Grimme et al. 2017).

Even if one accepts the extent to which social bots are active, especially on a micro-blogging service like Twitter, this does not mean that automation is responsible for the substance of manipulation online. Instead, social bots are likely to feed the metrics of the online attention economy; they influence trending topics through following, linking, and sharing, but manipulative content is mostly produced by real users. Moreover, in terms of distribution, Vosoughi et al. (2018: 1150) point out that it is not bots but human users that make “false news spread [...] farther, faster, deeper, and more broadly than the truth”. This finding serves as another argument for shifting our attention from alleged attackers and the offensive toolboxes at their disposal to the societies they target and their vulnerability to disinformation. With regard to countermeasures (as discussed in Haunschild et al. 2022), findings emphasize the need to address the attention economics in digital communication and to promote a new ethics of self-regulation in information consumption and provision instead of following restrictive regulatory approaches geared toward the supply side of information.

Conclusion

In this article, I presented an overview of theoretical approaches to and empirical studies of disinformation with a focus on its effects and potential threats to democracy. As to the general debate and current threat perceptions, I argued for

a re-orientation toward knowledge-oriented perspectives and social constructivist research strategies. This was followed by an analysis of policy documents that helped identify recurrent threat frames. These were then contrasted with empirical evidence as produced by different strands of empirical research. With a three-layered distinction between a micro, a meso, and a macro level of analysis, at which distortion and influence can potentially be measured, I included two less commonly studied perspectives that might well be at least equally important for threat assessment with regard to disinformation. Moreover, with echo chambers and automation, I also briefly referred to phenomena that are widely associated with the structural transformation of the digital public sphere and are assumed to be facilitating factors for the spread of disinformation. As with disinformation in general, their existence and more importantly their alleged effects are very difficult to prove, thus adding to the overall uncertainty conditions that political actors find themselves in.

Against the backdrop of alarmed publics and politicians, countermeasures need to be drafted carefully with a better understanding of the overall challenge (see Haunschild et al. 2022; Omand 2018; Pope 2018; Thornton and Miron 2019). Empirical research thus needs to be intensified and improved in order to produce a greater body of knowledge on the actual challenge and appropriate solutions. Before reacting to alarmist voices, we should in particular study the long-term effects of alleged campaigns on the macro level of public discourse. Above, I compared the spread of disinformation with the exploitation of certain discursive vulnerabilities. This comparison illustrates that such vulnerabilities cannot – at least not sufficiently – be reduced by conventional regulatory or even security measures, but need to be addressed by political or journalistic means (Omand 2018: 20). Vulnerabilities must be reduced from within. If social media for instance serve as catalysts for the spread of disinformation, the profit-oriented mechanisms of their particular attention economics need to be addressed by regulation (Bakir and McStay 2018) without restricting informational freedom. Civil society engagement as practiced in fact-checking regimes might also be part of more adequate solutions, if designed and implemented with caution and with actors being held accountable.

All in all, a reflective and critical approach to the study of disinformation seems appropriate. For if the most likely and only achievable goal of malevolent actors on the international scene is correctly understood as the generation of paranoia and chaos among democratic publics and the sowing of distrust in democratic institutions (Chen 2016; Farrel and Schneier 2018), alarmist discourses and governmental overreactions might directly lead us into this trap.

References

All links checked on August 20, 2021.

- Allcott, H. and Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31(2): 211–236.
- Bakir, V. and McStay, A. (2018). Fake News and the Economy of Emotions. *Digital Journalism*, 6(2): 154–175.

- Bakshy, E., Messing, S. and Adamic, L. A. (2015). Political Science. Exposure to Ideologically Diverse News and Opinion on Facebook. *Science*, 348(6239): 1130–1132.
- Ball, P. and Maxmen, A. (2020). The Epic Battle against Coronavirus Misinformation and Conspiracy Theories. *Nature*, 581(7809): 371–374.
- Bennett, W. L. and Livingston, S. (2018). The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions. *European Journal of Communication*, 33(2): 122–139.
- Bessi, A., Zollo, F., Del Vicario, M., Scala, A., Caldarelli, G. and Quattrociocchi, W. (2015). Trend of Narratives in the Age of Misinformation. *PloS One*, 10(8): e0134641.
- Chen, A. (2016, 27 July). The Real Paranoia-Inducing Purpose of Russian Hacks. *The New Yorker*. Retrieved from: <https://www.newyorker.com/news/news-desk/the-real-paranoia-inducing-purpose-of-russian-hacks>.
- Dubois, E. and Blank, G. (2018). The Echo Chamber is Overstated: The Moderating Effect of Political Interest and Diverse Media. *Information, Communication and Society*, 21(5): 729–745.
- Dunn Cavelty, M. (2008). *Cyber-Security and Threat Politics: US Efforts to Secure the Information Age. CSS Studies in Security and International Relations*. London: Routledge.
- Egloff, F. J. (2020). Contested Public Attributions of Cyber Incidents and the Role of Academia. *Contemporary Security Policy*, 41(1): 55–81.
- European Commission. (2018a). *A Multi-Dimensional Approach to Disinformation: Report of the Independent High Level Group on Fake News and Online Disinformation. Communications Networks, Content and Technology*. Luxembourg: Publications Office of the European Union.
- European Commission (2018b, December 5). *Action Plan against Disinformation: European Commission Contribution to the European Council (No. JOIN(2018) 36 Final)*. Brussels. Retrieved from: https://ec.europa.eu/info/sites/default/files/eu-communication-disinformation-euco-05122018_en.pdf.
- Farrell, H. J. and Schneier, B. (2018). Common-Knowledge Attacks on Democracy. *SSRN Electronic Journal*. Advance online publication. <https://doi.org/10.2139/ssrn.3273111>.
- Ferrara, E. (2017). Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election. *SSRN Electronic Journal*. Advance online publication.
- Fletcher, R. and Nielsen, R. K. (2017). Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *Journal of Communication*, 67(4): 476–498.
- Freedom House (2018). *Freedom on the Net 2018: The Rise of Digital Authoritarianism*. New York/Washington. Retrieved from: https://freedomhouse.org/sites/default/files/FOTN_2018_Final.pdf.
- Gallwitz, F. and Kreil, M. (2019). Die Mär von “Social Bots”. *Tagesspiegel*. Retrieved from: <https://background.tagesspiegel.de/digitalisierung/die-maer-von-social-bots>.
- Gentzkow, M. and Shapiro, J. M. (2011). Ideological Segregation Online and Offline. *The Quarterly Journal of Economics*, 126(4): 1799–1839.
- Goldsmith, J. (2016, 9 October). The DNC Hack and (the Lack of) Deterrence. *Lawfare Blog*. Retrieved from: <https://www.lawfareblog.com/dnc-hack-and-lack-deterrence>.
- Goldsmith, J. and Russell, S. (2018). Strengths Become Vulnerabilities: How a Digital World Disadvantages the United States in Its International Relations. *Aegis Series* (1806): 1–22. Retrieved from: <https://www.hoover.org/sites/default/files/research/docs/381100534-strengths-become-vulnerabilities.pdf>.

- Grimme, C., Preuss, M., Adam, L. and Trautmann, H. (2017). Social Bots: Human-Like by Means of Human Control? *Big Data*, 5(4): 279–293.
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. and Lazer, D. (2019). Fake News on Twitter during the 2016 U.S. Presidential Election. *Science*, 363(6425): 374–378. <https://doi.org/10.1126/science.aau2706>.
- Guess, A., Lyons, B., Nyhan, B. and Reifler, J. (2018). Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to Like-Minded Political News is Less Prevalent than You Think. Knight Foundation. Retrieved from: https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/133/original/Topos_KF_White-Paper_Nyhan_V1.pdf.
- Guess, A., Nyhan, B. and Reifler, J. (2018). Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign. Retrieved from: <http://www.dartmouth.edu/~nyhan/fake-news-2016.pdf>.
- Harsin, J. (2015). Regimes of Posttruth, Postpolitics, and Attention Economies. *Communication, Culture and Critique*, 8(2): 327–333.
- Haunschild, J., Kaufhold, M.-A. and Reuter, C. (2022). Cultural Violence and Peace in Social Media: Technical and Social Interventions. In: M. Dunn Cavelti and A. Wenger (eds). *Cyber Security: Socio-Technological Uncertainty and Political Fragmentation*. London: Routledge, pp. 48–63.
- Hegelich, S. and Janetzko, D. (2016). Are Social Bots on Twitter Political Actors?: Empirical Evidence from a Ukrainian Social Botnet. *International AAAI Conference on Web and Social Media*. Retrieved from: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13015/12793>.
- Hindman, M. S. (2009). *The Myth of Digital Democracy*. Princeton, NJ: Princeton University Press.
- Howard, P. N. and Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2798311.
- Jacobson, S., Myung, E. and Johnson, S. L. (2016). Open Media or Echo Chamber: The Use of Links in Audience Discussions on the Facebook Pages of Partisan News Organizations. *Information, Communication and Society*, 19(7): 875–891.
- Jamieson, K. H. (2018). *Cyberwar: How Russian Hackers and Trolls Helped Elect a President: What We Don't, Can't, and Do Know*. New York: Oxford University Press.
- Keersmaecker, J. de and Roets, A. (2017). 'Fake News': Incorrect, but Hard to Correct. The Role of Cognitive Ability on the Impact of False Information on Social Impressions. *Intelligence*, 65: 107–110.
- Keller, T. R. and Klinger, U. (2019). Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications. *Political Communication*, 36(1): 171–189.
- King, G., Pan, J. and Roberts, M. E. (2017). How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument. *American Political Science Review*, 111(3): 484–501.
- Lanoszka, A. (2019). Disinformation in International Politics. *European Journal of International Security*, 4(2): 227–248.
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S., Sunstein, S., Thorson, E., Watts, D. and Zittrain, J. L. (2018). The Science of Fake News. *Science*, 359(6380): 1094–1096.
- Lupia, A. and McCubbins, M. D. (1998). *The Democratic Dilemma. Can Citizens Learn What They Need to Know?* Cambridge: Cambridge University Press.

- Maréchal, N. (2017). Networked Authoritarianism and the Geopolitics of Information: Understanding Russian Internet Policy. *Media and Communication*, 5(1): 29–41.
- Marwick, A. and Lewis, R. (2017). Media Manipulation and Disinformation Online. *Data and Society*. Retrieved from: https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf.
- McQuail, D. (2008). *McQuail's Mass Communication Theory* (5th ed., reissued.). London: SAGE Publications.
- Morgan, S. (2018). Fake News, Disinformation, Manipulation and Online Tactics to Undermine Democracy. *Journal of Cyber Policy*, 3(1): 39–43.
- Omand, D. (2018). The Threats from Modern Digital Subversion and Sedition. *Journal of Cyber Policy*, 3(1): 5–23.
- Pennycook, G., Cannon, T. D. and Rand, D. G. (2018). Prior Exposure Increases Perceived Accuracy of Fake News. *Journal of Experimental Psychology. General*, 147(12): 1865–1880.
- Pomerantsev, P. (2014, 5 May). How Putin Is Reinventing Warfare. *Foreign Policy*. Retrieved from: <https://foreignpolicy.com/2014/05/05/how-putin-is-reinventing-warfare/>.
- Pope, A. E. (2018). Cyber-Securing Our Elections. *Journal of Cyber Policy*, 3(1): 24–38.
- Schünemann, W. J. (2014). *Subversive Souveräne: Vergleichende Diskursanalyse der gescheiterten Referenden im europäischen Verfassungsprozess. Theorie und Praxis der Diskursforschung*. Wiesbaden: Springer VS.
- Schünemann, W. J. (2018). Skad Analysis of European Multi-Level Political Debates. In R. Keller, A.-K. Hornidge, and W. J. Schünemann (eds). *Routledge Advances in Sociology. The Sociology of Knowledge Approach to Discourse: Investigating the Politics of Knowledge and Meaning-Making*. Abingdon, Oxon, New York: Routledge, pp. 91–111.
- Shackelford, S., Schneier, B., Sulmeyer, M., Boustead, A., Buchanan, B., Craig, A., Deckard, T. H. and Malekos Smith, J. (2016). Making Democracy Harder to Hack: Should Elections Be Classified as ‘Critical Infrastructure?’. *University of Michigan Journal of Law Reform*, 50(3): 629–668. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2852461.
- Shirky, C. (2008). *Here Comes Everybody: The Power of Organizing without Organizations*. New York: Penguin Books.
- Stier, S., Schünemann, W. J. and Steiger, S. (2018). Of Activists and Gatekeepers: Temporal and Structural Properties of Policy Networks on Twitter. *New Media and Society*, 20(5): 1910–1930.
- Strick, B., Robinson, O. and Sardarizadeh, S. (2020, May 28). Inside the Pro-China Network Targeting the US. BBC. Retrieved from: <https://www.bbc.com/news/blogs-trending-52657434>.
- Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton, NJ: Princeton University Press.
- Tago, A. (2018). Public Diplomacy and Foreign Policy. In C. G. Thies (ed.). *Oxford Encyclopedia. The Oxford Encyclopedia of Foreign Policy Analysis*. New York: Oxford University Press, pp. 453–464.
- Thornton, R. and Miron, M. (2019). Deterring Russian Cyber Warfare: The Practical, Legal and Ethical Constraints Faced by the United Kingdom. *Journal of Cyber Policy*, 4(2): 257–274.
- Tucker, J., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D. and Nyhan, B. (2018). Social Media, Political Polarization, and Political Disinformation: A

- Review of the Scientific Literature. *SSRN Electronic Journal*. Retrieved from: <https://doi.org/10.2139/ssrn.3144139>.
- UK Parliament (2017, December 20). *Intelligence and Security Committee of Parliament, Annual Report 2016–17 (No. HC 655)*. London. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/727949/ISC-Annual-Report-2016-17.pdf.
- US White House (2017, December). *National Security Strategy of the United States of America*. Washington, DC. Retrieved from: <https://trumpwhitehouse.archives.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf>.
- Vosoughi, S., Roy, D. and Aral, S. (2018). The Spread of True and False News Online. *Science*, 359(6380): 1146–1151.
- Zettl, K. (2019). Lesson Learned? Demokratische Resilienz gegenüber digitaler Wahlbeeinflussung in den USA und Deutschland. *Zeitschrift für Außen- und Sicherheitspolitik*, 12(4): 429–451.