If you see something, then say something to others: A simple contagion of misinformation and a complex contagion of socially-contingent corrections

Hyunjin Song

Department of Communication, University of Vienna, Austria

Word count: 3284 words Draft date: February 14, 2018

Draft in progress. Please do not cite without permission.

Please direct any questions and inquiries to hyunjin.song@univie.ac.at

Author Note

Hyunjin (Jin) Song is currently an assistant professor ("Universitätsassistent, post-doc") in the Department of Communication at the University of Vienna, and also a member of the Vienna Computational Communication Science Lab.

Abstract

As the citizens' news consumption is increasingly driven by online sources, the propagation of misinformation and so-called "fake news" on those platforms become an increasing concern for the public and policy makers. Our goal in this contribution is to offer a more systematic assessment of underlying mechanisms of misinformation spreading and its correction, combining a macro social contextual factor and individuals' cognitive basis of adopting misinformation into a more integrated, dynamic system model perspective. We first review existing evidence concerning individuals' cognitive basis of adopting such misinformation, and social context of which exposure to misinformation and its corrections are received. Next, adopting a well-known class of an epidemic model of virus infection and recovery, we combine this micro and macro dynamics into comprehensive, integrated model of misinformation diffusion on social networks. We do so by focusing on the distinction between simple contagion of misinformation vs. complex contagion of adopting corrective messages. Relying on Agent-based simulations, we further explore various boundary conditions of such dynamics, aiming to uncover how and when such misinformation propagates into the public, as well as what factors facilitate or hinder such diffusion process.

Keywords: Misinformation, fake news, correction, simple contagion, complex contagion, agent-based simulations

If you see something, then say something to others: A simple contagion of misinformation and a complex contagion of socially-contingent corrections

Citizens across the worlds are experiencing major changes in their news environments with the development of digital media. One of the most dramatic changes in the news environment in recent decades involves the role social networking sites (SNS) such as Facebook and Twitter play as a primary source of news outlets. Not only citizens' news consumptions are increasingly driven by such online sources (Shearer & Gottfried, 2017), but it also appears that citizens themselves are actively participating in news dissemination on those platforms by sharing news contents with their peers (e.g., Lee & Song, 2017; Shearer & Gottfried, 2017).

An effective deliberation among public is regarded as a keystone of thriving democracies, and modern political systems squarely depend on informed decisions of citizens in that regard (Delli Carpini & Keeter, 1996). Yet, a propagation of rumors, misinformation, and so-called "fake news" on those platforms becomes an increasing concern for the public and policy makers alike (Allcott & Gentzkow, 2017; Lazer et al., 2017), as evidenced in recent 2016 U.S. presidential election (Allcott & Gentzkow, 2017; Giglietto, Iannelli, Rossi, & Valeriani, 2016; Guess, Nyhan, & Reifler, 2018) and in Brexit votes (The New York Times, 2017). While a wide circulation of factually dubious information is not entirely new to political arena, a growing trend of digitally disseminated rumors and misinformations – often termed as a "fake news" phenomenon – is increasingly recognized as a serious threat to liberal democratic societies (Allcott & Gentzkow, 2017; Lazer et al., 2017). Either based on unsubstantiated rumors or based on factually wrong beliefs, many of the misinformed behave differently than those who are accurately informed (Kuklinski, Quirk, Jerit, Schwieder, & Rich, 2000). They often disagree about basic facts about many political and public issues (e.g., Nisbet, Cooper, & Garrett, 2015), and continue to believe and rely on such false information when making political judgments (Nyhan & Reifler, 2010; Thorson, 2016).

Along with these trends, there has been an growing interest among scholars on how people process and maintain factually false (or at least factually dubious) information from the perspectives of an individual's cognitive processes (Kuklinski et al., 2000; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Weeks, 2015). These studies have generated a valuable

insights of how individuals often maintain factually false beliefs, and further, how corrections to false beliefs are received and processed (Garrett, Weeks, & Neo, 2016; Lewandowsky et al., 2012; Thorson, 2016). However, despite growing interest and continued research effort to better understand the nature and its exact mechanism, what we know about the spread of misinformation and fake news specifically on online social networks is largely based on limited evidence due to its complex nature of the problem.

Against this backdrop, our goal in this contribution is to offer a more systematic assessment of underlying mechanisms of misinformation spreading and its correction, focusing on one's *social contexts* in which such (mis)information and corrective messages are received and processed. We argue that while *exposure* to (mis)information is likely to follow a simple contagion process, *changes* in one's beliefs regarding such (mis)information – which ultimately *the* goal of corrective messages – is likely to be, in Centola and Macy's (2007) term, a "complex contagion" where such changes require multiple sources of affirmation and reinforcement compared to simple contagion process. As a result, the effects of fact-checking and corrective messages are likely to be highly *socially* contingent, yet studies only now begin to consider this possibility more seriously (Bode & Vraga, 2017; Margolin, Hannak, & Weber, 2017).

In what follows, we first review existing evidence regarding political misperceptions and the effects of fact-checking (i.e., correction) messages. We advance our perspective by combining an individual-level cognitive and affective basis of adopting such misinformation with a social context of which misinformation and corrections are received. Based on a well-known class of an epidemic model of virus infection and recovery, we propose an integrated model of misinformation diffusion and socially-contingent corrections on social networks, with a special focus on the differences between a *simple contagion* of misinformation and a *complex contagion* of corrections and fact-checking messages. Relying on Agent-Based Model (ABM) simulations, we robustly explore boundary conditions of such dynamics, aiming to uncover how and when such misinformation propagates into the public.

A Psychology of Fake News, Misperceptions, and Corrections

Following Allcott and Gentzkow's (2017) definition, we define *fake news* as "distorted signals uncorrelated with the truth" (p. 212). This encompasses several related concepts, such as

misinformation, rumors, and disinformation. Literature on this topic generally maintain loosely defined, but at the same time highly interrelated, conceptualizations of those related terms. For instance, (political) rumors are often defined as "unsubstantiated claims about candidates and issues that are often false" (Weeks & Garrett, 2014, p. 401). Similarly, misinformation (or misperceptions) are defined as factual information (or beliefs) "that are false or contradict the best available evidence in the public domain" (Flynn, Nyhan, & Reifler, 2017, p. 128). In relation to this, *dis*information campaigns often denote organized, strategic efforts that trying to sway public opinion using rumors and misinformation (Garrett, 2017; Lewandowsky et al., 2012). Understood in this way, fake news often exclude unintentional reporting mistakes, parodies and satires, or unverifiable conspiracy theories (Allcott & Gentzkow, 2017). While term *fake news* often than not additionally entail specific pseudo-journalistic styles that mimic legitimate news sources to intentionally deceiving audiences (Egelhofer & Lecheler, 2017), we use term "fake news" somewhat loosely, denoting any type of misinformation – information that is not supported by best-available evidence – that is deliberatively circulated among publics.[1]

Literature on misinformation and its persistence often converges to the observation that publics' exposure to and acceptance of misinformation are largely driven by one's motivated consistency needs. That is, people disproportionately gravitate toward information that conforms to their partisan priors (Garrett et al., 2016; Weeks & Garrett, 2014), and more likely to accept and endorse such messages (Guess et al., 2018; Nyhan & Reifler, 2010). A mounting evidence – largely based on Kunda's (1990) or on Taber and Lodge's (2006) motivated reasoning framework – suggests that citizens tend to evaluate attitudinally congruent information as more convincing and valid *regardless of its veracity*, while attitudinally inconsistent information is likely to be perceived as weak and therefore likely to be rejected. Therefore, it is perhaps not surprising to find that most of the prior studies based on a motivated reasoning framework document that fact-checking messages (sometimes denoted as "corrective" or "debunking" messages in the literature) have only limited effects due to inherent tendency of humans to directionally process politically relevant information (Flynn et al., 2017; Taber & Lodge, 2006; Thorson, 2016). Even worse, corrective messages may backfire, may induce higher level of endorsements of false beliefs than actually lower them (e.g., Nyhan & Reifler,

2010; but see Wood & Porter, 2018).

Another line of studies based on a dual process theory of human cognitive processing and memory suggests that attitudinally-congruent misinformation creates automatic and strong affective responses – therefore automatically and effortlessly activated in one's memory – whereas attitudinally incongruent correction messages rarely produce such responses. Due to such asymmetrical nature, people have to rely on more deliberative, strategic search processes (which require significant cognitive resources) to recall attitudinally inconsistent correction messages and incorporate them into relevant judgments (Lewandowsky et al., 2012; Thorson, 2016). Also, since misinformation tends to form a coherent mental model based on one's partisan schema and stereotypes (e.g., Garrett, Nisbet, & Lynch, 2013), people tend to fill any gaps caused by corrections (that invalidate some parts of the existing mental model) with flawed but attitudinally congruent misinformation that is still readily accessible in their memory (Lewandowsky et al., 2012). Studies also find that this effect is much more likely when correction messages do not update the initial mental model that justifies misinformation (Chan, Jones, Jamieson, & Albarracín, 2017), when the perceived veracity of initial misinformation is high (due to fluency bias in one's cognitive processing: Lewandowsky et al., 2012), or when individuals can generate counter-arguing reasons in support for initial misperceptions (Chan et al., 2017; Garrett et al., 2013). Most importantly, due to limitations of strategic memory search processes (i.e., it requires more effortful processing), people may still rely on negated misinformation in subsequent reasoning *even when they remember such information is factually incorrect* (Lewandowsky et al., 2012). Therefore, even in the face of seemingly effective corrections, the effect of misperceptions lingers and continue to exert influence (Thorson, 2016).

Under certain situations, it seems that citizens *indeed* can adhere factual information based on correction messages despite their perpetual partisan bias (e.g., Nyhan, Porter, Reifler, & Wood, 2017; Wood & Porter, 2018). Yet as Margolin et al. (2017) note, it appears that such effects often require special *social context*. This observation is indeed much warranted, as most of the previous studies concerning misinformation and the effect of fact-checking messages are conducted in an experimental context with a single-shot, *asocial* correction message from media professionals and fact-checking organizations (e.g., Garrett et al., 2013; Nyhan & Reifler, 2010;

Weeks, 2015). Much of the literature on partisan selective exposure and political discussion networks already point that social contexts of which an individual is exposed to counter-attitudinal messages may have a powerful consequence on how such messages are interpreted and processed (Levitan & Visser, 2008; Messing & Westwood, 2014). There is also a suggestive evidence that fact-checking and corrective messages from one's peers in their social networks – what we would call a "social correction" – are more likely to, if not equally, be effective in reducing misperceptions (e.g., Bode & Vraga, 2017; Margolin et al., 2017). In what follow, we review several theoretical accounts of such *socially-based* correction messages on misinformation and fake news.

A Social Context of Misperceptions and Corrections: Simple vs. Complex Contagion

People's perceptions and behaviors are likely to be shaped by their social contacts (Centola & Macy, 2007; Lazer, Rubineau, Chetkovich, Katz, & Neblo, 2010), and therefore perceptions and behaviors may spread within social networks (e.g., Bond et al., 2012). Indeed, a non-negligible number of prior accounts of partisan misperceptions and fake news on social networks connects this idea to possible mechanisms of misinformation *spreading* (e.g., Bakshy, Rosenn, Marlow, & Adamic, 2012; Del Vicario et al., 2016). The most simplest form of those accounts posits that online social networks provide one of the ideal settings of a spread of partisan misinformation and fake news. Many of the partisan rumors and fake news tend to be richer in their "novelty" (Wu & Huberman, 2007) regardless of its veracity and informational value. As Giglietto et al.'s (2016) observation suggests, such non-redundant and "novel" information (false and unverified information being one of them) tend to be engaging, and spread faster via weaker ties (Granovetter, 1977). You can easily share and post such news with little to no effort, and your friends on your social networks are easily exposed to such information, in turn they also share such news to their own peers, and so on – creating what is called a *information cascade* of misinformation (Del Vicario et al., 2016). Also, individuals maintain lots of weak ties in social network platforms, which is thought to diversify the information flow. This creates a particularly congenial scenario for a (mis)information propagation and dissemination (Bakshy et al., 2012; Granovetter, 1977). Indeed, Guess et al.'s (2018) investigation of fake news consumption during the 2016 U.S. presidential election

suggests that Facebook was likely to be the focal gateway for visiting websites that propagates fake news, while Allcott and Gentzkow's (2017) study also reveals that such dubious stories are indeed widely shared on Facebook during the election.

Often, this process of spreading (mis)information within a social network can be described as a *simple contagion* process – the process of which a single contact with an "infected" individual (in this case, those who spread misinformation to their peers) is sufficient for such misinformation to be spread to another individual (Centola, 2010; Mønsted, Sapieżyński, Ferrara, & Lehmann, 2017; Siegel, 2009). While a spread of any human behavior requires a minimum threshold of one, for the diffusion of information itself – much like communicable diseases – "the threshold is almost always exactly one" (Centola & Macy, 2007, p. 706; also see Centola, 2010). Once your peer shares news, you become (almost automatically) aware of such news, and you do not require someone else to keep asking about the same news in order to be *aware* of it.

Moreover, due to inherent partisan motivated directionality, the threshold of actually *adapting* such a false yet pro-attitudinal claim (i.e., believing attitudinally congruent misinformation) may also exhibit similar low-threshold properties, although the actual threshold for believing misinformation would be bit higher compared to that of mere exposure and awareness (e.g., Mønsted et al., 2017), and also subject to some individual differences (e.g., Weeks, 2015). As previously suggested, people tend to directionally process political information (Taber & Lodge, 2006), and they tend to easily remember and recall attitudinally-consistent information than uncongenial ones. Therefore, at least for citizens who find a given misinformation to be congenial to their prior partisan identities, *adapting* such false yet pro-attitudinal claims also does not require high number of repeated exposure nor independent, multiple exposure to different sources supporting such claims. Indeed, many of the prior empirical studies support this perspective (e.g., Flynn et al., 2017; Garrett et al., 2016; Kuklinski et al., 2000; Nisbet et al., 2015).

In contrast, there are reasons to believe why the effectiveness of fact-checking and corrective messages would be different from that of a simple contagion of (mis)information. An adoption of a new perspective that contradicts to one's priors (such as an adoption of

counter-attitudinal fact-checking messages) is likely to be, in Centola and Macy's (2007) term, a "complex contagion," where it requires *exposure to multiple sources*, rather than *multiple exposures*, in order to be accepted and further spread into a given network. This is because many political attitudes and subsequent actions (such as politically motivated misperceptions) are likely to be deeply rooted in one's social identities and values, therefore changes in one's political beliefs and attitudes require multiple sources of affirmation and reinforcement from multiple *contacts* compared to simple contagion cases (e.g., González-Bailón, Borge-Holthoefer, Rivero, & Moreno, 2011; Larson, Nagler, Ronen, & Tucker, 2016; Siegel, 2009).

Moreover, such complex contagion dynamics surrounding correction messages are also likely to be dependent upon the attitudinal composition of one's local network; the number of one's neighbors who are *not* activated (e.g., those who still *believe* misinformation) tend to discourages their neighbors to adopt corrections, whereas the number of one's neighbors who are already activated would increase one's susceptibility to adopt new perspective given exposure to correction (i.e., a "*contested*" contagion: Centola & Macy, 2007; also see Friedkin, 2001). Indeed, many social contagions are often perceived to be lack of credibility and legitimacy until adopted by one's neighbors, therefore relative distribution of one's neighbors (in terms of supporters vs. opponents of the adoption) critically influence one's decision to adopt a controversial innovation (Friedkin, 2001; González-Bailón et al., 2011; Larson et al., 2016). This further implies that locally-defined social dynamics may drive specific adoption behaviors of counter-attitudinal information, and consequently, such process would non-trivially interact with a structure and its attitudinal composition of a given network (Friedkin, 2001). Indeed, there exits a considerable support for this perspective, showing that different network topologies (e.g., Centola, 2010; Siegel, 2009) or attitudinal makeup of social networks[2] (e.g., González-Bailón et al., 2011; Larson et al., 2016; Levitan & Visser, 2008) may produce different attitudinal and behavioral consequences, further suggesting that some network structures are more prone to generating cascades and adoptions than others.

There are at least several other reasons why socially-based correction messages, especially from one's peers, might be more effective than a single-shot, *asocial* correction message from more distanced sources (such as fact-checking organizations or mere strangers). First, people

may evaluate information coming from their peers to be more credible and trustworthy (e.g., Metzger, Flanagin, & Medders, 2010), and more willing to deliberate with their close social contacts (Morey, Eveland, & Hutchens, 2012). Research suggests that while individuals rather maintain much flexible attitudes as long as their social affiliation goals are met (Levitan, 2017), reputational risks and social accountability of rejecting corrective messages run high for more close social contacts compared to more distant sources such as strangers (Margolin et al., 2017).

Second, previous studies concerning citizens' political discussion network suggest that an individual's network construction is not likely driven by overt partisan considerations (Lazer et al., 2010; Song, 2015), therefore there exists a considerable degree of exposure to disagreement in citizens' everyday political interactions (e.g., Bakshy, Messing, & Adamic, 2015; Morey et al., 2012). Under such a situation, attitudinally heterogeneous networks trigger more systematic processing of available information, which make individuals to be more responsive to argument strengths (Levitan & Visser, 2008) or social utility (Messing & Westwood, 2014), therefore makes them less resistant to corrective messages from their peers.

All in all, prior empirical evidence and theoretical perspectives convincingly suggest that a *simple* contagion of misinformation and a *complex* contagion of corrective messages would exhibit different properties for population-level contagion dynamics. Structurally weak, but bridging ties such as distant contacts in social networks may provide sufficient means for a simple (mis)information – much like communicable disease – to be spread, while ideologically-driven directionally motivated reasoning may provide sufficient psychological grounds for partisans to easily adopt and believe such misinformation. In contrast, an adoption of corrective message requires independent and multiple reinforcements from many social contacts due to its counter-attitudinal and "contested" nature, critically dependent upon a structure of network and its attitudinal composition – similar to controversial innovations (Centola, 2010; Centola & Macy, 2007; González-Bailón et al., 2011). Aforementioned perspectives therefore undoubtedly point to the possibility that adoptions of fact-checking messages are likely to be highly socially-contingent, and under certain cases, a social correction would be much more effective than an isolated correction message as typically have been considered in previous experiment contexts (e.g., Garrett et al., 2013; Nyhan & Reifler, 2010).

Analytical Challenges in Studying (Mis)information Diffusion Within Social Networks

If the diffusion of (mis)information and adoption of corrective messages may non-trivially dependent upon a structure of a given network and its attitudinal composition, then how a typical (online) social network is structured in terms of its topological features, and how pervasive is "attitudinal" homophily on such a social network? How such structural features affect the overall diffusion dynamics empirically? This is indeed an important question to ask, since the flow of information (either misinformation or its correction) and its adoption are ultimately structured by how individuals are connected with each other in a given network.

A frequent and recurring theme for (attitudinal) makeup of citizens' social networks and its consequences, among scholars and general publics alike, is that most of the citizens today are put into a "echo chamber" or a "filter bubble" that insulate themselves from competing viewpoints and attitude-discrepant information (e.g., Bessi et al., 2015; Del Vicario et al., 2016; Lewandowsky, Ecker, & Cook, 2017). Yet, as Garrett (2017) puts it, "there is ample evidence that *exposure* [emphasis added] echo chambers are not a typical part of Internet users' experience" (p. 370). Most of citizens are appear to be embedded in sufficiently diverse social networks, either online (e.g., Bakshy et al., 2015; Messing & Westwood, 2014) or offline (Lazer et al., 2010; Song, 2015), documenting a substantial level of exposure to political difference in citizen's everyday political interactions. This suggests that popular claims of echo chambers or filter bubbles are less warranted than it is often assumed.[3] However, one should also bear in mind that evidence of exposure to a counter-attitudinal messages does not provide enough evidence for how such messages are actually interpreted and adopted (e.g., Garrett, 2017; Nyhan et al., 2017). Indeed, many of the prior studies on this topic relies on observational evidence regarding message *exposure patterns* alone, either based on participants' self-reports or based on messages posted (sometimes along with engagement data such as "likes" or "shares") in online social networks (e.g., Bakshy et al., 2015), not based on how an individuals actually process and interpret such information. As suggested by elsewhere, attitudinal composition of one's network may fundamentally alter how (identical) messages are interpreted and received by an individual (Garrett, 2017; Levitan & Visser, 2008). As experimental evidence to date suggests, the mere presence of ideologically diverse *exposure* does not automatically translate

into the possibility that such exposure would promote more balanced judgments. In contrast, experimental evidence provide more detailed picture of how such (counter-attitudinal) messages are actually processed and adopted (e.g., Messing & Westwood, 2014; Nyhan et al., 2017; Wood & Porter, 2018). Yet as previously suggested, a typical experimental approach tends to be misspecified in terms of important social dynamics surrounding the simple vs. complex contagion of misinformation and (counter-attitudinal) corrective messages (Centola, 2010; Margolin et al., 2017). In addition, designing a realistic experiment involving real-world social interactions often involves significant practical (e.g., Bond et al., 2012) and ethical challenges (e.g., Kramer, Guillory, & Hancock, 2014) for researchers.

Similarly to attitudinal composition of citizen's social networks, the question of how exactly the structure of citizens' interaction exhibit certain topological properties have attracted considerable interests among scholars. Prior observations on this topic tend to suggests that large-scale online networks exhibit "small-world"-like properties (Kumar, Novak, & Tomkins, 2010; Ugander, Karrer, Backstrom, & Marlow, 2011). A small world is characterized by high connectivity and relatively short average distance among most nodes – most nodes can be reached from every other node by a small number of hops or steps – compared to pure random graphs. This is due to the fact that a small minority of "hub" nodes concentrates a disproportionate number of links, providing a global bridge between smaller, more strongly interconnected local clusters (Barabási, 2004). This structure is known for expediting simple information to be spread globally (Bakshy et al., 2012; Granovetter, 1977), yet it also tends to facilitate complex contagion since the existence of tightly knit neighborhoods supports complex-type behavioral propagation that is difficult or considered to be run counter to prevalent norms and values (Centola, 2010).

References

Allcott, H. & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*, 211–36.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. doi:10.1126/science.aaa1160

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on world wide web* (pp. 519–528). WWW '12. Lyon, France: ACM. doi:10.1145/2187836.2187907

Barabási, A. L. (2004). *Linked: How everything is connected to everything else and what it means for business, science, and everyday life*. New York: Plume.

Bessi, A., Petroni, F., Del Vicario, M., Zollo, F., Anagnostopoulos, A., Scala, A., … Quattrociocchi, W. (2015). Viral misinformation: The role of homophily and polarization. In *Proceedings of the 24th international conference on world wide web* (pp. 355–356). WWW '15 Companion. Florence, Italy: ACM. doi:10.1145/2740908.2745939

Bode, L. & Vraga, E. K. (2017). See something, say something: Correction of global health misinformation on social media. *Health communication*, *Advanced online publication*, 1–10. doi:10.1080/10410236.2017.1331312

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, *329*(5996), 1194–1197. doi:10.1126/science.1185231

Centola, D. & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, *113*, 702–734.

Chan, M. S., Jones, C. R., Jamieson, K. H., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, *28*, 1531–1546. doi:10.1177/0956797617714579

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., … Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559.

Delli Carpini, M. X. & Keeter, S. (1996). *What americans know about politics and why it matters*. Yale University Press.

Egelhofer, J. L. & Lecheler, S. (2017, September). Conceptualizing "Fake News" for political communication research: A framework and research agenda. Paper presented at The Third Annual IJPP Conference, Oxford, UK.

Flynn, D., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*, *38*(S1), 127–150.

Friedkin, N. E. (2001). Norm formation in social influence networks. *Social networks*, *23*, 167–189.

Garrett, R. K. (2017). The "Echo chamber" distraction: Disinformation campaigns are the problem, not audience fragmentation. *Journal of Applied Research in Memory and Cognition*, *6*, 370–376. doi:10.1016/j.jarmac.2017.09.011

Garrett, R. K., Nisbet, E. C., & Lynch, E. K. (2013). Undermining the corrective effects of media-based political fact checking? The role of contextual cues and naïve theory. *Journal of Communication*, *63*, 617–637.

Garrett, R. K., Weeks, B. E., & Neo, R. L. (2016). Driving a wedge between evidence and beliefs: How online ideological news exposure promotes political misperceptions. *Journal of Computer-Mediated Communication*, *21*, 331–348.

Giglietto, F., Iannelli, L., Rossi, L., & Valeriani, A. (2016). Fakes, news and the election: A new taxonomy for the study of misleading information within the hybrid media system. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2878774

González-Bailón, S., Borge-Holthoefer, J., Rivero, A., & Moreno, Y. (2011). The dynamics of protest recruitment through an online network. *Scientific reports*, *1*, 197.

Granovetter, M. S. (1977). The strength of weak ties. In *Social networks* (pp. 347–367). Elsevier.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 us presidential campaign. Retrieved from https://www.dartmouth.edu/~nyhan/fake-news-2016.pdf

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788–8790. doi:10.1073/pnas.1320040111

Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics*, *62*(3), 790–816.

Kumar, R., Novak, J., & Tomkins, A. (2010). Structure and evolution of online social networks. In *Link mining: models, algorithms, and applications* (pp. 337–357). Springer.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Larson, J., Nagler, J., Ronen, J., & Tucker, J. (2016). Social networks and protest participation: Evidence from 93 million twitter users.

Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K., Hobbs, W., & Mattsson, C. (2017). Combating fake news: An agenda for research and action. *Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy*, *2*.

Lazer, D., Rubineau, B., Chetkovich, C., Katz, N., & Neblo, M. (2010). The coevolution of networks and political attitudes. *Political Communication*, *27*, 248–274.

Lee, J. & Song, H. (2017). Why people post news on social networking sites: A focus on technology adoption, media bias, and partisanship strength. *Electronic News*, *11*, 59–79.

Levitan, L. C. (2017). Social constraint and self-doubt: Mechanisms of social network influence on resistance to persuasion. *Political Psychology*.

Levitan, L. C. & Visser, P. S. (2008). The impact of the social context on resistance to persuasion: Effortful versus effortless responses to counter-attitudinal information. *Journal of Experimental Social Psychology*, *44*, 640–649.

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*, 106–131. doi:10.1177/1529100612451018

Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "Post-Truth" era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369.

Margolin, D. B., Hannak, A., & Weber, I. (2017). Political fact-checking on twitter: When do corrections have an effect? *Political Communication*, *Advanced online publication*, 1–24.

Messing, S. & Westwood, S. J. (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research*, *41*, 1042–1063.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of communication*, *60*(3), 413–439.

Mønsted, B., Sapieżyński, P., Ferrara, E., & Lehmann, S. (2017). Evidence of complex contagion of information in social media: an experiment using twitter bots. *PLOS ONE*, *12*, 1–12. doi:10.1371/journal.pone.0184148

Morey, A. C., Eveland, W. P., Jr., & Hutchens, M. J. (2012). The "who" matters: Types of interpersonal relationships and avoidance of political disagreement. *Political Communication*, *29*(1), 86–103.

Nisbet, E. C., Cooper, K. E., & Garrett, R. K. (2015). The partisan brain: how dissonant science messages lead conservatives and liberals to (dis) trust science. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 36–66.

Nyhan, B., Porter, E., Reifler, J., & Wood, T. (2017). Taking corrections literally but not seriously? The effects of information on factual beliefs and candidate favorability. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2995128

Nyhan, B. & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*, 303–330.

Shearer, E. & Gottfried, J. (2017, September). News use across social media platforms 2017. Retrieved from http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/

Siegel, D. A. (2009). Social networks and collective action. *American Journal of Political Science*, *53*(1), 122–138.

Song, H. (2015). Uncovering the structural underpinnings of political discussion networks: Evidence from an exponential random graph model. *Journal of Communication*, *65*, 146–169.

Taber, C. S. & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*, 755–769.

The New York Times. (2017, November). Signs of russian meddling in brexit referendum. Retrieved from https://www.nytimes.com/2017/11/15/world/europe/russia-brexit-twitter-facebook.html?smid=tw-share

Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, *33*, 460–480. doi:10.1080/10584609.2015.1102187

Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503*.

Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication*, *65*, 699–719.

Weeks, B. E. & Garrett, R. K. (2014). Electoral consequences of political rumors: Motivated reasoning, candidate rumors, and vote choice during the 2008 us presidential election. *International Journal of Public Opinion Research*, *26*, 401–422.

Wood, T. & Porter, E. (2018). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, 1–29. doi:10.1007/s11109-018-9443-y

Wu, F. & Huberman, B. A. (2007). Novelty and collective attention. *Proceedings of the National Academy of Sciences*, *104*(45), 17599–17601.

Footnotes

[1] Often, the term *fake news* is used as derogatory, rhetorical label to attack political opponents. While such a use of the term as a *label* is an important conceptual dimension to consider, this aspect of *fake news* is beyond the scope of this manuscript. See Egelhofer and Lecheler (2017) instead for a detailed conceptualization involving this distinction.

[2] Here, we use the term "attitudinal composition" to denote a relative distribution of supporters vs. opponents regarding a given attitude or a behavior being spread in a network.

[3] Indeed, in a completely segregated network where cross-ideological links are not present, a simple diffusion of partisan misinformation is not likely to saturate the entire network, since at least some segments of populations are never exposed to such information due to the lack of cross-ideological exposure. In light of our discussion, an "exposure" echo chamber actually *prohibits* global-scale (mis)information propagation.