How Misinformation Spreads

July 30, 2020

Introduction

Among the most worrisome aspects of disinformation and misinformation in the digital age are the number of people they might reach in a short time, and the persistence of their narratives in online spaces. In part because of the trust relationships between social media “friends,” social media are extremely effective at spreading dis- and misinformation (Amoruso et al. 2017). Celebrity death hoaxes are a good example of the speed of social media. Such hoaxes existed well before the internet—like one about the late nineteenth-century Ottoman merchant known as Far Away Moses—but they can now move much more quickly and spread more widely on social media. Due in part to the competitive pressures of the constant news cycle, professional mainstream media outlets sometimes pick up and amplify such hoaxes and other misinformation narratives (Funke 2018, 2019; Nansen et al. 2019; Winick 2017). As we discuss below, the role of traditional media in amplifying online dis- and misinformation is sometimes overlooked, but that amplification can be critically important in the lifecycle of misinformation. It can be so important, in fact, that such amplification by traditional media is often the end-goal of many disinformation producers (Wardle and Derakhshan 2017).

Many social scientists have adopted a definition of disinformation as false information that is spread deliberately, with an intention to harm a target population or mobilize allies. Misinformation is often defined as false information that is spread unintentionally, but tends to harm. (For further discussion of these terms and some problems with these definitions, see our research review Defining “Disinformation.”) This research review is primarily concerned with the spread of misinformation, but our review “Producers of Disinformation” focuses on the conditions under which disinformation narratives propagate, and the financial and political motivations underlying them.

Broadly speaking, social science researchers have taken two distinct approaches to answering questions about how misinformation spreads online. One is a cognitive social science approach in which researchers attempt to understand how individuals evaluate information, how they decide whether or not to share it, and the role of biases and prior
beliefs in shaping how people receive and transmit messages. The other is a computational social science approach that focuses on online networks, mapping the spread of links and content between and among groups of social actors. By necessity, each of these approaches must also consider the technological aspects of online platforms—the affordances—that both facilitate and constrain the ways that humans communicate information online. We discuss affordances at greater length below.

Measuring the extent and effects of misinformation

Understanding the scale of the dis- and misinformation problem is one of the most basic—and simultaneously most challenging—aspects of dis- and misinformation studies. Most social media platforms do not allow outside researchers access to the data that show what users consume, and what they spread (Lazer et al. 2017, 2018; Freelon 2018). With closed messaging products like WhatsApp, it’s impossible to know what people are seeing without labor-intensive ethnographic techniques like sitting next to them or interviewing them. Many research projects only focus on one social media platform and fail to capture how misinformation moves between 8chan, Reddit, Facebook, Twitter, cable news, and back again (Krafft and Donovan 2020). Well-intentioned social media employees tend to focus only on their own platforms’ ecosystems.

Due in part to the challenges that keep us from assessing the scale of misinformation, some scholars fear that the most harmful effect of misinformation is not that we might believe falsehoods, but that we might cease to trust any new information, or that politicians may come to believe there is no cost to lying or hypocrisy (Karpf 2019; Wardle and Derakhshian 2017). For example, Chesney and Citron (2018) have proposed the concept of the “liar’s dividend,” that as the public becomes more aware of the potential for deepfake audio and video, liars will be better able to shield themselves from consequences by denouncing real evidence as fake. Similarly, Vaccari and Chadwick (2020) find that deepfakes tend to create uncertainty more than they mislead, and may contribute to greater cynicism. Lastly, boyd (2017) has cautioned that US media literacy curricula promoting skepticism may have backfired by undermining students’ trust in media without proposing an alternative framework for understanding the world.

While we can gain some insight into how many people might encounter a given piece of content, we have a much harder time understanding its effects (if any) on those people (Lazer et al. 2018). Knowing how far misinformation penetrates our societies will make it easier for practitioners, funders, and educators to invest resources. At the same time, because concerns over online misinformation have become so prominent in many contexts,
researchers, journalists, and citizens may be at risk of overstating the problem. For example, news reports that accurately describe tens of thousands of bots amplifying disinformation on Twitter rarely put those numbers in context with Twitter’s roughly 160 million active daily users. In that light, online disinformation on social media looks much, much less pervasive (Sides, Tesler, and Vavreck 2018), and some projects have found that misinformation sharing is quite rare in some contexts (Guess, Nagler, and Tucker 2019), and in other contexts, many users simply ignore “fake news” content (Tandoc, Lim, and Ling 2020).

If we overstate the problem of online dis- and misinformation, we run the risk of unnecessarily undermining our faith in the institutions and knowledge sources that we depend on, like news outlets, public health authorities, and government statisticians. Causing us to lose that faith is one of the key goals of the actors trying to increase divisions in democratic societies and make them less resilient. In other words, if we overestimate dis- and misinformation, it does not need to be effective at misinforming people in order to have effects in our societies. As Karpf (2019) writes, “The rise of disinformation and propaganda undermines some of the essential governance norms that constrain the behavior of our political elites. It is entirely possible that the current disinformation disorder will render the [United States] ungovernable despite barely convincing any mass of voters to cast ballots that they would not otherwise have cast.” In other words, if politicians come to believe that there are no consequences for lying or breaking promises—if they stop fearing that voters will hold them accountable, or think that they can dupe partisans into believing anything—then we cannot keep them from becoming corrupt.

However, if we understate the scale of information disorder (Wardle and Derakhshan 2017) and fail to research and implement effective mitigation strategies, then we allow openings for extremist groups, racists, scammers, trolls, vaccine skeptics, climate deniers, and other threats to democratic discourse and public health. As demonstrated by the “infodemic” surrounding the outbreak of novel coronavirus in 2019 and the global Covid-19 pandemic in 2020, unchecked mis- and disinformation can have serious consequences to public health and civil order (Gottlieb and Dyer 2020; Starbird, Spiro, and Koltai 2020), and understanding how misinformation spreads is vital to mitigating its effects.

**Media diet**

In order to answer questions about how misinformation spreads, we need facts about audiences’ media diets. This information provides vitally important context to the misinformation discussion, but individuals’ media consumption is notoriously difficult to research (Ang 2006). Currently, in the US, Pew Research Center, which gathers its data
through surveys, stands out as the most comprehensive public source for data on where Americans get their news both online and off.

Both Twitter and Facebook are important sources of news for US adults, and WhatsApp is second to Facebook in the UK (Chadwick and Vaccari 2019). As of 2018 in the United States, about 70 percent of adults used Facebook, 73 percent used YouTube, and 22 percent used Twitter. Just under 40 percent of adults used Instagram, which Facebook owns. Moreover, 43 percent of US adults said they got news from Facebook. Facebook is popular across demographic groups in the US, though it is now far more popular with adults than teens, who have migrated to platforms like YouTube, Instagram, and Snapchat (Gramlich 2019).

We have scant information about what people see on social media worldwide. It’s practically impossible to study apps like WhatsApp at a large scale—what we know about them comes from researchers joining small groups and talking with individuals who use those apps. Because of platform policies, it’s also very hard for independent researchers to see what people are consuming on Facebook, Instagram, WeChat, VK, and others. As a result, much of what we think we know comes from Twitter data, and there are important implications from this that we discuss below.

We also can’t ignore the ongoing, important role of traditional media in the diffusion of information. While viewership numbers are declining, local television remains the most-used news source for Americans, beating out both cable and network coverage (Pew Research Center 2019). Further, there may be correlations between traditional media consumption and social media habits that we are only beginning to understand. For example, Chadwick, Vaccari, and O’Loughlin (2018) found that UK Twitter users who shared tabloid news were significantly more likely to share misinformation. We know that narratives migrate back and forth between social media and traditional media, but we are only beginning to understand how those migrations take place in different global contexts and at different points on political spectrums.

As we mentioned above, it is extremely difficult to get research data from social media platforms about how their users engage with content and each other (Freelon 2018). Facebook has made some attempts to open its data to researchers, but legal tangles and privacy considerations have largely stalled those moves.[1] Even if researchers were given access to messaging apps like WhatsApp (which Facebook also owns), such apps are notoriously difficult for researchers to study because of the way the platform and its groups are designed for private messaging. Google, which owns YouTube and has incredible advertising reach across the internet, also doesn’t have transparent, public mechanisms for allowing outside research access to its data. Twitter is the major exception to this trend—it
allows researchers to access data through an API. Importantly, Tromble, Storz, and Stockman (2017) warn that these Twitter data are not necessarily random and representative, and as such may bias scientific findings.

Regardless, access to Twitter has had the inevitable result that much of the information we have about social media behavior comes from Twitter. A great deal of this research is good and insightful, but the problem is that Twitter and its users are very different from, say, Facebook and its users. A study by Pew Research Center found some broad similarities between Twitter users and the overall US population, but Twitter users are more likely to be young, more likely to identify as Democrats, and diverge from the population on certain social issues. Importantly, among the US adult population in 2018, the most prolific 10 percent of Twitter users generated 80 percent of tweet traffic, which is a significant imbalance (Wojcik and Hughes 2019). There are other differences between Twitter and Facebook in the US. On Twitter, for example, users are more likely to follow people they do not know and form a wider net of connections. On Facebook, most users typically follow people they know personally. While political content is equally prevalent on both platforms (Duggan and Smith 2016), these differences in connection behavior may mean that mitigation strategies also need to be different.

The upshot is that there are major challenges to generalizing from studies based on Twitter data.

**Affordances**

This problem of generalizability is further compounded if we try to apply findings from single-platform, US-based research studies to other countries where audiences have different preferences, usage patterns, and cultural sensibilities. In India, for example, WhatsApp (which is owned by Facebook) has more than 200 million users. For those users, being on WhatsApp is nearly synonymous with using a smartphone. People in India use WhatsApp differently than users in other markets, forwarding content as much as they send private messages (Bengali 2019). In addition to the generalization problem, this indicates that tech platforms cannot rely on one-size-fits-all, global solutions to address local problems.

These issues are tied to the idea of affordances, a key concept in communications and technology studies. Every technology—satellites, toasters, unicycles—has affordances, but each set of affordances is different. Technologies and their affordances can even influence the ways we look at the world, and the ways we consider our places in it. Developed by Hutchby (2001) from earlier theories, the idea of affordances refers to the things that technologies allow us to do—such as post pictures of our restaurant meals to Instagram.
Crucially, however, the concept also incorporates constraints as much as freedoms. Instagram allows us to share colorful pictures of food, but that trend has prompted restaurants to change recipes and invest in gaudier décor to make themselves more attractive to influencers, and thus more economically competitive (Petter 2017). Facebook both permits and constrains a different set of options than Twitter, and their respective affordances have made them attractive to different audiences. Understanding affordances is key to understanding how misinformation may spread differently on different platforms and what happens when it crosses platforms. For example, engagement algorithms that present extreme content and mechanisms that encourage sharing content without reading it influence the way information spreads. Users of sites with little or no content moderation and high levels of anonymity, like 4chan, 8chan, and Reddit, also post on more public platforms like Facebook (Hine et al. 2017; Decker 2019). Krafft and Donovan (2020) argue that the decontextualization of information that occurs as it crosses platforms also aids in the spread of misinformation.

**Individual factors**

With those limitations in mind, what does recent research tell us about how and why individuals spread misinformation online? As Chadwick and Vaccari (2019) note, we still know very little about why people choose to share news on social media. A few studies suggest some potential considerations, however. A study by Gallup and the Knight Foundation found that “most people wanted to share an article for social or personal reasons, not because they were skeptical of the story,” and that wanting to share was associated with having a high level of trust in the article they intended to share (2018, 2). In a 2018 article, Vosoughi, Roy, and Aral present an intriguing finding that false information is more likely to spread quickly than truthful information. They suggest that we might attribute this finding to novelty, and argue that people are more likely to pay attention to, value, and share information that is new to them. The authors found that false rumors were seen as more novel, and that human users were more likely to retweet false rumors. In contrast, bots spread true and false information at equal rates. The upshot, the authors said, was that interventions focusing exclusively on bots may be misguided, and that human behavior drives the spread of false information. However, the potential conclusions from both of these studies are true—that people share because they trust, but also value and share what is new—then more research will be needed to illuminate what leads people to trust new information and to further explore the relationships between sharing behavior and information credibility. It’s likely that there are multiple conditions that promote information sharing—trust and novelty being two among them—and further understanding those motivations will be essential to developing effective strategies to counter misinformation.
Some recent findings are beginning to bring focus to this complex area. In a working paper, Pennycook et al. (2019) point to a disconnect between people’s ability to make accurate judgements about content and their intentions to share it. In other words, even though people in their experiment said they value only sharing accurate news, and were able to differentiate between high- and low-quality news content in the experiment, some also indicated they were willing to share false headlines that aligned with their political beliefs. If these findings are supported by future research, and if scholarly consensus builds along these lines, then that would suggest that people are not sharing bogus content to be deliberately inflammatory or because they cannot tell the difference between truth and lies, but because they do not stop to consider the accuracy of content. The authors argue that this finding suggests there are internet affordances—such as social media’s tendency to mix serious content alongside cat videos, and the platforms’ profit-driven priority of engagement over critique—that distract us from making the kinds of decisions that we value, and that interventions aimed at getting people to consider accuracy might have potential (Pennycook et al. 2019, 12–13).

The way humans react to emotional messaging seems to have an effect on the way they spread information. In a study of social media campaigns by advocacy organizations for people on the autism spectrum, Bail (2016) suggests that emotion plays a role in how likely information is to go viral, while Hasell and Weeks (2016, 653) in a panel study found evidence that “partisan media may encourage political information sharing by arousing anger in its audience.” Looking at words like “punish,” “fighting,” or “greed,” Brady et al. (2017, 7313–18) show how the language of moral emotions—“those that are most often associated with evaluations of societal norms”—increases the spread of messages, and that moral emotion language might partly explain why some political messaging spreads further.

There are other factors in individual identity and behavior that may bear on how misinformation spreads online, though again, at least some of those factors seem highly contextual. While not a study of misinformation per se, work by Kwon, Chadha, and Wang (2019) suggests that geographical proximity to an event—in this case, the 2017 Quebec mosque shooting—is related to the tone of social media conversations that people have. In the US, Barberá (2018, 2) found that “age and partisanship were the two most predictive factors” for sharing behaviors, with people 65 and older “nearly five times more likely to share false news stories on Twitter than those ages 18-25.” Registered Republicans were more likely than Democrats to share misinformation, he added, though that could be explained by the preponderance of right-wing misinformation circulating during his study period (see also Allcott and Gentzkow 2017). Barberá’s findings have more recently been supported by Guess, Nagler, and Tucker (2019, 1) using a combination of survey data and participants’ Facebook sharing history. Guess, Nagler, and Tucker found that users older than 65 “shared nearly seven times as many articles from fake news domains as the
youngest age group.” They found similar disparities in conservative users sharing more bogus content than liberal or moderate users, but also cautioned that that could be related to the prevalence of pro-Trump bogus news, and not a causal factor.

In the UK, Chadwick and Vaccari (2019) found that men shared more news than women by a wide margin, and that younger users shared more news than people over 45. Survey respondents said that informing others and expressing their feelings were their most important motivations. When it came to users sharing false information, Chadwick and Vaccari, after an informative discussion of how hard that behavior is to measure in surveys, found that more than 40 percent of users acknowledged sharing what the authors call “problematic” news, or content that is sensationalized, exaggerated, disreputable, or otherwise skirts the truth. Some of those who recalled sharing such content said they knew it was inaccurate or exaggerated at the time, while others said they learned later it was inaccurate or exaggerated. Men shared more intentional disinformation than women, young people shared more unintentional misinformation and intentional disinformation than older users—which runs counter to Barberá’s findings from the US—and Conservative supporters were more likely than Labour supporters to share misinformation.

There is evidence that a constellation of individual factors, such as age, political orientation, gender, and level of digital literacy, plays a role in what information people decide to share online, compounded with factors such as the novelty of the information. There is also a vein of research in political psychology that explores factors like individuals’ degrees of need for order, fear of uncertainty, comfort with new information, and dislike of ambiguity in relationship to their tendencies toward political conservatism or liberalism (Jost et al. 2003; Jost, Nosek, and Gosling 2008; Shook and Fazio 2009; Tullett et al. 2016). We are only aware of a limited number of studies that attempt to connect this line of inquiry with information-sharing behavior (see Jost et al. 2018), and further work along those lines could make a valuable contribution to our existing knowledge.

Networks

While it is important to remember that networks are designed by and composed of individuals with their own identities and motivations, and that even fully automated social media bots are programmed and activated by humans, looking at networks is the other logical place to investigate how misinformation spreads. Researchers have been intently investigating the speed and diffusion of bogus information on social media for years. As we mentioned above, despite a growing body of research, it’s still not clear how prevalent misinformation is on social media networks. Recent work by Guess, Nagler, and Tucker (2019) found that sharing “fake news” is relatively rare for the Facebook users in their
study, but that “visits to Facebook appear to be much more common than other platforms before visits to fake news articles in web consumption data, suggesting a powerful role for the social network” (2019, 1). Bovet and Makse (2019) found that 10 percent of the tweets linking to news sites in their Twitter sample went to a “fake news” or conspiracy site. Allcott, Gentzkow, and Yu (2019) found that interactions with “fake news” sites rose on both Twitter and Facebook leading up to the 2016 election, but then declined on Facebook while rising on Twitter. They cautiously suggest that their findings are “consistent with the view that the overall magnitude of the misinformation problem may have declined, possibly due to changes to the Facebook platform following the 2016 election.” However, in line with Guess, Nagler, and Tucker, they continue to say that Facebook “has played an outsized role” in the spread of misinformation, and “that the absolute quantity of interactions with misinformation on both platforms remains large” (2019, 1–2, 4). It remains to be seen if their findings that the magnitude of misinformation may have declined will hold true during the 2020 “infodemic” accompanying the spread of Covid-19.

The “information cascade” is a key concept for understanding network-based diffusion patterns (Murthy et al. 2016). Definitions vary somewhat, but the term describes an effect in which individuals base their decisions on what others do before them. Buying a popular stock would be one example, as would joining a long line for a food truck because you assume it must be better than one with no line. Some economists have proposed that riots and revolutions can be at least partly explained as information cascades (Banerjee 1992; Ellis and Fender 2011). However, as Easley and Kleinberg note, while people in a network are in a sense imitating the decisions of their predecessors, they are not doing so mindlessly. Instead, they are “drawing rational inferences from limited information,” that limited information being what people before them have done (2010, 425–26).

In studies of social media ecologies, researchers sometimes use the term cascade, or “retweet cascade,” to describe expanding patterns of content shares or discuss bot behavior. In research on a Brexit botnet aligned with the “Leave” movement, Bastos and Mercea (2019) found that the bots pushed hyperpartisan content and were able to rapidly trigger small and medium retweet cascades. They were not able to trigger large cascades, though, and the authors found no evidence to support the idea that bots can cause significant shifts in overall online political conversations.

In another study of cascades and social media sharing, Del Vicario et al. (2016) found that science content and conspiracy content on public Facebook pages had different cascade dynamics within homogenous clusters—so-called echo chambers of likeminded opinions (for more on the contested existence of echo chambers, see Contexts of Misinformation). Science information diffused more quickly. Conspiracy rumors were “assimilated more slowly and show a positive relation between lifetime and size.” In another study of rumor
content in shared photographs, Facebook researchers found that true rumors were more viral and saw larger cascades than false rumors. The rumor cascades overall, though, ran “deeper in the social network than reshare cascades in general.” The researchers also found that individual rumors experience bursts of popularity, and that rumors change over time (Friggeri et al. 2014).

Other researchers have found that misinformation spreads within dense clusters of US social media users (Shao et al. 2018), and point to an asymmetry on the political spectrum. Noting an increase in the proportion of “junk news” versus professional content shared since 2016, Marchal et al. (2018, 5) found that on Facebook, a far-right cluster and a mainstream conservative cluster “shared the widest array of junk news sources identified in our sample” and that the far-right pages “contributed the most to the spread of junk news.” (For more on the prevalence of disinformation on the political right in the US, see Barrett 2019a, 2019b; Benkler, Faris, and Roberts 2018; Bennett and Livingston 2018; Faris et al. 2017; Marwick 2018; Nithyanand, Schaffner, and Gill 2017.) In a study of news diffusion on Twitter during the 2016 US election, Bovet and Makse (2019) suggested that “fake” and extremely biased news traveled differently than center and left-leaning news. They found that misinformation sites were clustered with right-leaning media outlets, while center- and left-leaning news was primarily diffused by journalists and other influential users. The spread of misinformation and biased news seemed to result from more collective activity instead of influential users’ activity, leading the authors to suggest that those kinds of content had different diffusion mechanisms. However, Starbird’s (2017) study of conspiratorial alternative explanations for mass shooting events complicates any direct left-versus-right correlation with misinformation. While she also found a dense cluster of misinformation sites with political agendas, they aligned around antiglobalist themes, and some content supported Russian interests.

**Amplification**

Viral diffusion of information from one user to a handful of others is not the only mechanism by which information spreads. Broadcast—the mass dissemination of content from a single source to many others—remains enormously important in media systems.

This leads us to the concept of amplification, which researchers use to describe how narratives move from fringe publications and social media to mainstream, professional news outlets. Wardle and Derakhshan (2017) argue that “without amplification, dis-information goes nowhere,” and numerous organizations, such as Wardle’s First Draft News, are working to train journalists to be more cognizant of their roles in amplifying misinformation even as they report on its falseness.
The “Momo” hoax is an excellent, if disturbing, example of how amplification works. The rumor involved “false claims that a mysterious character [was] using WhatsApp messages to encourage children to kill themselves” (Waterson 2019), and it was amplified by a 2018 Indonesian newspaper report about a suicide (Alexander 2019). The urban legend circulated with little notice in the dim alleys of the internet until one person posted about the supposed “suicide challenge” in a small-town community Facebook group in the UK. From there, one regional newspaper ran a story, and soon both large UK media outlets and police forces were issuing breathless warnings. As it became clear that the entire narrative was a hoax, children’s organizations told schools, media, and police to stop issuing alerts, warning that the moral panic was frightening children (Waterson 2019). The Momo misinformation is an example of a moral panic or urban legend, not a deliberate disinformation campaign, but achieving that sort of mainstream media amplification is a goal for many disinformation producers (Phillips 2018; Wardle and Derakhshan 2017).

[1] In early 2020, Facebook provided research access to a dataset it had promised roughly 20 months earlier. As of July 2020, it remained to be seen if the dataset matches what was promised, and how useful it will be to researchers.

Our grateful acknowledgement to Dhiraj Murthy and Rebekah Tromble for their assistance in this research review.

Works Cited


Amoruso, Marco, Daniele Anello, Vincenzo Auletta, and Diodato Ferraioli. 2017. “Contrasting the Spread of Misinformation in Online Social Networks.” In AAMAS.


