SOCIAL MEDIA PLATFORMS' ROLE IN ADDRESSING DISINFORMATION

Social media companies are developing new policies and tools to combat disinformation. These new approaches signal an evolution from the platforms' earlier views on speech. While these strategies are important progress, to truly tackle disinformation more effectively we need to look at the problem more holistically and complement these efforts with other strategies.

Emine Etili[*]

TPQ

Fall 2020

* Emine Etili is the former public policy manager of Turkey at Twitter.

For the ability of technology to better human life is critically dependent on a parallel moral progress in man. Without the latter, the power of technology will simply be turned to evil purposes and mankind will be worse off than it was previously.[1]

Orwell feared those who would deprive us of information. Huxley feared those who would give us so much that we would be reduced to passivity and egoism. Orwell feared that the truth would be concealed from us. Huxley feared the truth would be drowned in a sea of irrelevance.[2]

L ess than a decade ago social media was widely hailed as the great hope for democracy and an opportunity for the voiceless everywhere to finally have a voice. In the span of a few short years, it has become perceived in the public eye as a threat to democracy at best and, at worst, as a nefarious force consciously tearing apart not only our governments but societies at large. In fact, headlines, politicians, and civil society around the world point to social media as the primary antagonist in our collective descent into polarization, hate speech, and disinformation. How did we get here when just a few years ago these very same platforms held out hopes for more democratic and—by virtue—more tolerant discourse, societies, and governance?

The answers lie not just with a single culprit, but what might be called an almost perfect storm of elements that came together including nature of the medium, our own cognitive limitations, as well as the role of other institutions. While social media companies have developed new approaches, it will take an holistic, inclusive approach to diminish the amount of online disinformation, as well as its unintentional but also destructive cousin misinformation, to regain a more civil public discourse.

New Technology

Going as far back as the printing press, technology and media have always had a profound impact on political, economic, and social life. The information age, however, has resulted in a fundamental change in how we process and relay information. For the first time, the general public is able to generate its own content in a non-hierarchical fashion, leveraging and expanding networks of information and contacts. From this new technology, a new form of communication emerged: the social media platform. We take it for granted today, but the ability to come into contact with old and current friends or world leaders with just one click has been

---

[1] Francis Fukuyama, *The End of History and the Last Man* (New York: Free Press, 1992).

[2] Neil Postman, *Amusing Ourselves to Death: Public Discourse in the Age of Show Business* (New York: Penguin Books, 1986).

a radical watershed in how we communicate. The proliferation of mobile phones helped move social media from website to mobile app format. The change in format not only allowed more people to access social media but allowed it to be anywhere, any time.

Social media users have grown by more than 10 percent just in the last year, with 3.96 billion users at the beginning of July 2020.[3] Over half of the world's population now uses social media. Facebook has 2.5 billion monthly users while Twitter sees about 330 million visitors every month.[4] Approximately half a billion tweets are posted daily while over 50 billion photos have been shared on Instagram to date. [5]

> *"Having global reach meant that [companies] had to have teams in place that understood local political and cultural contexts as well as country-specific legal and regulatory environments."*

Platforms and Speech

Initially, social media companies viewed themselves not just as advocates but also facilitators of free speech by virtue of their technological structures and philosophical approaches. In 2009, a Google vice president stated that "openness" is the fundamental principle on which the company operates.[6] Twitter was referred to as "the free speech wing of the free speech party."[7] Internet law scholar Kate Klonick notes that:

> A common theme exists in all three of these platforms' histories: American lawyers trained and acculturated in First Amendment law oversaw the development of company content moderation policy. Though they might not have

---

[3] Simon Kemp, "Digital Use Around the World in July 2020," *We Are Social*, 21 July 2020, https://wearesocial.com/blog/2020/07/digital-use-around-the-world-in-july-2020

[4] Omnicore Agency, "Facebook by the Numbers: Stats, Demographics & Fun Facts," https://www.omnicoreagency.com/facebook-statistics/; Omnicore Agency, "Twitter by the Numbers: Stats, Demographics & Fun Facts," https://www.omnicoreagency.com/twitter-statistics/

[5] Omnicore Agency, "Twitter"; Omnicore Agency, "Instagram by the Numbers: Stats, Demographics & Fun Facts," https://www.omnicoreagency.com/instagram-statistics/

[6] Nathaniel Persily and Joshua A. Tucker (eds.), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge: Cambridge UP, 2020), p.294, https://www.cambridge.org/core/services/aop-cambridge-core/content/view/E79E2BBF03C18C3A56A5CC393698F117/9781108835558AR.pdf/Social_Media_and_Democracy.pdf?event-type=FTLA

[7] Josh Halliday, "Twitter's Tony Wang: 'We are the free speech wing of the free speech party'," *Guardian*, 22 March 2012, https://www.theguardian.com/media/2012/mar/22/twitter-tony-wang-free-speech

'directly imported First Amendment doctrine,' the normative background in free speech had a direct impact on how they structured their policies.[8]

In the early days, the platforms' own rules such as Facebook's Community Standards and Twitter's Rules— along with 1ˢᵗ Amendment principles— were thought to be sufficient to moderate speech. These albeit private norms provided guidance on harmful and violent content, abuse, impersonation, spam, and other forms of content that the platforms did not want to remain online. Interestingly, one frequently over-looked aspect is that these norms often also banned some forms of legal content, in order to make the platforms a healthier and thus more widely enjoyable setting for their users. In fact:

> Most such rules from the platforms go well beyond what is required by na-tional laws. Indeed, if such rules were legislated by the government in the United States, almost all would be declared unconstitutional by the courts.[9]

But as technology evolved and the scale and reach of the platforms grew, new chal-lenges emerged. The platforms wished to remain just that: platforms that allowed others to freely share within their framework. With the development of emerging issues (such as hate speech), which has increasingly had offline ramifications as well, companies were no longer able to rely on old methods. They had to modify their policies and engage new tools to tackle nuanced, tricky issues. Often they were stuck between designing policies that would be excessively wide-reaching or not encompassing enough, between those who felt they were over-censoring and those who felt they were not acting effectively and thoroughly enough. Furthermore, they also grappled with governments—authoritarian or otherwise—whose knee-jerk im-pulse was to censor content. Varying legal requirements required country-specific approaches such as withholding content in-country in order to respect local laws, while trying to balance universal values of free speech. Having global reach meant that they had to have teams in place that understood local political and cultural con-texts as well as country-specific legal and regulatory environments. But given the billions of posts and users around the world, the companies struggled to operate in the original spirit of freedom while addressing requests— legitimate and sometimes otherwise— from governments and civil society.

The Disinformation Problem

Then, the online disinformation issue came to light in sharp relief during the 2016

---

[8] Nabiha Syed, "Real Talk About Fake News: Towards a Better Theory for Platform Governance," *The Yale Law Jour-nal*, 9 October 2017, http://www.yalelawjournal.org/forum/real-talk-about-fake-news

[9] Nathaniel Persily, "The Internet's Challenge to Democracy: Framing the Problem and Assessing Reforms," https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/kaf_democracy_internet_persily_single_pages_v3.pdf

US elections. Initially viewed through the prism of Russian interference, the subsequent four years have shown that the use of disinformation by foreign state actors for geopolitical ends is only part of a wider story around disinformation and truth.

> **❝***Major social media companies are responding to disinformation by using a combination of technology, human moderation, in-house and outsourced means in their own ways.***❞**

In a 2019 report on computational propaganda by Oxford University, researchers found prevalent use of "cyber troops" to manipulate public opinion on social media in 70 countries, up from 48 countries in 2018 and 28 countries in 2017.[10] In 52 out of the 70 countries that Oxford researchers examined, cyber troops created memes, videos, fake news websites or manipulated media to mislead users, often targeting specific communities. Furthermore, such online actors use novel strategies such as trolling, doxxing, and harassing as a means to muffle opposition speech and to threaten human rights.[11] In 2019, 47 countries used cyber troops as part of a "digital arsenal" to troll voices they opposed.[12]

Bad actors have not only mobilized en masse in this manner but they have taken advantage of the very nature of social media's openness to achieve their malign objectives. As Zeynep Tüfekçi states, there is nothing necessarily new about propaganda, however, social networking technologies have changed the scale, scope, and precision of how information is transmitted in the digital age.[13] This is not the clumsy, blackout censorship we were used to. As Tüfekçi points out:

> The most effective forms of censorship today involve meddling with trust and attention, not muzzling speech itself. They look like viral or coordinated harassment campaigns, which harness the dynamics of viral outrage to impose an unbearable and disproportionate cost on the act of speaking out. They look like epidemics of disinformation, meant to undercut the credibility of valid information sources.[14]

[10] Samantha Bradshaw and Philip N. Howard, "The Global Disinformation Order 2019 Global Inventory of Organised Social Media Manipulation*," Oxford Internet Institute*, 2019, https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf

[11] Bradshaw and Howard, "The Global Disinformation Order 2019."

[12] Bradshaw and Howard, "The Global Disinformation Order 2019."

[13] Zeynep Tüfekçi, "It's the (Democracy-Poisoning) Golden Age of Free Speech," *Wired*, 16 January 2018, https://www.wired.com/story/free-speech-issue-tech-turmoil-new-censorship/

[14] Tüfekçi, *"*Golden Age of Free Speech."

In other words, disinformation is not only foreign-based but frequently used domestically against local "enemies." Furthermore, it does not just take the form of targeted lies but other strategies that fill social media with content intended to confuse and ultimately silence. This was a kind of obfuscation of truth that Huxley foresaw, per Neil Postman, when he "feared the truth would be drowned out in a sea of irrelevance."[15]

How Did Companies Respond?

The emergence of widespread abuse of platforms in this way has of course forced companies to reconsider the more laissez-faire approach of earlier years. They had believed that "organic discourse of many participants—the much vaunted 'wisdom of the crowds'—would help to weed out false information and produce a multifaceted representation of 'the truth.'"[16] This belief, which was at the very heart of the design of the platforms, unwittingly became one of its greatest weaknesses. As one of Twitter's founders Ev Williams said, "I thought once everyone could speak freely and exchange information and ideas, the world is automatically going to be a better place ... I was wrong about that."[17]

The major social media companies are responding to disinformation by using a combination of technology, human moderation, in-house and outsourced means in their own ways. Although it is tempting to lump them all together, nevertheless they are taking varying approaches, partially due to their differing models and corporate values.

Facebook has introduced new tools to combat misinformation, including identifying false news through community and fact-checking organizations, using machine learning to detect fraud and act against spam accounts, and making spam more difficult by increasing efforts to detect fake accounts. They are also working with partners and users to improve rankings, make it easier to report cases, and improve fact-checking.[18] Running up to the November 2020 election, Facebook prohibited new political ads though previous ones can remain up. It also applied labels to post that seek to undermine the outcome of the election, allege legal voting methods led to fraud, and also added a label if a candidate declares victory before the final outcome is declared.[19]

---

[15] Postman (1986).

[16] James Surowieki, *The Wisdom of Crowds* (New York: Anchor Books, 2004); Dario Taraborelli, "Seven years after Nature, pilot study compares Wikipedia favorably to other encyclopedias in three languages," *Wikimedia Foundation (blog)*, 2 August 2012 quoted in Persily and Tucker (2020), p. 279.

[17] David Streitfeld, "'The Internet is broken': @ev is trying to salvage it," *New York Times,* 20 May 2017 quoted in Persily and Tucker (2020), p. 279.

[18] Facebook, "Working to Stop Misinformation and False News," https://www.facebook.com/formedia/blog/working-to-stop-misinformation-and-false-news

[19] Vera Bergengruen, "'The Devil Will Be in the Details.' How Social Media Platforms Are Bracing For Election Chaos," *Time,* 23 September 2020, https://time.com/5892347/social-media-platforms-bracing-for-election/

Twitter similarly updated its "civic integrity policy" in early September and was the first company to act on political ads. The company states: "Consequences for violating our civic integrity policy depends on the severity and type of the violation and the accounts' history of previous violations."[20] Among the potential actions is tweet deletion, temporary or permanent suspension, and labeling to provide additional context. Although Facebook will not fact-check misinformation in politicians' posts or ads, Twitter will flag false claims.[21] The company's algorithm also will not promote it to others, even if it is a top conversation.[22] Furthermore, Twitter is developing a new product called "Birdwatch", which looks like it will try to tackle misinformation by allowing users to add notes to provide more context for tweets.[23]

> "*To what extent is each of us culpable in precipitating dissemination of disinformation?*"

YouTube has moved to take down content that has been technically manipulated in a way to mislead users, content specifically aimed to mislead people about voting or the census processes, and content that advances false claims related to the technical eligibility requirements for current political candidates and sitting elected government officials to serve in office. YouTube also acts on channels that "attempt to impersonate another person or channel, misrepresent their country of origin, or conceal their association with a government actor and artificially increase the number of views, likes, comments, or other metric."[24] The platform will also "raise up" authoritative voices for news and information. On election night, YouTube gave users previews of verified news articles in search results.[25]

None of these efforts are likely to be silver bullets, however, and will likely require frequent iteration. YouTube, for example, relying heavily on machine learning to sort through disinformation, has now taken a step back, as this method had mixed results. During that time, 11 million videos were taken down which was double the usual rates; a normal than higher proportion of these decisions were overturned on

[20] Yoel Roth and Nick Pickles, "Updating our approach to misleading information," *Twitter Blog*, 11 May 2020, https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html

[21] Bergengruen, "The Devil Will Be in the Details."

[22] Bergengruen, "The Devil Will Be in the Details."

[23] Sean Hollister, "Twitter's 'Birdwatch' looks like a new attempt to root out propaganda and misinformation," *The Verge*, 4 October 2020, https://www.theverge.com/2020/10/4/21500687/twitter-birdwatch-misinfo-tool-propaganda

[24] Leslie Miller, "How YouTube supports elections," *YouTube Official Blog,* 3 February 2020, https://blog.youtube/news-and-events/how-youtube-supports-elections

[25] Bergengruen, "The Devil Will Be in the Details."

appeal.[26] Companies will have to continue using human moderation to ensure careful consideration of context while leveraging machine tools to achieve greater scale.

Causes of Disinformation/Misinformation

While companies implement new policies and tools what may be the root causes of disinformation?

The Medium

Marshall McLuhan famously stated that, "The medium is the message." Perhaps the issue then is the medium itself? Jason Hannan examines how Neil Postman carried McLuhan's statement to the television era, which he then extends to the social media era. Hannan notes that Postman argues the form of television (entertainment) negates the seriousness of its ostensibly serious content (e.g., news, political debates). Furthermore, Postman observes that the more Americans watch television news, the more misinformed they become about the world, leading him to suggest that what television news disseminates is not information but "disinformation."

Hannan then argues that:

> If television turned politics into show business, then social media might be said to have turned it into a giant high school, replete with cool kids, losers and bullies. Disagreements on social media reveal a curious epistemology embedded within their design. Popularity now competes with logic and evidence as an arbiter of truth.[27]

He further states that popularity and tribal affinity supersede logic and evidence. It would therefore be naïve to think that fact-checking can somehow contain the problem of fake news. Rather, we need to look at what is driving the fake news.

There are qualities inherent to the platform and online communication that make them particularly susceptible to problems like disinformation. These include speed, virality, and anonymity. There are other structural issues that have emerged such as the role of echo chambers or "monopoly" of companies due to their size.[28] The latter are potentially resolvable through the use of additional tools or modifications to the product. The former, however, is harder to overcome because of their integral part in the communication model.

---

[26] Alex Barker and Hannah Murphy, "YouTube reverts to human moderators in fight against misinformation," *Financial Times*, 20 September 2020, https://www.ft.com/content/e54737c5-8488-4e66-b087-d1ad426ac9fa

[27] Jason Hannan, "Trolling ourselves to death? Social media and post-truth politics," *European Journal of Communication*, Vol. 33, No. 2 (2018), pp. 214-26.

[28] Persily, "The Internet's Challenge to Democracy."

Cognition

To what extent is each of us culpable in precipitating dissemination of disinformation? In place of the traditional media gatekeepers, we have all become creators and disseminators of content. The problem with disinformation is that it can quickly and unwittingly become misinformation in the hands of users who cannot authenticate the veracity of the content they shared. As such, peer-to-peer transmission unfortunately plays a much more significant role in how ideas are spread.[29]

> ❝*While social media companies can, and should, use both technology and policies to fight against disinformation, they cannot succeed alone.*❞

The human brain is wired to make sense of the world in the simplest way possible, especially when it is overwhelmed with a bombardment of information, which we are today. Wardle and Derakhsan note that even before the use of social media people used mental shortcuts to evaluate the credibility of a source or message: reputation (based on recognition and familiarity), endorsement (whether others find it credible), consistency (whether the message is echoed on multiple sites), expectancy violation (if the web site looks and behaves in the expected manner), and self-confirmation (whether a message confirms one's beliefs). They state that in light of these heuristics in a time when we are heavily reliant on social media as a source of information, the current age of mis- and dis-information becomes clear.[30] In other words, users look for what is familiar, and what others they know also find familiar.

Furthermore, another study by scholars at MIT found that of all the verified true and false rumors that spread on Twitter, false news spreads more pervasively than the truth online.[31] Not only that, but human beings rather than bots turned out to be the primary culprits. This finding has ramifications for both how we should consider user behavior and also the next steps in mitigating misinformation:

---

[29] Cailin O'Connor and James Owen Weatherall, "The Social Media Propaganda Problem Is Worse Than You Think," *Issues in Science and Technology*, Fall 2019, https://issues.org/wp-content/uploads/2019/11/OConnor-Weatherall-The-Social-Media-Propaganda-Problem-Fall-2019.pdf

[30] Claire Wardle and Hossein Derakhshan with research support from Anne Burns and Nic Dias, "Information Disorder: Toward an interdisciplinary framework for research and policymaking," *Shorenstein Center on Media, Politics and Public Policy*, 31 October 2017, https://shorensteincenter.org/information-disorder-framework-for-research-and-policy-making/#The_Three_Types_of_Information_Disorder

[31] Soroush Vosoughi, Deb Roy, and Sinan Aral, "The spread of true and false news online," *Science*, 9 March 2018, pp. 1146-51, https://science.sciencemag.org/content/359/6380/1146

This implies that misinformation-containment policies should also emphasize behavioral interventions, like labeling and incentives to dissuade the spread of misinformation, rather than focusing exclusively on curtailing bots. [32]

Role of Other Institutions

While social media companies do provide a platform for disinformation/misinformation to spread, the public does recognize the role of media and other institutions as sources. A January 2020 NPR/Marist poll found that despite blaming tech companies for spreading disinformation, respondents pointed to different institution to reduce its flow with 39 percent to the media, 18 percent to tech companies, 15 percent to the government, and 12 percent to the public itself. In fact, 54 percent of Republicans responded that it is the media's responsibility to stop the spread of disinformation.[33]

Elites also play a key role in the dissemination of disinformation/misinformation. A Reuters study report found that prominent public figures have a disproportionate role in spreading misinformation with regard to COVID-19. Due to their prominence and recognition, they have very high levels of engagement on social media platforms:

> In terms of sources, top-down misinformation from politicians, celebrities, and other prominent public figures made up just 20 percent of the claims in our sample but accounted for 69 percent of total social media engagement. [34]

A Harvard study from October 2020 also found that elites and mass media were the primary perpetrators in disinformation around mail-in ballots and the risk for voter fraud during the November 2020 US election. The authors statetheir findings "suggest that disinformation campaign…was an elite-driven, mass-media led process in which social media played only a secondary and supportive role." They go on to further suggest that:

> The primary cure for the elite-driven, mass media communicated information disorder we observe here is unlikely to be more fact checking on Facebook. Instead, it is likely to require more aggressive policing by

---

[32] Vosoughi, Roy, and Aral, "true and false news."

[33] Brett Neely, "NPR Poll: Majority Of Americans Believe Trump Encourages Election Interference," *NPR*, 21 January 2020, https://www.npr.org/2020/01/21/797101409/npr-poll-majority-of-americans-believe-trump-encourages-election-interference

[34] J. Scott Brennen et al., "Types, sources, and claims of COVID-19 misinformation," *Reuters Institute*, 7 April 2020, https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation

traditional professional media, the Associated Press, the television networks, and local TV news editors of whether and how they cover Trump's propaganda efforts, and how they educate their audiences about the disinformation campaign the president and the Republican Party have waged.[35]

Based on the information above there are several conclusions that we can reach. First, social media companies should not be lumped into one category; they have different models, cultures, and resources. Secondly, the problem of disinformation is unfortunately an outgrowth to the medium as well to limitation in human cognition. Furthermore, indeed, social media companies have played a role in the dissemination of disinformation, however, it is increasingly becoming clear that the media and elites have played a major role as well.

> **"***Countries with higher levels of distrust have lower media literacy scores, with a correlation between trust in scientists and journalists and higher levels of media literacy.***"**

Moving Forward

There are numerous additional steps that can be taken to combat disinformation.

Creating Friction

The new strategies and tools that social media companies have deployed are starting to bear fruit. While platforms are designed to make sharing as easy as possible, platforms should continue to explore ways to create "friction" that makes it more difficult to automatically share bad content in the fight against disinformation.[36] Instagram did this with a pre-post prompt to curtail bullying. A recent Harvard study found that asking participants to explain how they knew that a political headline was true or false decreased their intention to share false headlines.

However, this is an ongoing process. Bad actors will always find new ways to game the system. In a recent podcast, Jack Dorsey reflected that he wished more

---

[35] Benkler et al., "Mail-In Voter Fraud: Anatomy of a Disinformation Campaign," *Berkman Klein Center for Internet & Society,* 2 October 2020, https://cyber.harvard.edu/publication/2020/Mail-in-Voter-Fraud-Disinformation-2020

[36] Erin Simpson and Adam Conner, "Fighting Coronavirus Misinformation and Disinformation: Preventive Product Recommendations for Social Media Platforms," *American Progress*, 18 August 2020, https://www.americanprogress.org/issues/technology-policy/reports/2020/08/18/488714/fighting-coronavirus-misinformation-disinformation/#fn-488714-44

disciplines had been included in the design of the product, such as "a game theorist to just really understand the ramifications of tiny decisions that we make, such as what happens with retweet versus retweet with comment and what happens when you put a count next to a like button?"[37] Companies can also develop teams with different perspectives that include backgrounds in ethics, sociology, and other fields to help foresee the societal impact of certain features and emerging risks.

Collaboration among Stakeholders

While social media companies can, and should, use both technology and policies to fight against disinformation, they cannot succeed alone. It will take a collaborative approach between all stakeholders to stem the tide. These stakeholders should include government as well as civil society. Governments around the world, especially Germany, Brazil, and the US, are reconsidering intermediary liability regulation as a way of holding companies accountable for content on their platforms. Governments should however resist the urge to regulate this problem away both because of partisan implications involved in some regulations (such as with Section 230 in the US) and because of the potential of such legislations to unintentionally stifle free speech. Rather, working with companies to get out and, where needed, amplify good information will be a healthier strategy.

Media Literacy

Civil society and governments can work on further developing media literacy. In a recent Pew Research study, 59 percent of those surveyed reported that it is hard to tell the difference between what is factual and what is misleading information.[38] Users need tools to help them better determine reliable information and sources; developing better critical thinking and analytical skills will be crucial in this regard.

A recent Open Society Institute report finds that that there is a positive relationship between the level of education and resilience to "fake news." Countries with higher levels of distrust have lower media literacy scores, with a correlation between trust in scientists and journalists and higher levels of media literacy. Finland, Sweden, and the Netherlands started teaching digital literacy and critical thinking about misinformation to schoolchildren.[39] More countries should consider including media literacy as a core 21st century skill.

---

[37] Interview with Jack Dorsey on "The Daily" with host Michael Barbaro, *New York Times*, 7 August 2020 (updated 19 August 2020), https://www.nytimes.com/2020/08/07/podcasts/the-daily/Jack-dorsey-twitter-trump.html

[38] Neely, "NPR Poll."

[39] Marin Lessenski, "Just think about it. Findings of the Media Literacy Index 2019," *European Policies Program Open Society Institute*, November 2019, https://osis.bg/wp-content/uploads/2019/11/MediaLiteracyIndex2019_-ENG.pdf

Addressing Broader Societal Issues

As Kentaro Toyama noted, "technology magnifies the underlying intent and capacity of people and institutions. But it doesn't in and of itself changes human intent, which evolves through non-technological social force."[40] We need to better understand the geopolitical, economic, and social factors that are driving both individuals and other actors to create the disinformation and misinformation tsunami. While it is tempting to see online issues in a void, they begin in the offline world. The last decade has seen increased protest and discontent with existing political and economic structures worldwide. Clearly something big is not working for many. Joshua Yaffa offers this suggestion:

> The real solution lies in crafting a society and a politics that are more responsive, credible, and just. Achieving that goal might require listening to those who are susceptible to disinformation, rather than mocking them and writing them off. [41]

> **"***The use of friction to curb viral sharing seems to be a strategy that should become more widely used.* **"**

We also need to accept that key institutions such as the media and elites are at least partially responsible as sources of disinformation/misinformation and find ways to hold them accountable.

Concluding Remarks

At the time this article was written, the 2020 US election results indicated that Joe Biden had won the election. Regardless, President Trump continued to claim in speeches and on social media that election fraud had taken place. Unlike the 2016 election this one confirms that disinformation is no longer a foreign interference issue but one which any ill-intentioned actor will use through a variety of mediums. As one article put it, we are face to face with "the bizarre reality that the biggest threat to American democracy right now is almost certainly the commander-in-chief, and

---

[40] Kentaro Toyama, "Twitter: It Won't Start a Revolution, But It can Feed One," *The Atlantic*, January 2011, https://www.theatlantic.com/technology/archive/2011/01/twitter-it-wont-start-a-revolution-but-it-can-feed-one/70530/

[41] Joshua Yaffa, "Is Russian Meddling as Dangerous as We Think?" *New Yorker*, 7 September 2020, https://www.newyorker.com/magazine/2020/09/14/is-russian-meddling-as-dangerous-as-we-think

that his primary mode of attack is a concerted disinformation campaign."[42]

Initial responses by social media companies to halt the spread of disinformation appear to have been at least partially successful. The use of friction to curb viral sharing seems to be a strategy that should become more widely used. Twitter in particular was assertive in labeling President Trump's tweets: over 1/3 of Trump's tweets were labeled with a warning between 3-6 November.[43] According to Twitter's statistics from 27 October-11 November, the company took action on about 300,000 tweets, or 0.2 percent of tweets, labeled under its Civic Integrity Policy for content that was potentially misleading. The company also indicated that about 74 percent of those who viewed the tweets saw them after a label or warning message was applied and there was a 29 percent decrease in quote tweets of the labeled tweets.[44]

Mass media, including traditionally right-wing Fox News, too took a more careful approach to election coverage. As a result, Fox attracted the ire of Donald Trump who urged his supporters to follow more fringe outlets such as NewsMax. As social media companies and mass media channels attempted to curb the amplification of disinformation, Trump and his supporters found new avenues to use. Trump supporters flocked to conservative social media and self-declared "unbiased" app Parler, which on 7 November ranked seventh on the App Store and by 8 November had reached first place.[45] The move from bigger platforms to more segmented ones increases the risk of echo chambers that will further allow echo chambers and misinformation to proliferate.

Such platforms will unfortunately continue to exist. The goal should of course be to halt the spread of disinformation but in doing so we need to also work on creating societies that foster critical thinking, and can have conversations and respect for differing viewpoints. While it is tempting to point to social media companies as responsible for the chaos caused by disinformation and misinformation, they alone are not to blame. It has taken the platforms time to respond with adequate tools and policies to combat disinformation in a deliberative manner. However, there are major offline forces at the heart of the disinformation problem. Until we address the issue more holistically, social media companies' responses will be wholly effective in eradicating disinformation.

---

[42] Julia Carrie Wong, "'Putin could only dream of it': how Trump became the biggest source of disinformation in 2020," *Guardian*, 2 November 2020, https://www.theguardian.com/us-news/2020/nov/02/trump-us-election-disinformation-russia

[43] Brian Heater, "Twitter labeled 300,000 US election tweets — around 0.2%," *TechCrunch*, 13 November 2020, https://techcrunch.com/2020/11/12/twitter-labeled-300000-us-election-tweets-around-0-2/

[44] Vijaye Gadde and Kayvon Beykpour, "An update on our work around the 2020 US Elections," *Twitter Blog*, 12 November 2020, https://blog.twitter.com/en_us/topics/company/2020/2020-election-update.html

[45] Taylor Hatmaker, "'Free speech' social network Parler tops app store rankings following Biden's election win," *TechCrunch*, 9 November 2020, https://techcrunch.com/2020/11/09/parler-app-store-facebook/