

Misinformation on Social Media

Abigayle Wright

MI 304: Information and Society

Dr. Ruth Shillair

December 4, 2022

Misinformation on Social Media

As said by Dr. Stephan Lewandowsky, “The fundamental problem with misinformation is that once people have heard it, they tend to believe and act on it, even after it’s been corrected,” (Abrams, 2021). Misinformation, along with disinformation, has become alarmingly dangerous on the internet, compromising the reliability and trust for gathering news and information from social media platforms. It is important to note that while misinformation is “false information that is spread, regardless of whether there is intent to mislead,” it can be easily tied to disinformation, which is “deliberately misleading or biased information” that can be used to manipulate users into believing lies (*News: Fake News, Misinformation & Disinformation*, 2022). As social media continues to grow, the shaded difference between the two has grayed since it has become much harder to recognize the true intentions behind any piece of falsified information. Although social media platforms are taking steps to help prevent the spread, the continued growth and popularity of misinformation shows that their efforts have not been strong enough. Misinformation and disinformation will continue to grow and harm others until clear, effective policies and guidelines are put in place in order to condemn the spread of it.

As observed by psychologists from the American Psychological Association, misinformation and disinformation have been used as damaging weapons across social media, igniting well-deserved concerns behind the dangerous claims relayed to users everywhere (Abrams, 2021). Despite the fact that some creators use misinformation to inform others of what they believe is right, most of the time, the people behind spreading disingenuous propaganda are deliberately using it to benefit themselves

while harming and manipulating others. As a result of the belief behind these lies, misinformation spreaders continue to become more influential, while society is forced to deal with the repercussions. In addition to individuals using misinformation to manipulate others, other parties, such as bots, can use disinformation for harassment. Also known as “software bots”, the Office of Cyber and Infrastructure Analysis from Homeland Security says “they can be used on social media platforms to do various useful and malicious tasks while simulating human behavior. These programs use artificial intelligence, big data analytics, and other programs or databases to imitate legitimate users posting content,” (2018). Due to their “super-spreader” agenda of a large variety, they are hard to track and regulate. Bot-caused harassment is especially prevalent on sites like Twitter and Twitch, where they are known to “overwhelm the user’s account to the point of deactivation,” (Homeland Security, 2018). The lies, manipulation, and harassment surrounding misinformation have shown to already have lasting and detrimental effects on big issues, including the coronavirus, vaccines, political elections, and even the war between Russia and Ukraine. Since misinformation being spread on social media has been used to heavily influence elections, specifically in the United States, it has also damaged credibility behind research and expert opinions about the COVID-19 pandemic (Himelein-Wachowiak et al., 2021). The falsified pieces of misinformation about these topics are easily spread across social platforms, specifically on Twitter, Facebook, TikTok, and Twitch. The only way to win the war on misinformation is for social media platforms themselves to enforce policies and consequences for starting or encouraging the spread of it.

If misinformation continues to exponentially escalate as it is now, it is guaranteed to cause destructive real-life events and consequences. With this, there have already been scandals behind misinformation on several platforms, one of the biggest being on Twitter. After Elon Musk recently bought Twitter, his ultimate goal was to “restore free speech” and use his newfound power to cater to the demands of users. With this, Twitter Blue was a newly released feature when Musk joined the company, and, so far, it has not been very successful. Users pay an \$8 monthly subscription fee for enhanced Twitter features, most popular being a blue check mark to indicate a verified account (*About Twitter Blue*, n.d.). Essentially allowing just anyone to buy a verification mark without valid verification has led to many complications, even leading Musk to go as far as temporarily suspending the service. Users have negatively impersonated celebrities and large companies, Musk included, which has led to Twitter Blue being recognized as a failure for the company. Even more disinformation has been caught on Twitter as a result of the subscription service, specifically by users using the paid-for verification mark as a way to gain relevance. For example, chaos ensued after a pharmaceutical company, Eli Lilly, was impersonated after a user bought a Twitter Blue verification mark for a username that was easily mistaken as the official account for the company. The account tweeted out “We are excited to announce insulin is free now.” As a result, the actual Twitter account for the company tweeted out an apology for the disinformation, saying, “We apologize to those who have been served a misleading message from a fake Lilly account.” (Kolodny & Pitt, 2022). Economic damage and unreliability for Eli Lilly followed, in addition to the disarray caused by the false advertising of free insulin on the platform. Since the launch, Twitter has had to make several adjustments to this

feature, and despite the fact that Musk has taken responsibility and made some changes, it is not that simple to recover. As Nilay Patel from The Verge said in an article dedicated to Musk, “you can make all the promises about ‘free speech’ you want, but the dull reality is that you still have to ban a bunch of legal speech if you want to make money,” (2022). Along with this, his “free speech” agenda has called for the restoration of accounts of many people with large followings who have been known to spread misinformation and hate speech, including Donald Trump, Andrew Tate, and Kanye West. According to The Washington Post, “The mass return of users who had been banned for such offenses as violent threats, harassment and misinformation will have a significant impact on the platform, experts said,” (Lorenz, 2022). As more misinformation from believable parodies and impersonations arise, Twitter and all other social media will become even more corrupted, with politics already massively impacted. Political polarization has become a result from the misinformation spread on social media and live broadcasts, including news sources like Fox News and CNN. As America entered the 2020 presidential election, according to the Pew Research Center, “evidence suggests that partisan polarization in the use and trust of media sources has widened in the past five years,” (Jurkowitz et al., 2020). After a 2014 study also done by the Pew Research Center, the divide seems to have grown even more since then, finding that “Republicans have grown increasingly alienated from most of the more established sources, while Democrats’ confidence in them remains stable, and in some cases, has strengthened.” The study also found that Republicans, by far, are reliant on a single news source, Fox News, where Democrats have been found to use more of a variety of news sources (Mitchell et al., 2014). Since Republicans are more likely to

consume information from one news source, they put themselves at a disadvantage for avoiding misinformation, since they, most of the time, do not check facts alongside other reporting outlets, carelessly believing just about anything they hear. This can cause greater harm and division as elections continue to happen, because news sources, even if unintentionally, increase political polarization and distrust of those with opposing views. This polarization from misinformation will continue to divide political parties, especially in the United States, which could end up dividing society altogether.

Although they are not always successful, social media platforms have been credited with making valiant efforts to try to minimize the spread of misinformation and disinformation. Over many platforms, including Facebook, Instagram, Twitch, and Twitter, policies and features have been introduced to fight misinformation, which span across a spectrum from highly effective to not effective at all. Twitch, mostly known to have harassment and misinformation shared in live-stream chat boxes, has taken steps to disprove and remove harmful misinformation topics, along with live streamers who continuously spread it. It is emphasized that they only enforce consequences against streamers when all of the pieces of criteria are met, as mentioned in a post on Twitch's website. "Together, we've identified three characteristics that all of these actors share: their online presence – whether on or off Twitch – is dedicated to (1) persistently sharing (2) widely disproven and broadly shared (3) harmful misinformation topics, such as conspiracies that promote violence," (*Preventing Harmful Misinformation Actors on Twitch*, n.d.). However, since fighting misinformation across live broadcasts can be very difficult to control, Twitch's policies may not be as effective as intended, because as soon as misinformation is shared through clips or short snippets from streams, the

damage is already done. Content directly from the platform may be removed, but controlling these clips from being spread on other platforms may be very difficult to achieve. Considering all of this, Twitch seems to be making a conscious effort, yet another platform, Facebook, has added very progressive ways of fighting misinformation. For example, Facebook now works with “independent, third-party fact-checking organizations who are certified through the non-partisan International Fact-Checking Network (IFCN) to identify, review and take action on this content,” (*Facebook’s Enforcement of Fact-Checker Ratings*, n.d.). Facebook also has many other policies on misinformation, including several pages on removal, enforcements, consequences, and how they are fact-checking content. Shown through evidence collected from extensive management approaches and enforced policies, Facebook seems to be the most dedicated and successful platform to fight misinformation. “During the month of April, we put warning labels on about 50 million pieces of content related to COVID-19 on Facebook, based on around 7,500 articles by our independent fact-checking partners,” (Rosen, 2020). Platforms, even very small ones, should take inspiration from Facebook on misinformation; Facebook policies have shown to be very efficient and effective, and it is all laid out to view on their website. With this in mind, when comparing the efforts of Facebook, Twitter, and Twitch specifically, Facebook’s policies would rank toward the top, Twitch closer to the middle, and Twitter near the lower end of the spectrum. Even with all of the changes Twitter has tried to make, it is still falling short in monitoring the misinformation being posted. Though many could do more, it is encouraging to see that platforms are taking any initiative to fight misinformation.

When trying to relinquish the harms of misinformation and harassment, encouraging platforms to use social media as a way to spread positive, educational, and important news could be very effective. To start, platforms must be willing to use their power to emphasize and spread news from reliable and verified sources, rather than only discrediting misinformation. Along with this, making efforts to have social media become more community driven could allow individuals to more safely express their insight among those they share opinions with, leading to less harassment and danger over differing arguments. In addition to platforms taking responsibility, it is important for users to have adequate digital literacy, along with sufficient guidance and knowledge on how to distinguish real news from fake news. According to The American Library Association's digital-literacy task force, "Digital Literacy is the ability to use information and communication technologies to find, evaluate, create, and communicate information, requiring both cognitive and technical skills." Digital literacy is important to acknowledge as an internet user, and with this, improving digital literacy by providing education and awareness on how to avoid misinformation has been proven to be very effective. As noticed by Dr. Sam Wineburg after taking the initiative to include a lecture on online disinformation for his undergraduate students at Stanford University, he discovered that "Once a group of smart Stanford undergraduates sees how easily they are taken in, how easily they fall for this stuff, they suddenly perk up and realize they have something to learn. Nobody wants to be an easy mark," (Pappas, 2022).

Improving individual digital literacy is mostly the responsibility of the user, but since it is unrealistic to assume that all schools would suddenly add digital literacy into their curriculum, responsibility can be thrown back to platforms, who should take their own

measures to provide users with guides on how to filter out questionable information on their platform specifically. Overall, platforms should focus on highlighting reliability, safety, support, and help to make users feel safer and more likely to continue using their site.

After seeing what misinformation is capable of through studies, real-life events, and expert opinions, it is very clear that it causes more harm than good, not only for social media users, but for society in general. Manipulation through disinformation has become especially easy due to an overall lack of digital literacy across the internet, with average social media users falling victim every day. It is important for social media platforms to start taking more responsibility and action against the disinformation, harassment, and hate speech that is continuously spread on their sites every minute. Although many have already made efforts and have come out successfully, they could still be doing more to protect the safety of their users against misinformation and its harmful manipulative tactics. By the looks of it, misinformation will not go away on its own, so if popular platforms do not start taking serious action, it may become the catalyst of their inevitable downfall.

Works Cited

- About Twitter Blue*. (n.d.). Twitter Help Center. Retrieved December 1, 2022, from <https://help.twitter.com/en/using-twitter/twitter-blue>
- Abrams, Z. (2021). Controlling the spread of misinformation. *American Psychological Association*, 52(2), 44, <https://www.apa.org/monitor/2021/03/controlling-misinformation>
- Digital Literacy*. (n.d.). American Library Association. Retrieved December 1, 2022, from <https://literacy.ala.org/digital-literacy/>
- Facebook's Enforcement of Fact-Checker Ratings*. (n.d.). Facebook. Retrieved December 1, 2022, from <https://www.facebook.com/business/helpM/297022994952764?id=673052479947730>
- Himelein-Wachowiak, M., Giorgi, S., Devoto, A., Rahman, M., Ungar, L., Schwartz, H. A., Epstein, D. H., Leggio, L., & Curtis, B. (2021). Bots and Misinformation Spread on Social Media: Implications for COVID-19. *Journal of Medical Internet Research*, 23(5), e26933. <https://doi.org/10.2196/26933>
- Homeland Security. (2018, May). *Social Media Bots Overview*. https://niccs.cisa.gov/sites/default/files/documents/pdf/ncsam_socialmediabotsoverview_508.pdf?trackDocs=ncsam_socialmediabotsoverview_508.pdf
- Jurkowitz, M., Mitchell, A., Shearer, E., & Walker, M. (2020, January 24). *U.S. Media Polarization and the 2020 Election: A Nation Divided*. Pew Research Center. <https://www.pewresearch.org/journalism/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided/>
- Kolodny, L., & Pitt, S. (2022, November 11). *Twitter pauses paid verifications after users abuse service to impersonate brands and people*. CNBC. <https://www.cnn.com/2022/11/11/twitter-blue-subscription-disappears-from-app.html>
- Lorenz, T. (2022, November 24). 'Opening the gates of hell': Musk says he will revive banned accounts. The Washington Post. <https://www.washingtonpost.com/technology/2022/11/24/twitter-musk-reverses-suspensions/>
- Mitchell, A., Gottfried, J., Kiley, J., & Matsa, K. E. (2014, October 21). *Political Polarization & Media Habits*. Pew Research Center. <https://www.pewresearch.org/journalism/2014/10/21/political-polarization-media-habits/>
- News: Fake News, Misinformation & Disinformation*. (2022, September 29). Library Guides at University of Washington Libraries. Retrieved November 18, 2022, from <https://guides.lib.uw.edu/bothell/news/misinfo>
- Pappas, S. (2022). Fighting fake news in the classroom. *American Psychological Association*, 53(1), 87. <https://www.apa.org/monitor/2022/01/career-fake-news>
- Patel, N. (2022, October 28). *Welcome to hell, Elon*. The Verge. <https://www.theverge.com/2022/10/28/23428132/elon-musk-twitter-acquisition-problems-speech-mode-ration>
- Preventing Harmful Misinformation Actors on Twitch*. (n.d.). Twitch Safety Center. Retrieved November 18, 2022, from https://safety.twitch.tv/s/article/Preventing-Misinformation-Actors-from-Using-Twitch?language=en_US
- Rosen, G. (2020, April 16). *An Update on Our Work to Keep People Informed and Limit Misinformation About COVID-19*. Meta. <https://about.fb.com/news/2020/04/covid-19-misinfo-update/>