

**Literature Review:**

**Impact of retracted journal articles: online COVID-19 misinformation**

Andy Hu

Section M

Massachusetts Academy of Math and Science

Kevin Crowthers

December 7, 2020

## **Introduction**

The Information Age has allowed for the rapid dissemination of information throughout society, but it has not been without consequences. One such consequence is the increased spread of misinformation. The Information Age has provided misinformation with the same ability to spread quickly as it has given information in general. In fact, misinformation spreads faster than correct information on Twitter (Wang et al., 2019). As such, the topic of misinformation on social media is an area of increasing interest due to the increased prevalence of social media in modern day society (Ospina, 2019). One possible source of misinformation on social media is retracted journal articles. Through a trust in the institution of science retracted journal articles can spread misinformation by causing the general public to believe that they have the correct facts when they do not. Having the wrong facts on any given issue can cause a myriad of issues, but the one that this paper focuses on is health problems stemming from misinformation. Health problems that stem from misinformation include things like the anti-vaccination movement as well as the cloud of misinformation surrounding COVID-19: both of which relate in part to retracted journal articles that spread through social media. One way that the impact of health misinformation has been quantified has been through sentiment analysis. However, little research has been done looking at quantifying the impact of retracted journal articles on misinformation with sentiment analysis.

## Health Misinformation

Throughout this paper, the term misinformation will be used to mean information that is false and spreads. Misinformation is not dependent on the intentions of the propagator, unlike its close counterpart, disinformation. Though misinformation has a multitude of impacts on society as a whole, the one that this paper focuses on is how misinformation affects health. Even the term “health” indicates a broad field of studies and many different subjects. The primary interest of this paper is in the COVID-19 pandemic, and how misinformation has increased the influence of the virus. However, since the pandemic is ongoing, it is important to look at how misinformation has impacted health in the past in order to gauge some more information about how the pandemic might play out. One perfect example is the anti-vaccination movement.

### Vaccines

In 2010, a study by Andrew Wakefield linking vaccines to autism was retracted due to several flaws within the paper (Eggertson, 2010). Upon further review, Wakefield and his coauthors were found to have conducted ethical violations, scientific misrepresentation of data, and deliberate fraud (Rao & Andrade, 2011). The study was published in 1998 by *The Lancet* — a well renowned medical journal that ranks 6<sup>th</sup> in the field of medicine (Scimago Journal & Country, 2020). Just a year after the Wakefield study had been published, other researchers had already disproved its findings (Rao & Andrade, 2011). Despite this, it took 12 years for the Lancet to finally fully retract the article, and by then, it had already had a strong influence on society. For example, many parents in 2008 and 2009 decided not to vaccinate kids, and small pockets of measles outbreaks occurred in the UK, USA, and Canada (Rao & Andrade,

2011). Measles was thought to be more or less eradicated, but the anti-vaccination movement proved to have a large enough effect to bring back the deadly disease. In essence, the deaths and suffering caused by the measles outbreak were avoidable, if it were not for the misinformation that perpetuated itself with the Wakefield study. Vaccination is still under hot debate in 2020, showing how persistent misinformation can be. Misinformation about vaccines will even end up impacting COVID-19 vaccination efforts, showing a need for more studies to researching how to either correct misinformation once it spreads, or how to ultimately prevent misinformation from spreading in the first place.

### COVID-19

Misinformation about COVID-19 is so great, that the general effects of the disease have been described as not only a pandemic, but an infodemic as well. The virus hit hard and fast in its early days, and some governments responded better than others. However, all governments and people in general were mostly in the dark about the effects of the virus. Information was relatively scarce as COVID-19 was a novel disease. As such, researchers rushed to find information, to perform experiments, and to publish their results. The usually rigorous process of vetting journal articles for potential flaws lapsed a bit, causing the rates of retractions on COVID-19 far surpass the normal rate (Yeo-The & Tang, 2020). Due to the high-profile of the virus, media outlets and social media users in general latched on to the newest research that was being done. Some of this research, later found to be in need of retraction, had already spread far and wide, influencing the minds of many.

A good example of this is hydroxychloroquine — a drug touted by President Trump for being effective at treating COVID-19. His support for hydroxychloroquine did not appear out of thin air. To his credit, studies were being conducted that did show some promise of a possible effective treatment. However, his initial support for the drug was far more than what was warranted, and his continued support when scientific evidence denouncing the drug was produced showed a troubling problem with misinformation (Greenburg, 2020). The fact that COVID-19 misinformation had been able to make its way up to the presidential office shows how pervasive misinformation really is. Even worse, because of the power and influence of the office of the president, the misinformation spread only increased from there on. One of Trump's favorite ways of communicating to the public his ideas- one of the ways in which he spread misinformation rapidly- is his personal Twitter account. However, he is not alone accountable for the misinformation on online social media, and there are many things to look at when it comes to misinformation and social media.

### **Social Media**

#### **Misinformation propagation on social media**

Misinformation online can be spread through many mediums. For example, news outlets and webpages are both vectors through which misinformation spreads. However, news outlets have existed for a long time and books have existed for even longer. While these two do their fair share of spreading misinformation, they do not nearly compare to the ability of social media to spread misinformation.

Looking at Figure 1, from 2004 to 2018 social media use has grown rapidly, with Facebook alone garnering 2 billion users. As social media became more relevant, its

ability to spread misinformation became clearer. In particular, the 2016 election was

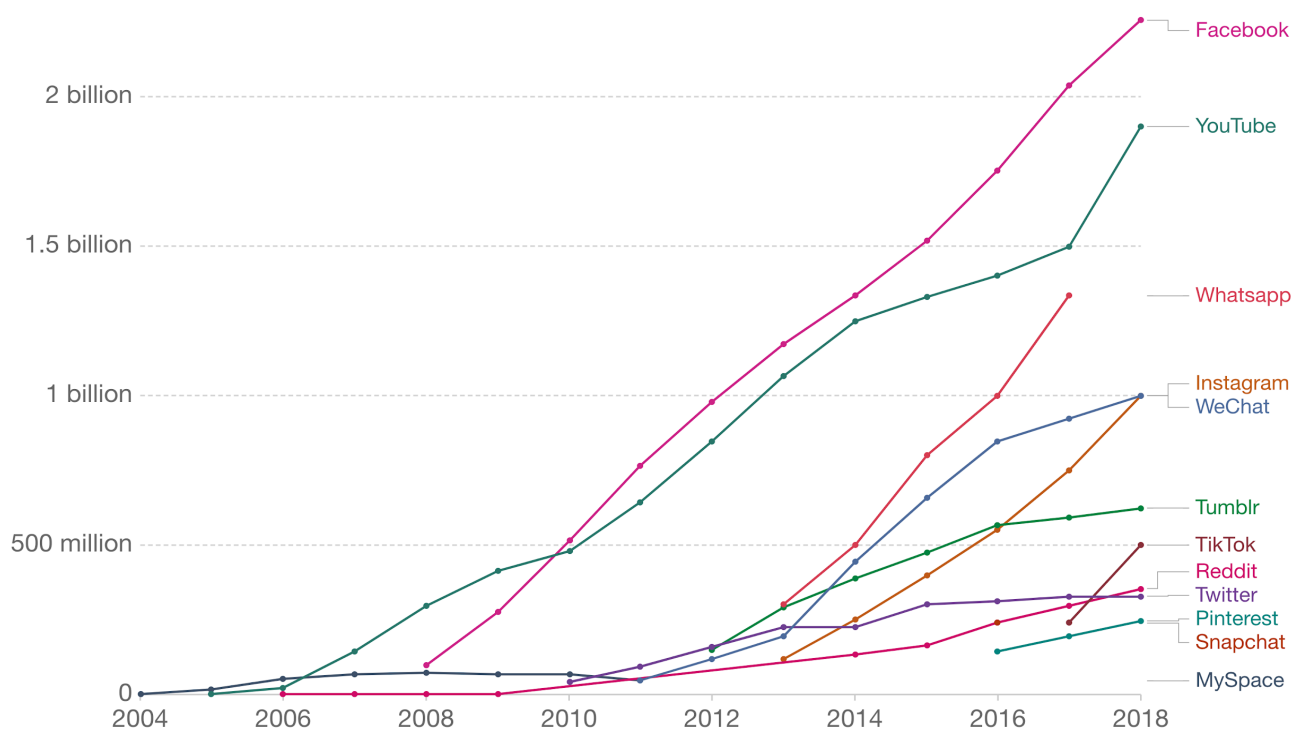
a

point of interest in looking at misinformation through social media. It is debatable

## Figure 1

### Number of people using social media platforms, 2004 to 2018

Estimates correspond to monthly active users (MAUs). Facebook, for example, measures MAUs as users that have logged in during the past 30 days. See source for more details.



Source: Statista and TNW (2019)

CC BY

Note: Various social media platforms show an overall rise in usage. Data is used for up to 2018.

2019 data not used because it was incomplete. From *The rise of social media*, by Ospina, E.O.,

2019, Our World in Data Results section, Retrieved November 15, 2020, from

<https://ourworldindata.org/rise-of-social-media>

exactly how much influence the Russian disinformation campaign was able to influence the 2016 election, but it is clear that social media — most notably Facebook — played a large role in allowing it to happen (Kurtzleben, 2018). Even beyond concerted attempts

by groups or individuals, misinformation is commonly shared in good faith — by people who believe that what they are sharing is the truth. In fact, people mostly share misinformation believing that it is correct (Buchanan, 2020). So, while concerted disinformation efforts are certainly troubling, the more pervasive issue is the seemingly benign accidental missteps which can lead to grave consequences – such as the anti-vaccination movement.

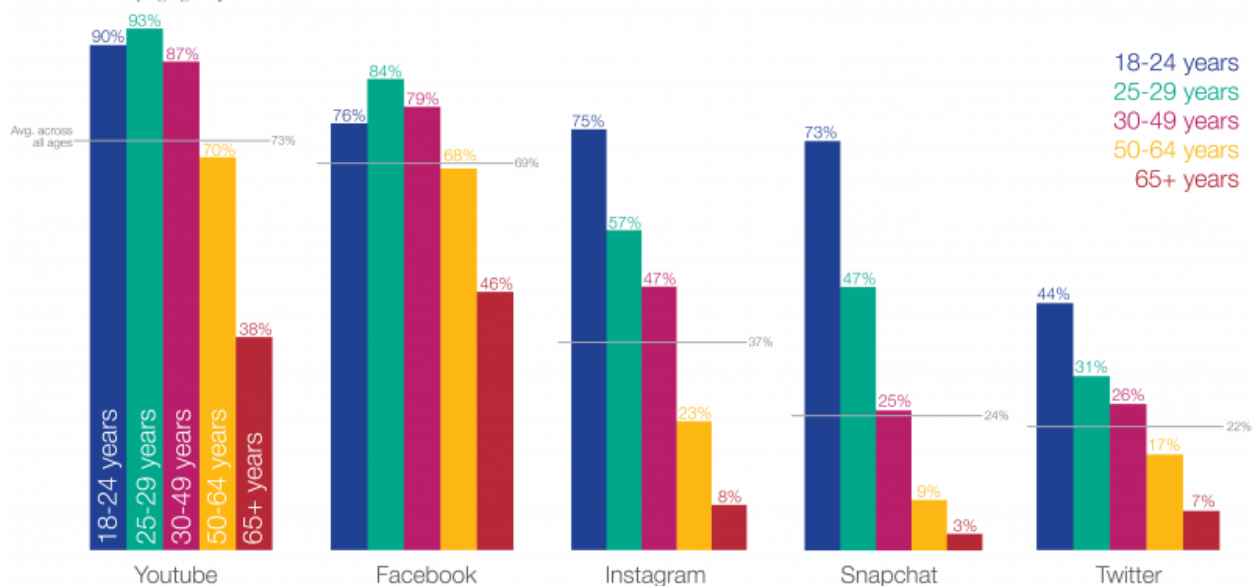
As shown in Figure 2, along with the fact that social media usage is growing larger year by year, those who use social media also tend to be in the younger generations. As time progresses and the older generations pass away, the younger generations will become the new older generation. Of course, because they have grown up with social media, they will likely continue to use it as long

**Figure 2**

### Use of social media platforms by age group in the US

The share of adults in the United States who say they ever use the following online platforms or social media apps in 2019. This is shown by age group.

Our World  
in Data



Data source: Pew Research Center (2019).

This is a visualization from OurWorldinData.org, where you find data and research to make progress against the world's largest problems.

Licensed under CC-BY by the author Esteban Ortiz-Ospina.

Note: Younger generations use all social media platforms more than older generations in general.

From *The rise of social media*, by Ospina, E.O., 2019, Our World in Data Results section, Retrieved November 15, 2020, <https://ourworldindata.org/rise-of-social-media>

as something new does not come along. Further evidence is how those under 18 show lots of usage of social media in general. What this means is that social media usage will continue to increase in the coming years, and the more that people get information from social media, the more misinformation that social media will end up spreading. Social media is a clear factor in allowing misinformation to spread such as the Wakefield study linking vaccines to autism. The anti-vaccination movement still has roughly 7 million individuals associated with it on Facebook alone, indicating that misinformation spread by social media will have a long-lasting effect on health information (Smith, 2017).

However, while overall social media use will increase in the coming years, digital literacy skills will as well. Schools teach courses on how to identify reputable sources from sources more likely to contain misinformation. Despite the fact that those older 65 are more likely to share misinformation on Facebook than other population groups, the coming years will still most likely result in less misinformation sharing overall (Mheidly & Fares, 2020). This may be due to the fact that older generations simply did not grow up with the internet and thus have not acquired the digital literacy that younger generations have. However, with retracted scientific journal articles, a problem still remains: a journal article may be purported as correct for a period of time before it is retracted, exactly what happened with Wakefield's paper. Thus, it is important to look at and analyze retractions in science to make sure that they are not overbearing sources of misinformation.



## **Retracted Scientific Journal Articles**

A journal article may be retracted for one of two reasons: misconduct or honest error. Misconduct includes unstated conflicts of interest, salami slicing (using one data set to make many studies), plagiarism, data fabrication, or ethical violations. Honest errors include sample or data errors, unverifiable information, unreproducible studies, and redundancies (Ewuoso, 2018).

In recent years, the proportion of retracted journal articles has increased. However, the number of retracted articles remains low in general, and the increase could possibly be attributed to better self-policing by scientific journals. This is supported by another study looking at retracted papers in general which found that for 714 retracted articles published from 1973 to 2002, retraction took an average of 49.82 months while for 1333 retracted articles published after 2002, retraction took 23.82 months (Steen, 2013). This implies that it is not necessarily the volume of retracted journal articles that are increasing, but rather better detection methods by the scientific community. Numerous studies have looked at and analyzed retracted articles in specific fields to see if there were any problems. However, the impact of retracted scientific journal articles outside of the scientific community has gone less studied, especially on online social media. This is in part because there are relatively low rates of retraction in general and because social media is still growing. However, as shown by Wakefield and his papers, retracted scientific journal articles have the possibility of having extreme negative effects on public health through social media. In terms of vaccines, one way of measuring feelings towards vaccinations has been by analyzing

the emotions conveyed by text on social media — or in other words, sentiment analysis (Raghupathi et al., 2020).

### **Sentiment Analysis**

#### **How is Sentiment Analysis Done?**

Sentiment analysis is the means by which the feelings conveyed by text are analyzed. At a high level, there are three broad categories of sentiment analysis: rule-based, automatic, and hybrid (Everything There Is to Know about Sentiment Analysis, 2020). A rule-based system uses rules created by humans to determine the sentiment of texts. The rules can be as simple as assigning words positive and negative and counting how often each word occurs in a sentence, or they can be complex, involving techniques like dependency parsing, lemmatization, stemming, and many others (Natural Language Processing (NLP) – What Is NLP & How Does it Work?, 2020). An automatic system relies more on machine learning with typical models such as linear regression and deep learning. Lastly, a hybrid model simply combines the best features of both the rule-based and automatic into one system that often results in higher accuracy (Everything There Is to Know about Sentiment Analysis, 2020).

#### **How has Sentiment Analysis Been Used?**

Sentiment analysis is often used to analyze feelings of customers for big businesses looking to maximize profits, though researchers have found other ways to use it (Everything There Is to Know about Sentiment Analysis, 2020). One such way is in analyzing user posts and comments on social medias to garner information on general sentiments towards a subject. For example, one particular study looked at the sentiments towards vaccines on Twitter for a set time period (Raghupathi et al., 2020).

Using a more rule-based approach, the results indicated that sentiments towards vaccinations were mostly positive, and many did not buy into anti-vaccination beliefs (Raghupathi et al., 2020). Other studies have looked at the sentiment towards infectious diseases on social media using a variety the aforementioned methods of sentiment analysis (Alamoodi. 2020). However, no studies to date were found that have analyzed sentiment on social media towards retracted scientific articles, thus pointing towards an area of further research.

### **Conclusion**

Due to the nature of the COVID-19 pandemic, misinformation has been allowed to spread on social media through retracted scientific journal articles. It is unclear how much these journal articles will affect health in the coming days, but research on the origin of the anti-vaccination movement does not give a positive indicator. Overall, more research is needed better ascertain whether or not retracted scientific journal articles are having a strong influence on public perception of the COVID-19 pandemic. One possible way of quantifying this is through sentiment analysis on social media posts and comments relating to retracted journal articles. However, even beyond the impacts towards the health of the general public, journals must adhere more strongly to the thorough peer-review that seems to have lapsed due to the pandemic in order to limit the spread of misinformation and protect the integrity of science as a whole.

### **References**

- Al-Ghareeb, A., Hillel, S., McKenna, L., Cleary, M., Visentin, D., Jones, M., Bressington, D., & Gray, R. (2018). Retraction of publications in nursing and

midwifery research: A systematic review. *International Journal of Nursing Studies*, 81, 8–13. <https://doi.org/10.1016/j.ijnurstu.2018.01.013>

Alamoodi, A., Zaidan, B., Zaidan, A., Albahri, O., Mohammed, K., Malik, R., . . .

Alaa, M. (2020). Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Systems with Applications*, 114155. doi:10.1016/j.eswa.2020.114155

Buchanan, T. (2020). Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *PLOS ONE*, 15(10), e0239666.

<https://doi.org/10.1371/journal.pone.0239666>

Eggertson, L. (2010). Lancet retracts 12-year-old article linking autism to MMR vaccines. *CMAJ: Canadian Medical Association Journal*, 182(4), E199–E200.

<https://doi.org/10.1503/cmaj.109-3179>

Everything There Is to Know about Sentiment Analysis. (2020). Retrieved from

<https://monkeylearn.com/sentiment-analysis/>

Ewuoso, O. C. (2018). Misconduct vs. Honest errors: Should honest errors in research be punished? *Biomedical Journal of Scientific & Technical*

*Research*, 8(5) <https://doi.org/10.26717/bjstr.2018.08.001715>

Greenburg, J. (2020). *PolitiFact - Trump wrong that hydroxychloroquine studies only gave drug to dying patients.* @politifact.

<https://www.politifact.com/factchecks/2020/may/20/donald-trump/trump-wrong-hydroxychloroquine-studies-only-gave-d/>

Kurtzleben, D. (2018) *Did Fake News On Facebook Help Elect Trump? Here's*

*What We Know*. (n.d.). NPR.Org. Retrieved November 23, 2020, from <https://www.npr.org/2018/04/11/601323233/6-facts-we-know-about-fake-news-in-the-2016-election>

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(3), 106–131.

Lucas, R. (2017, November 1). *How Russia used Facebook to organize 2 sets of protesters*. NPR.org. <https://www.npr.org/2017/11/01/561427876/how-russia-used-facebook-to-organize-two-sets-of-protesters>

Mheidly, N., & Fares, J. (2020). Leveraging media and health communication strategies to overcome the COVID-19 infodemic. *Journal of Public Health Policy*, 41(4), 410–420. <https://doi.org/10.1057/s41271-020-00247-w>

Natural Language Processing (NLP) – What Is NLP & How Does it Work? (2020). Retrieved from <https://monkeylearn.com/natural-language-processing/#techniques>

Ospina, E.O. (2019). *The rise of social media*. Our World in Data. Retrieved November 15, 2020, from <https://ourworldindata.org/rise-of-social-media>

Raghupathi, V., Ren, J., & Raghupathi, W. (2020). Studying Public Perception about Vaccination: A Sentiment Analysis of Tweets. *International Journal of Environmental Research and Public Health*, 17(10), 3464.

doi:10.3390/ijerph17103464

- Rao, T. S. S., & Andrade, C. (2011). The MMR vaccine and autism: Sensation, refutation, retraction, and fraud. *Indian Journal of Psychiatry*, 53(2), 95–96.  
<https://doi.org/10.4103/0019-5545.82529>
- Scimago Journal & Country. (2020). *Journal rankings on medicine (miscellaneous)*.  
Scimago Journal & Country Rank.  
<https://www.scimagojr.com/journalrank.php?category=2701>
- Smith T.C. (2017). *Vaccine Rejection and Hesitancy: A Review and Call to Action* | *Open Forum Infectious Diseases* | Oxford Academic. Retrieved November 23, 2020, from <https://academic.oup.com/ofid/article/4/3/ofx146/3978712>
- United Nations. (2020). *Scientists optimistic about COVID-19 vaccines for all*. UN News. <https://news.un.org/en/story/2020/12/1079322>
- Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Social Science & Medicine*, 240, 112552.  
<https://doi.org/10.1016/j.socscimed.2019.112552>
- Yeo-Teh, N. S., & Tang, B. L. (2020). An alarming retraction rate for scientific publications on coronavirus disease 2019 (COVID-19). *Accountability in Research*, 1-7. <https://doi.org/10.1080/08989621.2020.1782203>