

---

# Misinformation and Hate Speech: The Case of Anti-Asian Hate Speech During the COVID-19 Pandemic

Jae Yeon Kim and Aniket Kesari

---

**Abstract.** Donald Trump linked COVID-19 to Chinese people on March 16, 2020, by calling it the Chinese virus. Using 59,337 US tweets related to COVID-19 and anti-Asian hate, we analyzed how Trump’s anti-Asian speech altered online hate speech content. Trump increased the prevalence of both anti-Asian hate speech and counterhate speech. In addition, there is a linkage between hate speech and misinformation. Both before and after Trump’s tweet, hate speech speakers shared misinformation regarding the role of the Chinese government in the origin and spread of COVID-19. However, this tendency was amplified in the post-Trump tweet period. The literature on misinformation and hate speech has been developed in parallel, yet misinformation and hate speech are often interwoven in practice. This association may exist because biased people justify and defend their hate speech using misinformation.

---

## 1 Introduction

On March 16, 2020, Donald Trump, then president of the United States, linked Asian Americans to COVID-19 by calling the virus the Chinese virus.<sup>1</sup> As many studies have documented (Relia et al. 2019; Chan, Kim, and Leung 2021; Jeung and Nham 2020; Hswen et al. 2021; Lu and Sheng 2020; Gover, Harper, and Langton 2020; Roberto, Johnson, and Rauhaus 2020), the speech made “the Chinese virus” trend on social media and worsened the anti-Asian climate in the United States. In its first four weeks of operation, the Stop AAPI Hate website (Choi and Kulkarni 2020) received 1,497 anti-Asian hate crime reports from March 19 to April 15, 2020. These incidents targeted not just Chinese but also Filipino, Thai, and Korean Americans (Jeung et al. 2021) because prejudiced Americans do not always distinguish these Asian subgroups (Le Espiritu 1992).<sup>2</sup> We ask the question of whether Trump’s characterization of COVID-19 as the

---

1. Specifically, he used this language in a March 16 tweet that has since been deleted after his account was suspended, but the text can be seen here: <https://www.theatlantic.com/notes/2020/03/2020-time-capsule-5-the-chinese-virus/608260/>. That same day, he also gave a speech detailing the federal government’s response to the pandemic, referred to the virus as “starting in China,” and described his own administration’s actions as making a “lifesaving move with early action on China.” Full transcript available [here](#).

2. A case in point is the murder of Vincent Chin, a Chinese American man beaten to death by two White men in 1982. These White men were autoworkers who assumed that Chin was of Japanese descent and blamed

“Chinese virus” in association with laying blame on the Chinese government shifted the context of that phrase in online discourse.

This article adds to this growing literature on anti-Asian bias during the COVID-19 pandemic by showing the connection between misinformation and hate speech behind the anti-Asian bias that appeared on Twitter. Specifically, we focus on how Trump’s anti-Asian speech altered online hate speech content, as measured by changing word vectors. By analyzing 59,337 tweets created by US users and related to COVID-19 and Asian hate (Ziems et al. 2020), we track the change in the content of anti-Asian hate speech and counterhate speech before and after Trump’s tweet. We find that Trump’s tweet coincided with an increase in hateful content. The vector representation of the phrase “Chinese virus” became more closely associated with criticisms of the Chinese government. Although the work on misinformation (e.g., Swire-Thompson and Lazer 2019; Jerit and Zhao 2020; Gallotti et al. 2020; Kouzy et al. 2020) and hate speech has been conducted in isolation (e.g., Howard 2019; Coates 2020), we argue that in practice, these two forms of online speech are interwoven. This association may exist because biased people<sup>3</sup> justify and defend their hate speech using misinformation. Despite a recognition that terms like “Chinese virus,” “Wuhan virus,” and “Kung Flu,” are racially problematic, their use in conjunction with misinformation suggests that misinformation may be used to deflect such criticisms. The rest of the article explains how hate speech speakers connect hate speech and misinformation. Next, we explain the data source and empirical strategies. We then present findings and discuss their implications and limitations.

## 2 Misinformation and Hate Speech

The COVID-19 virus is new, but anti-Asian attitudes in the United States are not. The history of discrimination against Asian communities runs deep in US history. For example, the US government banned Chinese immigration in the late nineteenth century (1882) and incarcerated Japanese people, including US-born citizens, during World War II (1942-1946) (Kim 1999; Junn 2007; Ngai 2014; Nguyen, Carter, and Carter 2019).

Nevertheless, history does not repeat itself in the same way. Compared to the late nineteenth and early twentieth centuries, attitudes toward Asian immigration and Asian American communities are now much more liberal (Skrentny 2009; Schickler 2016). In addition, the size of Asian populations in the US has increased substantially. Revealing explicit and outward bias toward Asian Americans, therefore, comes with political costs and social sanctions.

Under these circumstances, biased people might be incentivized to justify and defend their anti-Asian hate speech by citing misinformation. To avoid the social sanctions associated with overt bias, individuals may suffer from a cognitive dissonance between needing to superficially uphold progressive social norms and justifying racially motivated biased attitudes and actions (Festinger 1957; Elliot and Devine 1994). Misinformation provides an alternative cognitive space for biased people, where their attitudes and actions can be justified. Thus, even though they are accused of xenophobia and racism, these biased people might not feel a strong sense of guilt or shame because they, and their group, believe they still hold the moral high ground (Kinder and Kam 2010). For instance, as we shall see, anti-Asian hate speakers justified calling COVID-19 the Chinese virus

---

him for the decline of the US auto industry. At that time, Japanese automakers were becoming competitive in the US market.

3. Biased people are defined as individuals who exhibit bias or animosity toward Asians or Asian Americans, as demonstrated by their hate speech content.

because they believed that Chinese government propaganda aimed to dissociate China from the virus. In their view, if people want to make China accountable for its accused misconduct, calling the virus the Chinese virus is preferable over calling it COVID-19. This is a conspiracy theory because it is impossible to falsify the claim (Harris 2018, 243). More importantly, the reason that health authorities recommend not connecting a virus to a particular region, nation, or people is beyond political; it is to avoid causing harm to innocent people, as evident in the rise of anti-Asian hate crimes. Yet, psychologically, this tactic is useful because it helps protect its users' social identity and status by avoiding the accusation that they are motivated by hate or bias—they can instead use criticisms of the Chinese government to deflect this critique (Tajfel 1970; Tajfel and Turner 1979; Bail 2021; Kim et al. 2021).

### 3 Hypotheses

We draw two descriptive hypotheses from the earlier theoretical discussions.

- H1: Trump's tweet was associated with an uptick in both anti-Asian hate speech and counterhate speech.
- H2: Both before and after Trump's tweet, hate speech speakers shared misinformation regarding the role of the Chinese government in the origin and spread of COVID-19. However, this tendency was amplified in the post-Trump tweet period.

### 4 Data

We investigate how the public discourse on anti-Asian hate changed before and after Trump's tweet. We looked for a large-scale Twitter dataset, where tweets are annotated as hate speech and counterhate speech. Using this criteria, we chose the COVID-HATE dataset, the largest Twitter dataset of anti-Asian hate and counterhate speech.<sup>4</sup> Ziems et al. (2020) created this dataset by collecting over 30 million tweets from January 15 through April 17, 2020.<sup>5</sup> They automatically classified 891,204 hate tweets and 200,198 counterhate tweets based on the hand-labeled dataset of 2,400 tweets.<sup>6</sup> These tweets were divided into hate speech, counterhate speech, neutral, and other categories.<sup>7</sup> They used BERT (Bidirectional Encoder Representations from Transformers) embeddings for feature engineering and a logistic regression model as a classifier. The logistic regression model was selected after testing its performance against that of random forest and support vector machine classifiers.

Our analysis has a geographical focus. For this reason, we limited the data to the geocoded tweets ( $n = 5,109,492$ ).<sup>8</sup> Ziems et al. (2020) identified these locations using either the OpenStreetAPI or the tweet users' profile location. We used this sample to conduct descriptive analysis. For the word embedding analysis, we used the subset of

4. For more information, see <http://claws.cc.gatech.edu/covid>.

5. For this task, they used a collection of 42 keywords and hashtags related to COVID-19 (e.g., #coronavirus), hate speech (e.g., #chinavirus), and counterhate speech (e.g., #RacismIsVirus).

6. These tweets were randomly sampled in multiple rounds. Ziems et al. (2020) intentionally oversampled tweets that include Asian, hate, and counterhate speech terms.

7. In their definition, anti-Asian COVID-19 hate speech is COVID-19-related antagonistic speech that is directed towards Asians and is abusive and derogatory. Counterhate speech either opposes such comments or defends Asian communities. Neutral tweets are related to COVID-19 but do not contain such attitudes or information.

8. The target dataset (dehydrated tweets) includes 6,548,324 tweet identifiers. Twenty-two percent of these tweets were unable to be extracted using the Twitter API on September 7, 2021, possibly due to the deletion of the original tweets.

tweets created in the US ( $n = 2,337,924$ ) and classified either as hate or counterhate speech ( $n = 59,337$ ).

## 5 Empirical Strategies

To test hypotheses, we track changes in how the “Chinese virus” was discussed between groups (hate speech speakers and counterhate speech speakers) and over time (before and after Trump’s inflammatory speech). We start with descriptive statistics that explore these relationships. We then examine how the context of specific words changed. Word embedding is a suitable tool to perform this task because it is designed to estimate a linguistic context around a focal word by looking at its neighboring words (Firth 1957). In this study, we used a recent variant of the word embedding method called embedding regression that helps social science researchers make descriptive inferences (Rodriguez, Spirling, and Stewart 2021b). This method addresses a key problem with standard word embedding approaches: the focal words of interest to social scientists could be rare in the corpus they study. If so, the quality of the resulting representation will be far from ideal. Our case also falls into this category, as only four percent of the tweets in our data mention “Chinese virus.” This method resolves this problem by taking embeddings pretrained on large corpora (e.g., word2vec or Glove), then combining these with a small sample of example texts for a focal word and creating a new customized embedding for the focal word. Another strength of this framework is that it allows researchers to apply embedding to a multivariate regression framework. This feature is useful because many social science questions can be easily formulated in the regression framework. For instance, the above hypotheses could be formalized as estimating the coefficients ( $\beta_1$  and  $\beta_2$ ) in the following regression model.  $Y$  represents contextual differences around a focal word,  $Group$  indicates hate speech or counterhate speech speakers, and  $Time$  points to the period before and after Trump’s inflammatory tweet.

$$Y = \beta_0 + \beta_1 Group + \beta_2 Time + error$$

Using the embedding regression method, we estimated these betas by measuring the distance (the Euclidean norms) from the origin (the vector of all dimensions) to the coefficients. We also estimated the standard errors of these coefficients via a non-parametric bootstrap. Using “text2vec” (Selivanov, Bickel, and Wang 2020) and “context” R packages (Rodriguez, Spirling, and Stewart 2021a), we trained our model based on Glove embeddings using 50 dimension,<sup>9</sup> six window size, and ten minimum frequency of words. The tweets were preprocessed so that they did not include any non-alphabetical characters, words with only one or two letters, excessive white space, stopwords (based on the Snowball dictionary),<sup>10</sup> and typical swear words (e.g., f\*\*\*). We also removed HTML tags and URLs, and replaced contractions with short forms using the “textclean” R package (Rinker 2018).

## 6 Results

### 6.1 Growth in Asian Hate Speech and Counterhate Speech

The COVID-HATE dataset shows that Trump’s anti-Asian tweet increased hate and counterhate speech among the US Twitter users. To begin with, we plot the Twitter trends in

---

9. The higher dimensions, such as 100, tend to overfit the data, so we settled on using 50 dimensions.

10. For more information, see Porter (2001).

the days and weeks surrounding Trump’s tweet. Figure 1 shows that there was a sharp increase in the number of tweets mentioning phrases like “Chinese virus” immediately after Trump’s tweet. Interestingly, this change was driven by an increase not just in hate speech but also in counterhate speech (see Figure 1). We visualize this change in both standard and logged scales, with the former showing the dramatic uptick in tweets. Here we show that Trump’s actions increased hateful rhetoric toward Asian Americans and that other Twitter users responded. In short, Trump may have increased the instances of hate speech, but largely by centering himself in the existing discourse and amplifying the existing tension between hate and counterhate speakers.

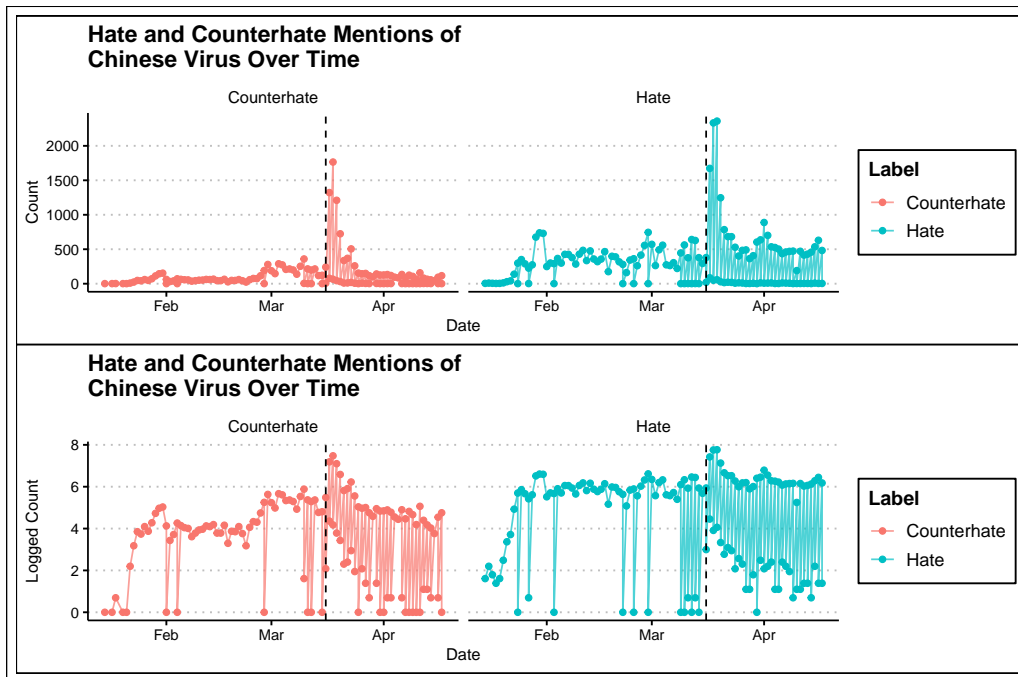


Figure 1: Raw (top) and logged (bottom) frequency of hate and counterhate tweets over time for “Chinese virus.”

Figure 2 shows a larger context that, in terms of the absolute volume, between the pre- and post-Trump tweet periods the number of hate speech increased from 5,810 to 9,663, and that of counterhate speech increased from 20,761 to 23,103. Based on this metric, Trump’s anti-Asian speech backfired more than it garnered support. To make this point clearer, we calculate the relative increase and decrease of hate and counterhate speech tweets. The proportion of hate speech tweets decreased from 78% to 71%, whereas that of counterhate speech tweets increased 22% to 29% between the two periods. These differences are highly statistically detectable based on the two-tailed t-tests ( $p$ -value < 0.01).

## 6.2 Evolving Context

Not only did the frequency of tweets discussing terms like “Chinese virus” increase in the days following Trump’s tweet, but so did the actual context surrounding them. Here, we analyze the variations in the contexts in which the term “Chinese virus” was discussed between groups and over time. We used word embedding to achieve this goal. The key insight with our word embedding regression is drawn from the adage that “you shall know a word by the company it keeps” (Firth 1957, 11). Word embedding techniques

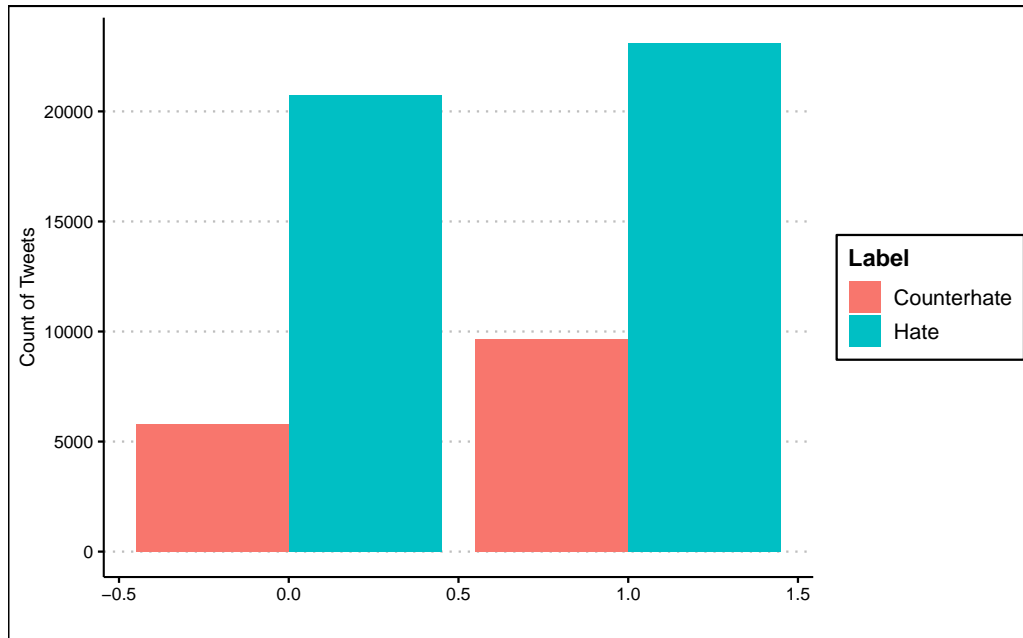


Figure 2: Total tweet count change before and after Trump’s Chinese virus tweet.

leverage this idea to produce vectors that represent particular words. If two vectors are close to one another, we can say that the two words are similar in that they share a similar context.

Figure 3 demonstrates the idea of changing context following Trump’s tweet. We show an embedding regression that regresses the term “Chinese virus” on the group (hate and counterhate speech) and time (pre- and post-Trump tweet period) variables. The y-value shows the distances of the coefficients of the group and time variables. The error bars indicate the bootstrapped 95% confidence intervals. The “Hate label” coefficient is the average difference of the contexts between hate speech and counterhate speech speakers, controlling for the time variation. The “Trump tweet” coefficient is the average difference of the contexts between the two time periods, controlling for the speaker group variation. Because these point estimates involve slight sampling variance, we reran the same embedding regression model 1,000 times. We then calculated the mean and standard deviation of the ratios between the coefficients of the Trump tweet and those of the hate speech label. The result shows that, on average, the coefficient of the Trump tweet is 444% greater than that of the hate speech label. The standard deviation of these data points is 0.36.

### 6.3 Contextual Hate and Counterhate Speech

While the frequency of hate and counterhate speech both increased following Trump’s tweet (see Figure 2), the content changed in different ways. Specifically, we find that following Trump’s tweet, among hate speech speakers, vectors for words that could be associated with suspicion and animosity toward the Chinese government became closely associated with the vector for “Chinese virus.” In contrast, counterhate speakers’ tweets did not exhibit this same change in vector representations. We investigate our hypothesis that Trump’s tweet was associated with a shift in context from conspiracy theories about the origins of the virus toward placing blame on the Chinese government.

Figure 4 provides more concrete information about these contextual changes. This figure

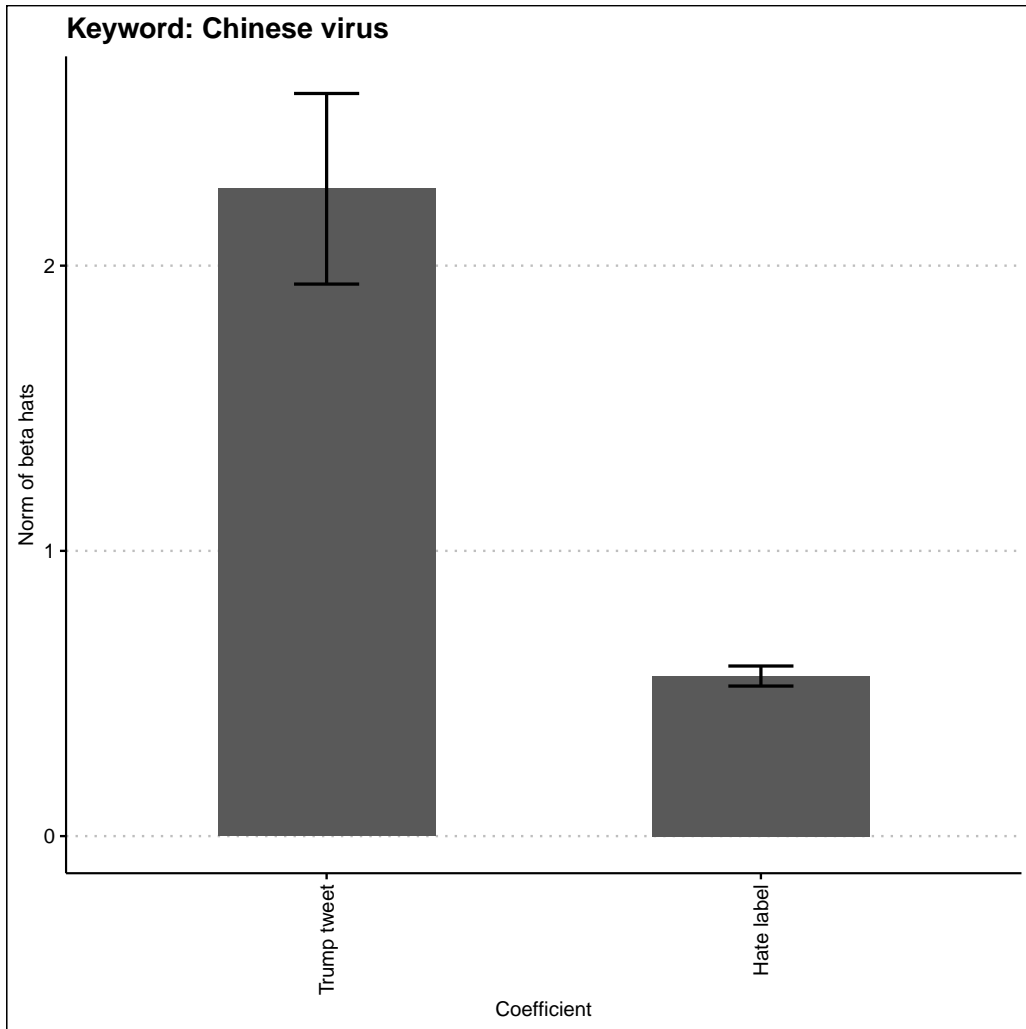


Figure 3: Cross-sectional and longitudinal cotextual differences.

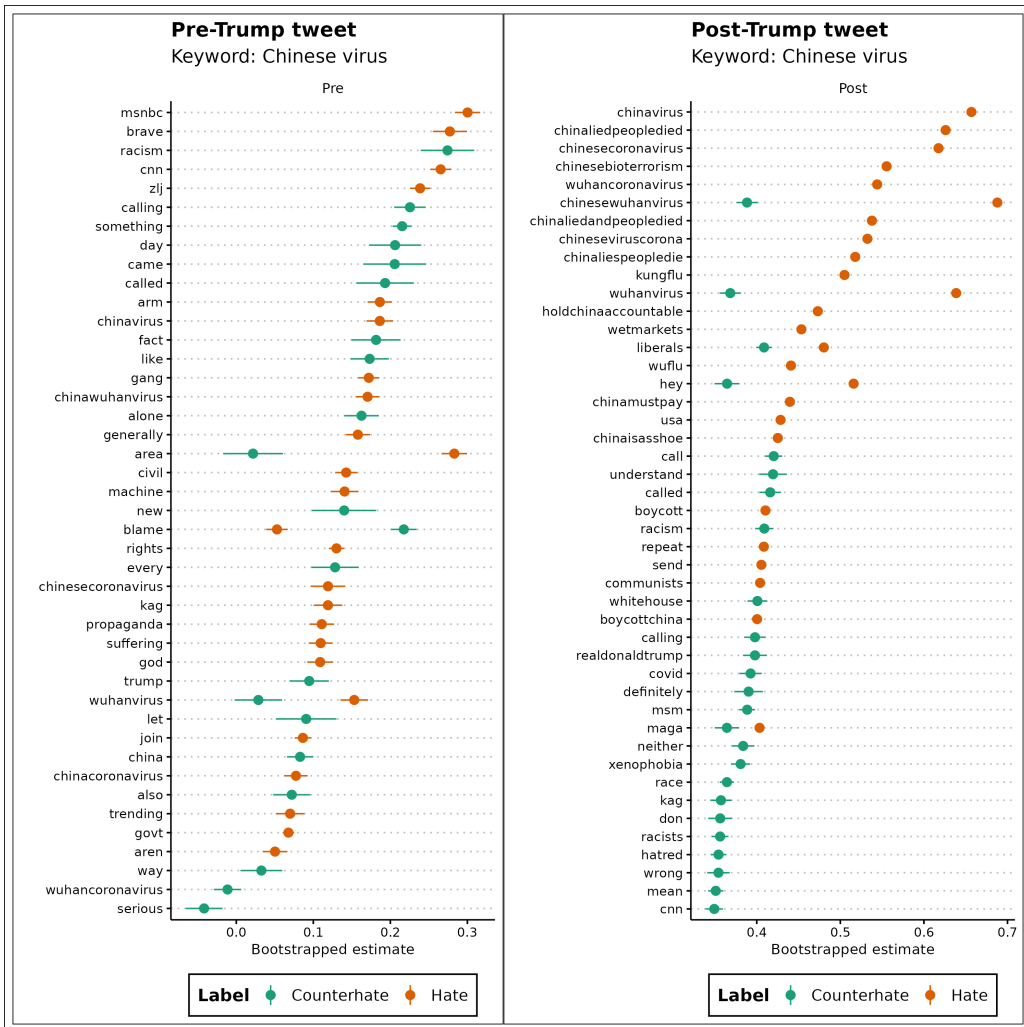


Figure 4: Top neighboring words.



contrasts the relatively close neighboring words (i.e., nearest neighbors) to “Chinese virus” between hate and counterhate speech speakers and tracks how this pattern changed over time. We estimate the distance between these neighboring words and the focal word using bootstrapping. Then, we rank the neighboring words based on their average cosine similarities. The error bars indicate their bootstrapped 95% confidence intervals.

Figure 4 illustrates how the two sides engaged with each other. In the pre-Trump tweet period (left panel), hate speech speakers attacked the liberal media (e.g., MSNBC and CNN) when they mentioned the Chinese virus, whereas counterhate speakers criticized these hate speech speakers’ blatant racism. Here, it is important to note that the “machine” that appears in the left panel indicates the Chinese propaganda machine. Hate speech speakers were likely aware that “Chinese virus” is a racially charged term. Nevertheless, they insisted on using the term because dissociating China from the virus is the goal of Chinese propaganda. For instance, in response to MSNBC’s political commentator Chris Hayes, Twitter user @Cryptominerz posted the following tweet on March 8, 2020:

Every news agency in the world was calling it the Wuhan Virus until the Chinese government propaganda machine kicked in to disassociate itself from being the origin of the virus. And the liberal media ate it right up.

The same figure also illustrates how Trump’s anti-Asian tweet divided his supporters and critics further. In the right panel (the post-Trump tweet period), we can see that hate speech speakers amplified Trump’s tweet by trending hashtags similar to the Chinese virus, such as #chinavirus, #wuhancoronavirus, #wuhanvirus, #chinesewuhanvirus, #kungflu, #chineseviruscorona, and #ccpvirus (CCP stands for the Chinese Communist Party). If the implications of these terms were rather implicit, hashtags like #chinaliedpeopledied sent a more direct message to the audience. For hate speech speakers, the hashtag #chinaliedpeopledied became their rallying cry. Counterhate speech speakers accused these hate speech speakers of racism and xenophobia. Twitter user @LeslieCackles criticized Trump’s speech on March 24, 2020:

Calling it the Chinese virus is causing people to not eat Chinese food from restaurants and causing even more racism than is already happening in this country. Even if it did start in China it doesn’t mean they are bad people. They are wonderful and dying too!

When accusations of racism happened, hate speech speakers accused the counterhate speech speakers as China sympathizers, because, in the minds of hate speech speakers, China must pay for the virus, which they started and covered up. Once again, here one can see that these individuals argued that disassociating China from the virus is a goal of the Chinese government’s propaganda. Twitter user @SundevilSal pushed back against the counterhate speech by commenting that “anyone says the #WuhanVirus did not come from China and tells you not to call it that is a Chinese communist sympathizer.”

## 7 Conclusion

This article examines how misinformation and hate speech are linked using the case of anti-Asian hate speech during the COVID-19 pandemic. The linkage between misinformation and hate speech is crucial for effective online content moderation. If users defend their hate speech based on misinformation, we cannot effectively reduce hate speech without addressing misinformation.

Designing interventions that can tackle these two problems simultaneously requires a

deeper and more precise understanding of the relationship between information and hate speech at the behavioral level. As our study is based on observational data and provides only descriptive inferential evidence, experimental research could be useful to disentangle the relationship between these two variables. For instance, future studies can investigate this by experimentally testing whether the effect of counterhate speech on moderating hate speech is reduced when hate speech speakers are exposed to misinformation. This experimental test is a useful next step, as our study is based on observational data and provides only descriptive inferential evidence.

## References

- Bail, Chris. 2021. *Breaking the Social Media Prism*. Princeton University Press.
- Chan, Nathan, Jae Yeon Kim, and Vivien Leung. 2021. "COVID-19 and Asian Americans: How Elite Messaging and Social Exclusion Shape Partisan Attitudes." *SocArXiv*. September.
- Choi, Cynthia, and Manjusha Kulkarni. 2020. *In One Month, STOP AAPI HATE Receives almost 1500 Incident Reports of Verbal Harassment, Shunning and Physical Assaults*. [asianpacificpolicyandplanningcouncil.org/wp-content/uploads/Press\\_Release\\_4\\_23\\_20.pdf](https://asianpacificpolicyandplanningcouncil.org/wp-content/uploads/Press_Release_4_23_20.pdf).
- Coates, Melanie. 2020. "Covid-19 and the Rise of Racism." *BMJ* 369. <https://doi.org/10.1136/bmj.m1384>. <https://www.bmj.com/content/369/bmj.m1384>.
- Elliot, Andrew J, and Patricia G Devine. 1994. "On the Motivational Nature of Cognitive Dissonance: Dissonance as Psychological Discomfort." *Journal of Personality and Social Psychology* 67 (3): 382.
- Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford University Press.
- Firth, John R. 1957. "A Synopsis of Linguistic Theory, 1930-1955." *Studies in Linguistic Analysis*, 1–32.
- Gallotti, Riccardo, Francesco Valle, Nicola Castaldo, Pierluigi Sacco, and Manlio De Domenico. 2020. "Assessing the risks of 'infodemics' in response to COVID-19 epidemics." *Nature Human Behaviour* 4 (12): 1285–1293. <https://doi.org/10.1038/s41562-020-00994-6>.
- Gover, Angela R, Shannon B Harper, and Lynn Langton. 2020. "Anti-Asian Hate Crime During the COVID-19 Pandemic: Exploring the Reproduction of Inequality." *American Journal of Criminal Justice* 45 (4): 647–667.
- Harris, Keith. 2018. "What's Epistemically Wrong with Conspiracy Theorising?" *Royal Institute of Philosophy Supplements* 84:235–257.
- Howard, Jeffrey W. 2019. "Free Speech and Hate Speech." *Annual Review of Political Science* 22:93–109.
- Hswen, Yulin, Xiang Xu, Anna Hing, Jared B Hawkins, John S Brownstein, and Gilbert C Gee. 2021. "Association of "# Covid19" Versus "# Chinesevirus" with Anti-Asian Sentiments on Twitter: March 9–23, 2020." *American Journal of Public Health* 111 (5): 956–964.
- Jerit, Jennifer, and Yangzi Zhao. 2020. "Political Misinformation." *Annual Review of Political Science* 23:77–94.
- Jeung, R, and K Nham. 2020. "Incidents of Coronavirus-related Discrimination." *Asian Pacific Policy and Planning Council*.
- Jeung, Russell, Aggie Yellow Horse, Tara Popovic, and Richard Lim. 2021. "Stop AAPI Hate National Report." *Stop AAPI Hate*.
- Junn, Jane. 2007. "From Coolie to Model Minority: U.S. Immigration Policy and the Construction of Racial Identity." *Du Bois Review: Social Science Research on Race* 4 (2): 355–373.
- Kim, Claire Jean. 1999. "The Racial Triangulation of Asian Americans." *Politics & Society* 27, no. 1 (March): 105–138.

- Kim, Jin Woo, Andrew Guess, Brendan Nyhan, and Jason Reifler. 2021. "The Distorting Prism of Social Media: How Self-Selection and Exposure to Incivility Fuel Online Comment Toxicity." *Journal of Communication* (September).
- Kinder, Donald R, and Cindy D Kam. 2010. *Us against Them*. University of Chicago Press.
- Kouzy, Ramez, Joseph Abi Jaoude, Afif Kraitem, Molly B El Alam, Basil Karam, Elio Adib, Jabra Zarka, et al. 2020. "Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter." *Cureus*.
- Le Espiritu, Yen. 1992. *Asian American Panethnicity: Bridging Institutions and Identities*. Temple University Press.
- Lu, Runjing, and Yanying Sheng. 2020. "From Fear to Hate: How the Covid-19 Pandemic Sparks Racial Animus in the United States." *Available at SSRN 3646880*.
- Ngai, Mae M. 2014. *Impossible Subjects*. Princeton University Press.
- Nguyen, Jenny, J. Scott Carter, and Shannon K. Carter. 2019. "From Yellow Peril to Model Minority: Perceived Threat by Asian Americans in Employment." *Social Science Quarterly* 100 (3): 565–577.
- Porter, Martin F. 2001. *Snowball: A Language for Stemming Algorithms*. <http://snowball.tartarus.org/>.
- Relia, Kunal, Zhengyi Li, Stephanie H Cook, and Rumi Chunara. 2019. "Race, Ethnicity and National Origin-based Discrimination in Social Media and Hate Crimes Across 100 US Cities." In *Proceedings of the International AAAI Conference on Web and Social Media*, 13:417–427.
- Rinker, Tyler W. 2018. *textclean: Text Cleaning Tools*. Version 0.9.3. Buffalo, New York. <https://github.com/trinker/textclean>.
- Roberto, Katherine J, Andrew F Johnson, and Beth M Rauhaus. 2020. "Stigmatization and Prejudice During the COVID-19 Pandemic." *Administrative Theory & Praxis* 42 (3): 364–378.
- Rodriguez, Pedro L, Arthur Spirling, and Brandon M Stewart. 2021a. *conText: a la Carte on Text (ConText) Embedding Regression*. R package version 0.1.0. <https://github.com/prodriguezsosa/EmbeddingRegression>.
- . 2021b. "Embedding Regression: Models for Context-Specific Description and Inference." *Working paper*, <https://arthurspirling.org/documents/embedregression.pdf>.
- Schickler, Eric. 2016. *Racial Realignment: The Transformation of American Liberalism, 1932–1965*. Princeton University Press.
- Selivanov, Dmitriy, Manuel Bickel, and Qing Wang. 2020. *text2vec: Modern Text Mining Framework for R*. R package version 0.6. <https://CRAN.R-project.org/package=text2vec>.
- Skrentny, John David. 2009. *The Minority Rights Revolution*. Boston, MA: Harvard University Press.
- Swire-Thompson, Briony, and David Lazer. 2019. "Public Health and Online Misinformation: Challenges and Recommendations." *Annual Review of Public Health* 41:433–451.
- Tajfel, Henri. 1970. "Experiments in Intergroup Discrimination." *Scientific American* 223 (5): 96–103.

Tajfel, Henri, and John C Turner. 1979. "An Integrative Theory of Intergroup Conflict." In *Organizational Identity: A Reader*, edited by William G Austin and Stephen Worchel, 33:56–65. 47. Oxford University Press.

Ziems, Caleb, Bing He, Sandeep Soni, and Srijan Kumar. 2020. "Racism Is A Virus: Anti-Asian Hate and Counterhate in Social Media During the Covid-19 Crisis." *arXiv preprint arXiv:2005.12423*.

## **Authors**

**Jae Yeon Kim** is an Assistant Research Scholar at the SNF Agora Institute, Johns Hopkins University. Kim is the corresponding author and can be reached at [jkim638@jhu.edu](mailto:jkim638@jhu.edu).

**Aniket Kesari** is a Research Fellow at the Information Law Institute, New York University.

The two authors contributed to this article equally.

## **Data Availability Statement**

Replication files are available at: [https://github.com/jaeyk/asian\\_hate\\_misinformation](https://github.com/jaeyk/asian_hate_misinformation)

## **Funding Statement**

Not applicable.

## **Keywords**

Hate speech; Misinformation; Asian Americans; COVID-19