

Supplemental Online Content

Ayers JW, Chu B, Zhu Z, et al. Spread of misinformation about face masks and COVID-19 by automated software on Facebook. *JAMA Intern Med*. Published online June 7, 2021.
doi:10.1001/jamainternmed.2021.2498

eAppendix. Automated software targeting original research spreads COVID-19 misinformation on Facebook

This supplemental material has been provided by the authors to give readers additional information about their work.

eAPPENDIX

Automated software targeting original research spreads COVID-19 misinformation on Facebook

Data acquisition

CrowdTangle is a Facebook-owned content discovery and social analytics tool that measures how content spreads online. Initially used to help newsrooms and media publishers stay up-to-date, it has more recently become available for academic use, particularly on misinformation, elections, racial justice, and well-being (<https://help.crowdtangle.com/en/articles/4302208-crowdtangle-for-academics-and-researchers>). CrowdTangle users can search across public Facebook groups for keywords, phrases or URLs, identify posts that match, and download entire posts as well as metadata (including timestamps and the Facebook group where the post was made).

We investigated data from CrowdTangle for the 5-day period after the publication of the Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection (DANMASK-19) randomized clinical trial (11/18/2020-11/22/2020). Specifically, we queried CrowdTangle for the exact URL pointing to the DANMASK-19 study on the *Annals of Internal Medicine* website (<https://www.acpjournals.org/doi/10.7326/m20-6817>), obtaining the names of Facebook groups (N = 563) where a link was posted, the time stamp when the post was made, and the content of the post (unless already deleted prior to our download, which was rarely the case as detailed in the research letter).

While CrowdTangle precludes us from sharing the raw data, the code used to assemble the data we analyzed is available upon reasonable request from the corresponding author (Dr. Ayers; ayers.john.w@gmail.com).

Data analysis

Identifying Facebook groups most and least influenced by automation

Our goal was to identify posts to Facebook groups sharing the DANMASK-19 study that were likely made by automated software that allows individuals to generate automated content and share it via counterfeit accounts (or “bots”). Background material on automated posting to Facebook is available in the cited summaries (1,2).

Although software can be used to automate the content of a post, the links will remain the same. As a result, standard techniques for discovering automated software study “coordinated link sharing behavior” (3-5). We modeled our analysis after these studies.

First, we downloaded all posts (N=299,925) made to Facebook groups that shared the DANMASK-19 study over the same period (11/18/2020-11/22/2020). CrowdTangle restricts downloads to 300,000 posts. All posts made during the study period were obtained and no posts were omitted by this restriction.

Second, we identified all URLs (including those about and not about DANMASK-19) in these posts (N=251,656). Of these 32,455 (23%) were shared at least twice. We then calculated the frequency that identical links were posted to pairs of Facebook groups and the corresponding elapsed time between those posts for all links as a marker for potential coordination or automation across groups. An example of this data structure is shown in Table S1.

Table S1. Example of the URL data structure		
URL (or link) ^a	Facebook group pairs ^b	Time elapsed between

		shares (seconds) ^c
https://www.facebook.com/nicolasdupontaignan/videos/702882607326543/	La Voix Du Peuple & Didier Raoult professeur Marseille	161
https://www.americanthinker.com/articles/2020/11/gavin_newsom_manufactures_a_covid_crisis_to_make_california_miserable_again.html	California Republicans Make Our Votes Count & Stuart Varney - Conservative Economic Journalist	238,152
https://www.facebook.com/andredefosses1964/videos/10157914957478031/	Regroupement contre le port du masque obligatoire & Résistance Québec	257
https://www.acpjournals.org/doi/10.7326/M20-6817	Trump51 - Vermont & Trump51 - Massachusetts	5
https://www.facebook.com/TheHipHopPatriot/videos/2574728796101102/	Trumpians & New Yorkers and friends that hate Cuomo	1,567
https://youtu.be/M7ybbIZ14KA	All About Hanover (AAH) & The Ingraham Angle	22,808
^a Is the URL that was shared across Facebook groups. ^b Is the pair of Facebook groups where the URL was shared. ^c Is the number of seconds between shared URL, rounded to whole seconds.		

Third, to define the subsets of Facebook groups most influenced by automation we relied on Giglietto et al.’s operationalization of automation (3): A pair of Facebook groups that (a) hosted identical links ≥ 5 times and (b) if at least half of these links were posted within < 10 seconds would be considered the most influenced by automation. This operationalization is an agreed upon (albeit conservative) definition of the likely presence of automation. Facebook groups most influenced by automation averaged 4.28 seconds (SD=3.02) between shares of identical links.

Fourth, as a comparator Facebook groups falling within the top 90th percentile of total time elapsed between when identical links were posted were considered the least influenced by automation. Facebook groups least influenced by automation averaged 4.35 hours (SD=11.71) between shares of identical links.

We then described the rate of which Facebook groups the most or least influenced by automation shared a link to DANMASK-19.

Evaluating misinformation in DANMASK-19 link shares

We investigated 2 distinct types of misinformation. (1) The primary conclusion of DANMASK-19 was misrepresented. This was defined as a post claiming that mask-wearing harms the wearer, since this is a willful misrepresentation of the trial's conclusions. (2) Conspiratorial claims were made about DANMASK-19. This was initially defined through open-ended coding and ultimately resolved to claims of covert political/corporate control of the DANMASK-19 trial and/or its dissemination. A separate outcome for not including either form of misinformation was computed. Two coders disagreed on 3.9% of labels (Cohen's Kappa = 0.76) and resolved disagreements through unanimous deliberation with a third coder who was also the senior author. Analyses of these data are fully described in the research letter.

References

1. N. Gleicher, "Coordinated Inauthentic Behavior Explained." Facebook, 6 December 2019. Available at <https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/>
2. Weedon, J.; Nuland, W.; and Stamos, A. 2017. Information operations and Facebook. Technical report, Facebook Security. Available at: <https://www.mm.dk/wp-content/uploads/2017/05/facebook-andinformation-operations-v1.pdf>.

3. Giglietto, F., Righetti, N., Rossi, L., & Marino, G. (2020). It takes a village to manipulate the media: coordinated link sharing behavior during 2018 and 2019 Italian elections. *Information, Communication & Society*, 23(6), 867-891.
4. Leonardo Nizzoli, Serena Tardelli, Marco Avvenuti, Stefano Cresci, and Maurizio Tesconi. Coordinated behavior on social media in 2019 UK general election. arXiv:2008.08370, 2020.
5. Giglietto, F.; Righetti, N.; Rossi, L.; and Marino, G. 2020a. Coordinated link sharing behavior as a signal to surface sources of problematic information on Facebook. In *International Conference on Social Media and Society*, 85–91.