Post-Spotlight Posts: The Impact of Sudden Social Media Attention on Account Behavior

20 May 2022, University of Washington (UW) Undergraduate Research Symposium

Joseph S. Schafer
UW Center for an Informed Public, Paul G. Allen School of Computer Science & Engineering
schaferj@uw.edu

Mentors:
Professor Kate Starbird
UW Center for an Informed Public, UW Human-Centered Design & Engineering
Professor Emma Spiro
UW Center for an Informed Public, UW iSchool
Professor Jevin West
UW Center for an Informed Public, UW iSchool

Abstract

Social media influencers often have a disproportionate importance to the spread of online information and mis/disinformation. One method of utilizing this influence is through 'spotlighting' smaller accounts, directing unusually high amounts of attention to accounts not used to having a large influence. Using a mixed-methods approach we examine the impacts that this spotlight has on the subsequent behavior of accounts who were spotlighted for Twitter posts in the early stages of the Covid-19 pandemic. We found that accounts that were spotlighted were more likely to post after being spotlighted, though this effect was not persistent and did not boost original content production. We also found that spotlighted accounts were significantly more likely to alter their self-presentation through updating their "bio" after a spotlight, and in a case study of a repeatedly-spotlighted account that these changes tended to increase specific verifiable associations and topic-relevant descriptors while more general descriptors were dropped.

Introduction

Misinformation and disinformation on social media present a significant challenge to 21st-century society, weakening our responses to crises in public health (Puri et al., 2020) and climate change (Lewandowsky, 2021), while weakening trust in democracy (Center for an Informed Public et al., 2021). Significant amounts of disinformation have previously been shown to originate from a handful of highly influential accounts (Benkler et al., 2020; Center for an Informed Public et al., 2021; Center for Countering Digital Hate, 2021). One common dynamic that influencers also use to shape online conversations is by 'spotlighting' small accounts, bringing them disproportionate amounts of attention. In this study, we examine social media posts on Twitter during the early stages of the coronavirus pandemic to see what the effects of this spotlighting behavior are on spotlighted accounts, which is essential for understanding the implications of such behavior by influencers.

After an overview of the relevant literature, this paper explains its three central research questions and the mixed-methods approach used in this study. Then the paper explains the results of each of the primary research questions and discusses the implications of these results and what they might mean for future research in this area.

Background

Contextual Event Background

This study examines posts made during the early stages of the global coronavirus (Covid-19) pandemic. The coronavirus (SARS-CoV-2) pandemic started in the final months of 2019, with an initial outbreak in the city of Wuhan, China. The pandemic spread worldwide and has killed millions of people. Misinformation and disinformation about the coronavirus have spread widely since the beginning of the pandemic (Fleming, 2020). Often, misinformation came from influential leaders, such as U.S. President Donald Trump (Niburski & Niburski, 2020) and Brazilian President Jair Bolsonaro (Ricard & Medeiros, 2020). The World Health Organization declared an 'infodemic' accompanying their declaration of the SARS-CoV-2 pandemic, recognizing the extensive problem misinformation about this topic presents (World Health Organization et al., 2020). Furthermore, given our examination of accounts that receive sudden increases of attention, studying an information environment where many are not sure of authoritative sources to follow (since many were likely not closely following public health and epidemiology experts prior to the coronavirus pandemic) is particularly interesting for this research. Note, this aspect of the information environment being uncertain is somewhat similar to those seeking information in crisis events (Vieweg et al., 2010), although a potentially notable difference is the ongoing nature of this crisis contrasts with acute crisis events like an earthquake or a terrorist attack (Starbird et al., 2014)..

Platform Background

Twitter is a major social media/microblogging platform built around short text posts, which optionally include hyperlinks, images, or videos. Accounts can interact with each other by "favoriting"/"liking" other accounts' content, "retweeting," i.e. reposting other accounts' posts without commentary, "quote-tweeting," i.e. reposting other accounts' posts with additional commentary, or "replying," i.e. commenting on another account's post. Within posts, accounts can also interact with informal communities and other accounts via hashtags (prefaced with #) and @-mentions, respectively (Burgess & Baym, 2020). Twitter is often used for quickly attaining information about current events, and has also historically been used for information-seeking in a variety of highly sensitive cases including natural disasters (Vieweg et al., 2010).

Research Background

Social media influencers have been a fruitful area of study for years. Recent research has shown that social media and influencers are critical information sources for many (Brady et al., 2017; Igwebuike & Chimuanya, 2021). Further, significant amounts of research both within misinformation contexts and outside of misinformation contexts have shown that influencers significantly impact what information spreads on social media (Igwebuike

& Chimuanya, 2021; Lofft, 2020). These can be quite negative and misinformation-spreading, as has been shown in election (Center for an Informed Public et al., 2021) and vaccine contexts (Center for Countering Digital Hate, 2021). However, other influencers do attempt to combat misinformation and spread true, reliable information online (Ezzat, 2020; Lavorgna et al., 2018). Influencer endorsement of information can also confer legitimacy, as information from known sources appears more credible and truthful (Brady et al., 2017; Igwebuike & Chimuanya, 2021).

This work is focused most similarly to the recent paper by Gurjar et. al. on popularity shocks (Gurjar et al., 2022). Like this work, we are interested in seeing how account behavior is altered after sudden bursts of attention. This work found that on their multimedia platform of study, that attention shocks only changed user behavior in the short term, and that users had difficulty in getting this increased attention to persist. They also found that users tended to engage more with the platform after an attention shock. The work here seeks to corroborate this study as to whether accounts also tend to engage with the platform more after being spotlighted, while also examining whether accounts make longer-lasting changes in their self-presentation after such an event.

Questions

Given the previous research on this topic, we identified four research questions that we were interested in studying, related to account behaviors after being spotlighted.

RQ1 [Engagement]: Do accounts post more after being spotlighted, in response to this sudden increase in attention?

RQ2 [Original Content]: After a spotlighting event, do those spotlighted create more original content?

RQ3 [Self-Presentation: Quantitative]: After a spotlighting event, do accounts alter how they present themselves on social media platforms?

RQ4 [Self-Presentation: Qualitative]: After being spotlighted, what sorts of changes might an account make to how it presents itself?

Methods

In order to study these questions, we used a mixed-methods approach of both quantitative and qualitative methods. For **RQ1, RQ2,** and **RQ3**, we used Bayesian modeling as described in (van de Schoot et al., 2021), in order to see whether being after a spotlight effect has a significant difference to the measured phenomena. More specifically, we rely on a Bayesian model of a negative binomial distribution as this is the most appropriate distribution for level of posting, since there is a long tail of posting behavior (some accounts post extremely high amounts relative to the normal engagement level of accounts), and it is not possible to have negative posts (meaning that a normal distribution would be inappropriate) for **RQ1** and **RQ2**. For **R3**, since we are modeling a probability function (i.e., the likelihood that an account would update their self-description), we use a beta distribution as our baseline distribution for our Bayesian model. For **R4**, a grounded, qualitative approach where the author read over each description used by the account and cataloged each change made was used (Charmaz, 2012).

Datasets

To get data for the study, two APIs provided by Twitter were used. First, since January of 2020 the Center for an Informed Public has collected a database of Twitter posts relating to the coronavirus pandemic, through using Twitter's Streaming API. Using this data, we looked at posts containing the case-insensitive substring "covid19" between April 15 and June 1, 2020, and constructed 3 separate sample groups. First, we created a random control group sample, *random*, which consisted of 500 random tweets containing "covid19" in that time period. We then created a group of highly-retweeted tweets from accounts which typically get large amounts of attention, *popular*. This consisted of the 500 most-retweeted tweets where the account's lowest recorded follower account was greater than half of the amount of retweets the post received. Finally, our population of most interest, *spotlight*, consisted of the 500 most-retweeted tweets where the number of retweets was over double the account's minimum number of followers in the time interval. Also, to make sure that our samples were roughly occurring in this time interval (rather than older, ultra-popular tweets which continued to be reposted long after their origin), we required at least 25% of the retweets of the post to have occurred in the interval. We use two controls here for our spotlight position, to make sure anomalies of this group in our findings are unique to spotlighted accounts both compared to the general account population on Twitter and compared to other accounts also getting attention.

While the data collected through this API was sufficient for answering **RQ3** and **RQ4**, we needed to collect more data in order to answer **RQ1** and **RQ2** since we did not have full platform engagement data for the accounts in each sample. From March-April of 2022, we used Twitter's Academic Historical API to collect the complete timelines (public posts) of each account in these sample populations: 325 *popular* accounts, 490 *spotlight* accounts, and 499 *random* accounts (these numbers are less than 500 for a variety of potential reasons including: duplicate tweets from the same account, accounts that were deleted, accounts that were suspended, or accounts that were since made private). Accounts were then filtered to remove accounts who were tweeting at highly abnormal rates (more than 200 tweets per day).

Limitations and Ethical Considerations

This study has several limitations which are important to acknowledge. First, we are only looking at accounts spotlighted in one context, that of the early stages of the coronavirus pandemic. Spotlighting occurs in a variety of contexts, but since we are only looking at spotlights in one event it is possible that the findings of spotlighted behavior here are not generalizable to spotlighted accounts in other contexts. Second, our study only focuses on Twitter, a platform which does not represent the broad spectrum of social media platforms (Tufekci, 2014). Third, while we are most interested in spotlighting caused by influencers, detecting this particular effect is at present challenging. Instead, we are using a proxy of tweets which receive disproportionate amounts of attention, which is often a result of influencer highlighting, but these are not exactly equivalent measurements.

To reduce the potential for harm as a result of this research, all accounts described in this paper are anonymized and almost all results are only reported in aggregate. The one exception, that of historical biographical data in answering **RQ4**, is a verified account which at the time of writing has significantly over 700,000 followers and is someone who is easily considered a public figure. Accounts which helped in spotlighting this account are similarly large, influential accounts who have been previously publicly written about in research contexts so including them here is minimally harmful.

Results

Engagement

Using the Bambi Python library (Capretto et al., 2022), we used a Bayesian, Negative Binomial GLM to estimate the effect of spotlight on engagement while controlling for activity prior to the spotlight and variation in individual account tendencies to tweet, across our three groups. We fit this Bayesian model to our engagement data for a time window of both 2 and 14 days before and after the spotlighting event (or corresponding centering tweet for our two control groups), from which we calculate the measured tweeting behavior change of each account group, in order to answer **RQ1**. Note, this and subsequent models of engagement data, we only included accounts which had been active at least that number of days before and after, in order to avoid biasing the data to overweight the after data due to some accounts being spotlighted or sampled very shortly after being created.

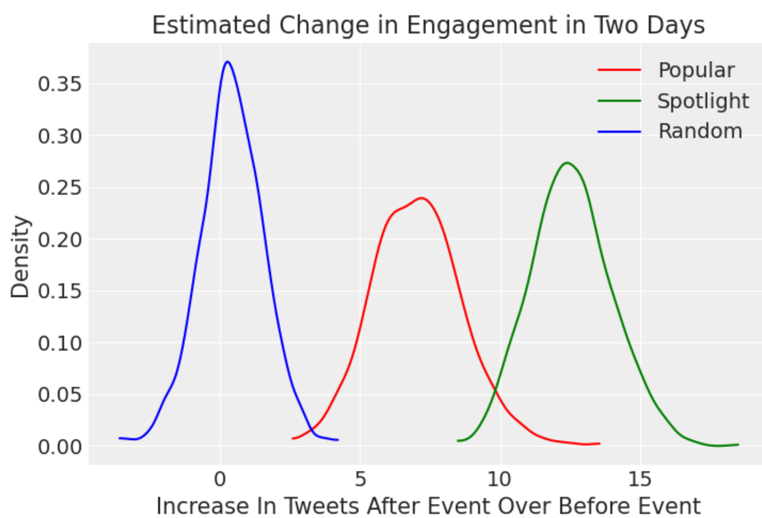

Figure 1: A density plot for the estimated increase in tweets in the two days after a spotlighting event for accounts in each sample population compared to the two days prior to the spotlighting event. This and subsequent density plots are generated with ArviZ (Kumar et al., 2019).

Sample Population	Mean	Standard Deviation	HDI-3%	HDI-97%
Popular	7.113	1.526	4.321	10.027
Spotlight	12.56	1.513	9.903	15.533
Random	0.368	1.11	-1.674	2.42

Table 1: A table of measurements for modeled values for the increase in tweeting behavior two days after the spotlighting event (or equivalent) for each sample. Negative values mean a decrease in engagement.

While we see that the 94% credible intervals of the *popular* and *spotlight* intervals slightly overlap, there is a significant increase of both groups over the random population in platform engagement in the two days after the spotlighting or corresponding event, and the mean estimate for the spotlighting sample is also significantly more of an effect than that of the popular sample. This would indicate that spotlighting does tend to increase in the short term platform engagement.

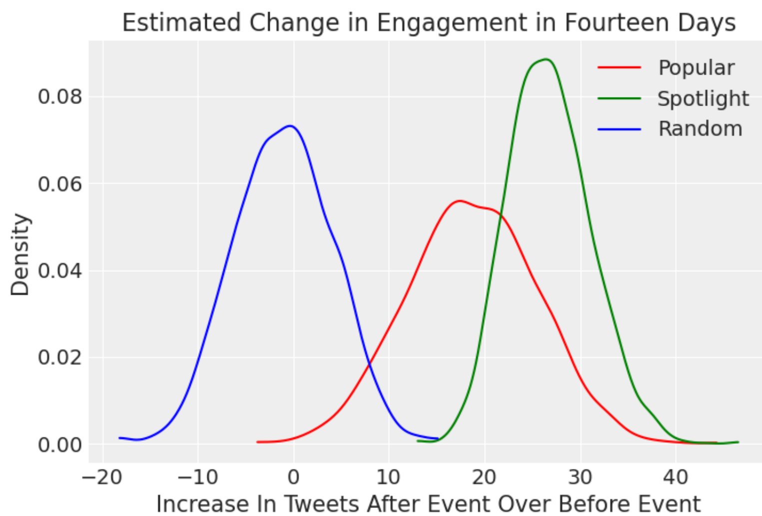

Figure 2: A density plot for the estimated increase in tweets in the fourteen days after a spotlighting event for accounts in each sample population compared to the two days prior to the spotlighting event.

Sample Population	Mean	Standard Deviation	CI-3%	CI-97%
Popular	18.81	6.767	6.858	31.835
Spotlight	27.101	4.321	19.945	36.113
Random	-1.006	5.23	-11.425	8.301

Table 2: A table of measurements for modeled values for the increase in tweeting behavior fourteen days after the spotlighting event (or equivalent) for each sample.

If we expand our interval to 14 days before and after the spotlighting event, we see this effect still exists but is reduced. Since the time interval is multiplied by a factor of 7, we would expect that the engagement difference would also be multiplied by 7 from the two-day model if this level of increased engagement persisted. Instead we see an increase of only slightly over a factor of two, likely indicating an approximate return in engagement level to that of their previous baseline. This corroborates the findings on a different platform of user behavior after popularity shocks, which showed these bursts of attention tended not to persist (Gurjar et al., 2022).

We also modeled using the same formula whether original posting on Twitter (i.e., non-replies and non-retweets) increased in the two days after a spotlighting event compared to the two days before, to answer **RQ2**.

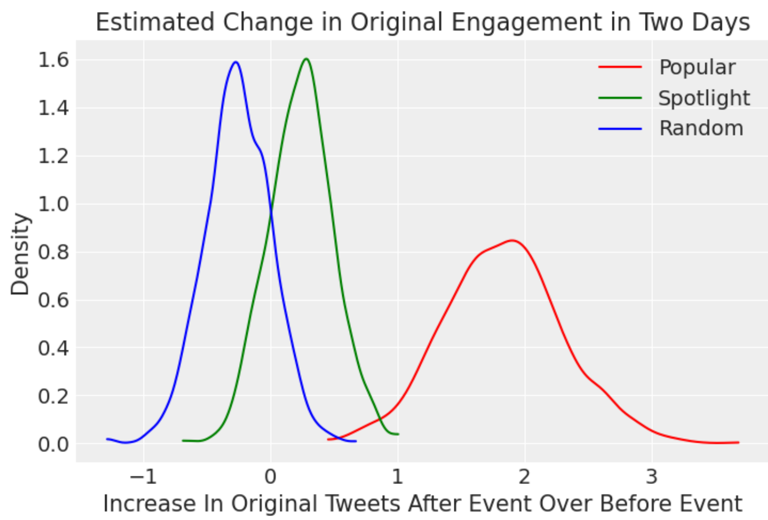

Figure 3: A density plot for the estimated increase in original tweets in the two days after a spotlighting event for accounts in each sample population compared to the two days prior to the spotlighting event.

Sample Population	Mean	Standard Deviation	CI-3%	CI-97%
Popular	1.834	0.471	0.941	2.744
Spotlight	0.238	0.254	-0.239	0.72
Random	-0.256	0.26	-0.729	0.23

Table 3: A table of measurements for modeled values for the increase in original tweeting two days after the spotlighting event (or equivalent) for each sample.

We do not see a significant bump for spotlighted accounts in terms of original engagement after being spotlighted. This would indicate that the majority of the earlier-measured bump is coming from non-original engagement, such as replies to other accounts (perhaps who already replied to them), or retweeting other accounts.

While omitted for length concerns, we observe a similar lack of change to original engagement behavior for spotlighted accounts on the 14-day interval.

Self-Description Changes

Again relying on Bayesian inference, we estimated the probability of a user updating their self-description before and after spotlighting for each of our three sample populations. Our model treated self-description (or "bio") changes in each group as binomially distributed, with a generic, weakly informative prior (Beta(2, 2)). We use a beta distribution as we are modeling the percentage of accounts in each group who change their bio in the time period, and train a model with this distribution using PyMC3 (Salvatier et al., 2016). Note, only accounts with at least one description captured in the two-week period before and after the event were included.

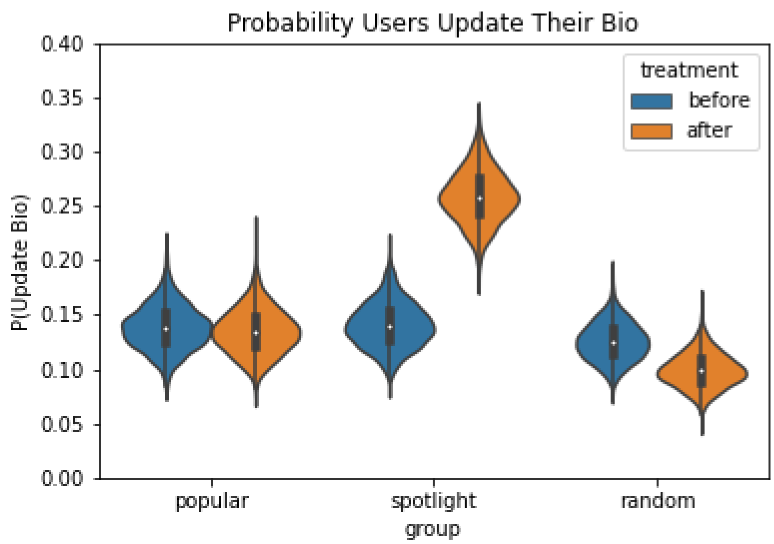

Figure 4: A violin plot of the estimated percentage of accounts in each group who on average would change their account self-description, or "bio," in the two week period before or after the spotlight/equivalent event. This violin plot was made with Seaborn (Waskom, 2021).

Sample	Total Number of Accounts	Changed before	Changed after	Mean Change Difference Modeled	Standard Deviation	CI-3%	CI-97%
Popular	277	37	36	-0.004	0.029	-0.062	0.045
Spotlight	300	41	77	0.118	0.031	0.059	0.177
Random	368	45	35	-0.026	0.023	-0.072	0.014

Table 4: Summary Statistics for the difference in accounts who changed their self-description in the two weeks after the spotlight event compared to the two weeks before.

The accounts that are spotlighted are significantly more likely to change their bio after the event than either corresponding control group, and this difference only occurs after the event. We augmented our Bayesian approach by additionally applying frequentist methods, and this approach yielded similar results. Specifically, the spotlight group is not significantly different from either control group prior to the event (popular-spotlight chi-square statistic of 0.0118 before, p-value of .91; spotlight-random chi-square statistic of .3049, p-value of .58), but is quite significantly different after the spotlighting event (popular-spotlight chi-square statistic of 14.68 after, p-value of .00013; spotlight-random chi-square statistic of 30.91, p-value < .00001). Notably, self-description changes, since a profile is persistent, might be a longer-lasting change for accounts that have been spotlighted, potentially giving these events a longer staying power.

Spotlight Event and Self-Description Case Study

To more deeply understand how an account might alter its self-description after being spotlighted, we conducted a case study of an account which was significantly spotlighted in the early stages of the coronavirus pandemic. In early January, this account stood at about 2,000 followers, which after a period of 48 hours grew to over 40,000 (a 20-fold increase) due to a spotlighted, highly-retweeted tweet.

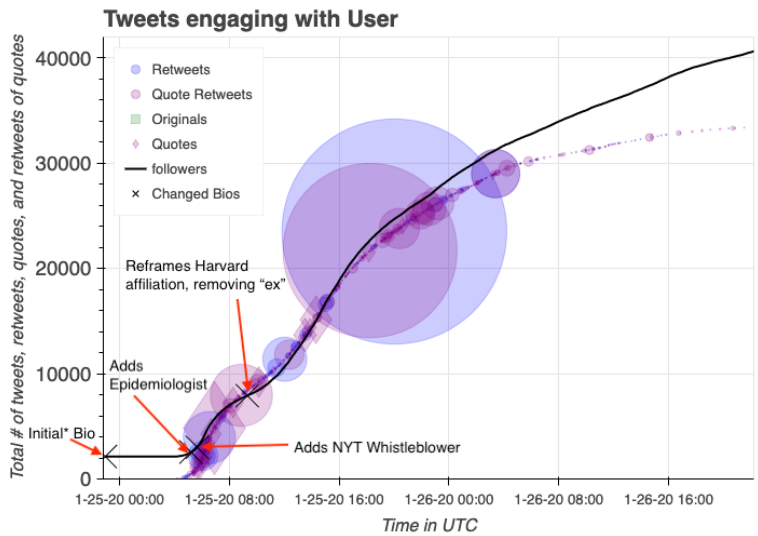

Figure 5: A cumulative plot of tweets engaging with this account's highly-spotlighted tweet about the infectiousness of Covid-19. The y-axis is the total number of tweets engaging with this post (retweets, quote tweets, and retweets of quote tweets), and individual entries are sized based on the follower count of the engaging account. Overlaid is a plot of the account's followers over time, and each time when the account changed its self-description. This visualization was generated with the Bokeh visualization library (Bokeh Development Team, 2018).

While the account has existed since 2009, up to that point the account had only accumulated around 2,000 followers. After posting a tweet which contained a quite high infectiousness estimate for the SARS-CoV-2 coronavirus, framed in a way designed to highlight the danger of such an infectious disease, the account began to get significant attention. Many influential accounts with hundreds of thousands of followers, including a well-known actor, a technology blogger, a national opinion editorial columnist, and a major online alt-right influencer helped to spotlight this account's post. As it began to get traction, the account made its first description alteration that we observed, adding "epidemiologist" to its description. Since the tweet that was getting spotlighted was in reference to public health and epidemiology (by covering the infectiousness of a disease which had the potential for becoming a pandemic), this self-description change helped make the account more relevant to its currently spotlighted content. Similarly, at this time the account added "still Harvard SPH [School of Public Health] scientist" - while it had previously listed Harvard as an ex-affiliation, the account added a signal that the account was still associated with a well-known institution, likely further attempting to add legitimacy for the account to comment on such issues. Approximately forty minutes later the account made its second observed change to its self-description, where it added "NYT featured whistleblower" to its description. Again, given that the original post was framed in such a way as to be raising alarm for an underappreciated crisis, this helped to give the account more credibility for its newly-spotlighted content.

Shortly after this we see the first major retweets of claims casting doubt on the original post, though content engaging with this post that is both supportive and critical continues to rapidly spread. Right after this first check, we also see the account make its final description change in the immediate interval of this first spotlighted tweet, where they further reframed their affiliation with Harvard to remove the "ex" descriptor entirely, again making this account potentially seem more credible. Notably, some of the accounts (in particular the alt-right influencer) then engaged with the post again more skeptically, amplifying criticisms.

The account continued to rapidly grow over the course of the pandemic, though at not quite the same breakneck pace as during this 48 hour period. By the end of 2020, the account had amassed well over 300,000 followers, and at the time of writing this was over 720,000. The account also kept making modifications to its self-description, as our lab recorded 150 changes made to the self-description over the course of 2020.

Also of note, the account continued to be amplified by influential accounts with millions of followers, but some of the largest, later accounts do not significantly increase the rate of engagement with the post. This could be due to a number of factors, including that the amplifying accounts tend to tweet in other languages than the account spotlighted, that the communities they reach had already been exposed to this content through other influencers, that the account's followers tend to not reliably engage with this account, among other explanations.

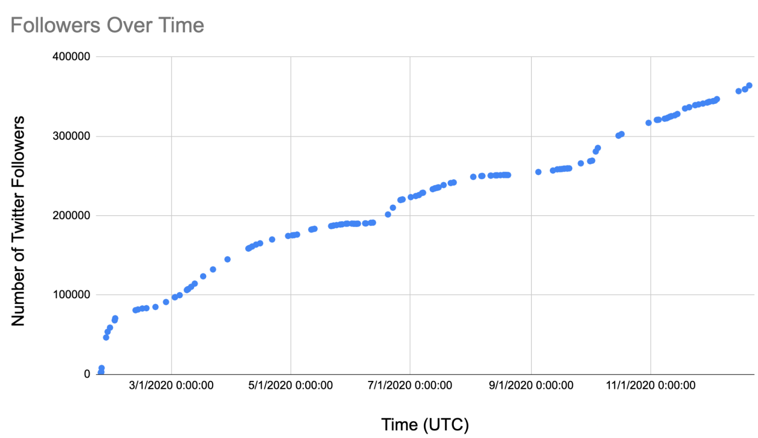

Figure 6: A graph of the account's follower count over the course of 2020. Each dot represents the first time that our research captured a new self-description for the account. Visualization generated in Google Sheets.

In qualitatively coding the account's self-description changes, we saw a variety of emergent categories of self-description the account utilized. These included broad career markers and descriptors (e.g. "Epidemiologist"), concrete institutional affiliations (e.g. "Harvard '07"), notable experiences of the account owner (e.g. "whistleblower"), and other markers of beliefs, identity, or miscellaneous experiences that the account felt important (e.g. "#Biden," or "RoomRater 10/10"). Understanding the reason for each shift is not something we were able to investigate, but as the account grew there was a general trend to reduce the number of broad descriptors and add more concrete, verifiable institutional affiliations.

Discussion

From this study, we see that accounts which see sudden spotlights, or abnormally large bursts of attention, do have some behavioral alterations. They are far likelier to update their self-description in this time, and increase the amount they engage with the platform in the short term. However, on average this increased engagement is a short-term phenomenon, and accounts return to their baseline fairly quickly. More research is needed to understand what sort of changes are made in self-presentation, beyond the illustrative case study provided. Further, while the effects of being spotlighted, for most accounts, might be relatively transient (as our engagement results illustrate), focusing on those for whom this does significantly alter subsequent behavior and influence (as in the case study above) should be a priority for research.

Understanding what makes some spotlights "work" is also an important question for future research. In the case study, we see the two largest amplifying accounts get involved quite late, with a comparatively small engagement boost after their amplification. Understanding why these amplifications, despite their large potential audiences, create a small effect compared to the earlier amplification by smaller (but still influential) accounts would be interesting to pursue in future research.

Additionally, platform trace data as used here is useful for understanding how spotlighting impacts account behavior, but it is insufficient for fully understanding the impact on users. People whose content has gained unexpectedly large amounts of attention on Twitter, both from content originating on and off the platform, have described the experience as disorienting and overwhelming (Armstrong, 2022; Cecire, 2015). These anecdotal examples were written about as mostly unwelcome, or at least not positive. However, there are other accounts who highlight when they are spotlighted by influential accounts, or, in the case study above, use their spotlighting as a way to kickstart their role as a highly influential account on topics of epidemiology and public health. Understanding what makes some users celebrate this experience of their accounts being spotlighted, while others find it unpleasant, would be valuable for understanding this dynamic further. More bluntly, this study's methods and findings are useful for studying *account* behaviors, as we are able to measure these on the platform directly; it is less useful for understanding the impact of spotlighting on *users*, the actual people who are behind the accounts being spotlighted, as getting to their emotions and intent is only possible through proxy data and impressions of which they leave digital records.

Finally, looking longer-term at the results of being spotlighted, particularly for accounts which do maintain an increased level of influence after this spotlight, would be a valuable area for future research. When an account is amplified by a far larger account, this necessarily exposes them to far larger audiences, which likely is not composed of the exact same makeup of communities that their prior audience was made up of. This is central to Schattschneider's arguments of conflict expansion in political science (Schattschneider, 1960), but on social media in this context it would be interesting to see the degree to which mutual reshaping of spotlighted accounts occurs, i.e. if a spotlighted account becomes more like the original beliefs of the audiences an account acquired while spotlighted, if the audience becomes more like the prior alignment of the spotlighted account, or some sort of mutual interaction.

Conclusion

In this mixed-methods study, we examined the impact that spotlighting, or sudden bursts of abnormally large amounts of attention accounts experience, has on their subsequent behavior on Twitter. We found that accounts are likely to post more on the platform after being spotlighted, but that increase only lasts for a short amount of time and is not present for original content. Additionally, spotlighted accounts are likely to update their self-presentation, and for a prominent, repeatedly-spotlighted account these changes including adding credibility signifiers for the topic of their spotlighted content and, over time, increasing the number of concrete, verifiable associations listed in the account's self-description. This helps researchers to better understand the impacts of spotlighting on modern social media systems, which will be essential for addressing issues like mis- and disinformation.

Acknowledgements and Disclosures

This research was supported by a Mary Gates Research Scholarship. The author has also been supported in studying online misinformation on related topics by the UW Center for an Informed Public, the National Science Foundation, the John S. and James L. Knight Foundation, Craig Newmark Philanthropies, and the Omidyar Network. The author would like to also thank Dr. Joseph B. Bak-Coleman for his support in this project.

References:

Armstrong, A. (2022, April 27). MAMMA MIA: On Accidentally Going Viral. *Audrey Armstrong, Writer*. https://audreyarmstrongwriter.wordpress.com/2022/04/27/mamma-mia-on-accidentally-going-viral/

Benkler, Y., Tilton, C., Etling, B., Roberts, H., Clark, J., Faris, R., Kaiser, J., & Schmitt, C. (2020). Mail-In Voter Fraud: Anatomy of a Disinformation Campaign. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3703701

Bokeh Development Team. (2018). *Bokeh: Python library for interactive visualization*. bokeh.org

Brady, J. T., Kelly, M. E., & Stein, S. L. (2017). The Trump Effect: With No Peer Review, How Do We Know What to Really Believe on Social Media? *Clinics in Colon and Rectal Surgery*, *30*(4), 270–276. https://doi.org/10.1055/s-0037-1604256

Burgess, J., & Baym, N. K. (2020). *Twitter: A Biography*. NYU Press. https://nyupress.org/9781479811069/twitter

Capretto, T., Piho, C., Kumar, R., Westfall, J., Yarkoni, T., & Martin, O. A. (2022). Bambi: A simple interface for fitting Bayesian linear models in Python. *ArXiv:2012.10754 [Stat]*. http://arxiv.org/abs/2012.10754

Cecire, N. (2015). Everybody's Authority. *PMLA*, *130*(2), 453–460. https://doi.org/10.1632/pmla.2015.130.2.453

Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory. (2021). *The Long Fuse: Misinformation and the 2020 Election* (v1.3.0). Stanford Digital Repository: Election Integrity Partnership. https://purl.stanford.edu/tr171zs0069

Center for Countering Digital Hate. (2021, March 24). *The Disinformation Dozen*. CCDH. https://www.counterhate.com/disinformationdozen

Charmaz, K. (2012). The Power and Potential of Grounded Theory. *Medical Sociology Online*, *6*(3), 2–15.

Ezzat, H. (2020). Social media influencers and the online identity of Egyptian youth. *Catalan Journal of Communication & Cultural Studies*, *12*(1), 119–133. https://doi.org/10.1386/cjcs_00017_1

Fleming, N. (2020). Coronavirus misinformation, and how scientists can help to fight it. *Nature*, *583*(7814), 155–156. https://doi.org/10.1038/d41586-020-01834-3

Gurjar, O., Bansal, T., Hitkul, Lamba, H., & Kumaraguru, P. (2022). Effect of Popularity Shocks on User Behavior. *International Conference on Web and Social Media*. https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/pubs/Shocks_ICWSM_2022.pdf

Igwebuike, E. E., & Chimuanya, L. (2021). Legitimating falsehood in social media: A discourse analysis of political fake news. *Discourse & Communication*, *15*(1), 42–58. https://doi.org/10.1177/1750481320961659

Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *Journal of Open Source Software*, *4*(33), 1143. https://doi.org/10.21105/joss.01143

Lavorgna, L., De Stefano, M., Sparaco, M., Moccia, M., Abbadessa, G., Montella, P., Buonanno, D., Esposito, S., Clerico, M., Cenci, C., Trojsi, F., Lanzillo, R., Rosa, L., Morra, V. B., Ippolito, D., Maniscalco, G., Bisecco, A., Tedeschi, G., & Bonavita, S. (2018). Fake news, influencers and health-related professional participation on the Web: A pilot study on a social-network of people with Multiple Sclerosis. *Multiple Sclerosis and Related Disorders*, *25*, 175–178. https://doi.org/10.1016/j.msard.2018.07.046

Lewandowsky, S. (2021). Climate Change Disinformation and How to Combat It. *Annual Review of Public Health*, *42*(1), 1–21. https://doi.org/10.1146/annurev-publhealth-090419-102409

Lofft, Z. (2020). When social media met nutrition: How influencers spread misinformation, and why we believe them. *Health Science Inquiry*, *11*(1), 56–61. https://doi.org/10.29173/hsi319

Niburski, K., & Niburski, O. (2020). Impact of Trump's Promotion of Unproven COVID-19 Treatments and

Subsequent Internet Trends: Observational Study. *Journal of Medical Internet Research*, *22*(11), e20044. https://doi.org/10.2196/20044

Puri, N., Coomes, E. A., Haghbayan, H., & Gunaratne, K. (2020). Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases. *Human Vaccines & Immunotherapeutics*, *16*(11), 2586–2593. https://doi.org/10.1080/21645515.2020.1780846

Ricard, J., & Medeiros, J. (2020). Using misinformation as a political weapon: COVID-19 and Bolsonaro in Brazil. *Harvard Kennedy School Misinformation Review*, *1*(3). https://doi.org/10.37016/mr-2020-013

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, *2*, e55. https://doi.org/10.7717/peerj-cs.55

Schattschneider, E. E. (1960). *The semisovereign people: A realist's view of democracy in America*. New York : Holt, Rinehart and Winston. http://archive.org/details/semisovereignpeo00scha

Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R. M. (2014, March 1). Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. *IConference 2014 Proceedings*. https://doi.org/10.9776/14308

Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 505–514.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: What twitter may contribute to situational awareness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1079–1088. https://doi.org/10.1145/1753326.1753486

Waskom, M. L. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

World Health Organization, United Nations, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, & IFRC. (2020, September 23). *Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation*. https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation