

Response to Referees Letter

Dear Editor and Reviewers:

We have received your reports on our manuscript "...".

In the following we have colored the Reviewers comments in blue and the text pasted from the revised manuscript in red to facilitate reading.

“ Reviewer #1:

The “Emergence of New Influencer Elites Challenges Traditional Polarization Analysis” is an exciting manuscript that vividly demonstrates how siloed and polarized political information has become in the last four years on Twitter. The rigorous analysis drives this point home again and again in a convincing fashion. My main concern about the paper is its framing. The authors contend that their analysis reveals the emergence of a new group of elite influencers. I disagree for the reasons I offer below. I believe the paper makes an important contribution without this argument, however, so the authors should be invited to revise and resubmit their manuscript. The authors really have two options for revising their article: 1) they can revise it so that they better define what a new influencer is and then reorganize the analysis so it proves their point more conclusively; or 2) they can drop the current frame and instead offer their paper as an exploration of just how polarized political discourse has become on Twitter and who is driving that polarization. If they move in the second direction, the paper also becomes a welcome addition to the literature showing the link between traditional media and social media.

I want to underscore that the analysis in this paper merits publication irrespective of the approach the authors take. I have not seen anything quite like it.
”

We go with option 2: Rewriting title/intro/significance/discussion (Stu, Hernan, Bolek)

“ Specific comments:

The main argument of this article is that the current definition of “elite” used by political science needs to be expanded to include individual “influencers” on social media. I do not find this argument convincing for a number of reasons:

- The authors do not demonstrate that they understand how political scientists currently define elites. They claim “past studies” have focused on three categories, including “elite legislators,” “elite celebrities,” and “elite traditional media” (p.1) but they do not offer any citations. While political scientists have traditionally focused on the first and third categories, they rarely discuss the second. Moreover, there has been a move of late to think of the parties, and therefore elites, as networks of “intense policy demanders” that include parties, activists, donors, think tanks, and partisan media outlets (Bawn et al. 2012). If the authors want to make this theoretical argument, they need to be clearer about how they define traditional elite and the “new influencers,” but they need to do so in a way that demonstrates they have a command of the literature to which they are contributing.

”

Should be solved by the reframing/rewriting of the intro. May need to better include the literature about elites (Stu).

“

- The analysis does not proceed in the way I expected given the argument the authors are trying to make. Based on the introduction, I expected the authors to offer a clear definition of a “new influencer” and a method for identifying them. Then the analysis would proceed by examining trends in polarization among new influencers and traditional elite in 2016 and 2020. Instead, the analysis shows the polarization of influencers (irrespective of their “newness”) between 2016 and 2020. If the authors want to make this argument, they need to make it easier to compare and contrast what is going on among old and new influencers.

”

Should also be solved by the reframing (Stu) and analysis of the differences between new influencers of 2020 and old influencers , i.e. Latent ideology (Alessandro) Fig. 8, and cosine similarity (James) Fig. 7.

“

- (A minor point) If I understand the study, the authors actually identify two types of new influencers, those who are central to the network in terms of the sharing of political information from media sources (“super spreaders”) and those who influence in other ways, perhaps by offering compelling original content. Most of the analysis is focused on identifying the former but you turn to the latter when you examine who the masses are they were ranked but different category in retweeting independent of news categories or URL source. I think this is a distinction worth highlighting

”

This is a misunderstanding. We only use one “super spreaders” in both cases. Should be clarified in the text (Alex)

“ P. 2-The authors are correct to underscore that they are “inferring” the effect of influencers by noting changes in the behavior of their followers. The authors should go further to offer alternative explanations for the behavior of the followers. In other words, Twitter is likely to be only one of many polarizing influences on these individuals. In fact, it is more likely that larger societal factors are causing the polarization of both Twitter influencers and their followers. ”

Good point to add in the discussion (Stu)

“ P. 2-I’d like to know more about how the reclassification of news outlets between 2016 and 2020 is affecting your analysis. I can’t help but think that the huge movement from the center to left and left-leaning in 2020 has to do with the recategorization CNN web news. Does that also explain why some of your news sources seem to come out of nowhere (they are unranked in 2016) in Figure 4? For example, CNN, Politico, and CNN politics all seem to be unranked in 2016 but that is just because they weren’t considered part of a particular news category that year. ”

Show classification of 2020 data with 2016 bias in SI (Alessandro). Need to clarify Fig. 4 (Brendan). “Unranked” was probably not a good word, maybe use “rank < X” instead?

“ P. 3—The authors’ discussion of bots seems strangely tacked on to the first paragraph. It seems like this discussion should come earlier. Moreover, can a bot be a “new influencer”? ”

We can clarify this (Alex), but I don’t think we should expand too much on bots because it’s tricky to define clearly and would require almost a whole another paper to study. Also, from what we see they do not seem to play a big role.

“ P.4—Can you change Fig. 2 so it has the more refined categories of news bias (e.g. right and right leaning instead of lumping them together)? ”

Sure (Zhenkun).

“ P. 4—I am thoroughly confused by the sentence where the authors classify Joe Biden as an individual influencer (end of first column). If Joe Biden is a not a member of the traditional elite, I’m not sure who is. I hope the authors

are not suggesting that he or Donald Trump is a “new influencer” yet this is precisely what they seem to be doing in Figure 3. ”

Need to clarify what we mean by ”individual” (James). We have to see if it makes sense to keep this analysis and maybe change it to influencers linked with media outlets vs individuals (see below).

“ P.4—The authors say, “some of these users are well connected to well-known news outlets, but a significant number are not...” What is a significant number? This is supposed to be the authors’ main argument but it is a very weak statement. ”

This is in relation to Fig. 4 (Brendan). It would help if we showed which influencers of Fig. 4 belongs (works for) a news outlet.

“ P. 5—The fact that so many “new influencers” have no staying power suggests that they are not really elites. Traditionally, elites are groups of people that maintain power over time either due to their wealth or education. Perhaps the authors should not try to redefine “elites” to include these influencers but rather should view the latter as individuals who have seized on a new medium to challenge the grip that traditional elites have on political discourse. Of course, I do not know how you square that with the fact that the political information they share is usually from traditional media and political elite sources. ”

Should be answered by the re-framing, possibly in the discussion (Stu).

“ P. 7—Figure 7 is phenomenal and super interesting. It drives home just how polarized political communities have become on Twitter. Is there any way to avoid using 1-5 in each category? It makes it harder to identify the individual nodes. In the caption you also say that the “direction of the links represents the flow of information” but nothing about the links indicates direction. They merely seem to indicate a link. ”

It seems like the reviewer is talking about Fig. 6 (not 7), i.e. the retweet networks (Brendan). I don’t understand what he is saying about not using 1-5? does he want that we use less or more? We can increase the size of the arrows to make the direction of the edges more visible (Brendan).

“ P. 8—In the discussion, the authors again say that, “some of the new influencers are associated with significant organizations but many are not.” ”

This is a good point. Fig 3 could show the proportion of influencers associated with news outlets instead of individual/organisations (James). This would answer this question and help with the narrative of the paper.

“ P. 8—The authors use causal language when they say, “These changes have increased polarization...” I recommend softening that language. ”

(Stu).

“ Minor points:

- P. 1-Why do the authors lump activists (“politically active and educated people”) and “contributors to parties” in with traditional news media?
- P. 1-“People in category B and C traditionally affect the direction of politics. . .” seems to suggest those in category A do not.
- P. 1-The authors appear to refer to two studies (14 and 15) and use the pronoun “we” but the authors are not authors of those papers.
- P. 1-“The impact of elites...during the election.” This claim needs a citation.
- P. 2-The authors say category C is “mainly traditional media and politicians” when earlier the latter was in category A.
- P. 2-Do you actually, “categorize the real-world presence” of the influencers as you say you do?

”

All of these should be addressed with the rewriting of the introduction (Stu).

“ **Reviewer #2:**

This is a difficult manuscript for me to review. It starts with an interesting angle: Existing conceptualizations of “elites” may not work too well in the modern media environment in which technological advances have significantly changed (and reduced) old media gatekeeping. The authors open by characterizing what they see as existing classifications of elite. I recognize (A), I didn’t know (B) was such a big deal, and (C) is described as “Traditional news media organizations like CNN, NY Times, Fox News, etc., politically active and educated people (opinion journalists, editorial boards and newsrooms of traditional media, contributors to parties, poll watchers).” I understand news media as elites, but why are they grouped with “educated people”? What are “poll watchers” and what makes them elite? This doesn’t sound like a prominent definition of “elite” based on previous work, as the authors suggest.

I think the main component of category (C) for the present article is traditional media organizations, so perhaps that should be the focus. As the research ends up investigating a social media network, it might be useful to note that “opinion leaders” in the classic work by Katz and Lazarsfeld can be seen as “elites” or “influencers.”

While the imprecision of this first paragraph is puzzling, the general question is still a good one: Has new media technology brought about a new kind of “elite”? (As I write this sentence, I’m not sure “elite” is such a helpful term; it is also not the same thing as “influencer.” It would be worthwhile to define exactly what quality is being studied here.)

”

Should be addressed by the reframing (Stu). Need to clearly define “influencers” and “elites”, also in relation by the definition of “opinion leaders” by Katz and Lazarsfeld (Stu, Bolek, Hernan).

“ The article then goes on to describe the content of tweets in 2016 and 2020. I am missing a connection: How do we get from defining elites to Twitter? Is this research about “Twitter elites” or elites in the country (the United States, in this case)? What is the mapping between those two roles? Why Twitter, not Facebook, YouTube, Instagram? Based on the research design, it seems to me that this article can speak only to “elite structure” on one specific platform. That’s a big step down from the opening paragraph. It is conceptually important for the authors to clarify if this paper aims to speak about platform-specific “elites” (and platform-specific influence on message spread). ”

Should be addressed in the abstract/intro/title (Stu).

“ The research presented in this paper is carried out on a large corpus of tweets collected during the 2016 and 2020 elections. (I wasn’t sure if these are all tweets that mentioned the candidates or just a lot of them.) I find it difficult to follow the details of the analysis, in part because the authors often refer to other work rather than explain (briefly) how they progress and why. There are a good number of scholars who have conducted similar research, but I am not one of them, so perhaps the methodological nuances are lost on me (and I can’t speak to their merit). ”

Need to clarify the Twitter analysis, why we do what we do (Alex).

“ To return to the broader theme, the authors conclude that “we see that many important influencers in 2020 were unimportant or not present in 2016.” I can’t help but wonder whether we should even call them “influencers” then, or discuss them in the context of “elites.” The authors themselves seem to be unsure, writing that “new entrants APPARENTLY have significant impact.” A future version of this manuscript should state clearly what constitutes “impact.” The “influence” or “impact” that can be detected in analyses of network structure such as this one is limited to the spread of messages. The influence of traditional elites (media or political) was (claimed to be) much more far-reaching than that. ”

Need to clarify in the intro/discussion that what we call “influencers” are nothing more than the nodes with the highest CI (Stu).

“ **Reviewer #3:**

This article conducts a descriptive analysis of the actors that most contribute to the diffusion of political news on Twitter, and the extent to which the networks in which they are embedded have become increasingly segregated by ideology in recent years. To address this question, the authors use a large-scale dataset of tweets with news URLs shared ahead of the 2016 and 2020 US presidential elections, and rely on network analysis methods to characterize the structure of the news sharing networks.

There is a lot to like about this study. It offers new evidence about an important question - how the rise of social media platforms is transforming the process by which people share political news. The comparison of two different elections is a particularly nice feature of the paper, especially since most past work generally focuses on much narrower periods. The finding that polarizing behavior has increased over time is similar to what Garimella and Weber find in their 2017 ICWSM article (cited here as well), although the authors here provide a more extensive analysis with a larger set of users and content types. Finally, the study is overall well-executed in terms of its research design and the application of the methodology it employs.

”

Thank you.

“ Despite these strengths, my main concern with the paper is that I don’t think the empirical evidence presented here justifies the sweeping conclusions about its impact on ”traditional polarization analysis” that the title and abstract implies.

There’s a few reasons for this. First, the idea that there are new actors that may have an impact on mass polarization beyond the elites that are referenced in the introduction is not novel. The proliferation of media options due to the emergence of internet and social media has been frequently pointed out as a potential driver for polarization in past literature (see for example, Prior 2013 ARPS or Lelkes et al 2015 AJPS). ”

Need to include these references in the introduction (Stu).

“ Second, the results in the paper show that traditional news sources and political actors still play a leading role in the distribution of political news on Twitter (Figure 6). And for the new actors that emerge, from the analysis we can’t tell if their influence goes beyond being news spreaders; for example if they have an ability to set the agenda or persuade/mobilize citizens, just to give an example of mechanisms by which past work has argued that elites can increase mass polarization. ”

It’s true. We should make this clear in the discussion (Stu).

“ Third, the study defines polarization as segregation in the networks of news diffusion on Twitter, which is quite different from standard definitions of political polarization in past research. It’s also limited to describing what is happening on Twitter and does not address whether these effects have offline consequences or, for example, whether users who share political news mostly from one ideological side (Figure 1) would report higher levels of polarization as a result of being in such echo chambers. This is an important question because there are studies that show that even in a fragmented news environment, most people still have moderate and ideologically diverse news diets (Guess 2021 AJPS, Bakshy 2015 et al Science). ”

Need to make these limitations clear in the intro (Stu).

“ In conclusion, while the article is strong in many aspects, I don’t think the article truly addresses ”the fundamental question of whether the rise of social media new influencers led to greater polarization” (p.2) or that the evidence is sufficient to make the strong causal claim that new elite types do ”increase the degree of polarization in American politics” (Significance Statement). ”

Agreed. Will be addressed with the new title/intro/abstract/significance statement and discussion (Stu, Bolek, Hernan).

“ Leaving this issue aside, my view is that the most important empirical claim from the study is that the networks of news diffusion on Twitter appear to have become more segregated in 2020 compared to 2016. As I mentioned earlier, this finding in itself is quite important. It’s consistent with the work of Garimella and Weber (ICWSM

2017), although here it is based on a much larger sample size and a broader set of content. My suggestion would be to frame the article around this finding.

If the authors were to incorporate this suggestion, what I would also advise is to do more to show that this is a robust finding. As it stands right now, there are other alternative explanations that could be driving this longitudinal change. For example, maybe Twitter in 2020 is attracting more polarized users or driving moderate users away, and perhaps the difference is not so much a change in behavior but a change in the types of people who are active on social media. One possible way to address this would be to replicate the analysis with users who were active in both years, as Garimella and Weber do. ”

Agreed. Reframing (Stu) and analysis of 2020 with only users that were also there in 2016: Latent ideology (Alessandro), cosine similarity (James).

“ The affordances of the platform may also play a role. As I understand it, the authors only analyze retweets (excluding quote tweets). This could mean 2016 vs 2020 is hard to compare because it’s possible that retweets have now become a feature that is not only used for non-critical resharing of information, which could lead to sampling bias in the 2020 analysis. Again, one solution here would be to see how robust the findings are to different sampling decisions. ”

Need to analyse quote network: Do Fig. 6 (Brendan), using the same influencers but with quotes instead of retweets. Fig. 7 (James): Compute cosine similarity based on quotes between the same influencers. Fig. 8 (Alessandro): Latent ideology based on quotes (Same influencers).

“ Another alternative explanation is that the results could be driven by changes in the ideological ratings of media outlets. The authors already explore this possibility in the paper when they show that changes in the positions of the outlets according to the latent space analysis are consistent with changes in the ratings. But they could do more to show this is not a concern; for example, by replicating Figure 1 with the old and new media ratings. ”

Replicate Fig. 1 with old and new media classification (Alessandro). This could then go in the SI.

“ One final concern that I believe is important relates to the implementation of the network analysis methods. The news diffusion networks appear to be built without taking into account tie strength (p.3), which means that an account that is retweeted 1,000 times by 10 users would have the same ”influence” (according to the metrics used here) as an account that is retweeted 1 time by 10 users. This seems an odd choice if the goal is to understand how each of these accounts are able to drive online conversations, but it’s possible that this is a limitation of the method used or there is another explanation that justifies it. ”

We should compare the CI ranking with a centrality that is weighted and add this to the SI. We could use CI weighted, although CI is not well defined for weighted network, so maybe something else for directed network (e.g. PageRank or Katz) (James, Hernan). Also we should stress that all the other network analysis take into account the weights (Alex).

“ **Reviewer #4:**

This is a really interesting paper. But it is also seriously unfocused and reads more like a series of (almost) unrelated graphs and figures from a common dataset. The text suggests that the paper wants to speak to debates about polarization in political science, and break new ground on the role of online “influencers”. It explicitly claims it will understand “the emergence of new influencers in the 2020 US presidential election.” And the abstract claims “influencers had changed the dynamics of political groupings and parties.”

First, the discussion of how the political science literature on “masses” and “elites” seems to miss the boat. The authors suggest that political science looks at: a) elite legislators; b) elite celebrities; c) elite traditional media. Other than rare cases, the political science literature really does not characterize “celebrities” as elites. Typical literature focuses on those in government, media, (the two ‘elite’ groups), and the ‘the masses’.

Address in the new introduction (Stu).

“ Second, so a question to focus on here is the claim to test the hypothesis “does the rise of new influencers lead to greater polarization.” Okay, so we need to see demonstrated that: 1) there has been a rise of new influencers; and 2) such influencers led to polarization (if we want to define that as a network structure with two distinct groups with few connections - okay). ”

Need to clarify that we study 1), but that the causality of 2) can not be directly proven here (Stu).

“ But then we see the introduction of the spread of fake news as a focus of analysis. This would have made sense if the authors clearly suggested that was a focus of the analysis, but it seems to exist as a (large) afterthought. Related to this, the suggestion of ‘bias’ is troubling. What is “bias”? Bias seems to be used to mean ideology that is distinct from a mythical “zero” point of unbiasedness. ”

This is a good point. We should introduce the fact that we also want to study the role of fake news in this increased polarisation (Stu).

“ Figure 4 seems to show how many users who were influential in 2016 were no longer influential in 2020, and how many influential users in 2020 had been very uninfluential in 2016. Maybe I am missing something here - but are we seeing a more thorough analysis of this that provides a coherent statistic of movement over time? Obviously the authors may have cherry-picked these cases. ”

Fig. 4 (Brendan) needs to be rethought. Also, the rise of new influencers will not be the main results of our paper anymore (Stu).

“ Now to the major point: Figure 8 seems to be the main finding in that it shows an amazing amount of polarization for 2020 compared to 2016. Ideology is estimated using a homophily measure based on the 100 top influencers from each type of news, and users who retweeted at least 3 different influencers from this set. There are two mysteries in this figure. First, it suggests that in 2016 influencers were retweeted by a FAR MORE homogenous set of retweeters than in 2020. This might be understandable for some ideologically extreme influencers. But did the set of users retweeting the NY Times, CNN, and ABC really become radically more homogenous in 2020 than 2016? And if so, can the authors explain what happened? For instance, no right-wing users retweet ABC in 2020 (as I THINK the graph suggests)? And, on the bottom two panels we see that the distribution of users is just amazingly bimodal in 2020 compared to 2016.

SO - why? Why does the same method applied to 2016 data and 2020 data produce such different distributions? It may of course be that there was much more homophily in 2020 than 2016. But was this centered on some small set of elite users? I am not suggesting the authors have an error in their code. But I just want to see some diagnostics here, and perhaps some analysis of a sample of users to see what was separating them so well? One thing that would generate the 2020 distribution is if all users EITHER retweet (Fox, Breitbart, Donald Trump) OR (MSNBC, Bernie Sanders, CNN). Is that what happened? I don’t see how one can publish such a stark finding without some analysis suggesting we should have faith in the interpretation that retweet networks of top influencers showed incredible homogeneity. In other words, I agree that Figure 8 suggests a huge increase in homophily in 2020. BUT - can we see why? What would this graph look like if one REMOVED the ‘new influencers’ in 2020, and just computed it based on the 2016 influencers? What would it look like if one computed it only for traditional media? ”

The Referee did not understand that the cosine similarity results (Fig. 7) is actually a robustness check of Fig. 8. But still, we should provide more robustness/diagnostics of the result of Fig. 8. We can compute some statistics of the ratio of cross-retweeting (from left to right and vis versa) compared to retweeting in the same group (Alessandro). Also, compute Fig. 8 without the "new influencers", which was also asked by Reviewer 3 (Alessandro).

" I think the analysis here is really interesting. But: 1) I'd slim the paper down to address the main points that the authors actually address in the intro; and 2) present more diagnostics in the body of the paper to explain their key finding in Figure 8 and try to explain it. "

Ok. I'm not sure we can remove part of the analysis. But we could maybe move Fig. 7 to the SI, but keep its results in the main text (Bolek, Hernan)? Fig. 5 also is not really used for the moment.

" SOME OTHER THINGS:

a) Why was figure 6 discussed before figures 3-5? Or, why was it just not numbered 3? Was this just some typesetting issue? "

I think it's because we moved the influencer table in the SI, so we referred to Fig. 6 instead of the Table (Alex).

" b) Figure 4 - which is probably the most important figure in the paper - was typeset in such a way as to make the individual users unreadable. Would it have been difficult to put the user names to the left of their rank in 2016? "

Need to improve Fig. 4 (Brendan).

" c) Figure 5 - I had a hard time understanding what was being shown with this. "

This is the follower count evolution. We should decide if we want to keep it as we don't really use it in the discussion for the moment (Stu, Hernan, Bolek).

" d) The authors seem to claim they have found echo chambers in Figure 1. Can they give some measure of that? "

We can add some measure of echo-chamber from Fig 1 C & D and better explain it (Alessandro, Alex).

" e) Saying that Figure 1 shows increased polarization from 2016 to 2020 because we seem to see a smaller proportion of tweets in the center may be right. But the bigger shift seems to be the shift left. "

Yes ok. Need to add this in the text (Alex).

List of tasks:

- Rewrite title/abstract/significance/intro/discussion taking into account the comments of the Reviewers. (Stu, Bolek, Hernan)
- Need to clearly define: influencers, elites, polarization. (Stu + everybody).
- Redo Fig. 4 (Influencers ranks change) taking into remarks. (Brendan)
- Should we change Fig. 3 to show proportion of influencers linked to media outlets instead of indiv/organisations ? This would help with the story. We can also show this information in Fig. 4 which would make it much better. (James, Bolek, Hernan, Alex)
- Redo some analysis of 2020 using only users present in both 2016 and 2020: Latent ideology and cosine similarity. (Alessandro, James)
- Investigate importance of quotes in 2016 and 2020: repeat cosine similarity (James), latent ideology (Alessandro) and influencers retweet network (Brendan) with quotes instead of retweets (but keeping the same influencers).
- Compare CI ranking from unweighted network with ranking with weighted network (James, Hernan).

- Redo Fig. 6 (retweet network) with more visible arrows (Brendan). Try to understand what the reviewer is asking here...
- Better explain/measure the echo-chambers from Fig. 1 C & D. (Alessandro, Alex)
- Show Fig. 1 for 2020 using 2016 classification. (Alessandro)
- Show Fig. 2 with more refined categories. (Zhenkun)
- Compute some diagnostics of Fig. 8 to show robustness (Alessandro, Alex).
- We should use the results of Fig. 5 (followers count) in the discussion (Stu, Bolek, Hernan).
- See if we can slim down the results, move some parts to SI. (Alex, Bolek, Hernan)
- Need to clarify several small points in the results text (Alex).
- Write answers to referees. (Stu, Bolek)