Color Codes for comments

Matt (measurement and modeling + news attraction)

Dan (trait v. state)

Trevor (text data questions and misc. responses)

Completed responses

**Response Letter**

Paper Title: News ‘Attraction’ and Digital Inequalities: Incidental News Exposure and the Equalization or Stratification of Political Information

ID: 223259527

**Editor’s Comments**

**Comment:** We have now received the reviewers’ reports on your paper “News ‘Attraction’ and Digital Inequalities: Incidental News Exposure and the Equalization or Stratification of Political Information." Your paper has more reviews than the usual two. We initially received divergent reviews, which led as to reach out another two reviewers to make a robust assessment of your piece. That said, we apologize for the long wait this paper had to go through.  
  
In general, the reviewers had positive Comments about the manuscript, but there is still work to do for this piece to be ready for publication. As such, we are offering you the opportunity to revise and resubmit your paper addressing the suggestions made by the reviewers as well as the Digital Journalism editorial team.

Reviewer 1 invites you to better visualize your results, while Reviewer 2 would like you to reflect on some concepts and measures in the study. Reviewer 3 requests a better justification for presenting competing hypotheses, as well as stronger arguments to back up some conceptual definitions (such as news engagement or news attraction). This reviewer also has many methodological observations. Reviewer 4 would like you to discuss the context of your study, and invites you to provide more information for the methods section regarding sampling, variables, and models. This reviewer also provides helpful insight to strengthen your discussion section.

**Response**: We would like to thank the editor’s for facilitating the review process. We also want to thank the reviewers for their time and attention, as their critical comments helped us revise the manuscript in meaningful ways. In this revision, we focused on clarifying our arguments around the concepts of news attraction and the state versus strait distinction in news exposure. In addition, we noticed several places where the reporting (especially related to the rolling-cross sectional design and nature of the ‘linkage’ between respondents and content data) lacked clarity. Finally, we updated the discussing section to integrate important issues around creating an informed audience.

**Comment:** In addition, the Digital Journalism editorial team would like you to respond to the following:

Sample and Data

The study relies on a cross-sectional survey of adult social media users in the US, but there is no explanation as of how these users were recruited. Did you hire a polling company, perhaps?

**Response**: Thank you for catching this. This is an oversight. We used Qualtrics, and the sample was balanced for age, race, gender, and census region. We added this important detail to the manuscript under the Survey Design and Sample sections on pages 10-11.

**Comment:** Also, the paper indicates that “survey responses were linked with social media content collected via Brandwatch (formerly Crimson Hexagon) and then validated by crosschecking content lists with CrowdTangle.” It’s not clear why survey responses needed to be validated, and how (what is that you validated, exactly?).

**Response**: This comment from the editors aligns with similar concerns from reviewers. We lacked precision in our reporting of the study design. In turn, this lack of clarity lead to some ambiguity about other aspects of the study, including the nature of the rolling cross-section (RCS) survey and the ‘linkage’ between respondents and social media content.

We have heavily revised our explanation of the study design in the Methods section to address these points on pages 10-12.

To be clear, the survey responses are not validated. Rather, we implemented a multi-step strategy to identify and verify the top news circulating on Facebook at the time of data collection.

We then embedded the top two news articles circulating on Facebook at the time (ranked by engagement metrics) into the survey as a cue for attention to news during the previous news cycle. The assumption is that posts with higher engagement metrics would be more likely to show up in people’s feeds, and therefore a respondent who only passively follows the news should have seen it.

To decide what counted as ‘top’ news organization, as well as the most widely circulated/top stories, we developed a criterion that employs third-party data brokers’ public and private ranking systems.

First, we used NewsWhip to identify the top 25 news organizations on Facebook (NewsWhip regularly release reports to the public). Next, we used Brandwatch (formerly Crimson Hexagon) to identify the top posts on Facebook based on engagement from those 25 outlets, and finally, CrowdTangle was referenced to verify that those top stories were indeed widely circulating on Facebook during four days prior to data collection. The author has accounts with Brandwatch and CrowdTangle, so, while the news posts were publicly available, the rankings for news engagement are only available to researchers.

The main benefit of this approach is that we can a) ID the top stories widely circulating on Facebook with a higher level of certainty than simply relying on the researcher’s intuition, and b) we can better take advantage of the RCS design. RCS designs are useful, because survey responses can be more closely tied media campaigns and election cycles. We followed the lead of previous published work in this area (see, De Vreese et al., 2017 for an overview and past uses of RCS to link media content with survey data). This approach is now rather common in election studies.

**Comment:** In addition, validating survey responses with social media data requires knowing the respondents’ usernames. Does this violate IRB concerns about anonymity/confidentiality? Please explain.

**Response**: No personally identifying information was collected from individual participants and we did not have access to respondents’ social media accounts. As outlined in the previous response, our discussion of the RCS and nature of the linkage study have been revised for clarity. We specifically address this concern in the revised manuscript. We note that no personally indefinable information was collected on page 11.

**Comment:** Is there any reason to measure age as intervals instead of a continuous ratio variable?

**Response**: We made a conscious decision to capture age in this manner. Ultimately, we traded parsimony over precision. Ratio-level of precision is not required to answer our research questions. While we agree that categorical variables can make interpretation of statistical results less straight-forward, we are comfortable with the eight categories offered here. This provides meaningful evidence of differences according to broadly defined age groups. Accuracy beyond this level of inference is not particularly useful or insightful for this study.

**Comment:** Why did you impute missing values in a 2,000-case sample? Did you have too many missing values? How big was your sample if you didn’t input missing data?

**Response**: Thank you for catching this. Indeed, readers may want to know about the nature of missingness. The sample size for our models before imputation is N = 1,731. We revised the sample reporting to include this information on pages 11-12.

We decided to impute data not for reasons of smaller sample sizes, but rather to protect against one of the drawbacks of the RCS design. That is, RCS typically means more waves (17 in our case) but fewer respondents per wave. Thus, there is an inherent potential for sample size variation between waves, which could theoretically bias statistical results (sampling bias). We took two steps to minimize this bias: multilevel modelling (as is common in designs that rely on multiple data collection points) and imputation of missing cases.

Based on analysis of missing patterns, we did not see any issue with data quality in this regard. The missingness ranges from 6.3% (Age) to 1% (Strength of Party) and only 9 of the 28 modelled variables showed missing cases. Here is the report (which we will omit from the manuscript for space considerations:

|  |  |  |
| --- | --- | --- |
| variable | number missing | % |
| age | 126 | 6.2749004 |
| ine2 | 106 | 5.27888446 |
| alg | 59 | 2.93824701 |
| eff.int | 22 | 1.09561753 |
| int | 22 | 1.09561753 |
| ideo | 19 | 0.94621514 |
| pid | 19 | 0.94621514 |
| ideo.ext | 19 | 0.94621514 |
| pid.str | 19 | 0.94621514 |
| ine1 | 0 | 0 |
| smnews | 0 | 0 |
| engage | 0 | 0 |
| know | 0 | 0 |
| nml | 0 | 0 |
| sml | 0 | 0 |
| int.freq | 0 | 0 |
| sm.freq | 0 | 0 |
| female | 0 | 0 |
| poc | 0 | 0 |
| edu | 0 | 0 |
| inc | 0 | 0 |
| sm.newsintent | 0 | 0 |
| smfollow | 0 | 0 |
| avoid | 0 | 0 |
| sm.ns | 0 | 0 |
| sm.div.occ | 0 | 0 |
| sm.div.tie | 0 | 0 |
| grp | 0 | 0 |

**Comment:** Exposure and Engagement; The methods section describes ‘trait-like’ and ‘state-like’ properties of news exposure, but none of this is explained in the front end, let alone the most adequate way to measure these properties. For instance, total exposure to political information was measured with Weeks et al., 2017’s selective exposure items, but it’s not clear why ‘trait-like’ properties should be measured as selective exposure. Please address the ‘trait-like’ and ‘state-like’ concepts in the lit review and explain the rationale to measure them.

**Response**: Thank you for raising this issue, as we view our theorizing and measurement of incidental exposure on both trait- and state-levels to be an important contribution. We have now offered clearer definitions of trait and state incidental exposure early in the paper (p. 4) and more consistently integrated this distinction throughout the introduction. These changes offer a stronger rationale for why we measure incidental exposure on these two levels (see p. 9).

**Comment:** Please indicate if you ran a factor analysis to create the high-effort engagement variable, to make sure the items loaded together.

**Response**: Thank you for raising this issue, as it gives us the opportunity to fine-tune our analysis. We performed an exploratory factor analysis and found that the items load onto a single factor. We had conceived of the high-effort variable as a conceptual robustness check, but considering the results of the factor analysis, we have decided to drop the variable from the paper completely. We would be happy to approach the issue differently at the suggestion of the editorial team.

**Comment:** Controls; What are the theoretical reasons to control for political ideology, party identity, and identity strength? These variables are significant in most of the models, so there is clearly something going on (especially if you are studying exposure to political information) but these significant findings are not discussed in the paper. You might want to give it some thought.

**Response**: We feel that no contemporary study on news audiences would be complete without including these important variables as covariates. This is because these characteristics have long-been associated with a range of civic and political behaviors, including attention to and engagement with political information. Theoretically, political identity (as we capture via ideology and party affiliation) haven been linked to mass media and news consumption via models of active citizenship; for example the ‘dutiful’ and monitorial citizens who pay attention to news during election cycles (Bennett et al, 2009; Shehata et al., 2015; Zaller, 2010), as well as theoretical models that address the nature of selective exposure and ideological news (e.g., Holbert et al., 2012; Stroud, 2011). Accordingly, party identity is also implicated in incidental exposure studies via the news attraction concept introduced by Thorson (as cited in the paper) and further explicated in the manuscript.

Therefore, we included these variables in the models to account for possible confounding influences on our variables of interest. For brevity’s sake, we omit this more detailed discussion from the literature review. But this comment did lead us to consider the connections between these antecedent characteristics and our news attraction concept, which we address in the Discussion section.

To address this comment, we added a justification for including these variables to the manuscript on page 16 (where we report on Controls). In addition, and in line with requests from the reviewers, we noted these findings in the Discussion section on page XX.   
  
**Comment:** Limitations; Finally, you mention self-reported measures as a limitation, pointing out this is an endemic issue to survey research. However, some of the self-reported measures in this study are particularly problematic. For instance, how reliable is asking respondents whether they clicked on a story, scanned the headline or read it entirely, Commented, discussed, etc.? These actions are not easy to remember. The same might happen with network size (is this something you could cross validate with Brandwatch or CrowdTangle?). Please elaborate on this issue.

**Response**: We agree that capturing the various practices and behaviors related to news consumption in contemporary media environment is a challenge. Thus, we chose the RCS design to better capture news attention and engagement behaviors. The benefit of this design is that we are cuing respondents by embedding actual news stories circulating on social media during the data collection waves. Each wave we adjusted the stories, based on third-party ranking metrics, to ensure that respondents received stories that were tied to that news cycle.

This practice is not uncommon in election studies (De Vreese et al., 2017) and we feel that it offers a higher level of external and internal validity. That is, for internal validity, respondents are answering questions based on whether they actually saw, or paid attention to, a specific story we embedded in the survey, and we filtered respondents out who could not recall that story. As opposed to more traditional measures that ask about a general recollection of their habits related to news media, this method offers more specificity. This design also offers a higher degree of external validity, as we are asking respondents to reflect on specific, real-time news stories that are circulating widely on social media within the last day or so according to more than one ranking metric (CrowdTangle and Brandwatch).

One shortcoming of the RCS design is that is significantly more labor intensive, more expensive, and sample sizes are smaller in each wave. This makes balancing the sample for demographic traits a challenge. We reasoned that these tradeoffs were worth the improved, if flawed, measures of news attention and engagement.

Finally, we also used established measures where possible and reported validity statistics via the Cronbach’s alpha. That is the case for social network size (a rather common operationalization, e.g., Barnidge & Xenos, 2021, as cited on page 15).

As noted above, in the previous version of the manuscript, there was some ambiguity about the nature of the linkage between respondents and the social media data we collected. To be clear again, no personally identifying information was collected from individual participants and we did not have access to respondents’ social media accounts. Thus, CrowdTangle/Brandwatch would not offer any data related to individuals, just that of publicly available posts from media organizations.

**Reviewer: 1**

**Comment:** The way in which the findings are presented in this paper will contribute significantly to the development of the field of incidental news exposure. The results are well structured, and the methodology used is clear. The limitations of the study are also declared. It is suggested to incorporate tables that allow faster visualization of the results in the different variables used.

**Response**: Thank you for your time in reading and responding to the manuscript.

**Reviewer: 2**

**Comment:** This manuscript investigates whether the use of digital media reduces or exacerbates inequalities in news exposure and engagement based on survey data. The paper is very well-written and deals with an important topic. The literature review section is thorough and makes logical sense. The analyses are rigorous. Limitations (e.g., cross-sectional nature of the data) are also well-noted. Overall, I highly value this paper and recommend this paper be published in this journal after addressing the following concerns.  
  
**Response**: Thank you for the close reading of the paper and we appreciate your helpful comments.

**Comment:** While the author's way of measuring IE is more sophisticated compared to the previous way of measuring this concept which relied on a single item, the author(s) still need to acknowledge that it is still very difficult to measure IE with the survey. Survey respondents are not good at distinguishing to what extent their exposure was "accidental" or "purposeful." The author(s) can mention this point in the limitation section.

**Response**: This is an excellent point. We made several direct efforts to both employ and build upon existing measures of IE (Nanz & Matthes, 2022; Weeks & Lane, 2020). However, we agree that these concepts are imperfect and inherently difficult to capture. We added this note on page XX of the Discussion section.

**Comment:** I don't think the author(s) explained what story they chose and why they chose this topic. I need more information.

**Response:** This comment relates to concerns raised by other reviewers as well as the editors. To address this point, we will simply paste the comment made to the editors above here:

“We lacked precision in our reporting of the study design. In turn, this lack of clarity lead to some ambiguity about other aspects of the study, including the nature of the rolling cross-section (RCS) survey and the ‘linkage’ between respondents and social media content.

We have heavily revised our explanation of the study design in the Methods section to address these points on pages 10-12.

To be clear, the survey responses are not validated. Rather, we implemented a multi-step strategy to identify and verify the top news circulating on Facebook at the time of data collection.

We then embedded the top two news articles circulating on Facebook at the time (ranked by engagement metrics) into the survey as a cue for attention to news during the previous news cycle. The assumption is that posts with higher engagement metrics would be more likely to show up in people’s feeds, and therefore a respondent who only passively follows the news should have seen it.

To decide what counted as ‘top’ news organization, as well as the most widely circulated/top stories, we developed a criterion that employs third-party data brokers’ public and private ranking systems.

First, we used NewsWhip to identify the top 25 news organizations on Facebook (NewsWhip regularly release reports to the public). Next, we used Brandwatch (formerly Crimson Hexagon) to identify the top posts on Facebook based on engagement from those 25 outlets, and finally, CrowdTangle was referenced to verify that those top stories were indeed widely circulating on Facebook during four days prior to data collection. The author has accounts with Brandwatch and CrowdTangle, so, while the news posts were publicly available, the rankings for news engagement are only available to researchers.

The main benefit of this approach is that we can a) ID the top stories widely circulating on Facebook with a higher level of certainty than simply relying on the researcher’s intuition, and b) we can better take advantage of the RCS design. RCS designs are useful, because survey responses can be more closely tied media campaigns and election cycles. We followed the lead of previous published work in this area (see, De Vreese et al., 2017 for an overview and past uses of RCS to link media content with survey data). This approach is now rather common in election studies.”

**Comment:** When measuring social media use, the author(s) used a single item. It is much more desirable to use multiple items when measuring this concept since a) different social media platforms have different functionalities and b) survey respondents may not take the term "social media" in the same way. For instance, some may think "YouTube" is not social media. Likewise, some may think of WhatsApp as social media (while some may not). So, the authors better use multiple items, or even if the authors end up using a single item, the authors need to at least add examples in the parenthesis such as "social media (e.g., x,y,z)."

**Response:** We agree with the premise of this comment and indeed rely on multiple-item measures, as well as more advanced reliability scales and techniques where appropriate (as evidenced by our approach to measurement of key independent and dependent variables).

Since this study was focused on answering a research question related IE and associated behavioral phenomena, we were not particularly concerned with controlling for overall social media use across multiple platforms.

But we were cognizant of the limitations of a single-item measure used in this way. This particularly survey item was adapted from a paper that recommends a similar measure for time spent on Facebook. Ernala and colleagues (2020) validity checked various measures of Facebook use in a large N sample from Facebook Research (15 countries and N = 49,934) that included both self-report and observed, live social media usage data. The single-item correlates best with real time usage. To our surprise, we omitted this reference.

To address this point, we included the citation for this measure in the manuscript on page 16.   
  
**Reviewer: 3**  
  
**Comment:** This paper addresses how incidental news exposure and news engagement relate to the concept of “news attraction.” The authors attempt to explicate and measure this concept and then test it to better understand if incidental exposure can help reduce informational inequalities. My primary concerns lie in the presentation of the hypotheses and in the operationalization of the “news attraction” variable. I have presented my concerns in the order they appeared in the paper:

**Response**: Thank you for the careful reading of the manuscript. These comments forced us to clarify key concepts and provide deeper rationale for our modelling and reporting decisions. We welcomed these critiques and hope the current version of the manuscript appropriately addresses these concerns.

**Comment:** The quotes on page 6 need correcting to capture which portions of the sentence are directly quoted.

**Response**: Thanks for catching this. We updated this sentence for the proper quotation placement.

**Comment:** I’m not following the paragraph that leads to H1a and H1b. Why would overall exposure be equal for those who are high and low in news attraction? It seems by definition that those high in news attraction would have greater levels of overall exposure.  
There’s a typo in H1a and H1b as well as H2a and H2b (“and” should be “in”).

I generally don’t like competing hypotheses. In some instances, competing hypotheses are used to ensure a supported hypothesis is possible no matter the results. I would prefer a stronger theoretical argument that leads to a specified prediction. Given that it’s difficult to do this ethically after data has been analyzed, I would like to see the authors do a better job justifying the decision to present competing hypotheses.

I would like to see a stronger argument for why “news engagement” matters in this context. For example, is engagement necessary for someone to learn from news content? It seems that simple exposure could prompt some knowledge gain. What exactly does engagement add to people’s knowledge gain process that is not available through exposure? Perhaps a stronger argument could be made for more in-depth processing (e.g., central route) occurring during “news engagement.”

**Comment:** The paragraph before H2a and H2b suggests interaction effects, but the hypotheses are not written to predict interaction effects. Please consider revising.

**Comment:** Was the trait-like scale for total exposure specific to social media use? It seems that specifying this information should come from social media would be important to this measure.  
Asking people to know if they came across information accidentally is asking a lot of their memory and accuracy. Are there prior studies that validate the use of this measure as an accurate way to know that someone encountered information accidentally? (The limitations of this self-report measure should be more thoroughly addressed in the paper – not just briefly touched on in the limitations paragraph.)

**Comment:** Why was total exposure multiplied by incidental exposure? Would the results change if these scales were not combined in this way (i.e., if the models were rerun with these variables assessed independently)?

**Comment:** Why was a separate “high-effort engagement” variable created? This needs more justification both theoretically and operationally.

**Comment:** It’s not clear to me how the second measure of news attraction, “self-reported interest,” is a news-related variable. Is political interest inherently part of news attraction? This measurement seems to suggest that someone who is interested in news and politics would necessarily get news on social media. Including interest as a covariate in the model makes sense, but it’s not clear to me why it is part of a measure of “news attraction.”  
Given the context of the study, I would encourage the authors to consider relabeling “news attraction” to “social media news attraction.” The crux of the argument surrounds exposure to news on social media.

**Comment:** Per my previous Comment, it seems entirely possible that someone could have high levels of political interest but only get news from non-social media sources.  
You address this previous point in the “regression analyses: exposure” section of the paper. I would like to see the overlap of non-social media news use and social media news use parsed out better in the literature review, specifically as it relates to “news attraction.” It wasn’t clear to me that “news attraction” would apply both to non-social media news as well as social media news, in part due to the way the variables in the news attraction variable were operationalized. I would like to see these decisions better justified (both conceptually and operationally).  
  
**Reviewer: 4**  
  
**Comment:** This paper elaborates on the concept of ‘news attraction’ to examine the extent to which digital media users are exposed to news intentionally or accidentally, and how incidental exposure is related to engagement with news. It uses data collected through a cross-sectional online survey fielded in the USA during the 2020 elections. The results provide evidence of a paradox: while the online environment can be an equalizer of exposure to news across social groups, it can also stratify engagement with the news.  
  
There are many things to like about the paper. It tackles a relevant topic, namely, whether digital media reduces or exacerbates inequalities in passive and active news use. It is theoretically rich, covering the key works in the literature on incidental exposure. The statistical analysis is more sophisticated than what is typical for papers using cross-sectional data, as it includes a latent class analysis, estimates hierarchical models, etc. The appendix is helpful, too, as it enables readers to assess the representativeness of the sample and the robustness of some findings.  
  
Having said that, I see several areas for improvement, especially regarding the context of the study, the methods used, and discussion sections. Let me elaborate on these shortcomings.

**Response**: Thank you for the supportive comments. We also appreciate you taking the time to outline your concerns. Your comments (along with those of the editors) made it clear to us on rereading the manuscript that some of the methodological reporting was either vague or incomplete. We revised those sections. We also made changes to the literature review to better explicate core concepts. We hope the current version of the manuscript sufficiently addresses these concerns.   
  
**Comment:** Somewhat ironically (as the authors make highlight the importance of studying the context of media exposure), I missed a discussion on the context of the study. Empirical findings are always bounded by cultural, temporal, and other forces. The polarized American media and political systems are rather unique in the world. I’m sure incidental exposure to political news and current events differs between election and nonelection years. All this is to say that it will greatly benefit the international audience of the journal if the authors include one or two paragraphs in the methods section about the particularities of the US case that are relevant for this particular study.

**Response**: Add this to the Discussion section and refer to media system/political culture as contextual and possible implications outside US.   
  
**Comment:** The methods section is lacking important pieces of information, and many of the authors’ choices are not justified or explained. This has the unfortunate consequence of making the statistical models less parsimonious and the results harder to follow.

**Response**: Thank you for raising this legitimate concern. As discussed elsewhere in this letter, we heavily revised the Methods section (pages 10-17) to better justify our research design. We also further explicated core concepts related to the trait/state distinction (page 4). Details are included in the responses below.

**Comment:** Why were the data collected using a rolling cross-sectional design (RCSD) instead of a single one-shot design? I’m asking considering that the longitudinal aspect of the survey does not seem of any relevance to this study.

**Response**: This comment echoed concerns raised above. For convenience, we will include previous answers here:

“We chose the RCS design to better capture news attention and engagement behaviors. The benefit of this design is that we are cuing respondents by embedding actual news stories circulating on social media during the data collection waves. Each wave we adjusted the stories, based on third-party ranking metrics, to ensure that respondents received stories that were tied to that news cycle.

This practice is not uncommon in election studies (De Vreese et al., 2017) and we feel that it offers a higher level of external and internal validity.”

And, from above:

“We lacked precision in our reporting of the study design. In turn, this lack of clarity lead to some ambiguity about other aspects of the study, including the nature of the rolling cross-section (RCS) survey and the ‘linkage’ between respondents and social media content.

We have heavily revised our explanation of the study design in the Methods section to address these points on pages 10-12.

To be clear, the survey responses are not validated. Rather, we implemented a multi-step strategy to identify and verify the top news circulating on Facebook at the time of data collection.

We then embedded the top two news articles circulating on Facebook at the time (ranked by engagement metrics) into the survey as a cue for attention to news during the previous news cycle. The assumption is that posts with higher engagement metrics would be more likely to show up in people’s feeds, and therefore a respondent who only passively follows the news should have seen it.

To decide what counted as ‘top’ news organization, as well as the most widely circulated/top stories, we developed a criterion that employs third-party data brokers’ public and private ranking systems.

First, we used NewsWhip to identify the top 25 news organizations on Facebook (NewsWhip regularly release reports to the public). Next, we used Brandwatch (formerly Crimson Hexagon) to identify the top posts on Facebook based on engagement from those 25 outlets, and finally, CrowdTangle was referenced to verify that those top stories were indeed widely circulating on Facebook during four days prior to data collection. The author has accounts with Brandwatch and CrowdTangle, so, while the news posts were publicly available, the rankings for news engagement are only available to researchers.

The main benefit of this approach is that we can a) ID the top stories widely circulating on Facebook with a higher level of certainty than simply relying on the researcher’s intuition, and b) we can better take advantage of the RCS design. RCS designs are useful, because survey responses can be more closely tied media campaigns and election cycles. We followed the lead of previous published work in this area (see, De Vreese et al., 2017 for an overview and past uses of RCS to link media content with survey data). This approach is now rather common in election studies.”

**Comment:** The data was analyzed using multilevel modeling. According to the authors, this was justified: it helps to control for measurement invariance across the 17 sampling frames employed. But the authors should be aware that MLM are harder to interpret for the lay reader than a simple, straightforward OLS. Perhaps comparing the robustness of results across using different estimators would help answer the question of how dependent the results on MLM are.

**Response**: Thank you for raising this concern, as it gives us a chance to stress-test our analysis. We spot-checked robustness using Model 1 (DV: Incidental Exposure) and Model 3 (DV: Total Exposure) in Table 1. In both cases, model comparisons show that the MLM approach fits the data better than the OLS approach, with lower AICs and statistically significant log-likelihood tests (for Model 1, chi-square = 5.16, p = .023, indicating the MLM is a slightly better fit; for Model 3, chi-square = 29.17, p < .001, indicating MLM is a much better fit).

Furthermore, the coefficient estimates are similar in both cases. For Model 1 the key estimates are Low = 1.15 (MLM) vs. 1.14 (OLS); Medium—Unmotivated = 0.62 (MLM) vs 0.62 (OLS); Medium—Motivated = 0.36 (MLM) vs. 0.37 (OLS); and High = -0.16 (MLM) vs. -0.14 (OLS). The differences are slightly bigger for Model 3, but not big enough to alter the paper’s substantive conclusions: Low = 1.12 (MLM) vs. 1.09 (OLS); Medium—Unmotivated = 0.79 (MLM) vs. 0.82 (OLS); Medium—Motivated = 1.22 (MLM) vs. 1.30 (OLS); High = 1.59 (MLM) vs. 1.67 (OLS).

We would be happy to add a footnote to the manuscript describing these robustness checks if the reviewer feels this is appropriate and would clarify the findings for the readers.

**Comment:** How much data was missing to justify the use of multiple imputation using chained equations? Do the results change if missing data is not imputed?

**Response**: This is an important detail that we omitted to report in the previous version of the manuscript. There are two answers here. The first is related to our rationale to impute (along with a report of the missingness by variable) and the second is related to our approach to data treatment in general.

First, to repeat a response to the editors:

Thank you for catching this. Indeed, readers may want to know about the nature of missingness. The sample size for our models before imputation is N = 1,731. We revised the sample reporting to include this information on pages 11-12.

We decided to impute data not for reasons of smaller sample sizes, but rather to protect against one of the drawbacks of the RCS design. That is, RCS typically means more waves (17 in our case) but fewer respondents per wave. Thus, there is an inherent potential for sample size variation between waves, which could theoretically bias statistical results (sampling bias). We took two steps to minimize this bias: multilevel modelling (as is common in designs that rely on multiple data collection points) and imputation of missing cases.

Based on analysis of missing patterns, we did not see any issue with data quality in this regard. The missingness ranges from 6.3% (Age) to 1% (Strength of Party) and only 9 of the 28 modelled variables showed missing cases. Here is the report (which we will omit from the manuscript for space considerations:

|  |  |  |
| --- | --- | --- |
| variable | number missing | % |
| age | 126 | 6.2749004 |
| ine2 | 106 | 5.27888446 |
| alg | 59 | 2.93824701 |
| eff.int | 22 | 1.09561753 |
| int | 22 | 1.09561753 |
| ideo | 19 | 0.94621514 |
| pid | 19 | 0.94621514 |
| ideo.ext | 19 | 0.94621514 |
| pid.str | 19 | 0.94621514 |
| ine1 | 0 | 0 |
| smnews | 0 | 0 |
| engage | 0 | 0 |
| know | 0 | 0 |
| nml | 0 | 0 |
| sml | 0 | 0 |
| int.freq | 0 | 0 |
| sm.freq | 0 | 0 |
| female | 0 | 0 |
| poc | 0 | 0 |
| edu | 0 | 0 |
| inc | 0 | 0 |
| sm.newsintent | 0 | 0 |
| smfollow | 0 | 0 |
| avoid | 0 | 0 |
| sm.ns | 0 | 0 |
| sm.div.occ | 0 | 0 |
| sm.div.tie | 0 | 0 |
| grp | 0 | 0 |

Second, we used a multiple imputation algorithm (predictive mean matching) which uses full information (drawing from a group of similar, but complete cases) and thus should reduce bias as much as possible (see Van Buuren, 2018).

As a matter of data treatment, we did not run alternative models without the imputations. The overall number missing is rather low. This is also a matter of research ethics, as of course we would not choose results based on a favorable outcome relative to imputing missing data. Our decision is based solely on the nature of data collection, and we prefer to use the imputed datasets.

For the reviewer’s reference, here are the model statistics for the imputed versus non-imputed datasets:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 1 Models** | | | | |
|  | M1 | M2 | M3 | M4 |
| **Fit Statistics Imputed** |  |  |  |  |
| ICC | .01 | .02 | .03 | .02 |
| LL | -3,137.86 | -1,051.30 | -3,088.25 | -1,339.30 |
| Pseudo-*R*2 | .17 | .11 | .44 | .11 |
| **Fit Statistics List-Wise Deletion** | | | | |
| ICC |  |  |  |  |
| LL |  |  |  |  |
| Pseudo-*R*2 |  |  |  |  |

|  |  |  |
| --- | --- | --- |
| **Table 2 Models** | | |
|  | M1 | M2 |
| **Fit Statistics Imputed** |  |  |
| ICC | 01 | .02 |
| LL | -1,634.20. | -1,285.53 |
| Pseudo-*R*2 | .45 | 0.46 |
| **Fit Statistics List-Wise Deletion** | | |
| ICC |  |  |
| LL |  |  |
| Pseudo-*R*2 |  |  |

**Comment:** Do results change when using unweighted data? I’m asking because weights, while helping to address deviations from the population distribution, inflate standard errors, too.

**Response:** In this particular case, the weights do not inflate the standard errors. We spot-checked robustness using the same two models before, and the unweighted standard error estimates are very similar to the weighted estimates. We prefer to use the weighted models given the nature of the survey sample, and we would be happy to add a footnote explaining these robustness checks. Briefly, the key standard error estimates are (order = Low, Medium—Unmot, Medium—Mot, High) are:

Model 1 weighted: 0.05, 0.06, 0.08, 0.11

Model 1 unweighted: 0.05, 0.06, 0.08, 0.10

Model 3 weighted: 0.06, 0.05, 0.07, 0.10

Model 3 unweighted: 0.05, 0.06, 0.07, 0.10

**Comment:** More importantly, how were survey responses linked with social media content, exactly? And the validation of content lists with CrowdTangle, what was that and why was it necessary? And if most stories come from Fox News, what does this say about the representativeness of the media stimuli?

**Response:** We pasted the comments related to social media content above several times. Briefly, we relied on third-party rankings of top news circulating on Facebook at the time of data collection to ensure that respondents received stories that were tied to that news cycle. We looked at several ranking lists to validate that indeed, the stories we chose to embed in the survey were actually the ‘top’ posts. The assumption is that posts with higher engagement metrics would be more likely to show up in people’s feeds, and therefore a respondent who only passively follows the news should have seen it.

Regarding Fox News, they dominated the rankings for engagement metrics according to all three tracking companies we referenced (NewsWhip, Brandwatch, and CrowdTangle). This is why we chose two stories to embed and then randomize (each respondent saw only one story), one for the top story (which was always Fox) and the first non-Fox story. In this sense, we were not aiming for a ‘representative’ sample of news that circulated during a given survey wave, but rather as a point of external validity, we wanted to capture the news that was being pushed to most people based on the algorithms that favor engagement rankings.

Add note to discussion section here

**Comment:** I find it somewhat confusing why some IVs are described as covariates and others as control variables. Covariates predict variance in the DV that is clearly not attributable to the IVs of interest. That’s why socio-demographics are usually covariates. Control variables, however, are included to eliminate spurious relationships between the IV of interest and the DV that might otherwise be thought to be causal. Again, more explanation would be helpful.

**Response:** We may need to reconsider or be more careful about this language. Honestly, I do not think these are mutually exclusive categories of variables. But will check our wording on these.  
  
**Comment:** In the concluding section, I missed two central aspects. First, a discussion of how the study findings relate to the most important function of news: to produce informed readers. There is a larger debate on the consequences of the digitization of news and the rise of social media on citizen competence. Questions such as: what do people learn from they news? Why is social media news use related (or not) to misinformation? These are important questions, and I would like to know the implications of this study on those issues.

**Response:** Good questions. OFC not the focus of the study, but certainly there are implications for leaning. Need to add few sentences to the discussion.   
  
**Comment:** Second, for a paper that mentions repeatedly concepts such as datafication and algorithmic categorization, I expect some a more thorough discussion of what the study findings mean for the debate on the platformitization of news in the current media environment. The last paragraph of the paper hints at this. I’m sure the authors can elaborate more.

**Response:** We may need to reconsider or be more careful about this language.  
  
**Comment:** Last but not least, please proofread the manuscript. Some propositions are missing, there are spelling problems, etc.

**Response:** Do this and thank the author.

**References**

Bennett, W. L., Wells, C., & Rank, A. (2009). Young citizens and civic learning: Two paradigms of citizenship in the digital age. *Citizenship studies*, *13*(2), 105-120.

de Vreese, C. H., Boukes, M., Schuck, A., Vliegenthart, R., Bos, L., & Lelkes, Y. (2017). Linking Survey and Media Content Data: Opportunities, Considerations, and Pitfalls. *Communication Methods and Measures*, *11*(4), 221–244

Ernala, S. K., Burke, M., Leavitt, A., & Ellison, N. B. (2020). How Well Do People Report Time Spent on Facebook? An Evaluation of Established Survey Questions with Recommendations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14.

Holbert, R. L., Hmielowski, J. D., & Weeks, B. E. (2012). Clarifying relationships between ideology and ideologically oriented cable TV news use: A case of suppression. *Communication Research*, *39*(2), 194-216.

Shehata, A., Ekström, M., & Olsson, T. (2016). Developing self-actualizing and dutiful citizens: Testing the AC-DC model using panel data among adolescents. *Communication Research*, *43*(8), 1141-1169.

Stroud, N. J. (2011). *Niche news: The politics of news choice*. Oxford University Press on Demand.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press. https://stefvanbuuren.name/fimd/

Zaller, J. (2003). A new standard of news quality: Burglar alarms for the monitorial citizen. *Political Communication*, *20*(2), 109-130.