
Taylor & Francis
Taylor & Francis Group

Behaviour & Information Technology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tbit20

# Using website referrals to identify unreliable content rabbit holes

Kevin T. Greene, Mayana Pereira, Nilima Pisharody, Rahul Dodhia, Juan Lavista Ferres & Jacob N. Shapiro

BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2024.2352093



# Using website referrals to identify unreliable content rabbit holes

Kevin T. Greene[a], Mayana Pereira[b], Nilima Pisharody[a], Rahul Dodhia[b], Juan Lavista Ferres[b] and Jacob N. Shapiro[a]

[a]Empirical Studies of Conflict Project, Princeton University, Princeton, USA; [b]AI for Good Research Lab, Microsoft, Redmond, USA

**ABSTRACT**

Does the URL referral structure of websites lead users into 'rabbit holes' of unreliable content? Past work suggests algorithmic recommender systems on sites like YouTube lead users to view more unreliable content. However, websites without algorithmic recommender systems have financial and political motivations to influence the movement of users, potentially creating browsing rabbit holes. We address this gap using browser telemetry that captures referrals to a large sample of domains rated as reliable or unreliable information sources. Our results suggest the incentives for unreliable sites to retain and monetise users create rabbit holes. After landing on an unreliable site, users are very likely to be referred to another page on the site. Further, unreliable sites are better at retaining users than reliable sites. We find less support for political motivations. While reliable and unreliable sites are largely disconnected from one another, the probability of traveling from one unreliable site to another is relatively low. Our findings indicate the need for additional focus on site-level incentives to shape traffic moving through their sites.

**ARTICLE HISTORY**
Received 5 June 2023
Accepted 25 April 2024

**KEYWORDS**
misinformation; rabbit hole; web navigation

## 1. Introduction

Does the URL referral structure of websites lead users into 'rabbit holes' of unreliable content? Previous studies have often focussed on how YouTube's recommendation system may send users down 'rabbit holes' of unreliable content (Chen et al. 2022; Hussein, Juneja, and Mitra 2020; O'Callaghan et al. 2015; Papadamou et al. 2022; Santos, Lelkes, and Levin 2021; Tarantola 2021) or extreme content (Hosseinmardi et al. 2021; Ledwich and Zaitsev 2019; Ribeiro et al. 2020).

However, the formation of rabbit holes has not been extended to websites that do not rely on recommender systems. For these websites, there are also incentives for influencing the navigation of users. First, the attention economy offers sites the opportunity to profit off of their users, often through ad revenue (Davenport and Beck 2001; GDI 2019; Ryan et al. 2020) creating financial incentives for all types of sites to retain users. Importantly, if unreliable sites are better at retaining users (more 'sticky'), users will likely be exposed to more unreliable content. Rabbit holes may form as sites aim to retain users by embedding links to additional pages on their site.[1]

Second, while past work has found that political interests of sites shape their outbound traffic patters (Kaiser, Rauchfleisch, and Bourassa 2020; Sehgal et al. 2021; Starbird et al. 2018) the interests of the website operators can also lead to rabbit hole formation. These interests can be served by linking to other sites with similar orientations and avoiding sites that provide contrasting information. Referring users to external unreliable sites provides another outlet for the formation of rabbit holes.

Importantly, these incentives exist for both reliable and unreliable information sources.[2] However, if as some have noted unreliable sites are partially partisan (Acerbi 2019) or motivated by monetisation (Herasimenka et al. 2022), leading to the formation of link-based rabbit holes, the downstream implication is that users are likely to encounter content that is of low journalistic quality.

We address the formation of rabbit holes using data on the reliability of a large sample of websites along with traffic referral patterns between these sites. We measure the reliability of media outlets using ratings created by NewsGuard. We have reliability ratings for more than 1000 sites. We assess the potential for unreliable content rabbit holes using anonymous privacy-preserving aggregated data from a large sample of Edge web browser users who opted to share anonymized diagnostic information with Microsoft.[3] Our data capture every outbound href referral from our list of sites. Href links are often presented as clickable links that move users to another page on the same site or another site entirely.

**CONTACT** Kevin T. Greene kg2082@princeton.edu
Supplemental data for this article can be accessed online at http://dx.doi.org/10.1080/0144929X.2024.2352093.

Importantly, href links are an editorial decision made by the site/author and thus actively push users to specific types of sites. Our work makes a novel contribution to our understanding of browsing rabbit holes, by looking beyond the role of recommendation systems, to include the incentives of website operators.

Our results suggest the financial incentives of sites to retain users to drive revenue lead users into unreliable content rabbit holes. After landing on an unreliable site, the next referral is very likely to lead to an additional page on the site, meaning users are likely to view additional unreliable content. While this financial incentive exists for all sites, we find that unreliable sites are especially good at retaining users.

We find only limited support that the political incentives of sites lead to unreliable content rabbit holes. We do observe that reliable and unreliable sites are largely disconnected from one another. However, the probability of traveling from one unreliable site to another is relatively low. We reach the same conclusions when looking at the embedded links within sites, site referral traffic, and after applying statistical matching techniques to adjust for differences in the total traffic of reliable and unreliable sites. Finally, we find that reliable sites rarely drive traffic to unreliable sites, suggesting that fact-checking drives little traffic to unreliable content. However, we do observe a small set of reliable sites directing users to unreliable sites. Over 90% of these referrals come from four right-wing sites, the *Drudge Report*, *Townhall*, *Hotair*, and *Real Clear Politics*.

## 2. Related work

Past work has focussed on how algorithmic systems, such as YouTube's recommendation system, can lead users down 'rabbit holes'. An unreliable content rabbit hole occurs as users are taken from one topic to more extreme or unreliable content, often without their awareness. The concern is that as recommendation systems aim to increase engagement by presenting additional content that is similar to what users viewed in the past, a user could unknowingly arrive at unreliable content. For example, a study by Papadamou et al. (2022) found that users were recommended more pseudoscience content based on their watch history. However, more recent work has questioned if the consumption of low-quality content is driven by algorithms. A study by Chen et al. (2022) found that most consumption occurred from users who subscribed to conspiracy theory channels.

Additional work suggests that the broader incentives of websites to shape the movement of users may also create the potential for rabbit holes, even without the use of algorithmic recommendations. First, a growing literature has made the case that there are more general financial incentives to shape the flow of online traffic. The 'attention economy' offers sites the opportunity to profit off of their users, often through ad revenue (Davenport and Beck 2001; GDI 2019; Ryan et al. 2020). One means of profiting is to retain users on your site, increasing views and engagements with ads.

Second, past work also makes the case that there may be political incentives to direct users to particular sites. Website operators may be motivated to direct users to sites that promote their political positions while steering them away from contrary evidence. Kaiser, Rauchfleisch, and Bourassa (2020) find that inlinks to far-right alternative media outlets predominantly came from other far-right sites. Sehgal et al. (2021) reach a similar conclusion finding that there are strong referral connections between unreliable sites. These motivations appear to expand beyond the left-right dimension, as others have found that reputable media organisations are unlikely to link to those that promote unreliable information. For instance, Hanley, Kumar, and Durumeric (2022) find that sites pushing conspiracy theories were rarely linked to by reliable media outlets, but were often linked to by other low-quality sites, while Starbird et al. (2018) find distinct clusters for the alternative media and mainstream media.

We build on these three previously unconnected literatures to evaluate if the broader financial and political incentives of websites lead to rabbit hole of unreliable content.

## 3. Materials and methods

To evaluate the evidence for navigational rabbit holes we first need a list of sites that have been identified as unreliable information sources, as well as a set of sites that are reliable information sources. Consistent with previous work we measure the reliability of the information at the publisher level, rather than at the story level (Grinberg et al. 2019; Lazer et al. 2018). To measure the reliability of domains we use the list of domains compiled by NewsGuard.[4] NewsGuard is a company that rates the journalistic quality of of thousands of news websites, accounting for a majority of traffic to news sites (NewsGuard 2021). Their evaluation is based on a set of nine criteria, including evaluating if sites publish false content, correct factual errors, and effectively separate news from opinion. Sites that receive an overall score that falls below 60 out of 100 are classified as unreliable information sources. Domains with a score of 60 or greater are classified as reliable information sources. We include both unreliable and reliable sites because, without a comparison group, it is impossible to know if our findings are indicative of unreliable sites, or a broader feature of online media. This is particularly important because modern media companies are set up to retain traffic, meaning that rabbit hole effects might just be online media effects. Figure 1 displays the distribution of NewsGuard scores for websites in our study.

To measure referrals within and between sites we collect one month of telemetry data from the Microsoft Edge browser. The Edge logs provide traffic information for roughly 10% of the desktop browsing activity in the United States (StatCounter 2021), accounting for millions of users. This data was collected from users who had opted to share diagnostic information.[5] Our analysis begins on February 23rd and ends on March 22nd. For each domain rated by NewsGuard, we collect its inbound and outbound traffic information. Inbound traffic measures the domains visited before the domain of interest, while outbound traffic measures the domains visited after the domain of interest. We collect all traffic that results from referrals from href links. Importantly, while direct referrals may not capture all the potential sites users could have arrived at (not all embedded links will be clicked), they do capture all direct navigations from one page to another. Additionally, we also opt for href referrals rather than all browsing navigation (back/forward navigation, favourites bar, etc.) as they might pick up incidental movement between sites, rather than a link intentionally aiming to move a user to another page or site.

For our study, there are three quantities of interest.

(1) The number of outbound referrals that are self-referrals. These represent instances where a user is retained on a site and are likely to view additional content. As noted previously, sites are incentivized to retain users to increase ad revenues. However, if unreliable sites are more sticky than reliable media sites, the consequence is users will likely be exposed to more unreliable content.
(2) The number of outbound referrals that lead to external unreliable sites. Unreliable sites may link to other unreliable sites because of their shared political views, while the expectation is that reliable sites will not frequently link to unreliable domains (Sehgal et al. 2021; Starbird et al. 2018).
(3) The number of outbound referrals that lead to external reliable sites. This is an important measure for understanding the potential for rabbit holes because reliable outlets generally represent a higher quality set of information. If unreliable sites provide few links to external reliable sites, users are more likely to continue engaging with unreliable content.

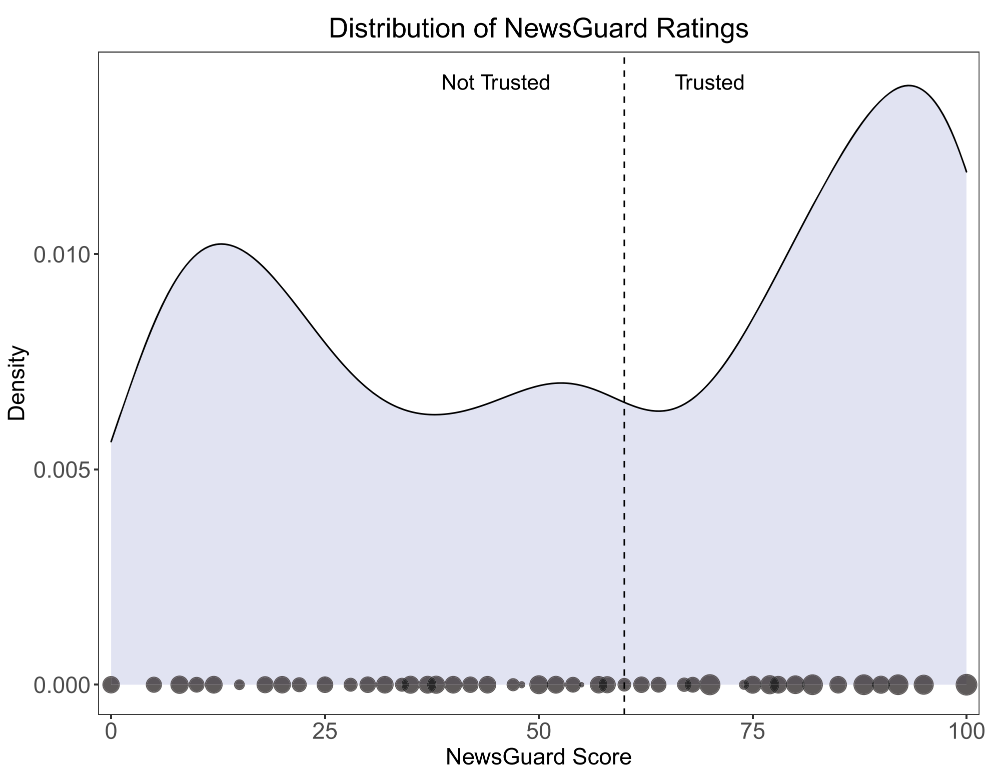


**Figure 1.** Distribution of NewsGuard scores in our study. The size of the points on the *x*-axis correspond to the logged web traffic for sites at that score. The dotted line delineates sites NewsGuard rates as trusted, from those they rate as not trusted.

## 4. Results

### *4.1. The stickiness of unreliable sites*

To begin our assessment of unreliable content rabbit holes, we evaluate if unreliable sites are more sticky than reliable sites (Quantity of Interest 1). If unreliable sites are better able to retain users, they are likely to view more unreliable content. While past work has tended to filter out internal site navigation in favour of evaluating the external sites that drive traffic to unreliable sites, internal navigation on unreliable sites leads users to additional unreliable content and serves as one means of forming unreliable content rabbit holes.

To assess this quantity for each site we calculate the proportion of its outbound referrals that are self-referrals (links leading to additional pages on the site). We find that, for our sample, roughly 68% of the total referrals to unreliable sites come from internal navigation (self-referrals).[6] For comparison, roughly 60% of referrals to reliable sites come from internal navigation. For additional context, the most trafficked unreliable site, Breitbart, receives roughly 96% of its inbound traffic from self-referrals.

To compare the overall differences in stickiness between unreliable and reliable domains, we calculate the proportion of incoming referrals that are self-referrals for each domain. For unreliable sites, the median proportion of self-referrals is roughly.98, while this value for reliable sites is around .75. We report a box plot of the self-referencing proportion of unreliable and reliable sites in Figure 2. As noted previously, the average level of self-referrals is considerably higher for unreliable sites than for reliable sites. The results of a t-test indicate that these differences are statistically significant ($t = -15.72$, $p<0.0001$, $n = 1265$). Because there are a variety of types of sites (aggregators, news sites, etc.) in Figure 2 (right), we re-estimate our results including only sites that engage in original reporting. This helps ensure that our results capture differences in site reliability rather than other site-level differences. The substantive results of the analyses are unchanged, unreliable sites are more likely to self-refer ($t = -14.97$, $p<0.0001$, $n = 1215$).

### *4.2. Referrals to unreliable sites*

The previous analysis establishes that unreliable sites are more likely to self-refer. The impact of this result is that once a user is on an unreliable site, they are likely to access additional unreliable content. This is one piece of evidence in support of the formation of unreliable content rabbit holes. Additional evidence of rabbit holes would be frequent referrals between unreliable sites and few referrals from unreliable to reliable sites (Quantities of Interest 2 and 3). Previous work finds support for this expectation when examining a smaller sample of sites (Kaiser, Rauchfleisch, and Bourassa 2020; Sehgal et al. 2021; Starbird et al. 2018).

To assess referrals to unreliable sites we calculate the proportion of outgoing referrals that lead to unreliable sites for each domain. We observe that the median level of referrals to unreliable sites is considerably higher for unreliable sites than for reliable sites. In fact, for reliable sites, the proportion of referrals to unreliable sites is close to zero. The results of a t-test indicate that these differences are statistically significant ($t = -120.94$, $p<0.0001$, $n = 1265$). These results are reported as violin plots in Figure 3. Violin plots illustrate that while the median proportion of referrals to unreliable sites is similar across reliable and unreliable sites after

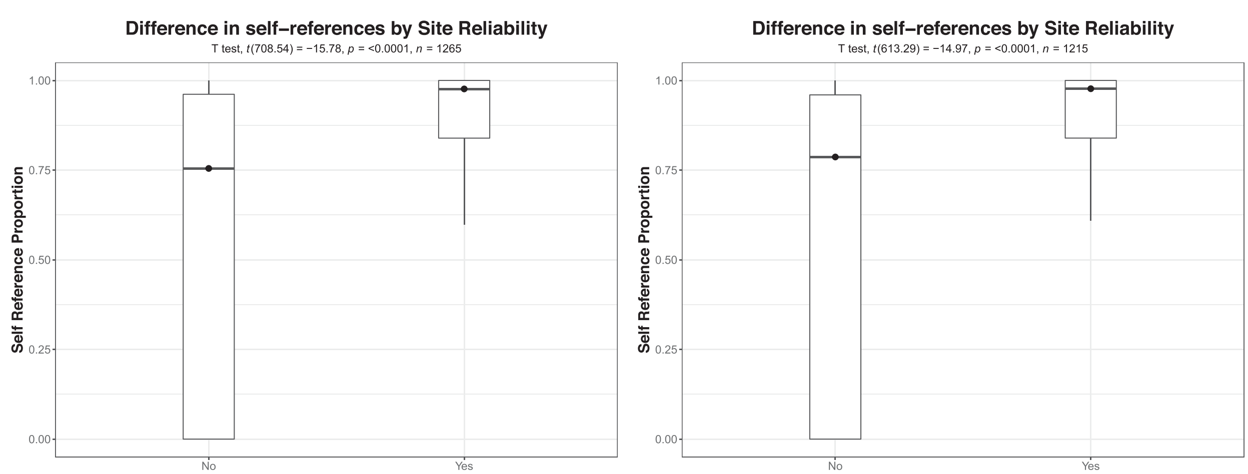


**Figure 2.** The proportion of incoming referrals that are self-referrals for unreliable and reliable domains. The subplot on the left uses all domains, while the subplot on the right uses only news sites that conduct original reporting.

removing self-references, the overall distributions are distinct. We also calculate the proportion of outgoing links that lead to reliable sites. In Figure 3 we see that reliable sites are much more likely to refer traffic to other reliable sites relative to unreliable sites. Again, we find that most referrals from reliable sites lead to other reliable sites. On the other hand, unreliable sites rarely refer traffic to reliable sites. The results of a t-test indicate that these differences are statistically significant ($t = -121.31$, $p<0.0001$, $n = 1265$). Our findings are consistent with previous work that finds that mainstream media sites are unlikely to link to conspiracy sites (Hanley, Kumar, and Durumeric 2022).

Next, we rerun the previous analysis, now removing referrals that are self-referrals (Figure 3(A,B), right). By excluding these observations, we get a measure of the likelihood of a site linking to another unreliable or reliable site. This allows us to see if sites are forming unreliable content rabbit holes by linking to other unreliable sites. As noted previously, most outbound referrals are self-referrals. We observe that while unreliable sites are more likely to link to other unreliable sites than reliable sites, the median levels are considerably lower when removing self-referrals. This suggests that on average, the formation of rabbit holes is likely to be based around unreliable sites retaining users on their site while limiting links to reliable media sites, rather than building a larger community of unreliable sites.

### 4.3. Analyzing embedded links in sites

Thus far we have relied on data that captures users moving within and between sites through referrals.

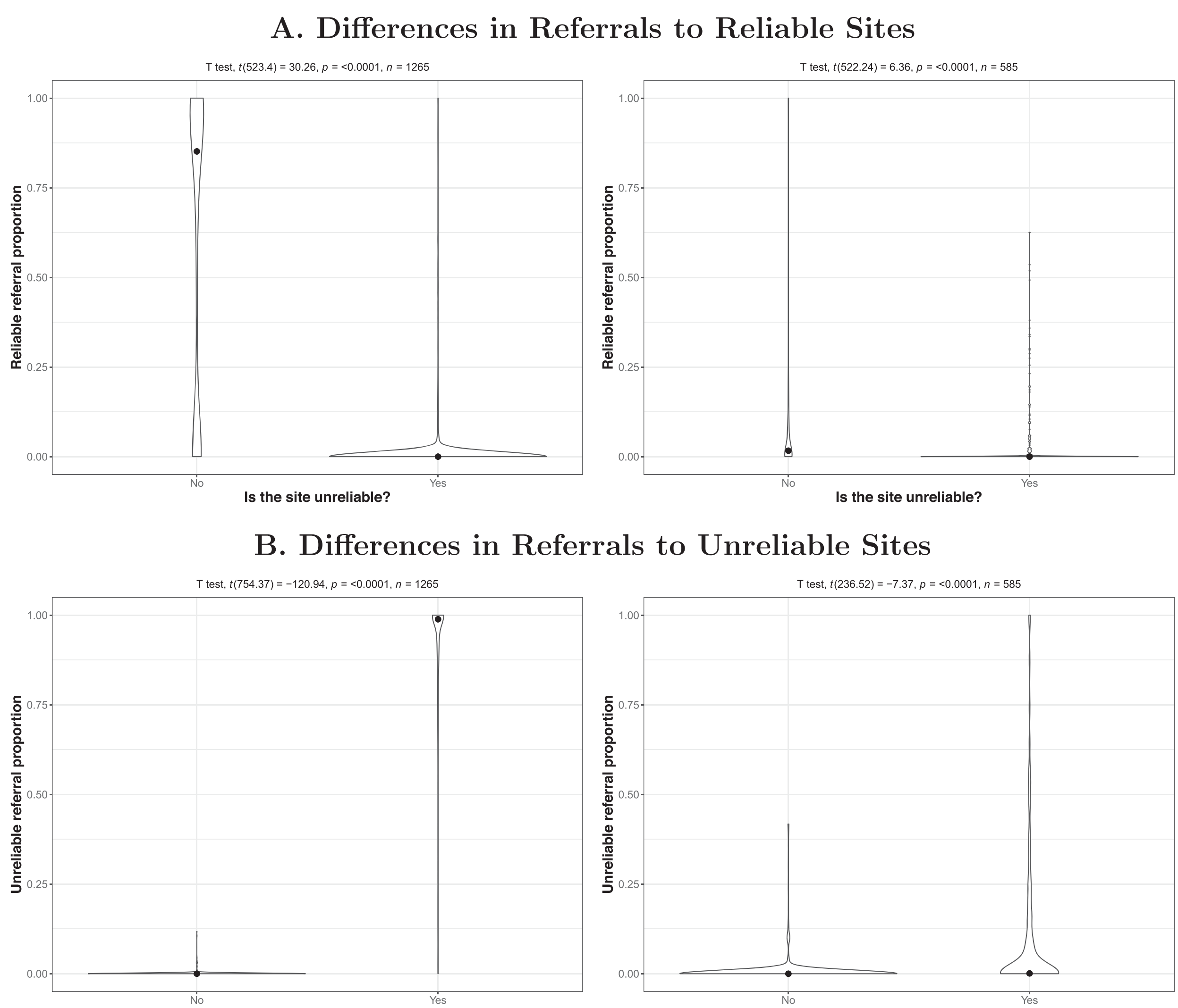


**Figure 3.** The proportion of outgoing referrals that lead to reliable sites (A) and unreliable sites (B). The subplots on the left include self-referrals, while the subplots on the right remove self-referrals.

While this measures the actualised network of connections between sites, it may not capture all the embedded links on a site. Unreliable sites might link to reliable media at the same rate as reliable sites, but users choose not to engage with these links. Given that unreliable sites are likely to be frequented by different types of users than more reputable sites (Guess, Nagler, and Tucker 2019) and recent findings that YouTube rabbit holes are partially explained by demand effects (Munger and Phillips 2022), addressing this limitation is an important one.

To this end, from our sample of NewsGuard sites, we take a random sample of 100 unreliable and 100 reliable sites and collect the links to each article on the site. For each site, we randomly select 25 articles and extract all the embedded URLs from the text of each article. Each URL is then checked against NewsGuard to identify if the domain is reliable or unreliable. As with our previous analyses, we calculate the proportion of links that lead to unreliable and reliable sources for each domain. Importantly our approach does not capture top and side panels which provide additional opportunities for a user to navigate to other popular articles on the site. Thus, we provide a conservative estimate of the potential avenues to arrive at unreliable content. We are also focussed on the links embedded in articles as these are links that are likely to direct users to additional content, rather than less visible links which may be added for search engine optimisation.

Our results using embedded links are highly similar to our previous results using traffic between sites (See Figure 4). First, we observe that reliable sites are likely to link to other reliable sites, while unreliable sites almost never do. The results of a t-test indicate that these differences are statistically significant ($t = -13.12$, $p<0.0001$, $n = 200$). Second, consistent with our previous results, we find that unreliable sites are very likely to provide outbound links to other unreliable sites, while this rarely happens for reliable sites. The results of a t-test indicate that these differences are statistically significant ($t = -15.22$, $p<0.0001$, $n = 200$). In sum, whether we investigate the navigational patterns between sites or the embedded URLs contained on sites, we observe a clear difference between reliable and unreliable sites. Once on an unreliable site, a user is likely to be presented with links to additional unreliable content and to be referred to this content.

### 4.4. Referrals to unreliable sites from reliable sites

As we have previously noted the likelihood of being referred from a reliable site to an unreliable site is extremely low. However, we do observe some reliable sites referring traffic to unreliable sites. These are worth exploring more closely for several reasons. First, reliable sites leading people to unreliable sites may be leading users to locations with widely different journalistic standards. Further, because users are being referred by reliable sites, this credibility may be conferred to unreliable sites. Second, different mechanisms might lead to referrals from reliable sites to unreliable sites. Popular sites such as *CNN* or the *The Washington Post* may be inadvertently leading users to unreliable sites by linking to these sites in fact-checks. Alternatively, reliable sites might be referring traffic to unreliable sites with whom they are ideologically aligned. Kaiser, Rauchfleisch, and Bourassa (2020) previously noted topical and hyperlink similarities between right-wing and far-right sites. Identifying the sites that are bridging the unreliable-reliable divide might help us to better understand how unreliable content makes its

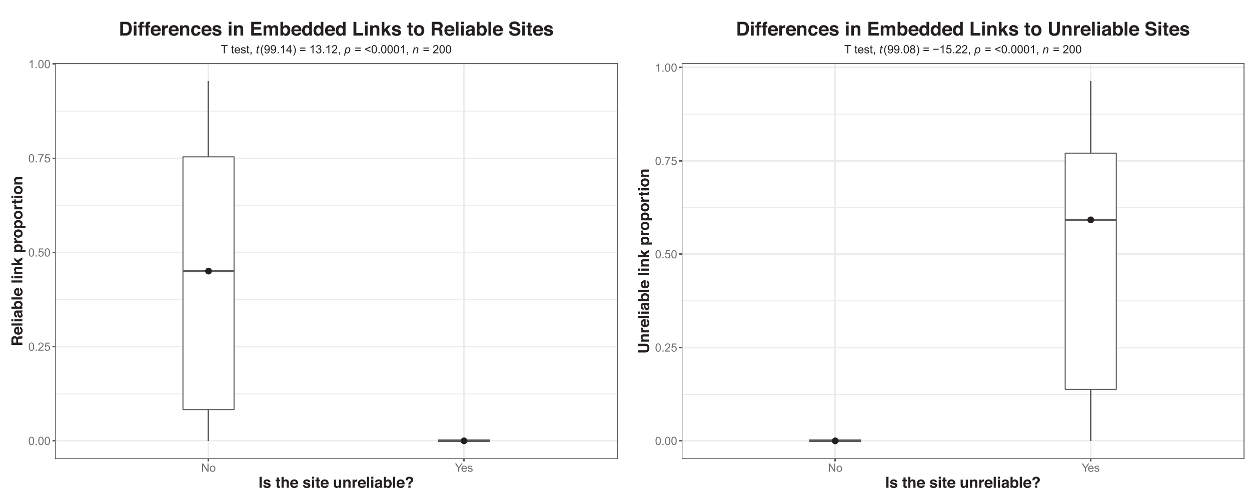


**Figure 4.** The proportion of embedded URLs that lead to reliable sites (left) and unreliable sites (right).

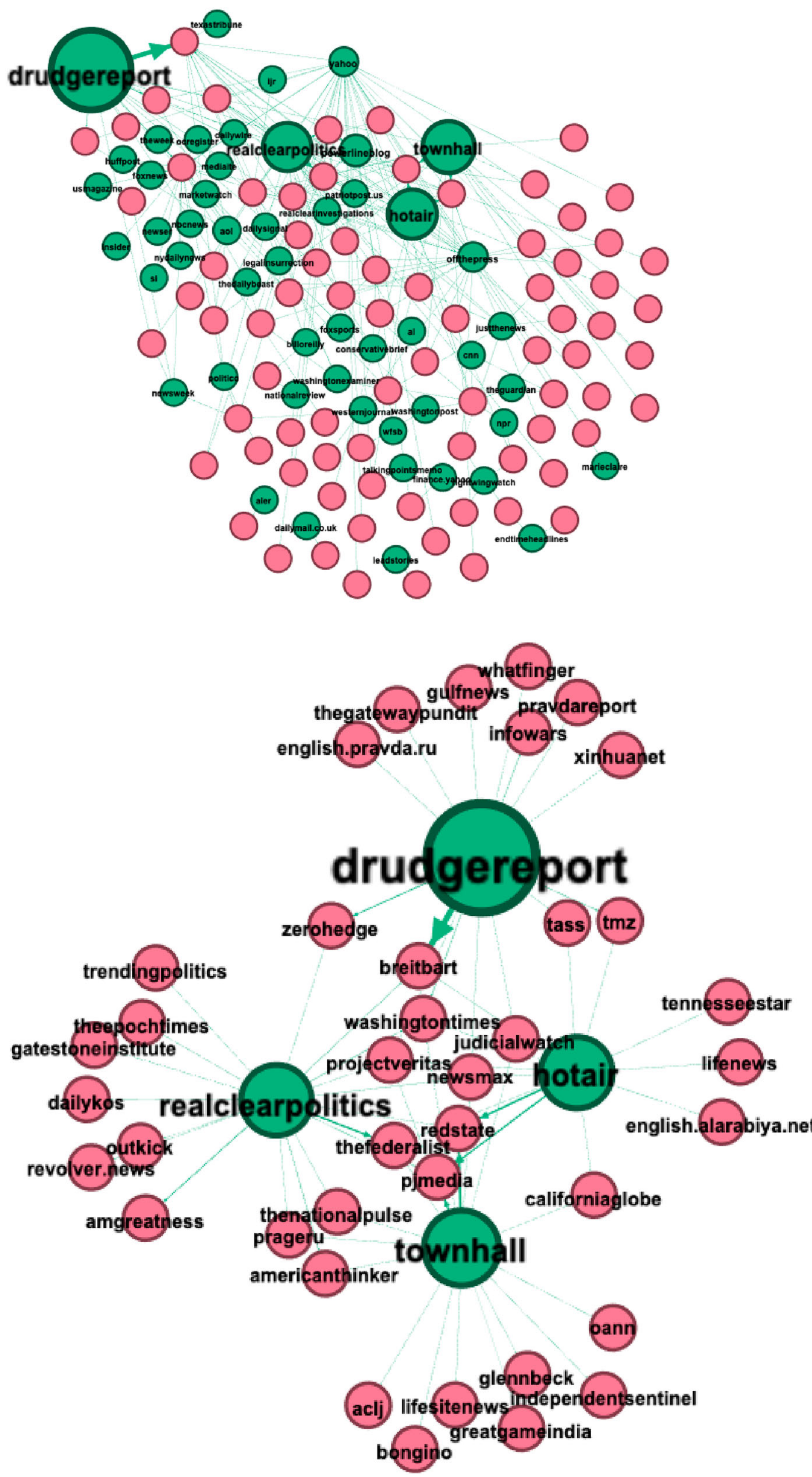


**Figure 5.** A directed graph presenting all the reliable sites referring traffic to unreliable sites. There are 51 in total (Top). The top four reliable sites referring to unreliable sites (Bottom). In total, these sites account for 90% of the referrals from reliable sites to unreliable sites. Green nodes represent reliable sites, while red nodes represent unreliable sites. The size of each node is based on its weighted out-degree.

way into the broader discourse and provide useful information for groups rating the information quality of web domains.

In Figure 5 we present each reliable site that referred traffic to unreliable sites. We find 51 reliable sites leading to over 200,000 referrals to unreliable sites. Our findings also reveal a clear pattern, the reliable sites that refer traffic to unreliable sites are overwhelmingly right-leaning. Roughly 90% of the referrals from reliable sites to unreliable sites come from four right-leaning sites, *the Drudge Report*, *Townhall*, *Hotair*, and *Real Clear Politics*. The unreliable sites being referred to are also overwhelming on the political right. Nine of the ten most referred sites are right-leaning and included outlets such as *Breitbart*, *Redstate*, *PJ Media*, and the *Federalist*.[7] Figure 5 displays all the reliable sites referring to unreliable sites, as well as all the unreliable sites being referred to by the top four referring sites.

## 5. Conclusion

Using browser telemetry data, we measure the traffic to and from unreliable and reliable sites to evaluate the potential for navigational rabbit holes. Our results uncover evidence consistent with unreliable content rabbit holes. First, we find that after landing on an unreliable site, successive referrals are highly likely to lead to additional unreliable sites. In part, this is due to unreliable sites providing links to other unreliable sites. However, most movements to unreliable sites come from internal site navigation, suggesting financial rather than political incentives. Our results indicate that unreliable sites are more sticky than reliable sites. In addition, unreliable sites provide few links and lead to little navigation to reliable sites. We find analogous results for reliable sites. Reliable sites often refer traffic to other reliable sites and seldom refer traffic to unreliable sites. The exceptions are a small number of right-wing sites that refer users to other ideologically similar unreliable sites.

## 6. Discussion

Our work builds on past efforts and suggests websites have additional ideological and financial incentives to direct users to specific locations (Davenport and Beck 2001; Kaiser, Rauchfleisch, and Bourassa 2020; Sehgal et al. 2021). We evaluate if these incentives in turn shape the types of links present on a site, and where users are referred, creating a potential rabbit hole of unreliable information. We find on average, unreliable sites present more links to additional pages on their site, helping to retain users and monetise their browsing activity. Further, we find that unreliable sites present users with links to additional unreliable sites, but very few links to reliable media sites. This suggests the need to consider not only the content created by unreliable sites but also how they retain users or push them to other dark parts of the internet. Given how sticky unreliable sites appear to be, an increased focus on the monetisation capabilities of unreliable sites also appears warranted.

Our findings could also be considered in relation to filter bubbles. Filter bubbles occur when users primarily engage with content that aligns with their preexisting beliefs (Areeb et al. 2023; Baeza-Yates 2018; Brown et al. 2022). When evaluating online platforms, the evidence for the existence (Eady et al. 2019; Guess 2021; Guess et al. 2018) and impact (Nyhan et al. 2023) of filter bubbles is mixed. However, when evaluating the evidence for filter bubbles past work has most often focussed on the political orientation of content. Our findings suggest that the quality of domains may be another dimension where users can enter filter bubbles. In particular, users who navigate to a low-quality site are likely to encounter many additional links to sites with similar standards and very limited exposure to domains with higher journalistic standards. Further, these potential filter bubbles can occur without the need for recommender systems, instead relying on the incentives of website operators to maintain traffic.

However, it is important to note several limitations of the study. First, much of the work on rabbit holes is about users being shown increasingly extreme content. As most news sites do not operate using similar recommender systems, our results cannot make any claims about sites radicalising individuals. Further, because of the nature of our data, we are not following individual users over time to track the evolution of their browsing behaviour. Second, there may be observable differences between unreliable and reliable sites that influence their referral patterns. In the Supplemental Files, we conduct additional analyses to account for the differences in total external referrals for unreliable and reliable sites. These results are consistent with those included in the main text. Third, our primary data measures referrals, rather than all the embedded links on a site. As the users that visit unreliable sites are likely different than those that visit reliable sites (Guess, Nyhan, and Reifler 2018; Guess, Nagler, and Tucker 2019; Osmundsen et al. 2021) we cannot rule out that our results are driven somewhat by user choice. However, our evaluation of all the embedded links in a sample of sites supports our primary findings. Further, others have observed that unreliable sites rarely provide outlinks to reliable media and vice-versa (Sehgal et al. 2021; Starbird et al. 2018). Fourth, while we do not find much evidence that users arrive at unreliable sites directly from reliable sources, reliable media may still play a role in spreading unreliable content (Watts and Rothschild 2017; Watts, Rothschild, and Mobius 2021). Fifth, one challenge with using web referral data is that domains may try to mask there traffic patterns using link redirection. However, for this to have an impact on our results we would need to assume that unreliable site use redirect preliminary when linking to reliable sites. However, we have no theoretical explanation for why unreliable sites would want to mask their track to reliable and not unreliable domains. Further, analyses of the URLs embedded on these domains, does not indicate considerable use of redirects. Finally, our analyses rely on lists of unreliable and reliable sites that are mostly viewed in the United States and are primarily in English and should be replicated in other contexts.

Ethical Considerations: As the study involves the use of user browsing data, we provide several safeguards to protect user privacy. First, all data were collected with user consent. Second, all personally identifying information was removed, so that the authors do not have access to individual users' data. Third, the user data was aggregated to the domain level. Thus the information available to the authors is traffic flows from one domain to another. This aggregation means that we are not able to observe the behaviour of a single user or group of users.

## Notes

1. Because past efforts have mostly ignored internal page navigation, this avenue for exposure has been largely overlooked.
2. Further, this does not mean that mainstream reliable media play no role in spreading unreliable information (Watts and Rothschild 2017; Watts, Rothschild, and Mobius 2021).
3. All personally identifiable information (PPI) was removed.
4. We use NewsGuard because it provides more detailed information on the type of reporting each site carries out than other raters, while being widely used in academic work (Aslett et al. 2022; Edelson et al. 2021; Guess et al. 2021) and very highly correlated with other rating agencies (Lin et al. 2023).
5. To ensure privacy and confidentiality, the data was anonymized and aggregated at the point of collection, thereby removing any personal or identifying information.
6. This is consistent with Chen et al. (2022), who find that internal navigation makes up much of the traffic to *The Gateway Pundit*.
7. Links to *PJ Media* and *Redstate* from *Townhall*, and *Hotair* are not surprising as they are all members of the Townhall Media group.


## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This project received generous support from Microsoft


## References


Acerbi, Alberto 2019. *Cultural Evolution in the Digital Age*. Oxford: Oxford University Press.

Areeb, Qazi Mohammad, Mohammad Nadeem, Shahab Saquib Sohail, Raza Imam, Faiyaz Doctor, Yassine Himeur, Amir Hussain, and Abbes Amira. 2023. "Filter Bubbles in Recommender Systems: Fact Or Fallacy–A Systematic Review." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*13 (6): e1512.

Aslett, Kevin, Andrew M. Guess, Richard Bonneau, Jonathan Nagler, and Joshua A. Tucker. 2022. "News Credibility Labels Have Limited Average Effects on News Diet Quality and Fail to Reduce Misperceptions." *Science Advances* 8 (18): eabl3844. https://doi.org/10.1126/sciadv.abl3844.

Baeza-Yates, Ricardo 2018. "Bias on the Web." *Communications of the ACM* 61 (6): 54–61. https://doi.org/10.1145/3209581.

Brown, Megan A., Jonathan Nagler, James Bisbee, Angela Lai, and Joshua A. Tucker. 2022. "Echo Chambers, Rabbit Holes, and Ideological Bias: How YouTube Recommends Content to Real Users."

Chen, Zhouhan, Haohan Chen, Juliana Freire, Jonathan Nagler, and Joshua A. Tucker. 2022. "Understanding How People Consume Low Quality and Extreme News Using Web Traffic Data." *Preprint*. arXiv:2201.04226.

Chen, Annie Y., Brendan Nyhan, Jason Reifler, Ronald E. Robertson, and Christo Wilson. 2022. "Subscriptions and External Links Help Drive Resentful Users to Alternative and Extremist Youtube Videos." *Preprint*. arXiv:2204.10921.

Davenport, Thomas H., and John C. Beck. 2001. "The Attention Economy." *Ubiquity* 2001 (May): 1–es. https://doi.org/10.1145/375348.376626.

Eady, Gregory, Jonathan Nagler, Andy Guess, Jan Zilinsky, and Joshua A. Tucker. 2019. "How Many People Live in Political Bubbles on Social Media? Evidence From Linked Survey and Twitter Data." *SAGE Open* 9 (1): 2158244019832705. https://doi.org/10.1177/2158244019832705.

Edelson, Laura, Minh-Kha Nguyen, Ian Goldstein, Oana Goga, Damon McCoy, and Tobias Lauinger. 2021. "Understanding Engagement with us (MIS) Information News Sources on Facebook." In *Proceedings of the 21st ACM Internet Measurement Conference*, 444–463.

GDI. 2019. "The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech?" *Global Disinformation Index*.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. "Fake News on Twitter During the 2016 US Presidential Election." *Science* 363 (6425): 374–378. https://doi.org/10.1126/science.aau2706.

Guess, Andrew M. 2021. "(Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets." *American Journal of Political Science* 65 (4): 1007–1022. https://doi.org/10.1111/ajps.v65.4.

Guess, Andy, Kevin Aslett, Joshua Tucker, Richard Bonneau, and Jonathan Nagler. 2021. "Cracking Open the News Feed: Exploring what Us Facebook Users See and Share with Large-Scale Platform Data." *Journal of Quantitative Description: Digital Media* 1: 1–48.

Guess, Andrew, Jonathan Nagler, and Joshua Tucker. 2019. "Less Than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook." *Science Advances* 5 (1): eaau4586. https://doi.org/10.1126/sciadv.aau4586.

Guess, Andrew, Brendan Nyhan, Benjamin Lyons, and Jason Reifler. 2018. "Avoiding the Echo Chamber About Echo Chambers." *Knight Foundation* 2 (1): 1–25.

Guess, Andrew, Brendan Nyhan, and Jason Reifler. 2018. "Selective Exposure to Misinformation: Evidence From the Consumption of Fake News During the 2016 US Presidential Campaign." *European Research Council* 9 (3): 4.

Hanley, Hans W. A., Deepak Kumar, and Zakir Durumeric. 2022. "No Calm in The Storm: Investigating QAnon Website Relationships." In *Proceedings of the International AAAI Conference on Web and Social Media,* Atlanta, Georgia. Vol. 16, 299–310.

Herasimenka, Aliaksandr, Yung Au, Anna George, Kate Joynes-Burgess, Aleksi Knuutila, Jonathan Bright, and Philip N. Howard. 2022. "The Political Economy of Digital Profiteering: Communication Resource Mobilization by Anti-Vaccination Actors." *Journal of Communication* 73 (2): 126–137.

Hosseinmardi, Homa, Amir Ghasemian, Aaron Clauset, Markus Mobius, David M. Rothschild, and Duncan J. Watts. 2021. "Examining the Consumption of Radical Content on YouTube." *Proceedings of the National Academy of Sciences* 118 (32): e2101967118. https://doi.org/10.1073/pnas.2101967118.

Hussein, Eslam, Prerna Juneja, and Tanushree Mitra. 2020. "Measuring Misinformation in Video Search Platforms: An Audit Study On YouTube." *Proceedings of the ACM on Human-Computer Interaction*4 (CSCW1): 1–27.

Kaiser, Jonas, Adrian Rauchfleisch, and Nikki Bourassa. 2020. "Connecting the (far-) Right Dots: A Topic Modeling and Hyperlink Analysis of (far-) Right Media Coverage During the US Elections 2016." *Digital Journalism* 8 (3): 422–441. https://doi.org/10.1080/21670811.2019.1682629.

Lazer, David M. J., Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, et al. 2018. "The Science of Fake News." *Science* 359 (6380): 1094–1096. https://doi.org/10.1126/science.aao2998.

Ledwich, Mark, and Anna Zaitsev. 2019. "Algorithmic Extremism: Examining YouTube's Rabbit Hole of Radicalization." *Preprint*. arXiv:1912.11211.

Lin, Hause, Jana Lasser, Stephan Lewandowsky, Rocky Cole, Andrew Gully, David Rand, and Gordon Pennycook. 2023. "High level of correspondence across different news domain quality rating sets." *PNAS Nexus* 2 (9): 1–8.

Munger, Kevin, and Joseph Phillips. 2022. "Right-Wing YouTube: A Supply and Demand Perspective." *The International Journal of Press/Politics* 27 (1): 186–219. https://doi.org/10.1177/1940161220964767.

NewsGuard. 2021. "Fighting Misinformation with Journalism, Not Algorithms." https://www.newsguardtech.com/wp-content/uploads/2021/12/NewsGuard-Whitepaper-on-Prebunking-With-Source-Labels-Dec-2021.pdf.

Nyhan, Brendan, Jaime Settle, Emily Thorson, Magdalena Wojcieszak, Pablo Barberá, Annie Y. Chen, Hunt Allcott, et al. 2023. "Like-Minded Sources on Facebook are Prevalent But Not Polarizing." *Nature* 620 (7972): 137–144. https://doi.org/10.1038/s41586-023-06297-w.

O'Callaghan, Derek, Derek Greene, Maura Conway, Joe Carthy, and Pádraig Cunningham. 2015. "Down the (white) Rabbit Hole: The Extreme Right and Online Recommender Systems." *Social Science Computer Review* 33 (4): 459–478. https://doi.org/10.1177/0894439314555329.

Osmundsen, Mathias, Alexander Bor, Peter Bjerregaard Vahlstrup, Anja Bechmann, and Michael Bang Petersen. 2021. "Partisan Polarization is the Primary Psychological Motivation Behind Political Fake News Sharing on Twitter." *American Political Science Review* 115 (3): 999–1015. https://doi.org/10.1017/S0003055421000290.

Papadamou, Kostantinos, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Sirivianos. 2022. "'It is Just A Flu': Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations." In *Proceedings of the International AAAI Conference on Web and Social Media,* Atlanta, Georgia. Vol. 16, 723–734.

Ribeiro, Manoel Horta, Raphael Ottoni, Robert West, Virgílio A. F. Almeida, and Wagner Meira Jr. 2020. "Auditing Radicalization Pathways on YouTube." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, Barcelona Spain, 131–141.

Ryan, Camille D., Andrew J. Schaul, Ryan Butner, and John T. Swarthout. 2020. "Monetizing Disinformation in the Attention Economy: The Case of Genetically Modified Organisms (GMOs)." *European Management Journal* 38 (1): 7–18. https://doi.org/10.1016/j.emj.2019.11.002.

Santos, Fernando P., Yphtach Lelkes, and Simon A. Levin. 2021. "Link Recommendation Algorithms and Dynamics of Polarization in Online Social Networks." *Proceedings of the National Academy of Sciences*118 (50): e2102141118. https://doi.org/10.1073/pnas.2102141118.

Sehgal, Vibhor, Ankit Peshin, Sadia Afroz, and Hany Farid. 2021. "Mutual Hyperlinking Among Misinformation Peddlers." *Preprint*. arXiv:2104.11694.

Starbird, Kate, Ahmer Arif, Tom Wilson, Katherine Van Koevering, Katya Yefimova, and Daniel Scarnecchia. 2018. "Ecosystem or Echo-System? Exploring Content Sharing Across Alternative Media Domains." In *Proceedings of the International AAAI Conference on Web and Social Media,* Stanford, California. Vol. 12.

StatCounter. 2021. "DesktopConnecting the Browser Market Share United States Of America." https://gs.statcounter.com/browser-market-share/desktop/united-states-of-america/2021.

Tarantola, A. 2021. "How Social Media Recommendation Algorithms Help Spread Hate." *Engadget*. https://www.engadget.com/how-social-media-recommendation-algorithms-help-spread-online-hate-180032029.html.

Watts, Duncan J., and David M. Rothschild. 2017. "Don't Blame the Election on Fake News. Blame it on the Media." *Columbia Journalism Review* 5:67–84.

Watts, Duncan J., David M. Rothschild, and Markus Mobius. 2021. "Measuring the News and Its Impact on Democracy." *Proceedings of the National Academy of Sciences* 118 (15): e1912443118. https://doi.org/10.1073/pnas.1912443118.