Exposure to Untrustworthy Websites in the 2020 US Election

Authors:

Ryan C. Moore*
(Corresponding Author)
PhD Candidate
Department of Communication
Stanford University
rymoore@stanford.edu

Ross Dahlke
PhD Student
Department of Communication
Stanford University
rdahlke@stanford.edu

Jeffrey T. Hancock
Harry and Norman Chandler Professor of Communication
Department of Communication
Stanford University
hancockj@stanford.edu

Abstract

Prior research has documented exposure to fake news and online misinformation using large-scale data on individuals' media use, which has provided important information about the scope and nature of people's exposure to misinformation online. However, most of this work has made use of data collected during the 2016 US election, and far fewer studies have examined how exposure to misinformation online has changed since 2016. In this paper, we examine exposure to untrustworthy websites in the lead up to the 2020 US election using a dataset of over 7.5 million passively tracked website visits from a nationally representative sample of American adults (N = 1,151). We find that a significantly smaller percentage of Americans were exposed to untrustworthy websites in 2020 compared to in 2016 (as calculated by Guess et al. [2020]). While exposure was concentrated among similar groups of people as it was in 2016, levels of exposure appear to be lower across the board. There were also differences in the role online platforms played in directing people to untrustworthy websites in 2020 compared to 2016. Our findings have implications for future research and practice around online misinformation.

Keywords: misinformation, 2020 election, web browsing, trace data

Exposure to Untrustworthy Websites in the 2020 US Election

Both concern over and research on misinformation has exploded in recent years (Weeks & Gil de Zúñiga, 2019). One focus of recent work has been estimating exposure to misinformation online, in part under the assumption that understanding the nature of exposure to misinformation will help guide the direction of future research and inform efforts to prevent misinformation's spread. For example, knowing which groups of people are more likely to be exposed to misinformation could direct research toward understanding the mechanisms underlying that group's susceptibility and suggest interventions to build resilience to misinformation among that group.

Some high-level findings from the extant work on exposure to online misinformation include the following. First, relatively few people are exposed to online misinformation, or at least fewer than many initially expected (Allcott & Gentzkow, 2017; Allen et al., 2020; Grinberg et al., 2019; Guess et al., 2019, 2020; Weeks et al., 2021). Second, exposure is highly concentrated (Grinberg et al., 2019; Guess et al., 2019, 2020; Nelson & Taneja, 2018). For example, Grinberg et al. (2019) found that 1% of Twitter users were exposed to 80% of the fake news on Twitter during the 2016 election. Third, some groups are significantly more likely to be exposed to misinformation online than other groups. For example, during the 2016 election, people 65 and older were twice as likely to be exposed to fake news on Twitter and seven times more likely to share fake news on Facebook than 18-29 year olds (Grinberg et al., 2019; Guess et al., 2019).

One feature of many of the studies investigating exposure to online misinformation is that they examine exposure during the 2016 US election. This focus may be due to the fact that concern over fake news rose during the 2016 election (Li & Su, 2020). However, an important limitation of previous research's focus on the 2016 election period is that we have little insight into how exposure to misinformation online has changed since 2016, noteworthy given the

myriad ways in which which the digital (mis)information ecosystem has changed since then (e.g., new platforms, new misinformation-generating world events, new modalities to disseminate misinformation; Ball & Maxmen, 2020; Lytvynenko, 2020).

Some work has examined changes in exposure over time. Guess, Lyons, et al. (2019) found a decline in the number of Americans exposed to fake news websites from 2016 to 2018. By contrast, Allen et al. (2020) found that exposure to fake news was “generally stable” from 2016 through the end of 2018. Discrepancies between these two findings may be in part explained by the authors using different lists to identify visits to fake news websites in their data. Allcott et al. (2019) also examined changes in exposure over time, finding that engagement with fake news on Facebook declined from 2016 through 2018 while engagement with fake news on Twitter rose over the same period.

Given changes in world events, the digital media landscape, and a significant societal effort to combat misinformation, an important question remains as to how misinformation exposure during the 2020 US presidential election compared to during the 2016 election, which served as the context for much of the existing research. In this project, we investigate exposure to online misinformation during the 2020 election and do so in a way that allows us to directly compare exposure in 2020 to exposure during the 2016 election. To do so, we collected web browsing data (i.e., URLs visited) from a nationally representative sample of American adults (N = 1,151) during the lead up to the 2020 election (total number of website visits = 7.5 million). To produce estimates of misinformation exposure directly comparable to those from 2016, we adopted the analytical approach of Guess et al. (2020), who studied exposure to online misinformation during the lead up to the 2016 election.

What might we expect about exposure to online misinformation in 2020? On the one hand, there are reasons to believe that exposure in 2020 was greater than exposure in 2016. 2020 was marked by a unique confluence of major national and global events that generated a considerable amount of misinformation to which people could have been exposed (e.g., the rise

of the COVID-19 pandemic, claims that the 2020 presidential election was fraudulent; Lytvynenko, 2020). Additionally, evidence suggests that the infrastructure available to disinformation campaigns and purveyors for disseminating and spreading false information across platforms has improved considerably in recent years (Alaphilippe, 2021; Wilson & Starbird, 2021). On the other hand, there are also reasons to believe that exposure in 2020 was lesser than that in 2016. Since the 2016 election, journalists and online platforms have taken a number of steps to curb the spread of misinformation. For example, social media platforms such as Facebook and Twitter have begun labeling and flagging untrustworthy content on their platforms (Mosseri, 2017; Roth & Achuthan, 2020) and journalists have intensified their efforts around accountability and transparency (Vu & Saldaña, 2021). Furthermore, tools and educational resources for helping the public identify online misinformation that have proliferated since 2016 may have improved internet users' ability to avoid misinformation (Fazio, 2020; Kozyreva et al., 2020).

Understanding exposure to online misinformation in 2020 is important for several reasons. First, it is relevant to the design of interventions to build resilience to misinformation. Understanding who is exposed to online misinformation -- and when and how they are exposed -- may help researchers identify how to make the content of interventions best suited for different types of users. Second, understanding the nature and level of misinformation exposure can inform policymaking efforts to mitigate misinformation (e.g., requiring social media platforms to implement certain flagging or fact-checking practices). Third, updates about misinformation exposure can inform public discourse around misinformation itself. Since 2016, misinformation has been a significant topic in communication from journalists, elites, scientists and among the general public (Li & Su, 2020), with implications for how the public trusts information sources. Indeed, greater amounts of elite discourse about misinformation has been shown to reduce people's trust in traditional news media (Van Duyn & Collier, 2019). If it is true that the vast majority of people's news diet is composed of reliable news (Allen et al., 2020), heightened

discourse around misinformation could cause people to be slightly more attuned to the presence of a small amount of information in their media diets while eroding trust in the vast majority.

Our analysis reveals that exposure to untrustworthy websites during the 2020 election was significantly lower than Guess et al. (2020) determined it to be during the 2016 election. Not only was the proportion of total individuals exposed to untrustworthy websites lower, but among the exposed, the average number of exposures was also lower. While we also found that the groups who were more likely to be exposed to untrustworthy websites in 2016 -- older adults, conservatives -- were still more likely in 2020, overall levels of exposure for these groups were lower than in 2016. We found these patterns despite using a database of untrustworthy websites over three times the size of that used in Guess et al. (2020), meaning that even with a considerably larger capacity for identifying misinformation in people's web browsing data, we still found less. Finally, we found changes in the ways people come to be exposed to untrustworthy websites, namely that people were more likely to arrive at them via Google and less likely to arrive via Facebook in 2020 as compared to 2016.

Review of Guess et al. (2020)

Guess et al. (2020) is a widely-cited study examining exposure to online misinformation during the 2016 US election. The study collected data from a national sample of American adults (N = 2,525) during the lead up to the 2016 election (October 7 - November 14, 2016; Election Day was November 8, 2016) using survey purveyor YouGov. Specifically, they collected two types of data from participants: (1) web browsing data collected via YouGov's Pulse browser plugin, which is installed by survey respondents on their web browsers and collects all URLs visited as they surf the internet and (2) demographic information collected via surveys. They then combined these data with a list of 490 web domains that "frequently publish

factually dubious or untrustworthy content" (Guess et al., 2020, p. 479) in order to identify visits to untrustworthy websites in their participants' web browsing data.[1]

Similar to most studies on misinformation exposure (Allen et al., 2020), Guess et al. take a domain-level approach to identifying misinformation. This approach means that rather than identifying specific articles which contain misinformation (e.g., www.obamawatcher.com/2020/03/michelles-fake-degrees), Guess et al. (2020) rely on a list of domains compiled by previous studies that are known to frequently publish untrustworthy content (e.g., www.obamawatcher.com). These untrustworthy websites represent their operationalization of online misinformation.

By matching their list of untrustworthy websites against their participants' web browsing data, Guess et al. found that a little less than half (44.3%) of their nationally representative sample were exposed to at least one untrustworthy website in the lead up to the 2016 election. They also found that supporters of Donald Trump, those with more conservative media diets, and those over the age of 65 were more likely to visit untrustworthy websites. Finally, they found that exposures to untrustworthy websites frequently occurred via Facebook.

Our primary goal was to adopt the analytic approach of Guess et al. (2020) to analyze data collected via the same source (YouGov Pulse) during the same period (4 weeks prior to Election Day and 1 week after Election Day) around the 2020 election in order to compare exposure to untrustworthy websites during the 2020 election to exposure during the 2016 election. One advantage of adopting Guess et al.'s analytical approach is that we can make relatively direct comparisons between our 2020 estimates and their 2016 estimates. For the sake of simplicity, throughout the rest of this paper we refer to findings from Guess et al. (2020) as representing the "2016 election" and our findings as representing the "2020 election."

[1] Guess et al. (2020) compiled this list of 490 sites from databases of untrustworthy websites created by Allcott & Gentzkow (2017) and Grinberg et al. (2019).

Methods

Participants

To measure exposure to untrustworthy websites during the 2020 election, we passively gathered web browsing data (across smartphones, laptops, and desktop computers) from 1,151 Americans using YouGov's Pulse browser plugin from October 2, 2020 to November 9, 2020 (Election Day was November 3, 2020). These participants also completed an online survey which complemented their web browsing data. All participants consented to the terms of the research and were compensated by YouGov for their participation. Of these participants, 58% (n = 670) supported Joe Biden in the 2020 election while 36% (n = 419) supported Donald Trump. Twenty-nine percent (n = 329) were aged 65+, 48% (n = 554) were 45-64, 15% (n = 171) were 30-44, and 8% (n = 97) were under 30. Forty-five percent (n = 522) reported identifying as male and 54% (n = 624) reported identifying as female. Eighty-nine percent (n = 1,029) said they follow politics most or some of the time and 30% (n = 348) were considered highly knowledgeable about politics according to Pew Research Center's (2018) civic knowledge questionnaire. In total, these participants 7.5 million websites over the course of our data collection period. YouGov weighted participants to match a nationally-representative sample and we use these weights in all subsequent calculations and results.[2]

Measures

In order to categorize whether a website that a participant visited is untrustworthy, we compiled a list of untrustworthy domains in two phases. First, we began with the list of untrustworthy domains used by Guess et al. (2020), which consisted of 490 unique untrustworthy websites collected by previous research. We supplemented this list of 490 websites with an additional 66 unique untrustworthy domains collected by Allcott et al. (2019).

[2] Anonymized survey data along with summary web traffic data used for the analyses in the paper are available at: https://osf.io/8fy2z/?view_only=0d63bc3fd4f24938867a04efa6084552. Full web traffic histories are not available in order to protect subject confidentiality.

Then, we augmented this list of 556 websites compiled by academic researchers with websites from NewsGuard, an organization of former journalists and news editors who manually rate the information quality of websites. We added the 1,240 domains that NewsGuard rated as "repeatedly publishing false content" to the Guess et al. (2020) + Allcott et al. (2019) list of untrustworthy websites, producing a total list of 1,796 unique untrustworthy domains.

Next, we match this list of untrustworthy domains to the URL-level web browsing data from our YouGov participants to identify visits to untrustworthy websites in our sample's web browsing behavior. Visits to hard news websites (e.g., www.nytimes.com, www.economist.com) also played a role in some of Guess et al. (2020)'s analyses. To identify visits to hard news websites in our sample's browsing, we use a database of 500 hard news sites compiled by Bakshy et al. (2015), the same method used by Guess et al. (2020).[3]

Results

Total Consumption of Untrustworthy Websites

During the 2016 election, Guess et al. (2020) found that 44.3% (95% CI = 40.8-47.7%) of Americans aged 18 or older were exposed to at least one untrustworthy website. By contrast, we find that during the 2020 election only 26.2% (95% CI = 22.5-29.8%) of Americans were exposed to at least one untrustworthy website. This decrease represents a significant reduction ($t = -11.12$, $p < .001$) in the percentage of Americans exposed to untrustworthy websites from 2016 to 2020.[4]

---

[3] Following Guess et al. (2020), we excluded Amazon, Twitter, and YouTube from Bakshy et al. (2015)'s list of hard news websites.

[4] The NewsGuard database allows us to conduct an additional comparison between the percentage of the population exposed to untrustworthy websites in 2020 and in 2016. In addition to rating sites on a binary criteria of whether or not they repeatedly publish false content, NewsGuard also provides a continuous "trust score" for each website they evaluate, ranging from 0 to 100 with 0 being the lowest credibility and 100 being the highest. These scores are based on a weighted sum of 9 different criteria of journalistic integrity (one of which being whether or not a site repeatedly publishes false content). Examples of sites with a score of 0 include fawkes-news.com and dmaga2020.xyz, sites with a score of 25 include oann.com and infowars.com, breitbart.com is an example of a site with a score of 50, sites with a score of 75 include nypost.com and thesun.co.uk, and sites with a score of 100 include theguardian.com and theatlantic.com. We can determine what percentage of our sample was exposed to at least one website below a NewsGuard score threshold, starting at 0 and moving all the way up to 100. In order to achieve the same 44.3% percentage of individuals exposed to at least one untrustworthy website that Guess et al. (2020) found for 2016, we would need to apply a NewsGuard score cutoff of 70 to our 2020 data, meaning that 44.3% of our sample was exposed to at least one site with a NewsGuard score lower than 70. Examples of sites with a NewsGuard score of 70 include msnbc.com and foxnews.com, which people may argue have partisan leanings but they are not repeated purveyors of false information and few would likely argue that

While a smaller percentage of Americans appeared to be exposed to any untrustworthy websites in 2020 compared to 2016, one possibility is that those who were exposed to untrustworthy websites in 2020 were exposed at greater rates than in 2016. However, our data suggest otherwise. First, the average number of untrustworthy website visits among those who visited at least one untrustworthy website in 2020 (22.8 visits; 95% CI = 15.2-30.4) was significantly lower than the average number in 2016 (32.1 visits; 95% CI = 26.3-37.9) ($t = -2.06$, $p = .04$). Second, the average length of time spent on untrustworthy websites was shorter in 2020 (38.6 seconds; 95% CI = 34.5-42.7s) than in 2016 (64.2 seconds[5]).

Comparing visits to untrustworthy websites in 2020 and in 2016, we find that overall consumption of untrustworthy sites was lower in 2020 than in 2016. Fewer Americans were exposed to any untrustworthy websites in the leadup to the 2020 election and those who were exposed visited fewer and were less engaged with the untrustworthy sites they visited. Next, we examine how the types of people more likely to visit untrustworthy websites in 2016 compare to the types of people who were more likely to visit them in 2020.

Who is Consuming Untrustworthy Websites?

In 2016, Guess et al. (2020) found that supporters of presidential candidate Donald Trump were significantly more likely to visit untrustworthy websites than supporters of his opponent Hillary Clinton, with 59.5% (95% CI = 54.5-64.5%) of Trump supporters visiting at least one untrustworthy website and 37.1% (95% CI = 31.9-42.2%) of Clinton supporters visiting at least one untrustworthy website. We find a similar difference between Donald Trump and Joe Biden supporters in 2020. 36.2% (95% CI = 29.7-42.6%) of Trump supporters visited at least one untrustworthy website while 17.8% (95% CI = 13.7-22.0%) of Biden supporters visited at least one untrustworthy website. However, for both Trump and Clinton/Biden supporters, the

they represent untrustworthy websites. In other words, to achieve levels of exposure in 2020 similar to those found in 2016, we would have to apply a criteria for untrustworthy websites that is unreasonable and includes several legitimate news outlets.

[5] Guess et al. (2020) did not calculate a CI for this quantity.

percentages of supporters exposed in 2020 were significantly lower than the percentages of supporters exposed in 2016 ($t_{Trump} = -8.20$, $p < .001$; $t_{Clinton/Biden} = -9.53$, $p < .001$).

Guess et al. (2020) also examined the association between political leaning and consumption of untrustworthy websites by classifying users based on the ideological slant of their news diet and examining exposure to untrustworthy websites among those with more liberal or conservative news diets. Once they estimated the average ideological position of participants' news diets (using estimates of the ideological leaning of hard news websites produced by Bakshy et al. [2015]), they split their participants into news slant deciles ranging from most liberal to most conservative. We implemented this same classification strategy on our own sample to examine how exposure to untrustworthy websites across people with news diets with different ideological slants in 2020 compares to 2016.

This comparison is presented in Figure 1. In 2016, there was a skewed V-shape in the relationship between media slant decile and the percentage of people exposed to untrustworthy websites. In 2016, roughly half of the people in the most liberal media slant deciles were exposed to untrustworthy websites. That percentage decreases moving toward more moderate media diets, and then it jumps back up among the most conservative diets, with 85.2% in the most conservative media diet decline being exposed to untrustworthy websites. In 2020, the pattern resembles something closer to a left-skewed distribution, where the percentage of people exposed to untrustworthy websites is relatively constant moving from the most liberal to more moderate media diets but then jumps dramatically among the most conservative diets. Due to these different patterns, the percentage of people exposed to untrustworthy websites in more liberal media diets in 2020 was considerably lower than the percentage in those diets in 2016. Among the most conservative media diets, the percentage of those exposed to untrustworthy sites was similar across both years.

[FIGURE 1 HERE]

Age also played an important role in exposure to untrustworthy websites during the 2016 election, with older individuals generally being more likely to visit untrustworthy sites. In 2016, those 65 and older were 1.6 times more likely to visit untrustworthy websites than adults under 30. During the 2020 election, it appears that the positive association between age and likelihood of exposure to untrustworthy websites persisted, with those 65 and older being 2.1 times more likely to visit untrustworthy websites than those under 30. However, all age groups in 2020 were less likely to be exposed to untrustworthy websites than those same age groups in 2016 (see Figure 2). For instance, while 56.2% (95% CI = 49.7-62.7%) of people 65 and older were exposed to untrustworthy websites in 2016, only 37.4% (95% CI = 29.6-45.2%) of people 65 and older were exposed in 2020.

[FIGURE 2 HERE]

To examine the association between a variety of individual differences and exposure to untrustworthy websites, we follow the strategy used by Guess et al. (2020) and estimate linear probability models for the 2016 and 2020 election periods. In these models, the dependent variable is a binary variable indicating whether an individual was (1) or was not (0) exposed to at least one untrustworthy website during the data collection period. The independent variables include presidential candidate support, political knowledge, political interest, level of education, gender, race, and age. The models for 2016 and 2020 are presented in Table 1. Across both 2016 and 2020, many of the same demographic variables were significantly associated with exposure to untrustworthy websites. Specifically, being a Trump supporter, possessing greater political knowledge, being more interested in politics, and being 65 years of age or older were associated with a significantly greater likelihood of being exposed to at least one untrustworthy website in both 2016 and 2020 (all p's < .05), while identifying as non-white was associated with

a significantly lower likelihood of being exposed to at least one untrustworthy website in both years ($p_{2016} < .001$; $p_{2020} < .05$).[6]

[TABLE 1 HERE]

Taken together, these findings suggest that the groups who were more likely to be exposed to untrustworthy websites in 2016 were largely the same groups more likely to be exposed in 2020. Critically, however, the likelihood of exposure in 2020 for these groups still appears to be lower than their rates in 2016. That is, while trends in who was exposed to misinformation in 2016 seem to persist in 2020, the amount to which all groups were exposed was lower across the board.

How are People Exposed to Untrustworthy Websites?

Guess et al. (2020) examined how individuals came to visit untrustworthy websites during the 2016 election by analyzing "referrers" to untrustworthy websites, or the sites that were within 30 seconds and one of the three previous sites before an untrustworthy website exposure. We repeated this analysis of referrers to untrustworthy websites for the 2020 election. The comparison in referrers between the 2016 and 2020 elections is presented in Figure 3.

In 2016, Facebook was a significant referrer to untrustworthy websites, with 15.1% of visits to untrustworthy websites being referred by Facebook. In 2020, however, only 6.6% of visits to untrustworthy sites were referred by Facebook, a significant reduction suggesting a potential decline in Facebook's role in directing people to untrustworthy websites ($\chi^2 = 457.5$, $p < .001$). Conversely, the share of untrustworthy websites referred by Google rose from 2016 to 2020, with 3.3% of visits to untrustworthy sites referred by Google in 2016 and 7.2% of visits referred by Google in 2020 ($\chi^2 = 304.6$, $p < .001$).

[FIGURE 3 HERE]

Discussion

[6] Substantively similar results are obtained when probit models are estimated (see Table S1).

Summary of Findings

Our primary goal in this paper was to examine how exposure to misinformation online during the 2020 election compared to exposure during the 2016 election. By adopting the analytical approach of Guess et al. (2020), who examined exposure to untrustworthy websites during the 2016 election, we assessed exposure to untrustworthy websites among a nationally representative sample during the 2020 election and compared it directly to 2016 exposure. Our web browsing data containing 7.5 million URLs represent observations of N = 1,151 Americans' real-world media usage during the course of real-world political events. Additionally, our analysis does not rely on self-reported measures of exposure to media content, which tend to be inaccurate compared to passively tracked behavioral measures of news exposure (Guess, 2015; Konitzer et al., 2021).

Overall, we found that a significantly smaller percentage of Americans were exposed to untrustworthy websites in 2020 (26.2%) compared to in 2016 (44.3%). We found this decrease in spite of using a database of untrustworthy sites over three times the size of the database used in Guess et al. (2020) to identify visits to untrustworthy websites in our participants' web browsing behavior, which increased our capacity to detect visits to untrustworthy websites. This decline runs contrary to expectations that the run up to the 2020 election would lead to record numbers of people being exposed to misinformation (e.g., Lytvynenko, 2020). We also observed that those who visited untrustworthy websites in 2020 tended to visit fewer untrustworthy sites overall and spent less time on average on each site than in 2016. These findings suggest that not only were fewer people exposed to untrustworthy websites in 2020, but that those who were exposed visited fewer of these sites and engaged for less time with them.

In 2016, certain groups were more likely than others to visit untrustworthy websites. Older adults were found to be more likely to visit untrustworthy sites than younger adults, supporters of Donald Trump were more likely than supporters of Hillary Clinton, and those with more ideologically extreme media diets were more likely than those with ideologically moderate

media diets. We found that these groups were still more likely to visit untrustworthy websites in 2020 (in 2020 Joe Biden supporters took the place of Hillary Clinton supporters), but the groups' levels of exposure were significantly lower in 2020 than in 2016. While it is encouraging that the likelihood of encountering untrustworthy websites appears to be declining over time for older adults, Trump supporters, and those with ideologically extreme media diets, the data reaffirm the need to support these groups in terms of future research and the provision of resources to build resilience to misinformation.

Finally, we found differences in how people came to visit untrustworthy websites in 2020 compared to in 2016. Specifically, Facebook referred significantly fewer visits to untrustworthy websites in 2020 than in 2016 while Google referred significantly more visits to untrustworthy sites. While further research is needed to elucidate these platforms' changing role in propagating misinformation (see below), our findings suggest that Facebook may have played a smaller role and Google may have played a larger role in directing people to misinformation on the web in the 2020 election compared to in 2016.

While our data suggest that exposure to untrustworthy websites declined at the population level, it is important to emphasize that our results should not be interpreted as indicating that misinformation is somehow less of a problem than it was previously. While exposure seems to have been lower in 2020 than in 2016, a sizable number of people were still exposed to untrustworthy sites. Extrapolating our results suggests that nearly 68 million Americans made a total of 1.5 billion visits to untrustworthy sites during the 2020 election. Furthermore, even if a smaller percentage of Americans were exposed to misinformation online in 2020, those exposures could have played a larger role in radicalization or participation in acts of political violence (e.g., the January 6, 2021 insurrection). All together, exposure of fewer people can still have serious consequences. While our data and approach are limited in their ability to speak to the consequences of exposure, it will be essential for future scholarship on

misinformation to consider both exposure to misinformation and the effect of that exposure at both the population and individual levels.

Contributions and Future Work

We make three key contributions to the misinformation literature. First, our study demonstrates the value of reapplying the same analytical approach of prior work in order to examine changes in a communication process (e.g., exposure to untrustworthy websites) longitudinally. By collecting the same data (URLs visited during web browsing) from the same source (YouGov Pulse) among the same population (nationally representative sample of American adults) for a similar period of time around the US Presidential election as Guess et al. (2020), we were able to make direct comparisons of exposure to untrustworthy websites between 2016 and 2020. The apples-to-apples comparisons afforded by this approach allowed us to precisely examine how the patterns in untrustworthy website exposure identified in 2016 changed in 2020. In addition, we incorporated improvements to Guess et al. (2020)'s approach into our analysis that accounted for differences between the 2016 and 2020 media environments. For example, we introduced NewsGuard's database of sites that repeatedly publish false content as an additional source to identify visits to untrustworthy websites in people's web browsing. One reason the introduction of NewsGuard was important is because the untrustworthy website databases used by Guess et al. (2020) were primarily based on websites circulating during the 2016 election, but fake news websites are often ephemeral (i.e., the domains go defunct after short periods of time; Lim et al., 2019). NewsGuard's database (which is updated weekly) allowed us to have more confidence that our database of untrustworthy sites was sufficiently up-to-date to match changes in the fake news ecosystem that have occurred since 2016.

Second, our findings indicate that the same groups who were more likely to visit untrustworthy websites in 2016 were largely the same groups more likely to do so in 2020. Older adults, Donald Trump supporters, and those with the most conservative media diets were

more likely to visit untrustworthy websites in both 2016 and 2020. The persistence of these trends highlights the importance of examining why older adults appear to be more susceptible to online misinformation than younger individuals and how they can be supported through interventions and other resources to build resilience to misinformation (for a review, see (Brashier & Schacter, 2020; Moore & Hancock, 2020). Our updated findings, which reveal that approximately one-third of the older adults in our sample were exposed to untrustworthy websites, make clear that it is important to continue studying the factors responsible for older adults' vulnerability.

In addition to this pattern among older adults, the 2016 pattern of conservatives being more likely than liberals to visit untrustworthy sites persisted in 2020. Research has begun to identify why conservative individuals appear more likely to engage with misinformation online (Freelon et al., 2020). Recent work suggests that the supply of misinformation may be greater on the ideological right than the left (Benkler et al., 2018; Garrett & Bond, 2021). Relatedly, it could be the case that more ideologically conservative media diets are more likely to expose users to misinformation via features like algorithmic curation and community structures relative to liberal media diets. Our results suggest an ongoing need for future work to investigate these potential causes.

Third, we provide evidence that Americans' visits to untrustworthy websites were significantly less likely to be referred by Facebook and more likely to be referred by Google in 2020 compared to 2016. This finding is significant given the amount of scrutiny from members of Congress and the American public towards these platforms' roles in the proliferation of misinformation (Auxier, 2020; Wakefield, 2021). Future work should attempt to better understand why Facebook's role in referring people to untrustworthy websites appears to be shrinking and Google's appears to be growing larger. For Facebook, does it indicate the efficacy with which they are implementing programs and policies to label or flag untrustworthy content? Or might it suggest a more fundamental behavioral change in how people are using Facebook

and other social media platforms? For instance, people may be less likely to click links to external websites now than in the past, preferring to stay on the platform. For Google, does it indicate that their results rankings algorithms were doing a worse job in 2020 than in 2016 at downranking misinformation? Or did people use Google queries more often to fact check misinformation results, as many digital media literacy training programs suggest? Unfortunately, as the recent Facebook Files controversy illustrates, without these platforms sharing data and becoming more transparent, it is difficult to understand these changes in referral patterns (The Wall Street Journal, 2021). We add our voices to those calling for increased data sharing and independent research of platform data (Edelson & McCoy, 2021; Persily, 2021).

Finally, our findings can inform policymaking and discourse around misinformation. Taken together with our findings reaffirming that certain groups (older adults and conservatives) are more likely to encounter misinformation, our results suggest a need for more focused and directed policy initiatives centering on groups with the greatest need for support in dealing with misinformation. Moreover, our findings join others suggesting that the attention and discussion which the media, politicians, and the public devote to fake news may be disproportionate to the extent that people are actually exposed to it (Weeks et al., 2021). Given evidence that a) the overwhelming majority of news consumed by the population is not misinformation (Allen et al., 2020), and b) that exposure to discourse around fake news can erode individuals' trust in news media (Van Duyn & Collier, 2019), we may need to consider the nature of the attention society pays to the problem of misinformation relative to other ongoing national and international crises. For example, the emphasis on misinformation present in the media and political discourse since 2016 may be partly to blame for the ongoing erosion of trust in media institutions occurring in the United States and around the world (Edelman, 2021). Our findings call for the need for more research, but also more grounded communication of that research to limit exaggeration of this phenomenon.

Limitations

Of course, our study should be interpreted in light of its limitations. The most noteworthy limitations relate to the URL logging methodology we used to estimate untrustworthy website exposure. First, untrustworthy websites were operationalized at the domain level. That is, in our data, untrustworthy websites were considered web domains that were rated by NewsGuard as repeatedly hosting false information (or imported from the database used by Guess et al. [2020]) (e.g., www.obamawatcher.com), rather than specific web pages or articles (e.g., www.obamawatcher.com/2020/03/michelles-fake-degrees). This operationalization is largely because most of the URLs captured by YouGov only contain domains and not full URLs to protect participant privacy. While many studies examining exposure to misinformation have taken this approach (Guess et al., 2020), there is undoubtedly misinformation that is hosted on domains which do not repeatedly publish false information but may otherwise occasionally publish false information. Domain-level measurements of exposure do not capture these specific instances of misinformation.

A second limitation of working with URL-based browsing data is that it only identifies content which leads to a URL being produced (Ellison et al., 2020). Crucially, this limitation means that for web pages which display content dynamically while maintaining a static URL, we only know that a person visited that static URL but not information about any of the content they saw while on that static URL. Take for example the Facebook Newsfeed. When a user visits www.facebook.com and is presented with their Newsfeed, that user can scroll through their Newsfeed and that does not result in the active URL in their web browser, www.facebook.com, changing. Thus, while an individual may be exposed to a variety of (mis)information while scrolling their Newsfeed (either in posts generated or links shared by others), we only observe instances in which individuals actually click on an external link which takes them away from their Newsfeed to a new website.

Additionally, these URL data were collected by participants installing plugins in their web browsers. Thus, we only capture individuals' behaviors that take place within web browsers.

Individuals' online behaviors which occur outside of web browsers, such as through the use of apps, do not appear in our dataset. This may be especially relevant for mobile internet use, which is more likely to occur via mobile apps than mobile web browsers (Wurmser, 2019). Indeed, among our own participants, we found that substantially more individuals were exposed to an untrustworthy website on desktop/laptop computers (34.2%; 95% CI = 29.4-39.1%) than on smartphones (13%; 95% CI = 8.9-17.2%).

It is important to note that these limitations affect both our 2020 data collection and Guess et al. (2020)'s 2016 data collection. Nevertheless, it is crucial for future research on misinformation exposure to contend with the limitations of web browsing data. Only social media companies ultimately possess the data on user behavior that could most accurately shed light on why their role in referring individuals to untrustworthy websites appears to be decreasing over time. Unfortunately, social media platforms rarely share data with misinformation researchers (Pasquetto & Swire-Thompson, 2020). While initiatives such as Social Science One have attempted to grant scholars access to data from social media platforms like Facebook (King & Persily, 2020), getting academic researchers access to platform data has proven quite difficult (Edelson & McCoy, 2021; Persily, 2021). The limitations of our study and those that other misinformation research face highlight the need for policymakers to consider requiring social media platforms to work with academic and other third-party researchers to better understand the complex dynamics of exposure to and engagement with (mis)information on their platforms.

That said, there are some current data collection approaches which could help fill in the gaps in online behavior missed by the URL-logging method. For example, Screenomics captures screenshots from individuals' devices to understand their moment-by-moment smartphone usage, including the use of apps and information contained within system notifications (Reeves et al., 2019, 2020). Future research on exposure to online misinformation

should triangulate across several data sources in order to gain a more complete portrait of individuals' online media use and the role that misinformation plays in it.

Explaining the Decline in Exposure from 2016 to 2020

Overall, our goal in this paper was to document exposure to untrustworthy websites in 2020 in such a way methodologically that we could compare 2020 exposure directly to 2016 exposure as estimated by Guess et al. (2020). We document a relatively across-the-board decline in exposure to untrustworthy websites from 2016 to 2020, but *why* does this decline occur? We offer a few candidate explanations, but future work and additional data will be needed to more directly test them.

First, exposure to misinformation could have been more likely to be displaced in 2020 than in 2016 to other locations outside of the web browser entirely, such as to text messaging or emergent social media apps like WhatsApp and TikTok. Indeed, people increasingly report getting their news regularly via WhatsApp (Newman et al., 2019) and social media platforms like Reddit and TikTok (Walker & Matsa, 2021), and there is significant concern about the spread of misinformation on these platforms (Delcker et al., 2020; Zadrozny, 2021).

Second, the time-frame of data collection used by Guess et al. (2020) and adopted by us (4 weeks prior to Election Day and 1 week after Election Day) may have examined different misinformation dynamics around the 2020 and 2016 elections. Specifically, the 2020 election was marked by a post-Election Day period in which sitting president Donald Trump made a series of claims about election fraud that caused him to receive less votes than his opponent Joe Biden, culminating in the announcement that Joe Biden was the winner of the election on November 7, which Donald Trump refused to concede. During this time, a significant amount of misinformation about election fraud spread online (Frenkel, 2020). Comparatively, the aftermath of the 2016 election may not have been as rife with online misinformation as the aftermath of the 2020 election. However, because Guess et al.'s (2020) (and thus our) data collection period included more time prior to the election than after, we may have missed some of the online

misinformation relevant to the 2020 election outcome that was not present in 2016. In our view, this point demonstrates that when comparing the effects of events on media consumption behaviors, even similar events (e.g., presidential elections) may feature different dynamics at different points in time.

Third, as we pointed out in the Limitations section, URL-tracking methods only log instances in which a URL is actually clicked and visited. Visits to untrustworthy websites may have decreased from 2016 to 2020 because people were increasingly exposed to untrustworthy website content within dynamic URLs, for example scrolling through the Facebook Newsfeed or Twitter Timeline. Indeed, evidence suggests that relatively few people click on news links posted on social media websites (Ju et al., 2014; Mitchell & Jurkowitz, 2014) yet may be influenced by the information in headlines nonetheless (Rapp & Salovich, 2018). This propensity to stay on the platform over clicking to visit external websites also may be increasing over time (Geeng et al., 2020). Such a change in user behavior could also help explain why we found that Facebook played a smaller role in referring people to untrustworthy websites in 2020 than in 2016.

Conclusion

We provide evidence that exposure to untrustworthy websites decreased from 2016 to 2020. More work is needed to understand the factors explaining this apparent change, but our results represent an important update to work on exposure to online misinformation. The same groups who were most likely to be exposed in 2016 are much the same groups who were more likely exposed in 2020, justifying a more focused approach to research on and support for those groups in dealing with misinformation. While one could interpret our findings as evidence that the problem of online misinformation is improving in some way, they could also be interpreted as evidence that the nature of the problem is changing. Our work provides some initial insights into where researchers can start looking to understand the changing dynamics of online misinformation exposure.

References

Alaphilippe, A. (2021, February 4). Disinformation is evolving to move under the radar. *Brookings*. https://www.brookings.edu/techstream/disinformation-is-evolving-to-move-under-the-radar/

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, *6*(2), 2053168019848554. https://doi.org/10.1177/2053168019848554

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, *6*(14), eaay3539. https://doi.org/10.1126/sciadv.aay3539

Auxier, B. (2020, October 15). 64% of Americans say social media have a mostly negative effect on the way things are going in the U.S. today. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2020/10/15/64-of-americans-say-social-media-have-a-mostly-negative-effect-on-the-way-things-are-going-in-the-u-s-today/

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Ball, P., & Maxmen, A. (2020). The epic battle against coronavirus misinformation and conspiracy theories. *Nature*, *581*(7809), 371–374. https://doi.org/10.1038/d41586-020-01452-z

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press. https://doi.org/10.1093/oso/9780190923624.001.0001

Brashier, N. M., & Schacter, D. L. (2020). Aging in an Era of Fake News. *Current Directions in Psychological Science*, *29*(3), 316–323. https://doi.org/10.1177/0963721420915872

Delcker, J., Wanat, Z., & Scott, M. (2020, March 16). *The coronavirus fake news pandemic sweeping WhatsApp*. POLITICO. https://www.politico.eu/article/the-coronavirus-covid19-fake-news-pandemic-sweeping-whatsapp-misinformation/

Edelman. (2021). *2021 Edelman Trust Barometer* (p. 58). https://www.edelman.com/trust/2021-trust-barometer

Edelson, L., & McCoy, D. (2021, September 22). *How Facebook Hinders Misinformation Research*. Scientific American. https://www.scientificamerican.com/article/how-facebook-hinders-misinformation-research/

Ellison, N. B., Triệu, P., Schoenebeck, S., Brewer, R., & Israni, A. (2020). Why We Don't Click: Interrogating the Relationship Between Viewing and Clicking in Social Media Contexts by Exploring the "Non-Click." *Journal of Computer-Mediated Communication*, *25*(6), 402–426. https://doi.org/10.1093/jcmc/zmaa013

Fazio, L. K. (2020). Recognizing the Role of Psychological Science in Improving Online Spaces. *Psychological Science in the Public Interest*, *21*(3), 99–102. https://doi.org/10.1177/1529100620972100

Freelon, D., Marwick, A., & Kreiss, D. (2020). False equivalencies: Online activism from left to right. *Science*, *369*(6508), 1197–1201. https://doi.org/10.1126/science.abb2428

Frenkel, S. (2020, November 23). How Misinformation 'Superspreaders' Seed False Election Theories. *The New York Times*. https://www.nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html

Garrett, R. K., & Bond, R. M. (2021). Conservatives' susceptibility to political misperceptions. *Science Advances*, *7*(23), eabf1234. https://doi.org/10.1126/sciadv.abf1234

Geeng, C., Yee, S., & Roesner, F. (2020). Fake News on Facebook and Twitter: Investigating

How People (Don't) Investigate. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376784

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378. https://doi.org/10.1126/science.aau2706

Guess, A., Lyons, B., Montgomery, J. M., Nyhan, B., & Reifler, J. (2019). *Fake news, Facebook ads, and misperceptions: Assessing information quality in the 2018 U.S. midterm election campaign* (p. 36).

Guess, A. M. (2015). Measure for Measure: An Experimental Test of Online Political Media Exposure. *Political Analysis*, *23*(1), 59–75. https://doi.org/10.1093/pan/mpu010

Guess, A., Nagler, J., & Tucker, J. (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, A., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, *4*(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x

Ju, A., Jeong, S. H., & Chyi, H. I. (2014). Will Social Media Save Newspapers? *Journalism Practice*, *8*(1), 1–17. https://doi.org/10.1080/17512786.2013.794022

King, G., & Persily, N. (2020). A New Model for Industry–Academic Partnerships. *PS: Political Science & Politics*, *53*(4), 703–709. https://doi.org/10.1017/S1049096519001021

Konitzer, T., Allen, J., Eckman, S., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2021). Comparing Estimates of News Consumption from Survey and Passively Collected Behavioral Data. *Public Opinion Quarterly*, *85*(S1), 347–370. https://doi.org/10.1093/poq/nfab023

Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools. *Psychological Science in the Public Interest*,

21(3), 103–156. https://doi.org/10.1177/1529100620946707

Li, J., & Su, M.-H. (2020). Real Talk About Fake News: Identity Language and Disconnected Networks of the US Public's "Fake News" Discourse on Twitter. *Social Media + Society*, *6*(2), 2056305120916841. https://doi.org/10.1177/2056305120916841

Lim, G., Maynier, E., Scott-Railton, J., Fittarelli, A., Moran, N., & Deibert, R. (2019). *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign* (Citizen Lab Research Report No. 118). University of Toronto. https://citizenlab.ca/2019/05/burned-after-reading-endless-mayflys-ephemeral-disinformation-campaign/

Lytvynenko, J. (2020, December 6). *In 2020, Disinformation Broke The US*. BuzzFeed News. https://www.buzzfeednews.com/article/janelytvynenko/disinformation-broke-us

Mitchell, A., & Jurkowitz, M. (2014, March 13). How Readers Get to News Sites: Social, Search and Direct. *Pew Research Center's Journalism Project*. https://www.pewresearch.org/journalism/2014/03/13/social-search-direct/

Moore, R. C., & Hancock, J. T. (2020). Older Adults, Social Technologies, and the Coronavirus Pandemic: Challenges, Strengths, and Strategies for Support. *Social Media + Society*, *6*(3), 2056305120948162. https://doi.org/10.1177/2056305120948162

Mosseri, A. (2017, April 7). *Working to Stop Misinformation and False News*. Working to Stop Misinformation and False News | Facebook Media. https://www.facebook.com/formedia/blog/working-to-stop-misinformation-and-false-news

Nelson, J. L., & Taneja, H. (2018). The small, disloyal fake news audience: The role of audience availability in fake news consumption. *New Media & Society*, *20*(10), 3720–3737. https://doi.org/10.1177/1461444818758715

Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2019). *Reuters Institute Digital News Report 2019* (SSRN Scholarly Paper ID 3414941). Social Science Research Network. https://papers.ssrn.com/abstract=3414941

Pasquetto, I., & Swire-Thompson, B. (2020). Tackling misinformation: What researchers could

do with social media data: Introduction. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-49

Persily, N. (2021). A Proposal for Researcher Access to Platform Data: The Platform Transparency and Accountability Act. *Journal of Online Trust and Safety*, *1*(1), Article 1. https://doi.org/10.54501/jots.v1i1.22

Pew Research Center, pew. (2018, April 26). 10. Political engagement, knowledge and the midterms. *Pew Research Center - U.S. Politics & Policy*. https://www.pewresearch.org/politics/2018/04/26/10-political-engagement-knowledge-and-the-midterms/

Rapp, D. N., & Salovich, N. A. (2018). Can't We Just Disregard Fake News? The Consequences of Exposure to Inaccurate Information. *Policy Insights from the Behavioral and Brain Sciences*, *5*(2), 232–239. https://doi.org/10.1177/2372732218785193

Reeves, B., Ram, N., Robinson, T. N., Cummings, J. J., Giles, C. L., Pan, J., Chiatti, A., Cho, M., Roehrick, K., Yang, X., Gagneja, A., Brinberg, M., Muise, D., Lu, Y., Luo, M., Fitzgerald, A., & Yeykelis, L. (2019). Screenomics: A Framework to Capture and Analyze Personal Life Experiences and the Ways that Technology Shapes Them. *Human–Computer Interaction*, *0*(0), 1–52. https://doi.org/10.1080/07370024.2019.1578652

Reeves, B., Robinson, T., & Ram, N. (2020). Time for the Human Screenome Project. *Nature*, *577*(7790), 314–317. https://doi.org/10.1038/d41586-020-00032-5

Roth, Y., & Achuthan, A. (2020, February 4). *Building rules in public: Our approach to synthetic & manipulated media*. Twitter. https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media

The Wall Street Journal. (2021, October 1). The Facebook Files. *Wall Street Journal*. https://www.wsj.com/articles/the-facebook-files-11631713039

Van Duyn, E., & Collier, J. (2019). Priming and Fake News: The Effects of Elite Discourse on

Evaluations of News Media. *Mass Communication and Society*, *22*(1), 29–48. https://doi.org/10.1080/15205436.2018.1511807

Vu, H. T., & Saldaña, M. (2021). Chillin' Effects of Fake News: Changes in Practices Related to Accountability and Transparency in American Newsrooms Under the Influence of Misinformation and Accusations Against the News Media. *Journalism & Mass Communication Quarterly*, *98*(3), 769–789. https://doi.org/10.1177/1077699020984781

Wakefield, J. (2021, March 25). Google, Facebook Twitter grilled in US on fake news. *BBC News*. https://www.bbc.com/news/technology-56523378

Walker, M., & Matsa, K. E. (2021, September 20). News Consumption Across Social Media in 2021. *Pew Research Center's Journalism Project*. https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/

Weeks, B. E., & Gil de Zúñiga, H. (2019). What's Next? Six Observations for the Future of Political Misinformation Research. *American Behavioral Scientist*, 0002764219878236. https://doi.org/10.1177/0002764219878236

Weeks, B. E., Menchen-Trevino, E., Calabrese, C., Casas, A., & Wojcieszak, M. (2021). Partisan media, untrustworthy news sites, and political misperceptions. *New Media & Society*, 14614448211033300. https://doi.org/10.1177/14614448211033300

Wilson, T., & Starbird, K. (2021). Cross-platform Information Operations: Mobilizing Narratives & Building Resilience through both "Big" & "Alt" Tech. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 345:1-345:32. https://doi.org/10.1145/3476086

Wurmser, Y. (2019, May 30). *US Time Spent with Mobile 2019*. Insider Intelligence. https://www.emarketer.com/content/us-time-spent-with-mobile-2019

Zadrozny, B. (2021, July 31). *How one man's viral rant shows TikTok's problem with misinformation*. NBC News. https://www.nbcnews.com/tech/tech-news/tiktok-audio-

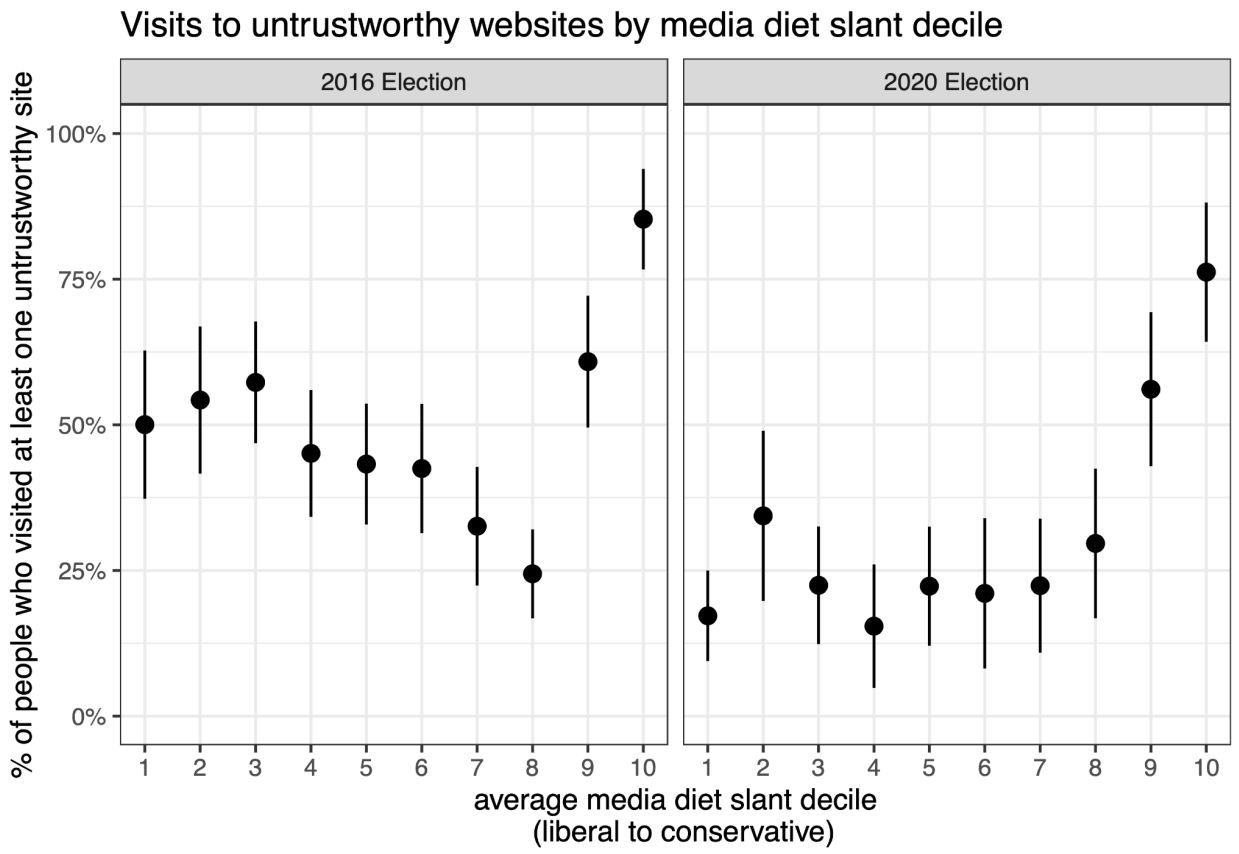
gives-new-virality-misinformation-rcna1393

Table 1

	Exposure to at least one untrustworthy website (binary)	
	2016	2020
Trump supporter	0.164***	0.138***
	(0.021)	(0.027)
Political knowledge	0.017**	0.049***
	(0.005)	(0.012)
Political interest	0.048***	0.043**
	(0.014)	(0.015)
College	0.040	0.049
	(0.022)	(0.028)
Female	−0.022	0.006
	(0.020)	(0.026)
Non-white	−0.094***	−0.071*
	(0.021)	(0.029)
Age 30-44 years	0.041	−0.007
	(0.032)	(0.041)
Age 45-64 years	0.064*	0.042
	(0.029)	(0.038)
Age 65+	0.129***	0.113**
	(0.033)	(0.043)
Constant	0.126**	−0.080
	(0.046)	(0.059)
Observations	2,514	1,151
R^2	0.092	0.105
Adjusted R^2	0.089	0.098
Residual Std. Error	0.466 (df = 2504)	0.421 (df = 1141)
F Statistic	28.265*** (df = 9; 2504)	14.827*** (df = 9; 1141)
Note:		*p<0.05; **p<0.01; ***p<0.001

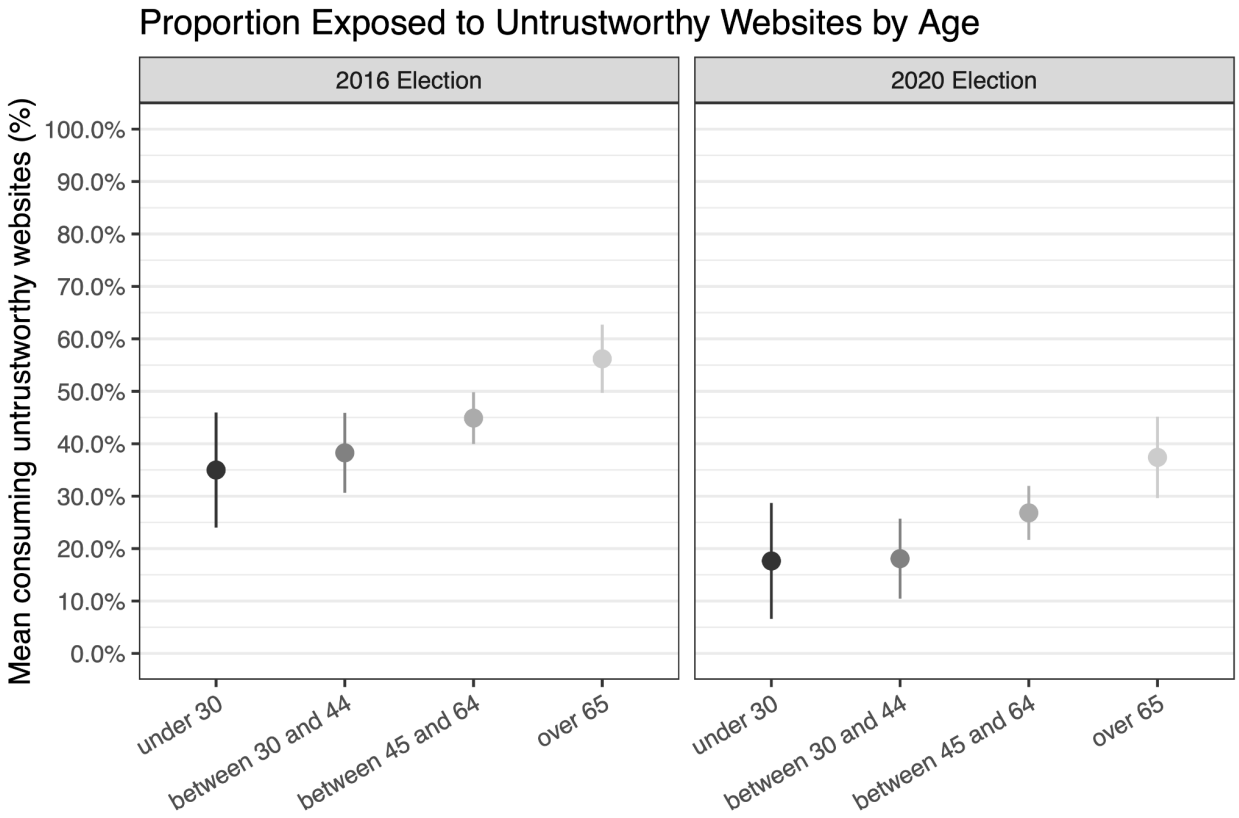
Note: OLS regression coefficients are shown with standard errors in parentheses (models estimated using survey weights). The model on the left represents the 2016 election and the model on the right represents the 2020 election. The dependent variable in both models is a binary variable indicating whether an individual was exposed to at least one untrustworthy website during the data collection period (1) or not (0). “Trump supporter” variable: intending to vote for Trump in 2020 election = 1; not intending to vote for Trump in 2020 election = 0. “Political knowledge” variable: variable ranging from 0-4 representing the number of questions in Pew Research Center’s civic knowledge questionnaire answered correctly out of 4. “Political interest” variable: variable ranging from 1-4 where 4 = people who say they pay attention to what’s going on in government and politics “most of the time” and 1 = those who pay attention “hardly at all”. “College” variable: 1 = college graduate; 0 = not a college graduate. “Female” variable: 1 = indicated identifying as a female, 0 = did not indicate identifying as a female. “Non-white” variable: 1 = indicated identifying as a race other than white; 0 = indicated identifying as white. P-values are two-sided.

Figure 1



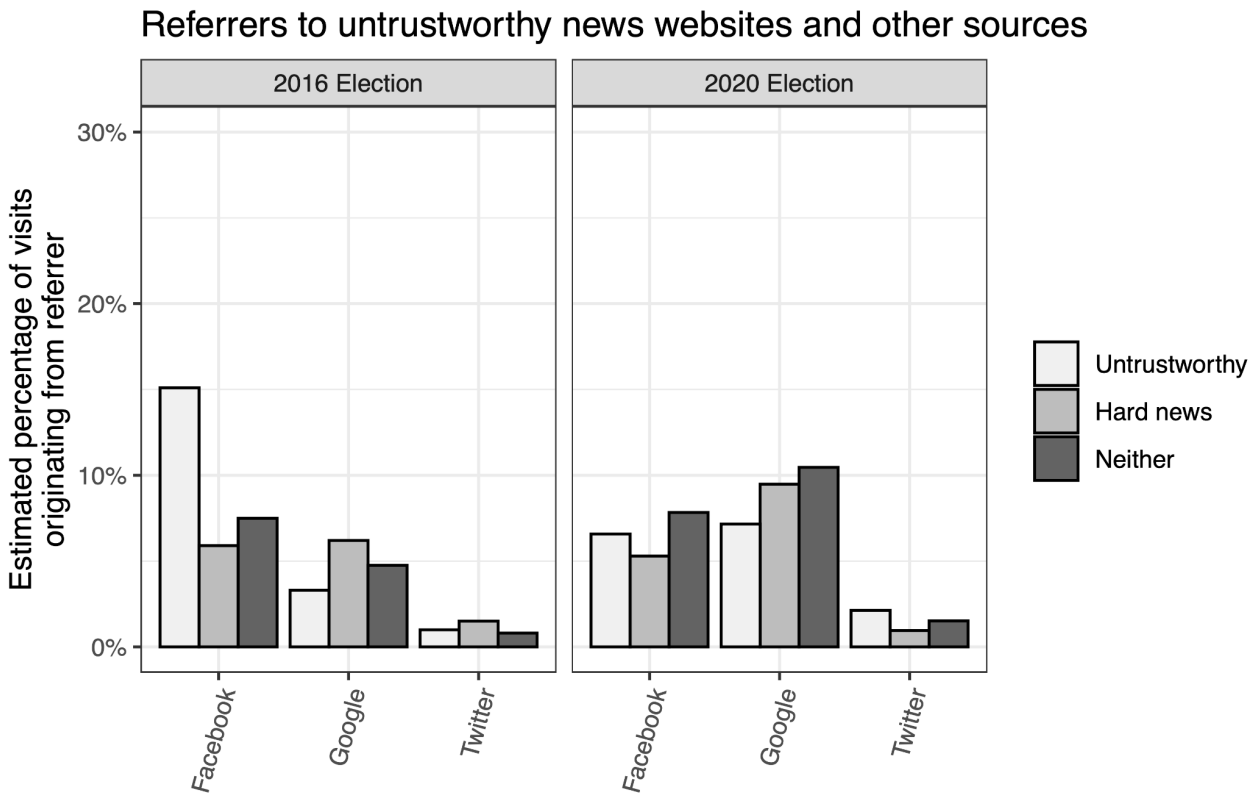
Note: Estimated percentage of people in media diets with different ideological slants who were exposed to at least one untrustworthy website. The panel on the left represents 2016 and the panel on the right represents 2020. Individuals' media diets were split into deciles based on the ideological slant of their news consumption as determined by ideological ratings of hard news websites from Bakshy et al. (2015). Along the x-axis, lower values represent more liberal media diets, middle values represent more moderate media diets, and higher values represent more conservative media diets. Along the y-axis, higher values indicate a greater estimated percentage of people in that media slant decile who were exposed to at least one untrustworthy website. Bars are 95% confidence intervals.

Figure 2



Note: Estimated percentage of people in different age groups (under 30, between 30 and 44, between 45 and 64, 65 and older) who were exposed to at least one untrustworthy website. The panel on the left represents 2016 and the panel on the right represents 2020. Higher values indicate a greater estimated percentage of people in each age group who were exposed to at least one untrustworthy website. Bars are 95% confidence intervals.

Figure 3



Note: Percentage of total visits to untrustworthy websites, hard news websites, and websites which were neither untrustworthy nor hard news that were referred by different platforms (Facebook, Google, Twitter). The panel on the left represents 2016 and the panel on the right represents 2020. Higher values indicate a greater percentage of visits to each type of website which were preceded by visits to a platform.