

Exposure to untrustworthy websites in the 2020 US election

Received: 9 February 2022

Ryan C. Moore^{1,2}✉, Ross Dahlke^{1,2} & Jeffrey T. Hancock¹

Accepted: 20 February 2023

Published online: 13 April 2023

 Check for updates

Research using large-scale data on individuals' internet use has provided vital information about the scope and nature of exposure to misinformation online. However, most prior work relies on data collected during the 2016 US election. Here we examine exposure to untrustworthy websites during the 2020 US election, using over 7.5 million website visits from 1,151 American adults. We find that 26.2% (95% confidence interval 22.5% to 29.8%) of Americans were exposed to untrustworthy websites in 2020, down from 44.3% (95% confidence interval 40.8% to 47.7%) in 2016. Older adults and conservatives continued to be the most exposed in 2020 as in 2016, albeit at lower rates. The role of online platforms in exposing people to untrustworthy websites changed, with Facebook playing a smaller role in 2020 than in 2016. Our findings do not minimize misinformation as a key social problem, but instead highlight important changes in its consumption, suggesting directions for future research and practice.

Both concern over and research on misinformation have exploded in recent years¹. In an effort to better understand and prevent the spread of misinformation, much extant research documents exposure to misinformation online using large-scale data on individuals' internet use. This work finds that relatively few people are exposed to misinformation online, or at least fewer than many initially expected^{2–7}. Exposure is also highly concentrated^{4–6,8}. For example, Grinberg et al.⁴ found that 1% of Twitter users were exposed to 80% of the fake news on Twitter during the 2016 election. Relatedly, certain individuals are more likely to be exposed to misinformation online than other groups. For example, during the 2016 election, people aged 65 years 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^{4,5}.

Many studies investigating exposure to online misinformation have leveraged data collected during the 2016 US election, perhaps because concern over fake news rose during the 2016 election⁹. However, a consequence of focusing on the 2016 election is that we have little insight into how exposure to misinformation online has changed since then. This limitation is important given the myriad ways the digital (mis)information ecosystem has changed (for example, new platforms, new misinformation-generating world events and new modalities to disseminate misinformation^{10,11}). Furthermore, in the

wake of 2016, online platforms such as Facebook have taken steps to mitigate their reputation as a purveyor of misinformation¹².

While most prior research focuses on 2016, some work has examined changes in exposure over time. For example, Guess et al.¹³ found a decline in the number of Americans exposed to fake news websites from 2016 to 2018. In contrast, Allen et al.³ found that exposure to fake news was 'generally stable' from 2016 to the end of 2018. Discrepancies between these two findings may partly be explained by the authors using different lists to identify visits to fake news websites in their data. Allcott et al. also examined changes in exposure over time, finding that engagement with fake news on Facebook declined from 2016 to 2018, while engagement with fake news on Twitter rose over the same period¹⁴.

Given the changes in world events, the digital media landscape and a considerable societal effort to combat misinformation, a vital question remains as to how misinformation exposure during the 2020 US presidential election compared to exposure during the 2016 election, which served as the context for much of the existing research. In this Article, we investigate exposure to online misinformation during the 2020 election in a way that allows us to compare exposure in 2020 with 2016 directly. To do this, we collected web browsing data (that is, uniform resource locators (URLs) visited) from a nationally representative sample of American adults ($N=1,151$) during the lead-up

¹Department of Communication, Stanford University, Stanford, CA, USA. ²These authors contributed equally: Ryan C. Moore, Ross Dahlke.

✉e-mail: rymoore@stanford.edu

to the 2020 election (7.5 million total website visits). We identify visits to misinformation websites in participants' web browsing using a database of websites known to repeatedly publish misinformation. To produce estimates of misinformation exposure directly comparable to those from 2016, we adopted the analytical approach of Guess et al.⁶, who studied exposure to misinformation during the lead-up to the 2016 election using web browsing data (for full methodological details, see Methods). While we stuck closely to the data sources and analytical techniques used by Guess et al.⁶, we made key updates that ensured our analysis was up to date for the media environment in 2020, such as expanding the database of misinformation websites from 490 sites to 1,796 sites and including mobile browsing data.

What might we expect about Americans' exposure to online misinformation in 2020? On the one hand, there are reasons to believe that exposure in 2020 was greater than in 2016. The 2020 election was marked by a confluence of major national and global events that generated a considerable amount of misinformation to which people could have been exposed (for example, the rise of the Coronavirus disease 2019 pandemic, claims that the 2020 presidential election was fraudulent¹⁴). Furthermore, the infrastructure available to disinformation purveyors for disseminating and spreading false information across platforms has improved considerably in recent years^{15,16}. On the other hand, there are also reasons to believe that exposure in 2020 was less than in 2016. Since the 2016 election, journalists and online platforms have taken several steps to curb the spread of misinformation. For example, social media platforms such as Facebook and Twitter have begun labelling and flagging untrustworthy content on their platforms^{12,17,18}, and journalists have intensified their efforts around accountability and transparency¹⁹. Furthermore, interventions and educational resources for helping the public identify online misinformation that have proliferated since 2016 may have improved internet users' ability to avoid misinformation^{20,21}. Finally, the few studies that have examined misinformation exposure post-2016 have identified a decline or no change in exposure^{3,13}.

Understanding exposure to online misinformation in 2020 is important for several reasons. First, it represents an effort to see whether the inferences drawn about the scope and nature of misinformation exposure from 2016 are valid in 2020, helping to enhance the 'temporal validity' of the knowledge base on online misinformation exposure (for a greater discussion of the concept and utility of temporal validity in the context of social media research, see Munger^{22,23}). Second, 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, and understanding the nature of misinformation exposure can inform policymaking efforts to mitigate misinformation (for example, requiring social media platforms to implement certain flagging or fact-checking practices). Finally, updates about misinformation exposure can inform public discourse around misinformation itself. Since 2016, misinformation has been a prominent topic in communication from journalists, elites, scientists and the general public⁹, with implications for how the public trusts information sources²⁴. Indeed, exposure to elite discourse about misinformation has been shown to reduce people's trust in traditional news media²⁵. If it is true that the vast majority of people's news diet is composed of reliable news³, then the heightened discourse around misinformation could erode trust in reliable news.

Overall, our analysis reveals that exposure to untrustworthy websites during the 2020 election was substantially lower than Guess et al.⁶ determined during the 2016 election. Not only was the percentage of total individuals exposed to untrustworthy websites lower, but the average number of exposures among the exposed and the average amount of time spent on each site were also lower. While we also found that the groups who were more likely to be exposed to untrustworthy

websites in 2016—older adults and conservatives—were still more likely in 2020, 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 used in Guess et al.⁶, meaning that even with a considerably larger capacity to identify misinformation in people's web browsing data, we still found less. Finally, we found changes in how people come to be exposed to untrustworthy websites, namely that people were less likely to arrive at them via Facebook and webmail in 2020 than in 2016.

Results

Total consumption of untrustworthy websites in 2020

During the 2016 election, Guess et al.⁶ found that 44.3% (95% confidence interval (CI) 40.8% to 47.7%) of Americans aged 18 years 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% to 29.8%) of Americans were exposed to at least one untrustworthy website. This decrease represents a significant reduction ($t(2, 2,493.6) = -11.12$, $P < 0.001$, Cohen's $d = 0.39$, 95% CI -19.8 to -16.5) in the percentage of Americans exposed to untrustworthy websites from 2016 to 2020 (for additional estimates of 2020 exposure using NewsGuard's 0–100 'trust score', see page 2 of Supplementary Materials).

While a smaller percentage of Americans appeared to be exposed to untrustworthy websites in 2020 compared with 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 to 30.4) was significantly lower than the average number in 2016 (32.1 visits; 95% CI 26.3 to 37.9) ($t(2, 638) = -2.06$, $P = 0.04$, $d = 0.12$, 95% CI -13.8 to -4.8). Second, the average length of time spent on untrustworthy websites was shorter in 2020 (38.6 s; 95% CI 34.5 to 42.7 s) than in 2016 (64.2 s; Guess et al.⁶ did not quantify uncertainty for this quantity).

Comparing visits to untrustworthy websites in 2020 and 2016, we find that overall consumption of untrustworthy sites was lower in 2020 than in 2016 (for an analysis of untrustworthy website consumption from 2016 to 2020 using available estimates from other studies, see also page 4 of Supplementary Materials). Fewer Americans were exposed to untrustworthy websites in the lead-up 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 compared with those who were more likely to visit them in 2020.

Who consumed untrustworthy websites in 2020?

In 2016, Guess et al.⁶ found that supporters of then-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% to 64.5%) of Trump supporters visiting at least one untrustworthy website and 37.1% (95% CI 31.9% to 42.2%) of Clinton supporters visiting at least one untrustworthy website ($t(2, 2,000.0) = 10.6$, $P < 0.001$, $d = 0.46$, 95% CI 10.3% to 24.5%). We find a similar difference between Donald Trump and Joe Biden supporters in 2020. Of Trump supporters, 36.2% (95% CI 29.7% to 42.6%) visited at least one untrustworthy website in 2020, while 17.8% (95% CI 13.7% to 22.0%) of Biden supporters visited at least one untrustworthy website ($t(2, 742.1) = 6.6$, $P < 0.001$, $d = 0.42$, 95% CI 15.6% to 21.1%). However, for both Trump and Clinton/Biden supporters, the percentages of supporters exposed in 2020 were significantly lower than the percentages of supporters exposed in 2016 ($t_{\text{Trump}}(2, 818.8) = -8.2$, $P < 0.001$, $d = 0.48$, 95% CI -26.2 to -20.5; $t_{\text{Clinton/Biden}}(2, 1,654.8) = -9.5$, $P < 0.001$, $d = 0.44$, 95% CI -21.3 to -17.2).

Guess et al.⁶ further examined the association between political leaning and consumption of untrustworthy websites by classifying

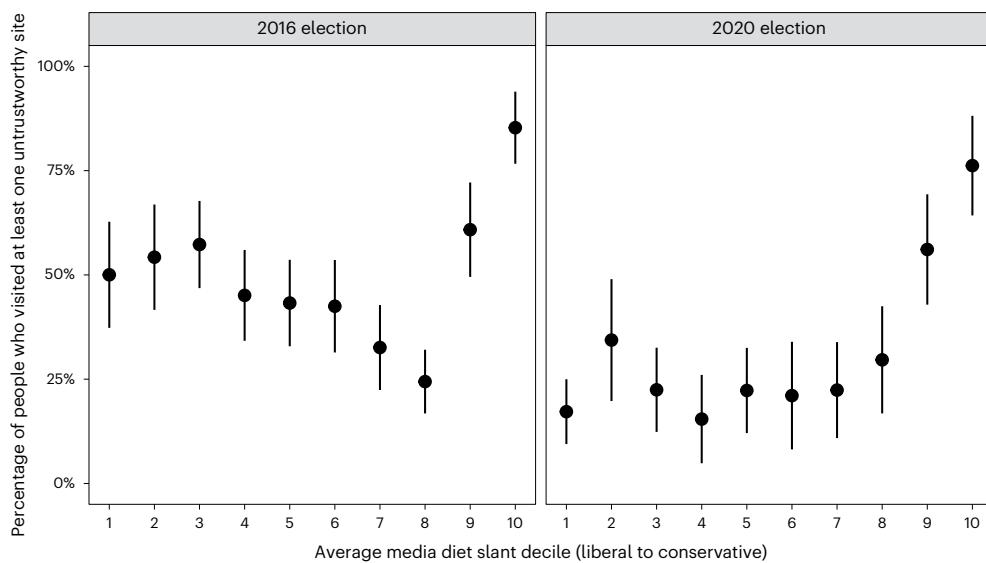


Fig. 1 | Visits to untrustworthy websites by media slant decile. Estimated percentage of people in media diets with different ideological slants who were exposed to at least one untrustworthy website ($N = 3,432$). The left represents 2016 and the right represents 2020. Individuals' media diets were split into deciles on the basis of the ideological slant of their news consumption, as determined by ideological ratings of hard news websites from Bakshy et al.²⁶. 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. Guess et al.⁶ calculated 2016 deciles separately for Trump and Clinton supporters; we recalculated them to be among all participants using their publicly available replication data. Data are presented as weighted mean values \pm 95% CIs.

users on the basis of the ideological slant of their news diets 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.²⁶), they split their participants into news slant deciles ranging from most liberal to most conservative. We implemented this same classification strategy on our sample to examine how exposure to untrustworthy websites across people with news diets with different ideological slants in 2020 compares with 2016.

This comparison is shown in Fig. 1 (Supplementary Table 7 contains the estimates comprising Fig. 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 towards more moderate media diets but then increases among the most conservative diets, with 85.2% (95% CI 76.7% to 93.9%) in the most conservative media diet 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 again increases dramatically among the most conservative diets. Due to these different patterns, the percentage of people exposed to untrustworthy websites in, for instance, the most liberal media diet in 2020 was considerably lower than in the most liberal media diet in 2016 ($t(2, 257.5) = -7.12, P < 0.001, d = 0.74, 95\% \text{ CI } -37.8\% \text{ to } -27.9\%$). Among the most conservative media diet, the percentage of those exposed to untrustworthy sites did not differ significantly across both years ($t(2, 159.9) = -1.88, P = 0.06, d = 0.23, 95\% \text{ CI } -13.9\% \text{ to } -4.3\%$). Put another way, these results indicate that decreases in untrustworthy website exposure from 2016 to 2020 were concentrated mainly among those with liberal media diets.

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 aged

65 years and older were 1.6 times more likely to visit untrustworthy websites than adults under 30 years. During the 2020 election, it appears that the positive association between age and likelihood of exposure to untrustworthy websites persisted, with those 65 years and older being 2.1 times more likely to visit untrustworthy websites than those under 30 years. However, all age groups in 2020 were less likely to be exposed to untrustworthy websites than in 2016 (Fig. 2). For instance, while 56.2% (95% CI 49.7% to 62.7%) of people 65 years and older were exposed to untrustworthy websites in 2016, only 37.4% (95% CI 29.6% to 45.2%) of people 65 years and older were exposed in 2020.

To examine the association between a variety of individual differences and exposure to untrustworthy websites, we followed the strategy used by Guess et al.⁶ and estimated a linear probability model for the 2020 election period. In this model, 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 (substantively similar results are obtained when probit models are estimated, Supplementary Table 1). 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, while identifying as non-white was associated with a significantly lower likelihood of being exposed to at least one untrustworthy website in both years.

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. 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 of misinformation that all groups were exposed to was lower across the board.

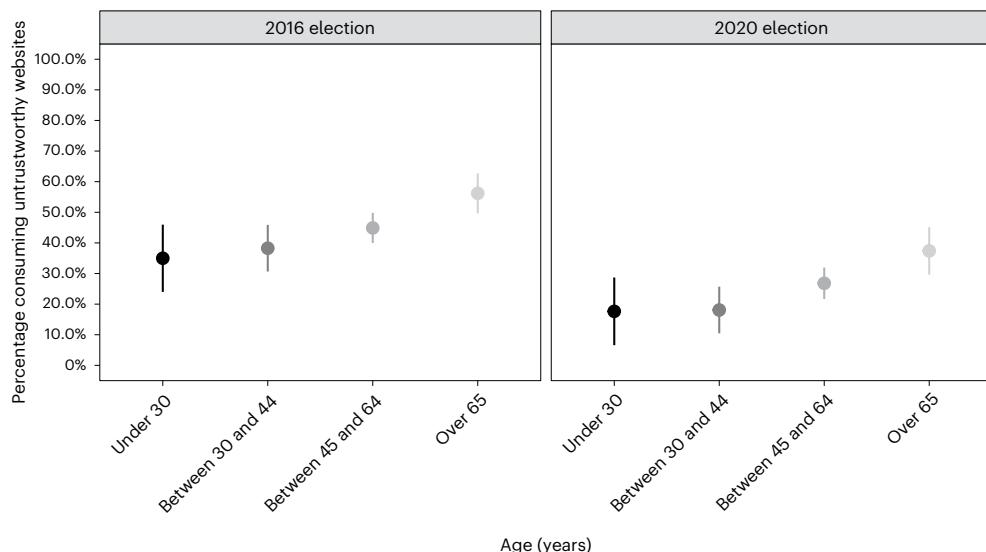


Fig. 2 | Proportion exposed to untrustworthy websites by age group.

Estimated average percentage of people in different age groups (under 30 years, between 30 and 44 years, between 45 and 64 years, and 65 years and older) who were exposed to at least one untrustworthy website ($N = 3,676$). The left

represents 2016 and 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. Data are presented as weighted mean values \pm 95% CIs. Lighter colours correspond to older age groups.

How were people exposed to untrustworthy websites in 2020?

Guess et al.⁶ examined how individuals came to visit untrustworthy websites during the 2016 election by analysing ‘referrers’ to untrustworthy websites, or the sites that were within 30 s 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 shown in Fig. 3.

In 2016, Facebook was a prominent referrer to untrustworthy websites, with 15.1% of visits to untrustworthy websites being referred by Facebook. In 2020, however, only 5.6% of visits to untrustworthy sites were referred by Facebook, a significant reduction ($\chi^2(1) = 573.4$, $P < 0.001$, $\phi = 0.10$, 95% CI -10.7 to -8.9). This suggests a potential decline in Facebook’s role in directing people to untrustworthy websites. 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 6.2% of visits referred by Google in 2020 ($\chi^2(1) = 181.2$, $P < 0.001$, $\phi = 0.05$, 95% CI 2.4 to 3.4). However, when looking at the average number of referrals to untrustworthy websites per exposed person (Supplementary Fig. 3), we see that while Facebook’s average number of untrustworthy website referrals per person fell from 6.4 in 2016 to 1.4 in 2020, Google’s average number of referrals per person stayed relatively constant, going from 1.4 in 2016 to 1.5 in 2020, suggesting caution in interpreting the proportional results indicating that Google’s role in referring people to untrustworthy websites increased from 2016 to 2020. Furthermore, Google referred a greater proportion of visits to credible, hard news websites in 2020 (8.5%) than in 2016 (6.2%) ($\chi^2(1) = 2,428.7$, $P < 0.001$, $\phi = 0.20$, 95% CI 2.3 to 2.4). In addition to Facebook, webmail (for example, Gmail) also played a smaller role in referring people to untrustworthy websites in 2020 compared with 2016, both in proportional (9.5% of visits to untrustworthy websites referred by webmail in 2016 versus 5.7% of visits in 2020; $\chi^2(1) = 130.4$, $P < 0.001$, $\phi = 0.10$, 95% CI -10.1 to -8.9) and average-per-person (an average of 4.0 referrals in 2016 versus an average of 1.4 referrals in 2020) terms.

Discussion

Our goal in this paper was to examine how exposure to misinformation online during the 2020 election compared with exposure during the

2016 election. By adopting the analytical approach of Guess et al.⁶, 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 with 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. Our analysis does not rely on self-reported media exposure measures, which tend to be inaccurate compared with passively tracked behavioural measures of news exposure^{27,28}.

Overall, we found that a significantly smaller percentage of Americans were exposed to untrustworthy websites in 2020 (26.2%) compared with 2016 (44.3%). We found this decrease despite using a database of untrustworthy sites over three times the size of the database used in Guess et al.⁶ to identify visits to untrustworthy websites in our participants’ web browsing behaviour, 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 misinformation¹¹ (but also for reasons to predict lower misinformation exposure in 2020, see refs. 17,19,21). We also observed that those who did visit untrustworthy websites in 2020 tended to visit fewer untrustworthy sites overall and spent less time on average on each site than in 2016.

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). However, 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 re-affirm 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 with 2016. Specifically, Facebook

Table 1 | Who chooses to visit untrustworthy news websites

Exposure to at least one untrustworthy website (binary)		
	2016	2020
Trump supporter	$\beta=0.164$ s.e. 0.021 $P<0.001$ 95% CI 0.122 to 0.206	$\beta=0.138$ s.e. 0.026 $P<0.001$ 95% CI 0.085 to 0.190
Political knowledge	$\beta=0.017$ s.e. 0.005 $P=0.002$ 95% CI 0.007 to 0.028	$\beta=0.049$ s.e. 0.012 $P<0.001$ 95% CI 0.026 to 0.072
Political interest	$\beta=0.048$ s.e. 0.014 $P<0.001$ 95% CI 0.022 to 0.075	$\beta=0.043$ s.e. 0.014 $P=0.004$ 95% CI 0.014 to 0.071
College	$\beta=0.040$ s.e. 0.022 $P=0.069$ 95% CI -0.003 to 0.082	$\beta=0.049$ s.e. 0.028 $P=0.079$ 95% CI -0.006 to 0.104
Female	$\beta=-0.022$ s.e. 0.020 $P=0.274$ 95% CI -0.061 to 0.017	$\beta=0.006$ s.e. 0.025 $P=0.815$ 95% CI -0.044 to 0.056
Non-white	$\beta=-0.094$ s.e. 0.021 $P<0.001$ 95% CI -0.136 to -0.052	$\beta=-0.071$ s.e. 0.029 $P=0.016$ 95% CI -0.128 to -0.014
Age 30–44 years	$\beta=0.041$ s.e. 0.032 $P=0.201$ 95% CI -0.022 to 0.104	$\beta=-0.007$ s.e. 0.041 $P=0.862$ 95% CI -0.088 to 0.074
Age 45–64 years	$\beta=0.064$ s.e. 0.029 $P=0.030$ 95% CI 0.006 to 0.122	$\beta=0.042$ s.e. 0.038 $P=0.275$ 95% CI -0.033 to 0.117
Age 65+ years	$\beta=0.129$ s.e. 0.033 $P<0.001$ 95% CI 0.064 to 0.195	$\beta=0.113$ s.e. 0.042 $P=0.009$ 95% CI 0.030 to 0.197
Constant	$\beta=0.126$ s.e. 0.046 $P=0.006$ 95% CI 0.036 to 0.216	$\beta=-0.080$ s.e. 0.058 $P=0.175$ 95% CI -0.196 to 0.036
R^2	0.092	0.105
N	2,514	1,151

Ordinary least squares regression coefficients are shown with standard errors (s.e.) 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 to 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 to 4, where 4 is people who say they pay attention to what is going on in government and politics 'most of the time' and 1 is those who pay attention 'hardly at all'. 'College' variable: 1: college graduate; 0: not a college graduate. 'Female' variable: 1: indicates identifying as a female; 0: did not indicate identifying as a female. 'Non-white' variable: 1: indicates identifying as a race other than white; 0: indicates identifying as white. P values are two-sided and not corrected for multiple comparisons.

and webmail referred significantly fewer visits to untrustworthy websites in 2020 than in 2016. While further research is needed to elucidate these platforms' changing role in propagating misinformation (see below), our findings suggest that Facebook and webmail may have played a smaller role in directing people to misinformation on the web in the 2020 election compared with in 2016.

While our data suggest that exposure to untrustworthy websites declined at the population level, our results should not be interpreted

as indicating that misinformation is somehow less of a problem than it was previously. While exposure was lower in 2020 than in 2016, many 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 influencing participation in acts of political violence (for example, the 6 January 2021 insurrection). Altogether, exposure to fewer people can still have serious consequences²⁹. 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 the population and individual levels.

We make several contributions to the misinformation literature. First, our study demonstrates the value of re-applying the same analytical approach of prior work to examine changes in a mediated process (for example, 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.⁶, 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.'s⁶ 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 that the untrustworthy website databases used by Guess et al.⁶ were primarily based on websites circulating during the 2016 election, but fake news websites are often ephemeral (that is, the domains go defunct after short periods of time^{30,31}). 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 more likely to do so in 2020. The persistence of these trends highlights the importance of examining why populations such as older adults appear to be more susceptible to online misinformation and how they can be supported through interventions and other resources to build resilience to misinformation^{32,33}. 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 online³⁴. One reason might be that the supply of misinformation is greater on the ideological right than on the left^{35,36}. Relatedly, it could be that more ideologically conservative media diets are more likely to expose users to misinformation via features such as algorithmic curation and community structures relative to liberal media diets. The left may also be more likely to circulate misinformation via modalities other than website links (for example, social media posts, memes and other image-based formats) that are more difficult to detect using URL-based web tracking methods. Our results suggest an ongoing need for future work to investigate these potential causes.

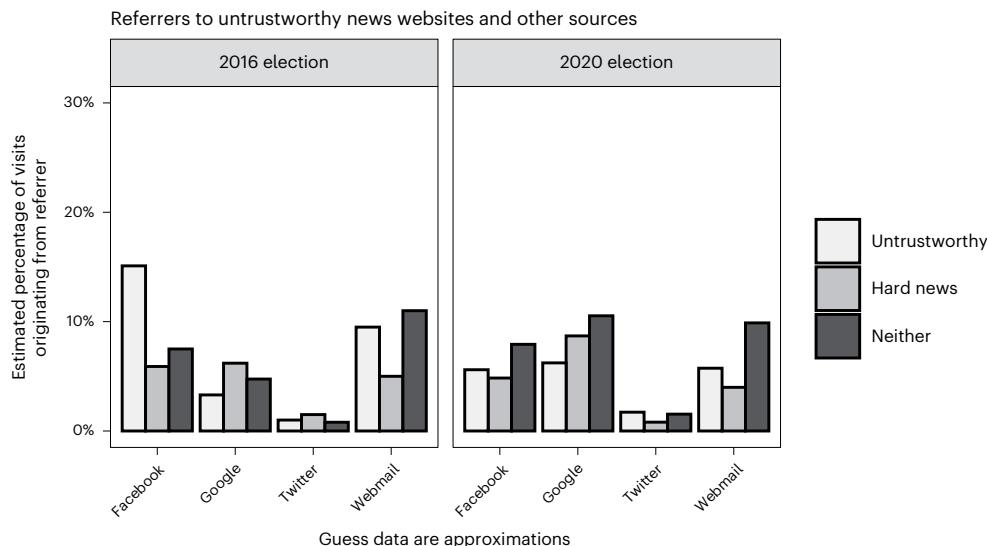


Fig. 3 | Referrers to untrustworthy news websites and other sources.

Percentage of total visits to untrustworthy websites, hard news websites and websites that were neither untrustworthy nor hard news that were referred by

different platforms (Facebook, Google, Twitter and webmail). The left represents 2016, and the right represents 2020. Higher values indicate a greater percentage of visits to each type of website that were preceded by visits to a platform.

Third, we provide evidence that Americans' visits to untrustworthy websites were less likely to be referred by Facebook in 2020 than 2016. This finding is noteworthy given the amount of scrutiny from members of Congress and the American public towards platforms, such as Facebook, roles in the proliferation of misinformation^{37,38}. Future work should endeavour to better understand why Facebook's role in referring people to untrustworthy websites appears to be shrinking. Does it indicate the efficacy with which Facebook is implementing programmes and policies to label or flag untrustworthy content? Or might it suggest a more fundamental behavioural change in how people use 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.

Finally, our findings can inform policymaking and public discourse around misinformation. Our findings re-affirm that certain groups of people are more likely to encounter misinformation, suggesting a need for more focused and directed policy initiatives centring on groups with the greatest need for support in dealing with misinformation. Moreover, our findings join others suggesting that the attention and discussion that the media, politicians and the public devote to fake news may be disproportionate to the extent that people are actually exposed to it^{7,24}. Given evidence that (1) the overwhelming majority of news consumed by the population is not misinformation³ and (2) that exposure to discourse around fake news can erode individuals' trust in news media²⁵, 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 worldwide³⁹. Our findings call for the need for more research but also more grounded communication of that research to appropriately contextualize this phenomenon.

Of course, our study should be interpreted in light of its limitations. The most noteworthy limitations relate to our URL logging methodology 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 NewsGuard rated as repeatedly hosting false information (or imported from the database used by Guess et al.⁶) (for example, www.obamawatcher.com), rather than specific web pages or articles (for

example, www.obamawatcher.com/2020/03/michelles-fake-degrees). This operationalization is largely because most 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 approach³, there is undoubtedly misinformation hosted on domains that do not repeatedly publish false information but may occasionally publish false information. Domain-level measurements of exposure do not capture these specific instances of misinformation.

A second limitation of URL-based browsing data is that it only identifies content that leads to a URL being produced⁴⁰. Crucially, this limitation means that for web pages that 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, Facebook's feed. When a user visits www.facebook.com and is presented with their feed, that user can scroll through their feed 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 feed (in either posts generated or links shared by others), we only observe instances in which individuals actually click on an external link that takes them away from their feed to a new website.

Additionally, participants collected these URL data by installing plugins in their web browsers. Thus, we only capture individuals' behaviours within web browsers. Individuals' online behaviours outside of web browsers, such as through apps, do not appear in our dataset. This nuance may be especially relevant for mobile internet use, which is more likely to occur via mobile apps than mobile web browsers⁴¹. Indeed, among our participants, we found that more individuals were exposed to an untrustworthy website on desktop/laptop computers (34.2%; 95% CI 29.4% to 39.1%) than on smartphones and tablets (13%; 95% CI 8.9% to 17.2%). As the consumption of news and political information increasingly occurs on mobile devices, it becomes all the more important for researchers to invest in methods that allow for the collection of mobile browsing and app-use data to increase the validity of inferences about online news exposure⁴².

It is important to note that these limitations affect both our 2020 data collection and Guess et al.'s⁶ 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 behaviour 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 researchers⁴³. While initiatives such as Social Science One have attempted to grant scholars access to data from social media platforms such as Facebook⁴⁴, getting academic researchers access to platform data has proven difficult^{45,46}. The limitations of our study and those that other misinformation research face highlight the importance of social media platforms working 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, other current data collection approaches could help fill in the gaps in online behaviour 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^{47,48}. The Screenomics method has been used to study exposure to political information⁴⁹. Future research on exposure to online misinformation should triangulate across several data sources to gain a more complete portrait of individuals' online media use and the role that misinformation plays in it.

Our findings indicate 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 test them more directly.

First, exposure to misinformation could have been more likely to be displaced in 2020 than in 2016 to other locations outside the web browser, such as text messaging or emergent social media apps such as WhatsApp and TikTok. Indeed, people increasingly report getting their news regularly via WhatsApp⁵⁰ and social media platforms such as Reddit and TikTok⁵¹, and there is concern about the spread of misinformation on these platforms^{52,53}.

Second, the time frame of data collection used by Guess et al. (2020) and adopted by us (4 weeks before 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 7 November, which Donald Trump refused to concede. During this time, much misinformation about election fraud circulated online⁵⁴. 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 before the election than after, we may have missed some of the misinformation relevant to the 2020 election outcome that was not present in 2016. This demonstrates that when comparing the effects of events on media consumption behaviours, even similar events (for example, 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 feed or Twitter timeline. Indeed, evidence suggests that relatively few people click on news links posted on social media websites^{55,56} yet can be influenced by the information in headlines⁵⁷. This propensity to stay on the platform over clicking to visit external websites also may be increasing over time⁵⁸. Such a change in user behaviour could also help explain why we found that Facebook played a smaller role in referring people to untrustworthy websites in 2020 than in 2016.

There are two explanations for the decline in exposure to untrustworthy websites from 2016 to 2020 that we can address with our data.

First, it is possible that the decline in untrustworthy website consumption from 2016 to 2020 reflected a broader decline in online news consumption overall, both credible and untrustworthy. Using the news websites contained in Bakshy et al.²⁶ and those rated by NewsGuard as not repeatedly publishing false content, we find that Americans' exposure to hard news websites in 2016 did not significantly differ from exposure to hard news sites in 2020. The number of Americans exposed to at least one hard news site, the average number of hard news sites accessed, and the average amount of time spent on hard news sites were similar in 2016 and 2020 (Pages 5–8 of the Supplementary Materials). These results suggest that untrustworthy news exposure uniquely declined from 2016 to 2020 while exposure to trustworthy news remained constant.

Another possibility is that consumption of untrustworthy websites fell from 2016 to 2020 because individuals' use of online fact-checking resources increased during that time. Exposure to fact-checking resources can reduce people's engagement with misinformation online^{59–63}, and using fact checks is a strategy commonly taught in effective digital media literacy interventions^{64,65}. During the 2016 election, Guess et al.⁶ found that 25.3% (95% CI 22.5% to 28.2%) of Americans visited a fact-checking website. During the 2020 election, we estimate that 13.1% (95% CI 11.2% to 15.2%) of Americans visited a fact-checking site, around half as many as in 2016 (Supplementary Fig. 4). In 2016, approximately 42% of those exposed to at least one untrustworthy website were exposed to at least one fact-checking website. In 2020, this number fell to less than 30%. In sum, it does not appear that the use of fact-checking websites increased from 2016 to 2020, either among those exposed to misinformation or the population as a whole, casting doubt on the idea that exposure to misinformation decreased from 2016 to 2020 because the use of fact-checking resources increased during that time.

In this paper, we provide evidence that exposure to untrustworthy websites decreased from 2016 to 2020. More research is needed to understand the factors explaining this change, but our results represent an important update to our understanding of exposure to online misinformation. The groups 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.

Methods

This study was approved by the Stanford University Institutional Review Board (protocol no. IRB-53941). Informed consent was obtained from all of the participants, who also received incentives from YouGov, the survey company that collected the data. All analyses were conducted using R (v4.1.2) (ref. 66).

Review of Guess et al.⁶ methods

Guess et al.⁶ is a widely cited study examining exposure to online misinformation during the 2016 US election. The authors collected data from a national sample of American adults ($N = 2,525$) during the lead-up to the 2016 election (7 October to 14 November 2016; election day was 8 November 2016) using the survey company 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. p. 479) to identify visits to untrustworthy websites in their participants' web browsing data.

Similar to most research on misinformation exposure³, Guess et al. take a domain-level approach to identifying misinformation. This approach means that, rather than identifying specific web pages that contain misinformation (for example, www.obamawatcher.com/2020/03/michelles-fake-degrees), Guess et al.⁶ rely on a list of domains that are known to frequently publish untrustworthy content (for example, www.obamawatcher.com).

By matching their list of untrustworthy websites against their participants' web browsing data, Guess et al. found that 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 years were more likely to visit untrustworthy websites. Finally, they found that exposures to untrustworthy websites frequently occurred via Facebook.

Our goal was to adopt the analytic approach of Guess et al.⁶ to analyse data collected via the same source (YouGov Pulse) during the same period (4 weeks before election day and 1 week after election day) around the 2020 election to compare exposure to untrustworthy websites during the 2020 election with exposure during the 2016 election. We do not repeat every analysis found in Guess et al.⁶ in this paper, although several additional analyses can be found in our Supplementary Materials. Similar to Guess et al.⁶, our data distributions were assumed to be normal but this was not formally tested. The central advantage of adopting Guess et al.'s⁶ analytical approach is that we can make relatively direct comparisons between our 2020 estimates and their 2016 estimates.

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 2 October 2020 to 9 November 2020 (election day was 3 November 2020). These participants also completed an online survey that 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 per cent ($n = 329$) were aged 65+ years, 48% ($n = 554$) were 45–64 years, 15% ($n = 171$) were 30–44 years and 8% ($n = 97$) were under 30 years. Forty-five per cent ($n = 522$) reported identifying as male and 54% ($n = 624$) reported identifying as female. Eighty-nine per cent ($n = 1,029$) 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 civic knowledge questionnaire⁶⁷. In total, these participants visited 7.5 million websites for our data collection period. YouGov weighted participants to approximate the demographic attributes of the US population and we use these weights in all subsequent calculations and results. 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 to protect participant confidentiality.

Unlike Guess et al.⁶, who included only desktop browsing data in their analysis of exposure to untrustworthy websites in the 2016 election, we chose to include both desktop and mobile browsing data in our analysis of exposure during the 2020 election for two reasons. First, the consumption of news and political information increasingly occurs on mobile devices, and thus it is increasingly important to document mobile exposure to (mis)information⁴². Approximately 20% more Americans reported using smartphones each week to access news in 2020 than did in 2016 (refs. 50,68), and nearly twice as many Americans report often getting their news on smartphones than on desktop/laptops⁶⁹. Second, a greater proportion of our sample had mobile browsing data available than did Guess et al.'s⁶ sample (the primary reason given by Guess et al. for not including mobile data

in their analyses⁶). In our sample, mobile browsing data was available for 30.1% of participants, while Guess et al. only had mobile browsing data for 19% of their participants⁶.

Measures

To categorize whether a website a participant visited is untrustworthy, we compiled a list of untrustworthy domains in two phases (of course, people can be exposed to misinformation outside of what is measurable in the websites they visit, for a greater discussion, see the limitations section). First, we began with the list of untrustworthy domains used by Guess et al.⁶, which consisted of 490 unique untrustworthy websites collected by previous research. We supplemented this list of 490 websites with 66 additional untrustworthy domains collected by Allcott et al.¹⁴. Then, we augmented this list of 556 websites 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.⁶ and Allcott et al.¹⁴ list of untrustworthy websites, producing a total list of 1,796 unique untrustworthy domains. Next, we matched this list of untrustworthy domains to our participants' URL-level web browsing data to identify visits to untrustworthy websites in our sample's web browsing behaviour (for an analysis of our 2020 browsing data limited to the list of 490 untrustworthy websites used in Guess et al.⁶, see pages 12–17 of Supplementary Materials).

Visits to hard news websites (for example, www.nytimes.com and www.economist.com) also played a role in Guess et al.'s analyses⁶. 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.²⁶, the same method used by Guess et al.⁶. To ensure that this list of hard news websites is up to date, we augmented this list of 500 sites with all 5,471 websites rated by NewsGuard as not repeatedly publishing false content.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available in the Open Science Framework at https://osf.io/8fy2z/?view_only=0d63bc3fd4f24938867a04efa6084552. Full web traffic histories are not available to protect participant confidentiality.

Code availability

The code that supports the findings of this study is available in the Open Science Framework at https://osf.io/8fy2z/?view_only=0d63bc3fd4f24938867a04efa6084552.

References

- Weeks, B. E. & Gil de Zúñiga, H. What's next? Six observations for the future of political misinformation research. *Am. Behav. Sci.* <https://doi.org/10.1177/0002764219878236> (2019).
- Allcott, H. & Gentzkow, M. Social media and fake news in the 2016 election. *J. Econ. Perspect.* **31**, 211–236 (2017).
- Allen, J., Howland, B., Mobius, M., Rothschild, D. & Watts, D. J. Evaluating the fake news problem at the scale of the information ecosystem. *Sci. Adv.* **6**, eaay3539 (2020).
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. & Lazer, D. Fake news on Twitter during the 2016 US presidential election. *Science* **363**, 374–378 (2019).
- Guess, A., Nagler, J. & Tucker, J. Less than you think: prevalence and predictors of fake news dissemination on Facebook. *Sci. Adv.* **5**, eaau4586 (2019).
- Guess, A., Nyhan, B. & Reifler, J. Exposure to untrustworthy websites in the 2016 US election. *Nat. Hum. Behav.* **4**, 472–480 (2020).

7. Weeks, B. E., Menchen-Trevino, E., Calabrese, C., Casas, A. & Wojcieszak, M. Partisan media, untrustworthy news sites, and political misperceptions. *New Media Soc.* <https://doi.org/10.1177/14614448211033300> (2021).
8. Nelson, J. L. & Taneja, H. The small, disloyal fake news audience: the role of audience availability in fake news consumption. *New Media Soc.* **20**, 3720–3737 (2018).
9. Li, J. & Su, M.-H. Real talk about fake news: identity language and disconnected networks of the US public's "fake news" discourse on Twitter. *Soc. Media Soc.* **6**, 2056305120916841 (2020).
10. Ball, P. & Maxmen, A. The epic battle against coronavirus misinformation and conspiracy theories. *Nature* **581**, 371–374 (2020).
11. Lytvynenko, J. In 2020, Disinformation broke the US. *BuzzFeed News* <https://www.buzzfeednews.com/article/janelytvynenko/disinformation-broke-us> (2020).
12. Lyons, T. New research shows Facebook making strides against false news. *Meta* <https://about.fb.com/news/2018/10/inside-feed-michigan-lemonde/> (2018).
13. Guess, A., Lyons, B., Montgomery, J. M., Nyhan, B. & Reifler, J. Fake news, Facebook ads, and misperceptions: assessing information quality in the 2018 US midterm election campaign <https://bpb-us-e1.wpmucdn.com/sites.dartmouth.edu/dist/5/2293/files/2021/03/fake-news-2018.pdf> (2019).
14. Allcott, H., Gentzkow, M. & Yu, C. Trends in the diffusion of misinformation on social media. *Res. Polit.* **6**, 2053168019848554 (2019).
15. Alaphilippe, A. Disinformation is evolving to move under the radar. *Brookings* <https://www.brookings.edu/techstream/disinformation-is-evolving-to-move-under-the-radar/> (2021).
16. Wilson, T. & Starbird, K. Cross-platform information operations: mobilizing narratives & building resilience through both 'big' & 'alt' tech. *Proc. ACM Hum. Comput. Interact.* **5**, 345:1–345:32 (2021).
17. Newton, C. Social networks solved their 2016 election problems, but their 2020 problems are bigger. *The Verge* <https://www.theverge.com/2020/11/6/21552763/election-2020-tech-social-platforms-facebook-twitter-youtube-misinformation> (2020).
18. Roth, Y. & Achuthan, A. 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 (2020).
19. Vu, H. T. & Saldaña, M. 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. *J. Mass Commun. Q.* **98**, 769–789 (2021).
20. Fazio, L. K. Recognizing the role of psychological science in improving online spaces. *Psychol. Sci. Public Interest* **21**, 99–102 (2020).
21. Kozyreva, A., Lewandowsky, S. & Hertwig, R. Citizens versus the internet: confronting digital challenges with cognitive tools. *Psychol. Sci. Public Interest* **21**, 103–156 (2020).
22. Munger, K. The limited value of non-replicable field experiments in contexts with low temporal validity. *Soc. Media Soc.* **5**, 2056305119859294 (2019).
23. Munger, K. Temporal validity as meta-science. Preprint at OSF <https://osf.io/hqkmr> (2022).
24. Jungherr, A. & Rauchfleisch, A. Negative downstream effects of disinformation discourse: evidence from the US. Preprint at SocArXiv <https://doi.org/10.31235/osf.io/3rzm> (2022).
25. Van Duy, E. & Collier, J. Priming and fake news: the effects of elite discourse on evaluations of news media. *Mass Commun. Soc.* **22**, 29–48 (2019).
26. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).
27. Guess, A. M. Measure for measure: an experimental test of online political media exposure. *Political Anal.* **23**, 59–75 (2015).
28. Konitzer, T. et al. Comparing estimates of news consumption from survey and passively collected behavioral data. *Public Opin. Q.* **85**, 347–370 (2021).
29. Dahlke, R. & Hancock, J. The effect of online misinformation exposure on false election beliefs. Preprint at <https://doi.org/10.31219/osf.io/325tn> (2022).
30. Lim, G. et al. Burned after reading: endless Mayfly's ephemeral disinformation campaign. *The Citizen Lab—University of Toronto* <https://citizenlab.ca/2019/05/burned-after-reading-endless-mayflys-ephemeral-disinformation-campaign/> (2019).
31. Dahlke, R., Kumar, D., Durumeric, Z. & Hancock, J. PIE Metrics: quantifying the systematic bias in the ephemerality and inaccessibility of web scraping content from URL-logged web-browsing digital trace data. Preprint at <https://doi.org/10.31219/osf.io/bkpqt> (2023).
32. Brashier, N. M. & Schacter, D. L. Aging in an era of fake news. *Curr. Dir. Psychol. Sci.* **29**, 316–323 (2020).
33. Moore, R. C. & Hancock, J. T. A digital media literacy intervention for older adults improves resilience to fake news. *Sci. Rep.* **12**, 6008 (2022).
34. Freelon, D., Marwick, A. & Kreiss, D. False equivalencies: online activism from left to right. *Science* **369**, 1197–1201 (2020).
35. Benkler, Y., Faris, R. & Roberts, H. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford Univ. Press, 2018).
36. Garrett, R. K. & Bond, R. M. Conservatives' susceptibility to political misperceptions. *Sci. Adv.* **7**, eabf1234 (2021).
37. Auxier, B. 64% of Americans say social media have a mostly negative effect on the way things are going in the US 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/> (2020).
38. Wakefield, J. Google, Facebook Twitter grilled in US on fake news. *BBC News* <https://www.bbc.com/news/technology-56523378> (2021).
39. Edelman. 2021 Edelman trust barometer. 58 <https://www.edelman.com/trust/2021-trust-barometer> (2021).
40. Ellison, N. B., Triệu, P., Schoenebeck, S., Brewer, R. & Israni, A. Why we don't click: interrogating the relationship between viewing and clicking in social media contexts by exploring the "non-click". *J. Comput.-Mediat. Commun.* **25**, 402–426 (2020).
41. Wurmser, Y. US time spent with mobile 2019. *Insider Intelligence* <https://www.emarketer.com/content/us-time-spent-with-mobile-2019> (2019).
42. González-Bailón, S. & Xenos, M. A. The blind spots of measuring online news exposure: a comparison of self-reported and observational data in nine countries. *Information Commun. Soc.* **0**, 1–19 (2022).
43. Pasquetto, I. & Swire-Thompson, B. Tackling misinformation: what researchers could do with social media data: introduction. *Harvard Kennedy School Misinformation Review* <https://doi.org/10.37016/mr-2020-49> (2020).
44. King, G. & Persily, N. A new model for industry-academic partnerships. *PS Polit. Sci. Polit.* **53**, 703–709 (2020).
45. Edelson, L. & McCoy, D. How Facebook hinders misinformation research. *Scientific American* <https://www.scientificamerican.com/article/how-facebook-hinders-misinformation-research/> (2021).
46. Persily, N. A proposal for researcher access to platform data: the platform transparency and accountability act. *J. Online Trust Saf.* <https://doi.org/10.54501/jots.v1i1.22> (2021).
47. Reeves, B. et al. Screenomics: a framework to capture and analyze personal life experiences and the ways that technology shapes them. *Hum. Comput. Interact.* **0**, 1–52 (2019).

48. Reeves, B., Robinson, T. & Ram, N. Time for the human screenome project. *Nature* **577**, 314–317 (2020).

49. Muise, D. L. *Temporality of Political Content Exposure via Smartphone Screens: A Computational Description of Rapidity and Idiosyncrasy*. Stanford Univ. (2022).

50. Newman, N., Fletcher, R., Schulz, A., Andi, S. & Nielsen, R. K. Reuters Institute digital news report 2020. *Reuters Institute at the University of Oxford* https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf (2020).

51. Walker, M. & Matsa, K. E. 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/> (2021).

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

53. Zadrozny, B. 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> (2021).

54. Frenkel, S. How misinformation 'superspreaders' seed false election theories. *The New York Times* (23 November 2020).

55. Ju, A., Jeong, S. H. & Chyi, H. I. Will social media save newspapers? *Journalism Pract.* **8**, 1–17 (2014).

56. Mitchell, A. & Jurkowitz, M. 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/> (2014).

57. Rapp, D. N. & Salovich, N. A. Can't we just disregard fake news? The consequences of exposure to inaccurate information. *Policy Insights Behav. Brain Sci.* **5**, 232–239 (2018).

58. Geeng, C., Yee, S. & Roesner, F. Fake news on Facebook and Twitter: investigating how people (don't) investigate. *Proc. of the 2020 CHI Conference on Human Factors in Computing Systems* 1–14 (Association for Computing Machinery, 2020).

59. Aslett, K., Guess, A. M., Bonneau, R., Nagler, J. & Tucker, J. A. News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions. *Sci. Adv.* **8**, eabl3844 (2022).

60. Guess, A. Fact-checking on Twitter: an examination of campaign 2014. *American Press Institute* <https://www.americanpressinstitute.org/wp-content/uploads/2015/04/Project-1B-Guess-updated.pdf> (2015).

61. Kauk, J., Kreysa, H. & Schweinberger, S. R. Understanding and countering the spread of conspiracy theories in social networks: evidence from epidemiological models of Twitter data. *PLoS ONE* **16**, e0256179 (2021).

62. Nekmat, E. Nudge effect of fact-check alerts: source influence and media skepticism on sharing of news misinformation in social media. *Soc. Media Soc.* **6**, 2056305119897322 (2020).

63. Vijaykumar, S. et al. How shades of truth and age affect responses to COVID-19 (mis)information: randomized survey experiment among WhatsApp users in UK and Brazil. *Humanit. Soc. Sci. Commun.* **8**, 1–12 (2021).

64. Brodsky, J. E. et al. Improving college students' fact-checking strategies through lateral reading instruction in a general education civics course. *Cogn. Res.* **6**, 23 (2021).

65. Wineburg, S., Breakstone, J., McGrew, S., Smith, M. D. & Ortega, T. Lateral reading on the open internet: a district-wide field study in high school government classes. *J. Educ. Psychol.* <https://doi.org/10.1037/edu0000740> (2022).

66. R Core Team. R: a language and environment for statistical computing. (2021).

67. Pew Research Center. Political engagement, knowledge and the midterms. *Pew Research Center—US Politics & Policy* <https://www.pewresearch.org/politics/2018/04/26/10-political-engagement-knowledge-and-the-midterms/> (2018).

68. Newman, N., Fletcher, R., Levy, D. A. L. & Nielsen, R. K. *Reuters Institute digital news report 2016*. *Reuters Institute at the University of Oxford* <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/Digital%2520News%2520Report%25202016.pdf> (2016).

69. Walker, M. Americans favor mobile devices over desktops and laptops for getting news. *Pew Research Center* <https://www.pewresearch.org/fact-tank/2019/11/19/americans-favor-mobile-devices-over-desktops-and-laptops-for-getting-news/> (2019).

Acknowledgements

The authors thank D. Adib-Azpeitia, T. Başar, J. Bayer, M. Gentzkow, A. Guess, C. Langbort, M. Regenwetter, J. Ugander and members of the Stanford Internet Observatory, Stanford Information Warfare Working Group, and Stanford Social Media Lab for feedback on this project. M. Farooqui provided research assistance. An earlier version of this paper was presented at the 2022 meeting of the International Communication Association in Paris, France and as part of the Stanford Cyber Policy Center's seminar series. Finally, we thank the participants who provided the data to make this study possible. This project was funded by a MURI award no. W911NF-20-1-0252 and support from Stanford's Freeman Spogli Institute for International Studies. R.C.M. is supported by a Stanford Interdisciplinary Graduate Fellowship. R.D. is supported by a graduate fellowship award from Knight-Hennessy Scholars at Stanford University. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

R.C.M. and J.T.H. conceived the study and collected the data. R.D. analysed the data and constructed the tables and figures. R.C.M. and R.D. wrote the first draft of the paper. R.C.M., R.D. and J.T.H. edited and revised the paper. All authors approved the final version of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41562-023-01564-2>.

Correspondence and requests for materials should be addressed to Ryan C. Moore.

Peer review information *Nature Human Behaviour* thanks Benjamin Lyons and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
 - Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
 - Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Data were collected by survey firm YouGov using their proprietary survey software (similar to commercially available software-as-a-service vendors Qualtrics and SurveyMonkey).

Data analysis Analyses were conducted in R (version 4.1.2).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

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.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study uses survey data linked to data about respondents' web traffic on desktop/laptop computers and smartphones. This study is quantitative
Research sample	Respondents were drawn from YouGov's proprietary panel and then matched and weighted to approximate a nationally representative sample of U.S. adults. This sample was chosen to be similar and thus comparable to the sample collected during the 2016 election by Guess et al. (2020).
Sampling strategy	We worked with YouGov to recruit a comparably nationally-representative US sample as that used in Guess et al. (2020), which was comprised of participants who both completed surveys and provided web browsing data.
Data collection	Survey data were collected online using YouGov's proprietary browser-based survey software. Respondents could complete the survey on any appropriate web- and browser-enabled device. Web traffic data were collected via YouGov's Pulse app/plugin, which respondents could install on their desktop/laptop computers and/or smartphones.
Timing	Survey responses were collected from September 18 - December 9, 2020 and web traffic were collected from October 2, 2020 - November 9, 2020. This period for web traffic data collection matches the timeframe used by Guess et al. (2020) (4 weeks prior to Election Day and 1 week after Election Day).
Data exclusions	Of the 1,515 total respondents from whom we collected survey responses, only 1,151 of those participants also had web traffic data. Our analyses were constrained to the 1,151 participants for whom we had both survey and web traffic data.
Non-participation	Recruitment of sample from YouGov's proprietary panel was handled by the vendor. We are unable to report a non-participation rate.
Randomization	Randomization did not play a role in our study as it was observational rather than experimental.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern		

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Of our 1,151 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. YouGov weighted participants to match a nationally-representative sample and we use these weights in our calculations and results.

Recruitment

Participants were recruited via email by the Internet survey firm YouGov to members of their proprietary panel. Respondents had to affirm their consent before participation. Those who participate in online survey panels may have different online behaviors than those who do not (see Muise, 2022). However, because we used the same data collection platform and panel as the Guess et al (2020) study, these potential differences in online behaviors are in expectation consistent across both of our studies, limiting the extent to which they may confound our comparisons of 2016 untrustworthy website exposure (as estimated by Guess et al.) and 2020 exposure (as estimated by us).

Ethics oversight

Stanford University Institutional Review Board (Protocol # IRB-53941)

Note that full information on the approval of the study protocol must also be provided in the manuscript.