



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Contents lists available at ScienceDirect

# Vaccine

journal homepage: [www.elsevier.com/locate/vaccine](http://www.elsevier.com/locate/vaccine)

## The sources and correlates of exposure to vaccine-related (mis)information online <sup>☆</sup>



Andrew M. Guess <sup>a,\*</sup>, Brendan Nyhan <sup>b</sup>, Zachary O’Keeffe <sup>c</sup>, Jason Reifler <sup>d</sup>

<sup>a</sup> Department of Politics, Princeton University, United States

<sup>b</sup> Department of Government, Dartmouth College, United States

<sup>c</sup> Department of Political Science, University of Michigan, United States

<sup>d</sup> Department of Politics, University of Exeter, United Kingdom

### ARTICLE INFO

#### Article history:

Received 11 June 2020

Received in revised form 1 October 2020

Accepted 7 October 2020

Available online 22 October 2020

#### Keywords:

Vaccine hesitancy  
Vaccine skepticism  
Online  
Information  
Social media  
Search

### ABSTRACT

**Objectives:** To assess the quantity and type of vaccine-related information Americans consume online and its relationship to social media use and attitudes toward vaccines.

**Methods:** Analysis of individual-level web browsing data linked with survey responses from representative samples of Americans collected between October 2016 and February 2019.

**Results:** We estimate that approximately 84% of Americans visit a vaccine-related webpage each year. Encounters with vaccine-skeptical content are less frequent; they make up only 7.5% of vaccine-related pageviews and are encountered by only 18.5% of people annually. However, these pages are more likely to be published by untrustworthy sources. Moreover, skeptical content exposure is more common among people with less favorable vaccine attitudes. Finally, usage of online intermediaries is frequently linked to vaccine-related information exposure. Google use is differentially associated with subsequent exposure to non-skeptical content, whereas exposure to vaccine-skeptical webpages is associated with usage of webmail and, to a lesser extent, Facebook.

**Conclusions:** Online exposure to vaccine-skeptical content is relatively rare, but vigilance is required given the potential for exposure among vulnerable audiences.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

Vaccine hesitancy, which has been named one of the top ten threats to global health by the World Health Organization [33], threatens to exacerbate the spread of communicable diseases in the U.S. and around the world. Many scholars identify the internet and social media as important vectors for the spread of information questioning the safety and effectiveness of vaccines, citing webpages (e.g., [35,20,1,21,23,19]), search engines such as Google

(e.g., [32,28,11]), content sharing sites such as YouTube (e.g., [22,4,9]), and social media sites such as Facebook and Twitter (e.g., [3,6,16,17,29]) as key platforms for its dissemination. Exposure to this type of dubious content may worsen people’s attitudes toward vaccines and be difficult to refute (e.g., [2,18,8,25]) – an especially important risk to understand in the context of expected future vaccines during the global COVID-19 pandemic.

Analyses that have been conducted to date overwhelmingly focus on the supply of online (mis)information about vaccines or the effects of exposure to that information. By comparison, relatively little is known about the vaccine information people actually consume online and its relationship to both their attitudes toward the topic and the communication formats and platforms that they use.

To effectively address the problem of vaccine hesitancy, it is essential to learn what vaccine-related information people actually see online and the means by which they are exposed to it. To do so, we conduct the most systematic measurement to date of vaccine-related information exposure online in the United States, analyzing survey and linked online behavioral data from seven large nationally representative samples collected from late 2016 to early 2019.

<sup>☆</sup> We thank the Institute for Healthcare Policy and Innovation at the University of Michigan, Democracy Fund, the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 682758), the Nelson A. Rockefeller Center at Dartmouth College, and the Weidenbaum Center on the Economy, Government, and Public Policy at Washington University in St. Louis for funding support and Levy for outstanding research assistance. We would also like to thank Gabrielle Levy, Will Schulz, and Shun Yamaya for helpful comments and Amelia Jamison, David A. Broniatowski, Mark Dredze, Zach Wood-Doughty, DureAden Khan, and Sandra Crouse Quinn for generously sharing their data.

\* Corresponding author.

E-mail address: [aguess@princeton.edu](mailto:aguess@princeton.edu) (A.M. Guess).

These data allow us to measure how much vaccine information is consumed online, what proportion of it is skeptical toward vaccines, and to examine how these consumption patterns relate to both people’s attitudes toward vaccines and their social media usage.

Our analyses consider three specific concerns about vaccine information online. First, we evaluate the quantity and type of vaccine-related information people consume online to identify how much of it is skeptical about the safety and efficacy of vaccines. Second, we consider the relationship between people’s attitudes or beliefs toward vaccines and the information they consume online. Many observers fear that people with strong views or misinformed beliefs will become trapped in “echo chambers” of like-minded sources online [30]. Though behavioral evidence suggests these fears are overstated [15,14,13], little is known about the extent of selective exposure to online information about vaccines. Finally, we consider the role of social media platforms, search engines, and email in exposure to vaccine-related webpages and the extent to which their use is associated with the consumption of vaccine-skeptical information.

## 2. Methods

We analyze the contents and correlates of people’s online exposure to vaccine-related information using survey responses linked with individual-level behavioral data detailing the websites that people visit. Our data come from seven nationally representative surveys we conducted among U.S. adults between October 2016 and February 2019 via the survey company YouGov. The data included survey responses as well as web browsing behavior data from the period around the survey, which was collected from participants with their informed consent. We analyze responses to the surveys for which corresponding behavioral data is available.

We also analyze the relationship between behavioral data and vaccine attitudes among the subset of respondents for whom prior survey measures of vaccine attitudes are available. The attitude data uses the mean response (adjusted for direction) to a ten-question battery adapted from Freed et al. [10] (see [Online Appendix A](#) for wording). These data were collected by one of the authors (Reifler) from a subset of respondents in separate surveys conducted from October 2–30, 2014 and July 30–September 28, 2015. Mean responses are calculated by respondent and include responses to both scales where available; “don’t know” responses are treated as missing.

The behavioral data we observe consists of a stream of URLs visited by respondents on their desktop and laptop computers during the period around each of the national surveys. These data come from YouGov survey panelists who, in addition to participating in surveys, have consented to also provide behavioral data for additional incentives. These data are linked by the vendor to these respondents’ survey data and provided to researchers for an additional fee. The data provided by the vendor to the researchers is anonymized.

We identify vaccine-related information from visited URLs using keyword filtering and human coding (see the [Online Appendix A](#) for a more detailed description). First, we attempted to scrape (i.e., extract) the text of the URLs that respondents visited during the periods in which we observe their online behavior. (Pages that were not available were likely either expired links or query results that we could not access.) Among the 74% of unique URLs whose text could be successfully scraped, we then identified all pages visited by respondents in which words with the stem “vacc” appeared three or more times in the text (a total of 3,508 unique URLs). Finally, we coded each of these pages for whether the information it provided about vaccines is skeptical about their safety and

efficacy or not, which we refer to below as “vaccine-skeptical” or “not vaccine-skeptical.” The coding protocol we employed (provided in the [Online Appendix A](#)) was developed inductively while reviewing pages from domains likely to contain vaccine-relevant content. After it was developed, an author and a research assistant then used the coding protocol to jointly classify a random subset of 100 documents as vaccine-skeptical or not; intercoder agreement was 100%. The research assistant then manually coded the remainder of the documents identified by the keyword filtering process.

Following Guess, Nyhan and Reifler [14], we then combine the URL exposure data and set of coded vaccine URLs to create two outcome measures for each survey respondent using the behavioral data associated with their survey response. First, we calculate binary indicators of whether they visited one or more vaccine-skeptical webpages and/or one or more non-skeptical webpages during the period around the survey. We also create count variables measuring how many pages of each type that they visited during the period in question. Finally, we assess the quality of these webpages using the list of untrustworthy websites compiled by Grinberg et al. [12] and the prevalence of .gov and .edu domains.

We use these data to better understand the process by which people encounter online content about vaccines. Though it is not possible to measure on-platform behavior directly, we estimate the prevalence of visits to prominent social media, email, and search websites immediately prior to visiting a vaccine-related webpage (i.e., if those sites were visited in the prior 30 s and were among the three preceding URLs). We follow an analogous approach to measure platform use immediately after a visit. We supplement this analysis by measuring visits to the Facebook pages of prominent anti-vaccine groups, an important potential source of vaccine skepticism (e.g., [7]). We construct a list of 22 of the most important anti-vaccine Facebook groups based on reporting by journalists and an analysis of anti-vaccine advertisers on Facebook [5,17,24,26,31,34]. We estimate linear regression models using ordinary least squares, pooling data across waves (i.e., combining responses from multiple surveys). Because some individuals participated in multiple surveys, it would be incorrect to treat each observation as statistically independent; we therefore use clustered standard errors to account for within-individual correlations in survey responses. In addition, we include panel fixed effects to account for any time-specific differences between survey waves. Finally, we apply survey weights constructed by YouGov so that the characteristics of our sample resemble those of the U.S. population to the maximum extent possible. (Corresponding models without survey weights are reported in the [Online Appendix A](#); generalized linear models are not estimated due to perfect separation in logistic regression models. Replication data and code will be made available online upon publication of this article.)

## 3. Results

### 3.1. Respondent characteristics

We analyze behavior data from 7320 unique individuals who participated in one or more of the YouGov surveys we conducted between October 2016 and February 2019. These individuals represent 7320 of the 12,017 unique individuals in our survey data and provided 13,494 of the 19,308 observed responses. We observe an average of 26.6 days of behavioral data per respondent among those for whom web traffic data is available.

[Table 1](#) provides specific fielding dates, sample sizes, and demographic characteristics of the participants for whom behavioral data is available and compares them to the full set of survey respondents. In general, the set of respondents for whom

**Table 1**  
Participant demographics by sample.

	Oct.–Nov. 2016	Oct.–Nov. 2017	June–July 2018	July–Aug. 2018	Oct.–Nov. 2018	Nov.–Jan. 2019	Jan.–March 2019	Full sample	Web data	Vaccine attitudes
Male	47.6%	47.6%	48.1%	48.8%	48.3%	48.7%	48.4%	48.2%	47.9%	51.9%
College	29.2%	27.2%	30.2%	29.3%	29.0%	28.8%	29.6%	29.0%	29.8%	20.0%
White	68.2%	69.1%	64.9%	65.2%	64.6%	64.1%	65.2%	65.7%	67.9%	61.2%
Parent	26.3%	28.4%	22.8%	20.7%	27.5%	26.6%	24.5%	25.7%	23.9%	21.1%
Democrat	37.2%	34.6%	38.4%	38.4%	39.5%	36.4%	34.9%	37.1%	37.8%	36.2%
Repub.	26.1%	25.0%	27.4%	26.0%	26.5%	26.8%	28.0%	26.5%	26.3%	23.5%
Median age	58	45	54	57	55	50	55	54	55	55
N	3251	2100	1718	2000	3332	4907	2000	19,308	13,494	1250
<i>Dates of data collection (mm/dd)</i>										
Surveys	10/21–10/31	11/2–11/8	6/25–7/24	7/26–7/30	10/19–10/24	11/20–12/28	2/6–2/25			
Web data	10/7–11/14	10/25–11/22	6/11–7/31	7/12–8/2	10/5–11/5	11/12–1/16	1/24–3/11			

All population proportions are calculated using sample weights except for age. Partisan statistics do not include respondents who report leaning toward one party. Vaccine attitudes data include only responses for which corresponding online behavior data is also available.

behavioral data are available is virtually indistinguishable from the total pool of survey respondents on observable demographic measures such as sex, education, race, parental status, party, and age.

We also measure prior vaccine attitudes for 1250 respondents for whom such measures are available. For this measure, the sample mean is 3.88 on a 1–5 scale where 5 indicates the most positive view of vaccines. For presentational purposes, we categorize respondents into three equally sized groups using a tercile split. The tercile of Americans whom we describe as having the “most favorable” vaccine attitudes has a mean response of 4.69; the middle tercile (which we call “somewhat favorable”) has a mean of 4; and the lowest tercile (which we call “least favorable”) has a mean of 2.95. The points of division between the groups are 3.61 and 4.37, which represent approximately the 33th and 67th percentiles of the vaccine attitudes distribution.

3.2. Prevalence and content of online vaccine information exposure

Our data indicate that 12.6% of respondents were exposed to at least one vaccine-related webpage during the period for which behavioral data is available. Assuming that exposure rate is uniform over time and respondents, it is equivalent to an annualized rate of 84% given the mean number of days of behavioral data we observe per response (26.59; see [Online Appendix A](#) for details on how we calculate this quantity). In other words, at the rate of exposure found in our data, we expect that most Americans are exposed to at least one vaccine-related webpage per year. However, exposure rates vary dramatically between vaccine-skeptical content and other types of vaccine-related information. We find that just 1.48% of respondents visited a vaccine-skeptical webpage during the periods when behavioral data was being recorded, which is equivalent to an annualized rate of 18.5% given the mean number of days of behavioral data we observe.

By contrast, the frequency of exposure to non-skeptical content is much higher — approximately 11.76% of respondents visited a webpage that was not classified as skeptical when behavioral data was being recorded, which is equivalent to an annualized rate of 82%. The volume of consumption of webpages related to vaccines is similarly skewed toward non-skeptical content. In total, 92.5% of the related pages that respondents viewed were not skeptical about vaccine safety or efficacy.

Visits to vaccine-related webpages are not common; the distribution is instead heavily skewed due to high levels of exposure among a small subset of respondents. In the period of observation, the 99th percentile of exposure to vaccine-skeptical content for an individual respondent is one page and the maximum is 72 pages. Similarly, the 95th percentile of observed exposure to

non-skeptical vaccine content is 3 pages, the 99th percentile is 10 pages, and the maximum is 249 pages.

Which sites with vaccine-related content are people visiting? [Table 2](#) lists the top non-vaccine-skeptical (left column) and vaccine-skeptical (right column) web domains in our data. Wikipedia dominates the non-skeptical list with 230 visits total from our sample, followed by the Centers for Disease Control’s CDC.gov with 64 visits. Also prominent on the list are trustworthy medical sites (Mayo Clinic, WebMD), mainstream news sources (*New York Times*, Axios, National Public Radio), and Petco, which is a source of information about pet vaccines. The vaccine-skeptical sources appear to be a combination of sites purporting to offer news and information about natural health (e.g., Mercola), blogging platforms where anyone can post (e.g., Wordpress), and untrustworthy news websites (e.g., YourNewsWire), though they also include general sites with news and opinion (e.g., azcentral.com).

To more systematically characterize the differences in informational quality between vaccine-skeptical and non-skeptical content, we measure the difference in the prevalence of websites that Grinberg et al. [12] identify as untrustworthy among the vaccine-related webpages that Americans visit. In total, we find that 21.5% of pageviews of vaccine-skeptical content come from sites that have been specifically identified as unreliable compared to only 0.2% among the non-skeptical content. Similarly, 12.4% of pageviews of non-skeptical content went to .gov or .edu domains (where informational quality is likely higher on average) compared to 0.8% for skeptical content.

**Table 2**  
Top domains for vaccine content.

Non-skeptical webpages		Vaccine-skeptical webpages	
Domain	N	Domain	N
wikipedia.org	230	mercola.com	34
cdc.gov	64	collective-evolution.com	32
isidewith.com	63	naturalnews.com	15
webmd.com	57	healthimpactnews.com	6
mayoclinic.org	51	greenmedinfo.com	5
nytimes.com	40	guacamoley.com	5
axios.com	39	wordpress.com	5
nih.gov	38	betrayalseries.com	4
petco.com	32	boredpanda.com	4
npr.org	28	deeprootsathome.com	4
cnn.com	24	nvic.org	4
medicinenet.com	24	vaccineimpact.com	4
latimes.com	21	yournewswire.com	4
nbcnews.com	20	azcentral.com	3
healthline.com	16	democracynow.org	3

Number of visits to vaccine-skeptical and non-skeptical webpages by domain in web tracking data.

### 3.3. Relationship between vaccine attitudes and information exposure

Do people seek out vaccine information that is consistent with their attitudes toward vaccines? Using prior measures of vaccine attitudes, we can see in Fig. 1 whether rates of exposure to skeptical and non-skeptical content vary by the vaccine attitudes measured in our surveys. We find a negligible difference in the overall share of visits to vaccine-skeptical pages by vaccine attitudes: 0.77% among respondents with the least favorable attitudes toward vaccines (those in the lowest tercile in the vaccine scale) compared to 0.89% among those who have the most favorable attitudes (those in the highest tercile). However, the difference is more stark for visits to non-skeptical vaccine content – nearly 19% of respondents with the most favorable attitudes toward vaccines were exposed to one or more vaccine-related webpages that were not skeptical about the safety or efficacy of vaccines compared to just 8% of respondents with the least favorable attitudes. In other words, people with more favorable attitudes toward vaccines were relatively more likely to be exposed to less skeptical information about the topic, which could be the result of people with more favorable attitudes seeking pro-attitudinal information, people with less favorable attitudes avoiding counter-attitudinal information, or some combination of the two (our data do not allow us to distinguish between these possibilities). This finding holds when we estimate a series of linear regressions controlling for demographic characteristics (Table 3).

### 3.4. Originators of online vaccine information

Following prior research on news intermediaries and incidental exposure, we explore the pathways by which people arrive at vaccine-related information online. Given concerns about the role of social media and especially Facebook in spreading vaccine skepticism and misinformation, one possible approach is to consider the association between Facebook usage and exposure to vaccine-skeptical information. In the Online Appendix A, we show that Facebook usage is associated with greater exposure to vaccine-skeptical websites conditional on basic demographic and online behavior covariates. However, this correlational approach cannot rule out the possibility that vaccine skepticism encourages greater Facebook use or that the relationship is an artifact of confounding by unobservables.

We therefore take a different approach that better leverages our behavioral data on online browsing behavior. Following Guess, Nyhan and Reifler [14], we compute the proportion of visits to prominent online intermediaries immediately before and after visits to vaccine-related or other types of websites. This approach allows us to infer the relationship between usage of these platforms and exposure to vaccine-related content. Fig. 2 breaks down the proportion of exposures to non-vaccine-skeptical webpages (light gray), vaccine-skeptical webpages (dark gray), or neither (black) that followed or preceded visits to Facebook (the most important social media platform), Google (the dominant search engine), and webmail services such as Gmail, Hotmail, and Yahoo

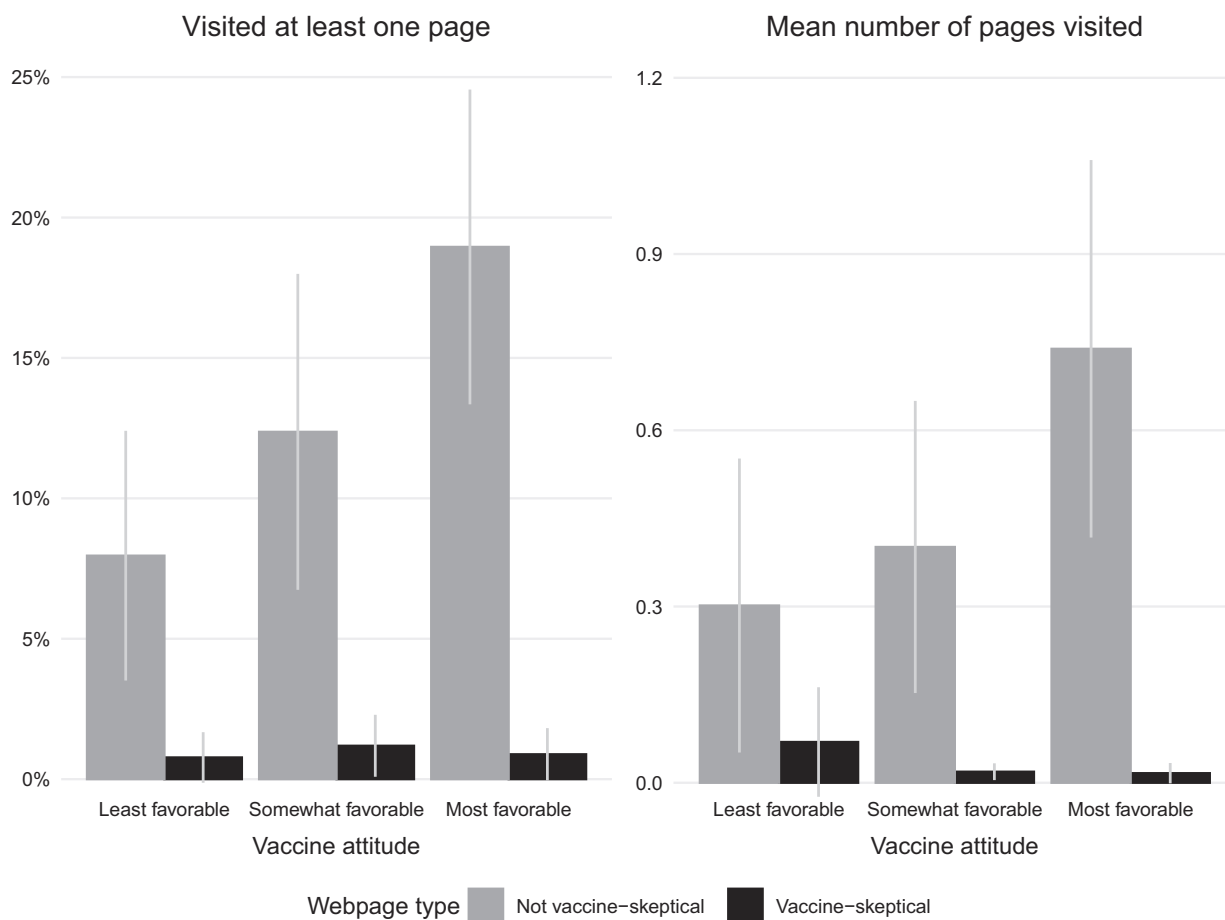


Fig. 1. Relationship between vaccine attitudes and online information exposure. Means incorporating survey weights with 95% confidence intervals. Vaccine attitudes measured on a 5-point scale and split into terciles. Online information exposure measured as binary indicators for visits to at least one webpage (left) and number of webpage visits (right). Webpage information content coded for vaccine skepticism among pages with 3 or more mentions of vaccine-related keywords (see Online Appendix A for details on coding protocol).



**Table 3**  
Correlates of exposure to online vaccine-related information (behavioral data).

	Non-skeptical webpages				Vaccine-skeptical webpages			
	Visited ≥ 1 page		Total pages visited		Visited ≥ 1 page		Total pages visited	
College graduate	0.020*	0.027	0.029	0.275	0.001	-0.008	0.003	-0.059
	(0.008)	(0.035)	(0.107)	(0.270)	(0.003)	(0.006)	(0.015)	(0.045)
Female	0.002	-0.014	0.009	0.060	0.003	0.012	0.024	0.077
	(0.009)	(0.029)	(0.110)	(0.174)	(0.003)	(0.007)	(0.017)	(0.045)
Nonwhite	-0.038***	0.007	-0.125	0.105	-0.009**	-0.017**	-0.028*	-0.085
	(0.010)	(0.033)	(0.124)	(0.213)	(0.003)	(0.006)	(0.014)	(0.050)
Age 30–44	-0.063***	-0.122*	-0.757**	-0.839	-0.009	-0.002	-0.032	0.019
	(0.016)	(0.058)	(0.236)	(0.515)	(0.005)	(0.004)	(0.021)	(0.040)
Age 45–59	-0.065***	-0.088	-0.925***	-1.021*	-0.003	0.002	-0.036	-0.030
	(0.017)	(0.059)	(0.232)	(0.507)	(0.006)	(0.004)	(0.021)	(0.033)
Age 60+	-0.081***	-0.034	-1.031***	-0.875	-0.005	0.012	-0.003	0.017
	(0.017)	(0.062)	(0.247)	(0.533)	(0.005)	(0.006)	(0.028)	(0.032)
Parent	-0.014	0.007	-0.034	0.042	-0.002	0.012	0.003	0.088
	(0.010)	(0.034)	(0.109)	(0.272)	(0.003)	(0.009)	(0.011)	(0.078)
Days in panel	0.002***	0.003	0.011	0.001	0.000	0.000	-0.001	0.001
	(0.000)	(0.002)	(0.006)	(0.008)	(0.000)	(0.000)	(0.000)	(0.001)
log(URLs visited)	0.021***	0.019	0.152***	0.192*	0.005***	0.002	0.021***	0.023
	(0.002)	(0.010)	(0.028)	(0.084)	(0.001)	(0.001)	(0.005)	(0.017)
Vaccine favorability		0.047**		0.146		0.000		-0.034
		(0.016)		(0.115)		(0.004)		(0.028)
Constant	-0.002	-0.163	0.053	-1.012	-0.010	-0.029	-0.060	-0.129
	(0.026)	(0.108)	(0.296)	(0.613)	(0.008)	(0.023)	(0.040)	(0.104)
R <sup>2</sup>	0.067	0.101	0.027	0.066	0.014	0.029	0.004	0.024
N	12,835	1177	12,835	1177	12,835	1177	12,835	1177

\*p < 0.05, \*\*p < .01, \*\*\*p < .001 (two-sided); OLS models with standard errors clustered by respondent and survey weights applied. All models include panel fixed effects. Data from YouGov panel participants with online traffic data available. Each observation represents behavioral web traffic data associated with responses to one of seven national surveys conducted from 2016–2019. Vaccine attitudes data were collected in prior surveys among a subset of respondents.

Mail. (We provide an equivalent figure including Bing and Twitter in the [Online Appendix A.](#))

The left panel of [Fig. 2](#) shows that Google appears to play an outside role in promoting exposure to non-skeptical content. Visits to Google immediately preceded 11.2% of visits to non-skeptical webpages related to vaccines compared with 6.3% of vaccine-skeptical webpages and 3.1% of unrelated content. (The most frequent search terms we observe and their frequency are listed in the [Online Appendix A.](#)) By contrast, we observe suggestive evidence of greater use of webmail sites such as Gmail immediately prior to visits to vaccine-skeptical webpages. In total, approximately 7.9% of vaccine-skeptical webpages were immediately preceded by visits to webmail sites compared to 3.8% of non-skeptical webpages (and 9.9% of unrelated webpages). Finally, we observe only modest evidence that Facebook use is associated with visits to vaccine-skeptical content. Specifically, 5.1% of visits to vaccine-skeptical webpages were immediately preceded by visits to Facebook compared to 2.8% of non-skeptical webpages (and 3.5% of unrelated webpages).

The right panel of [Fig. 2](#) presents a parallel analysis of visits to Facebook, Google, and webmail after exposure to vaccine-related webpages. We find that Google use is greatly diminished immediately after exposure to non-skeptical webpages compared to immediately before (4.9% compared to 11.2%), suggesting that it predominantly acts as a gateway to non-skeptical content. Patterns of Facebook and webmail use are more similar before and after exposure, suggesting that respondents may switch back and forth between these sites and specific pages.

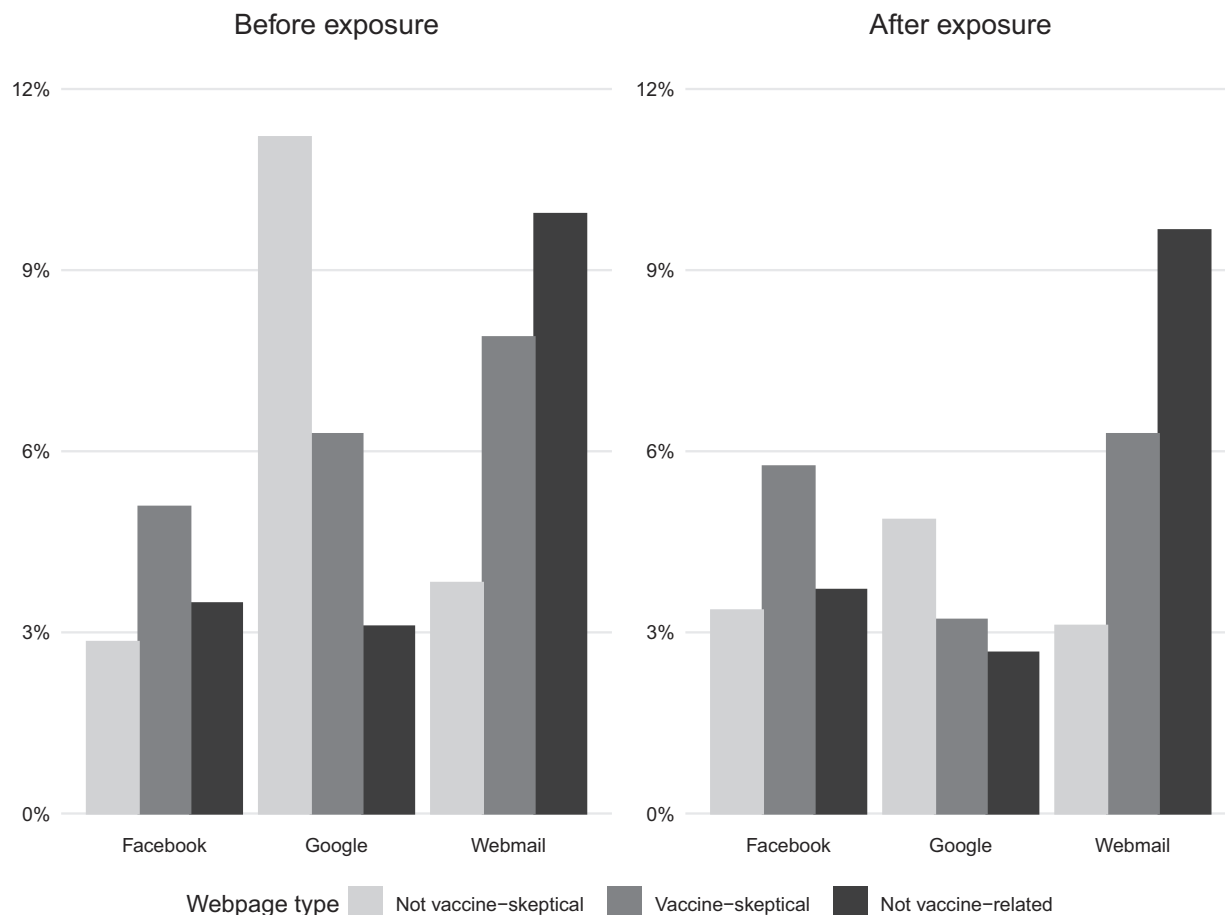
We buttress the evidence above by analyzing visits to the Facebook pages of prominent anti-vaccine groups and groups that publish anti-vaccine ads on Facebook (see the [Online Appendix A](#) for the list of groups). Though we cannot observe the contents of respondents’ news feeds, we only observe 11 respondents out of 13,494 visiting a page associated with one of these groups in their desktop or laptop web browser (i.e., loading a URL in which the group name or its Facebook ID appears in the URL of a public page or public post by

the group). Only one respondent in our data visited more than three such pages from anti-vaccine Facebook groups.

#### 4. Discussion

Our analysis of more than 134 million pageviews by our participants shows that visits to webpages that are skeptical of the safety and efficacy of vaccines are relatively uncommon. We estimate that only 18.5% of Americans visit one or more vaccine-skeptical webpages per year at an annualized rate. These skeptical webpages are more likely to be published by untrustworthy websites. Moreover, exposure rates are higher among people with the least favorable attitudes toward vaccines. Among this group, more than 18% of pageviews go to skeptical webpages compared to only 2% among people with the most favorable attitudes. Finally, we analyze the roles of social media, email, and search engines in promoting exposure to vaccine-related information. Our results indicate that Google is an especially important gateway to non-skeptical vaccine content and that exposure to vaccine-skeptical content appears differentially associated with webmail use and, to a lesser extent, Facebook use. These results indicate that both prior attitudes and online information sources play an important role in the online information that Americans consume about vaccines, providing an important pre-COVID baseline for the amount and type of vaccine information that people consume online.

Nonetheless, we face important limitations in our data that should be noted. First, we cannot observe respondent activity before, after, or between the data collection periods for our studies. Second, it is not possible for us to directly observe respondent exposure to information on Facebook or Twitter or on mobile devices. Mobile web tracking data is limited to domains only, preventing us from replicating our article-level coding procedure for those devices. In addition, mobile data is typically only available for a small share of tracking samples [14]. Future research should seek to fill in these gaps with more fine-grained data (e.g., [27]).



**Fig. 2.** Frequency of pre- and post-visit exposure to social media, search, and email. Share of exposure to pages coded as vaccine-skeptical, not vaccine-skeptical, or not vaccine-related that followed (left panel) or preceded (right panel) visits to Facebook, Google, or webmail. Webpage information content coded for vaccine skepticism among pages with 3 or more mentions of vaccine-related keywords (see [Online Appendix A](#) for details on coding protocol). Sites were identified as having been visited immediately prior to or after exposure if they were among the three visits before or after a given URL and if they appeared within 30 s of webpage exposure.

Third, we cannot measure exposure to information on mobile web browsers or within applications. Fourth, we lack data on our respondents’ offline information exposure or social connections, which may differ from what they consume online. Finally, our data do not extend to the COVID-19 pandemic; future research should examine online information consumption once a vaccine for the novel coronavirus vaccine is approved and widespread vaccination campaigns begin.

We also acknowledge potential limitations in our analytical approach. We relied on a keyword search and manual coding procedure that is feasible on a dataset with millions of observations and that allows us to achieve high face validity and intercoder reliability. However, this approach does not allow us to reliably measure every brief, subtle, or context-specific reference to vaccines.

Despite those limitations, our results identify unfulfilled opportunities for platforms to promote accurate information about vaccines. Incidental exposure to non-skeptical information about vaccines currently appears to be quite rare; many visits to non-skeptical pages appear to originate either from Google searches for vaccine-related information or purposeful visits to specific websites. Platforms could better promote incidental exposure to accurate information (e.g., via Facebook’s COVID-19 Information Center or Google’s COVID-19 information dashboard). We must also remain vigilant in monitoring the quality of vaccine-related information on the online platforms that now shape Americans’ information diets, especially as COVID-19 vaccine candidates approach regulatory approval.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Appendix A. Supplementary material**

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.vaccine.2020.10.018>.

**References**

- [1] Bean Sandra J. Emerging and continuing trends in vaccine opposition website content. *Vaccine* 2011;29(10):1874–80.
- [2] Betsch Cornelia, Renkewitz Frank, Betsch Tilmann, Ulshöfer Corina. The influence of vaccine-critical websites on perceiving vaccination risks. *J Health Psychol* 2010;15(3):446–55.
- [3] Betsch Cornelia, Brewer Noel T, Brocard Pauline, Davies Patrick, Gaissmaier Wolfgang, Haase Niels, et al. Opportunities and challenges of Web 2.0 for vaccination decisions. *Vaccine* 2012;30(25):3727–33.
- [4] Briones Rowena, Nan Xiaoli, Madden Kelly, Waks Leah. When vaccines go viral: an analysis of HPV vaccine coverage on YouTube. *Health Commun* 2012;27(5):478–85.
- [5] Broderick, Ryan. Here’s How One Of Facebook’s Biggest Anti-Vax Communities Built Its Massive Network. *BuzzFeed News*;2019, February 15, 2019. Downloaded January 21, 2020 from <https://www.buzzfeednews.com/article/ryanhatesthis/facebook-anti-vaccination-vaxxer-communities-promotions-ads>.

- [6] Chakraborty Priam, Colditz Jason B, Silvestre Anthony J, Reuel Friedman M, Bogen Katherine W, Primack Brian A. Observation of public sentiment toward human papillomavirus vaccination on Twitter. *Cogent Med* 2017;4(1):1390853.
- [7] Chiou Lesley, Catherine Tucker. Fake news and advertising on social media: a study of the anti-vaccination movement. National Bureau of Economic Research working paper; 2018. Downloaded January 21, 2019 from <https://www.nber.org/papers/w25223.pdf>.
- [8] Dunn Adam G, Leask Julie, Zhou Xujuan, Mandl Kenneth D, Coiera Enrico. Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: an observational study. *J Med Internet Res* 2015;17(6):e144.
- [9] Elkin Lucy E, Pullon Susan RH, Stubbe Maria H. Should I vaccinate my child? Comparing the displayed stances of vaccine information retrieved from Google, Facebook and YouTube. *Vaccine* 2020;38(13):2771–8. <http://www.sciencedirect.com/science/article/pii/S0264410X2030253X>.
- [10] Freed Gary L, Clark Sarah J, Butchart Amy T, Singer Dianne C, Davis Matthew M. Parental vaccine safety concerns in 2009. *Pediatrics* 2010;125(4):654–9.
- [11] Fu Linda Y, Zook Kathleen, Spoehr-Labutta Zachary, Pamela Hu, Joseph Jill G. Search engine ranking, quality, and content of web pages that are critical versus noncritical of human papillomavirus vaccine. *J Adolesc Health* 2016;58(1):33–9.
- [12] Grinberg Nir, Joseph Kenneth, Friedland Lisa, Swire-Thompson Briony, Lazer David. Fake news on Twitter during the 2016 US presidential election. *Science* 2019;363(6425):374–8.
- [13] Guess Andrew, Benjamin Lyons, Brendan Nyhan, Jason Reifler. Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to Like-minded Political News is Less Prevalent than you Think.” Knight Foundation report; 2018. February 12, 2018. Downloaded October 23, 2018 from [https://kf-site-production.s3.amazonaws.com/media\\_elements/files/000/000/133/original/Topos\\_KF\\_White-Paper\\_Nyhan\\_V1.pdf](https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/133/original/Topos_KF_White-Paper_Nyhan_V1.pdf).
- [14] Guess Andrew, Nyhan Brendan, Reifler Jason. Exposure to untrustworthy websites in the 2016 U.S. election. *Nature Human Behaviour*; 2020. Downloaded January 15, 2020 from <http://www.dartmouth.edu/bnyhan/fake-news-2016.pdf>.
- [15] Guess Andrew M. (Almost) everything in moderation: new evidence on Americans’ online media diets. *Am J Polit Sci* 2020.
- [16] Hoffman Beth L, Felter Elizabeth M, Chu Kar-Hai, Shensa Ariel, Hermann Chad, Wolynn Todd, et al. It’s not all about autism: the emerging landscape of anti-vaccination sentiment on Facebook. *Vaccine* 2019;37(16):2216–23.
- [17] Jamison Amelia M, Broniatowski David A, Dredze Mark, Wood-Doughty Zach, Khan DureAden, Quinn Sandra Crouse. Vaccine-related advertising in the Facebook Ad Archive. *Vaccine* 2020;38(3):512–20.
- [18] Jolley Daniel, Douglas Karen M. The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLoS One* 2014;9(2).
- [19] Jones Abbey M, Omer Saad B, Bednarczyk Robert A, Halsey Neal A, Moulton Lawrence H, Salmon Daniel A. Parents’ source of vaccine information and impact on vaccine attitudes, beliefs, and nonmedical exemptions. *Adv Prevent Med* 2012; 2012.
- [20] Kata Anna. A postmodern Pandora’s box: anti-vaccination misinformation on the Internet. *Vaccine* 2010;28(7):1709–16.
- [21] Kata Anna. Anti-vaccine activists, Web 2.0, and the postmodern paradigm—An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* 2012;30(25):3778–89.
- [22] Keelan Jennifer, Pavri-Garcia Vera, Tomlinson George, Wilson Kumanan. YouTube as a source of information on immunization: a content analysis. *J Am Med Assoc* 2007;298(21):2482–4.
- [23] Madden Kelly, Nan Xiaoli, Briones Rowena, Waks Leah. Sorting through search results: a content analysis of HPV vaccine information online. *Vaccine* 2012;30(25):3741–6.
- [24] Madrigal Alexis C. The Small, Small World of Facebook’s Anti-vaxxers; 2019. The Atlantic, February 27, 2019. Downloaded January 21, 2020 from <https://www.theatlantic.com/health/archive/2019/02/anti-vaxx-facebook-social-media/583681/>.
- [25] Nyhan Brendan, Reifler Jason, Richey Sean, Freed Gary L. Effective messages in vaccine promotion: a randomized trial. *Pediatrics* 2014;133(4):e835–42.
- [26] Pilkington Ed, Glenza Jessica. Facebook under pressure to halt rise of anti-vaccination groups; 2019. The Guardian, February 12, 2019. Downloaded January 21, 2020 from <https://www.theguardian.com/technology/2019/feb/12/facebook-anti-vaxxer-vaccination-groups-pressure-misinformation>.
- [27] Ram Nilam, Yang Xiao, Cho Mu-Jung, Brinberg Miriam, Muirhead Fiona, Reeves Byron, et al. Screenomics: A new approach for observing and studying individuals’ digital lives. *J Adolesc Res* 2020;35(1):16–50.
- [28] Ruiz Jeanette B, Bell Robert A. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. *Vaccine* 2014;32(44):5776–80.
- [29] Shah Zubair, Surian Didi, Dyda Amalie, Coiera Enrico, Mandl Kenneth D, Dunn Adam G. Automatically appraising the credibility of vaccine-related web pages shared on social media: a twitter surveillance study. *J Med Internet Res* 2019;21(11):e14007.
- [30] Sunstein Cass R. *Republic.com*. Princeton University Press; 2001.
- [31] Telford Taylor. Anti-vaxxers are spreading conspiracy theories on Facebook, and the company is struggling to stop them; 2019. Washington Post, February 13, 2019. Downloaded January 21, 2020 from <https://www.washingtonpost.com/business/2019/02/13/anti-vaxxers-are-spreading-conspiracy-theories-facebook-company-is-struggling-stop-them/>.
- [32] Wolfe Robert M, Sharp Lisa K. Vaccination or immunization? The impact of search terms on the internet. *J Health Commun* 2005;10(6):537–51.
- [33] World Health Organization. Ten threats to global health in 2019 Downloaded January 14, 2020 from <https://www.who.int/emergencies/ten-threats-to-global-health-in-2019>, ; 2019.
- [34] Zadrozny Brandy. Anti-vaccination groups still crowdfunding on Facebook despite crackdown; 2019. NBC News, October 11, 2019. Downloaded January 21, 2020 from <https://www.nbcnews.com/tech/internet/anti-vaccination-groups-still-crowdfunding-facebook-despite-crackdown-n1064981>.
- [35] Zimmermann Richard K, Wolfe Robert M, Fox Dwight E, Fox Jake R, Nowalk Mary Patricia, Troy Judith A, et al. Vaccine criticism on the world wide web. *J Med Internet Res* 2005;7(2):e17.