The sources and correlates of exposure to vaccine-related (mis)information online[†]

Andrew M. Guess[1], Brendan Nyhan[2], Zachary O'Keeffe[3], and Jason Reifler[4]

[1]Department of Politics, Princeton University
[2]Department of Government, Dartmouth College
[3]Department of Political Science, University of Michigan
[4]Department of Politics, University of Exeter

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.

[†]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 Gabrielle Levy for outstanding research assistance. We would also link 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.

Vaccine hesitancy, which has been named one of the top ten threats to global health by the World Health Organization (2019), 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., Zimmerman et al. 2005; Kata 2010; Bean 2011; Kata 2012; Madden et al. 2012; Jones et al. 2012), search engines such as Google (e.g., Wolfe and Sharp 2005; Ruiz and Bell 2014; Fu et al. 2016), content sharing sites such as YouTube (e.g., Keelan et al. 2007; Briones et al. 2012; Elkin, Pullon and Stubbe 2020), and social media sites such as Facebook and Twitter (e.g., Betsch et al. 2012; Chakraborty et al. 2017; Hoffman et al. 2019; Jamison et al. 2020; Shah et al. 2019) 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., Betsch et al. 2010; Jolley and Douglas 2014; Dunn et al. 2015; Nyhan et al. 2014) — 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. 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 (Sunstein 2001). Though behavioral evidence suggests these fears are overstated (Guess 2020; Guess, Nyhan and Reifler 2020; Guess et al. 2018), 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.

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. (2010) (see Online Appendix A for wording). These data were collected by one of the authors (redacted) 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 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) 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 (2020), 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. (2019) 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 seconds 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., Chiou and Tucker 2018). 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 (Broderick 2019; Jamison et al. 2020; Madrigal 2019; Pilkington and Glenza 2019; Telford 2019; Zadrozny 2019).

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; 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.)

Results

Respondent characteristics

We analyze behavior data from 7,320 unique individuals who participated in one or more of the YouGov surveys we conducted between October 2016 and February 2019. These individuals represent 7,320 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 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.

Table 1: Participant demographics by sample

	Oct.–Nov. 2016	Oct.–Nov. 2017	June–July 2018	July–Aug. 2018	Oct.–Nov. 2018	Nov. 2018– 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	3,251	2,100	1,718	2,000	3,332	4,907	2,000	19,308	13,494	1,250
Dates of data collection										
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.

We also measure prior vaccine attitudes for 1,250 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 33$^{\text{th}}$ and 67$^{\text{th}}$ percentiles of the vaccine attitudes distribution.

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. (2019) 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.

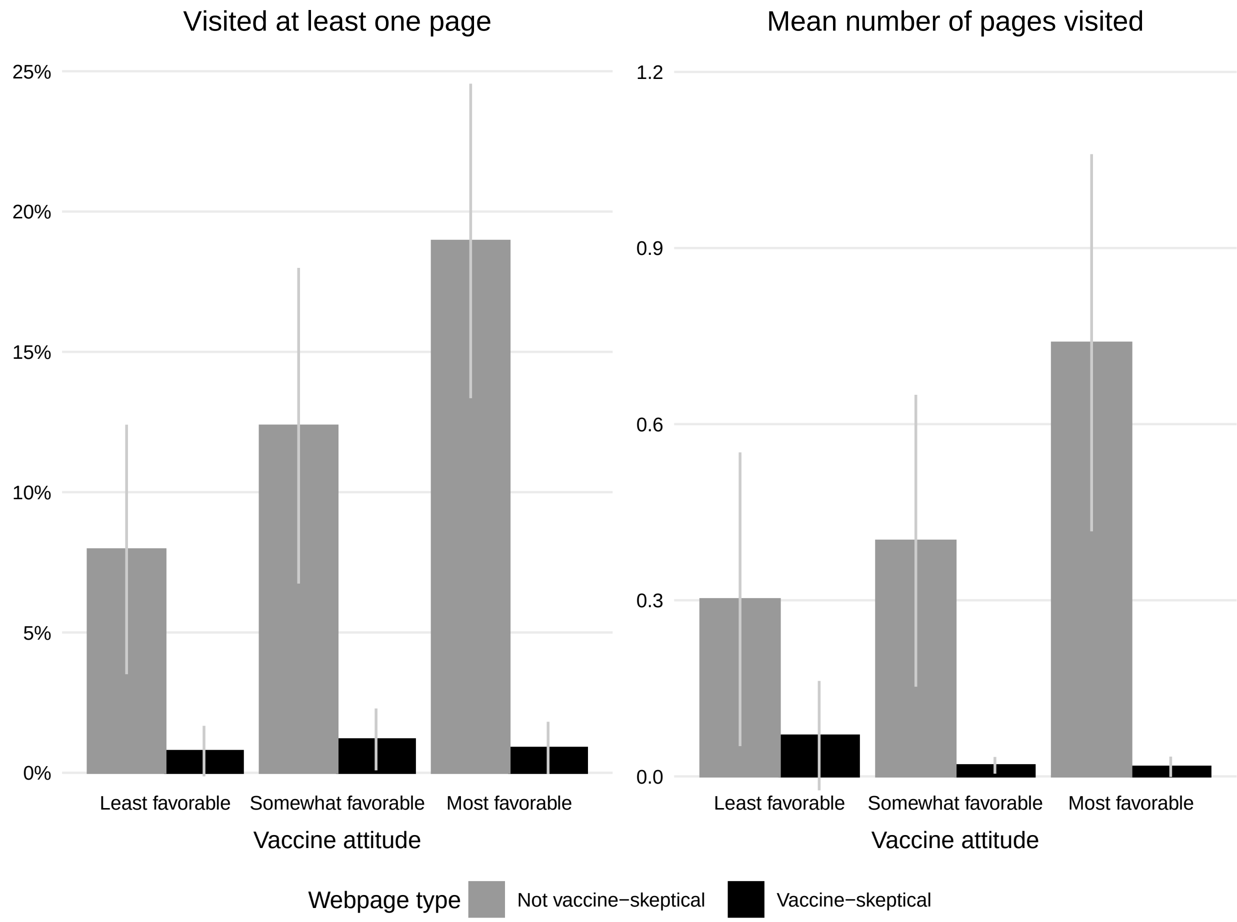
Table 2: Top domains for vaccine content

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

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 Figure 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).

Figure 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 for details on coding protocol).

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, 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.

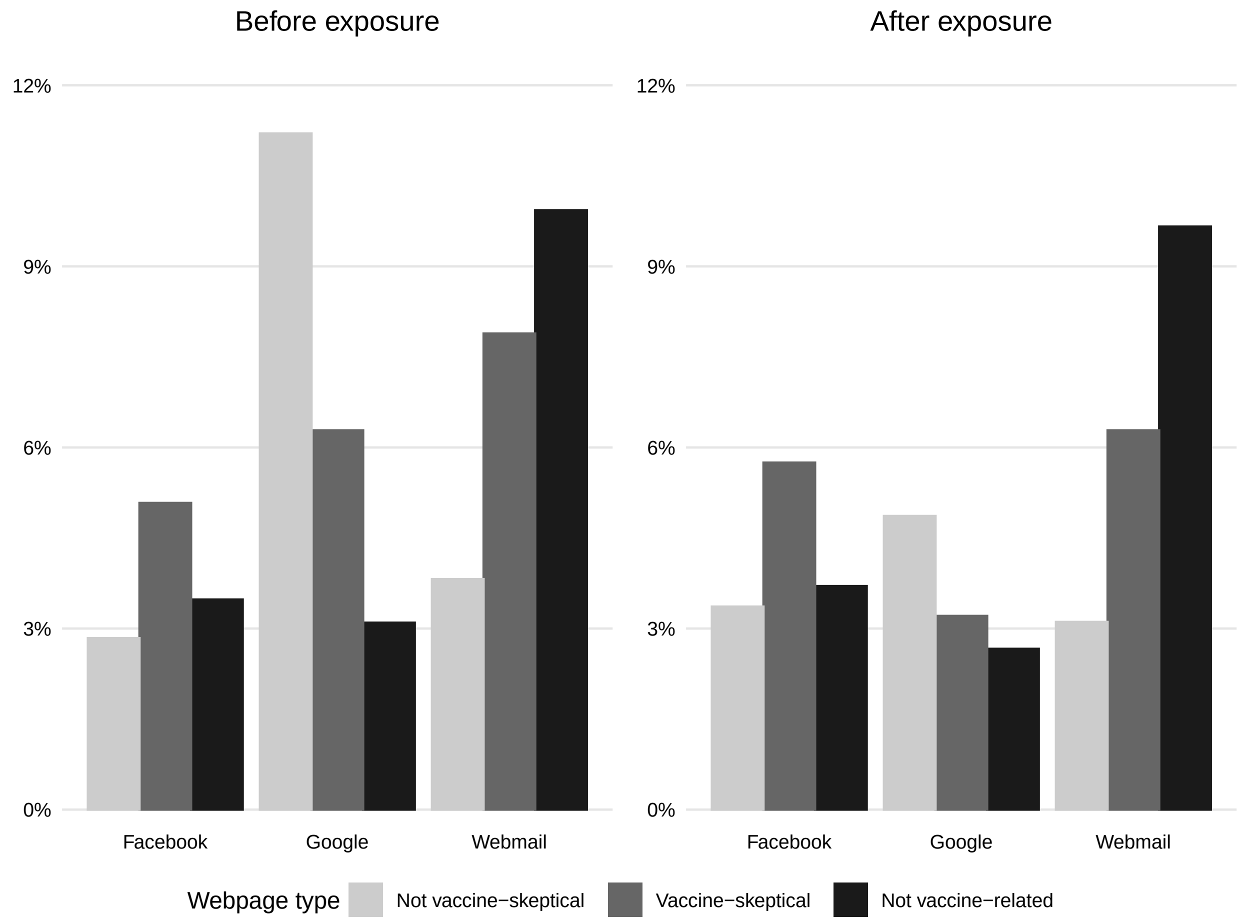
Table 3: Correlates of exposure to online vaccine-related information (behavioral data)

	Non-skeptical webpages				Vaccine-skeptical webpages			
	Visited $\geq$ 1 page		Total pages visited		Visited $\geq$ 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^2	0.067	0.101	0.027	0.066	0.014	0.029	0.004	0.024
N	12,835	1,177	12,835	1,177	12,835	1,177	12,835	1,177

$^{*}p < 0.05$, $^{**}p < .01$, $^{***}p < .001$ (two-sided); OLS models with standard errors clustered by respondent and survey weights applied. All models include panel fixed effects. Data from YouGov Pulse 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.

We therefore take a different approach that better leverages our behavioral data on online browsing behavior. Following Guess, Nyhan and Reifler (2020), 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. Figure 2 breaks down the proportion of exposures to non-vaccine-skeptical webpages (light gray), vaccine-skeptical webpages (dark gray), or neither (black)

Figure 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 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 seconds of webpage exposure.

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

The left panel of Figure 2 shows that Google appears to play an outsize 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.) 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 Figure 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 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.

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 (Guess, Nyhan and Reifler 2020). Future research should seek to fill in these gaps with more fine-grained data (e.g., Ram et al. 2020). 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.

References

Bean, Sandra J. 2011. "Emerging and continuing trends in vaccine opposition website content." *Vaccine* 29(10):1874–1880.

Betsch, Cornelia, Frank Renkewitz, Tilmann Betsch and Corina Ulshöfer. 2010. "The influence of vaccine-critical websites on perceiving vaccination risks." *Journal of Health Psychology* 15(3):446–455.

Betsch, Cornelia, Noel T. Brewer, Pauline Brocard, Patrick Davies, Wolfgang Gaissmaier, Niels Haase, Julie Leask, Frank Renkewitz, Britta Renner, Valerie F. Reyna, Constanze Rossmann, Katharina Sachse, Alexander Schachinger, Michael Siegrist and Marybelle Stryk. 2012. "Opportunities and challenges of Web 2.0 for vaccination decisions." *Vaccine* 30(25):3727–3733.

Briones, Rowena, Xiaoli Nan, Kelly Madden and Leah Waks. 2012. "When vaccines go viral: An analysis of HPV vaccine coverage on YouTube." *Health Communication* 27(5):478–485.

Broderick, Ryan. 2019. "Here's How One Of Facebook's Biggest Anti-Vax Communities Built Its Massive Network." BuzzFeed News, February 15, 2019. Downloaded January 21, 2020 from `https://www.buzzfeednews.com/article/ryanhatesthis/facebook-anti-vaccination-vaxxer-communities-promotions-ads`.

Chakraborty, Priam, Jason B. Colditz, Anthony J. Silvestre, M. Reuel Friedman, Katherine W. Bogen and Brian A. Primack. 2017. "Observation of public sentiment toward human papillomavirus vaccination on Twitter." *Cogent Medicine* 4(1):1390853.

Chiou, Lesley and Catherine Tucker. 2018. "Fake news and advertising on social media: A study of the anti-vaccination movement." National Bureau of Economic Research working paper. Downloaded January 21, 2019 from `https://www.nber.org/papers/w25223.pdf`.

Dunn, Adam G, Julie Leask, Xujuan Zhou, Kenneth D. Mandl and Enrico Coiera. 2015. "Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: an observational study." *Journal of Medical Internet Research* 17(6):e144.

Elkin, Lucy E, Susan R.H. Pullon and Maria H. Stubbe. 2020. "'Should I vaccinate my child?' Comparing the displayed stances of vaccine information retrieved from Google, Facebook and YouTube." *Vaccine* 38(13):2771 – 2778.
URL: *http://www.sciencedirect.com/science/article/pii/S0264410X2030253X*

Freed, Gary L., Sarah J. Clark, Amy T. Butchart, Dianne C. Singer and Matthew M. Davis. 2010. "Parental vaccine safety concerns in 2009." *Pediatrics* 125(4):654–659.

Fu, Linda Y, Kathleen Zook, Zachary Spoehr-Labutta, Pamela Hu and Jill G. Joseph. 2016. "Search engine ranking, quality, and content of web pages that are critical versus noncritical of human papillomavirus vaccine." *Journal of Adolescent Health* 58(1):33–39.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson and David Lazer. 2019. "Fake news on Twitter during the 2016 US presidential election." *Science* 363(6425):374–378.

Guess, Andrew, Benjamin Lyons, Brendan Nyhan and Jason Reifler. 2018. "Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to Like-minded Political News is Less Prevalent than you Think." Knight Foundation report, 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`.

Guess, Andrew, Brendan Nyhan and Jason Reifler. 2020. "Exposure to untrustworthy websites in the 2016 U.S. election." *Nature Human Behaviour* . Downloaded January 15, 2020 from `http://www.dartmouth.edu/~bnyhan/fake-news-2016.pdf`.

Guess, Andrew M. 2020. "(Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets." *American Journal of Political Science* .

Hoffman, Beth L., Elizabeth M. Felter, Kar-Hai Chu, Ariel Shensa, Chad Hermann, Todd Wolynn, Daria Williams and Brian A. Primack. 2019. "It's not all about autism: the emerging landscape of anti-vaccination sentiment on Facebook." *Vaccine* 37(16):2216–2223.

Jamison, Amelia M., David A. Broniatowski, Mark Dredze, Zach Wood-Doughty, DureAden Khan and Sandra Crouse Quinn. 2020. "Vaccine-related advertising in the Facebook Ad Archive." *Vaccine* 38(3):512–520.

Jolley, Daniel and Karen M. Douglas. 2014. "The effects of anti-vaccine conspiracy theories on vaccination intentions." *PloS one* 9(2).

Jones, Abbey M, Saad B Omer, Robert A Bednarczyk, Neal A Halsey, Lawrence H Moulton and Daniel A Salmon. 2012. "ParentsâĂŹ source of vaccine information and impact on vaccine attitudes, beliefs, and nonmedical exemptions." *Advances in preventive medicine* 2012.

Kata, Anna. 2010. "A postmodern Pandora's box: Anti-vaccination misinformation on the Internet." *Vaccine* 28(7):1709–1716.

Kata, Anna. 2012. "Anti-vaccine activists, Web 2.0, and the postmodern paradigm–An overview of tactics and tropes used online by the anti-vaccination movement." *Vaccine* 30(25):3778–3789.

Keelan, Jennifer, Vera Pavri-Garcia, George Tomlinson and Kumanan Wilson. 2007. "YouTube as a source of information on immunization: a content analysis." *Journal of the American Medical Association* 298(21):2482–2484.

Madden, Kelly, Xiaoli Nan, Rowena Briones and Leah Waks. 2012. "Sorting through search results: a content analysis of HPV vaccine information online." *Vaccine* 30(25):3741–3746.

Madrigal, Alexis C. 2019. "The Small, Small World of Facebook's Anti-vaxxers." *The Atlantic*, February 27, 2019. Downloaded January 21, 2020 from `https://www.theatlantic.com/health/archive/2019/02/anti-vaxx-facebook-social-media/583681/`.

Nyhan, Brendan, Jason Reifler, Sean Richey and Gary L. Freed. 2014. "Effective messages in vaccine promotion: A randomized trial." *Pediatrics* 133(4):e835–e842.

Pilkington, Ed and Jessica Glenza. 2019. "Facebook under pressure to halt rise of anti-vaccination groups." 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.

Ram, Nilam, Xiao Yang, Mu-Jung Cho, Miriam Brinberg, Fiona Muirhead, Byron Reeves and Thomas N Robinson. 2020. "Screenomics: A new approach for observing and studying individuals' digital lives." *Journal of Adolescent Research* 35(1):16–50.

Ruiz, Jeanette B. and Robert A. Bell. 2014. "Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information." *Vaccine* 32(44):5776–5780.

Shah, Zubair, Didi Surian, Amalie Dyda, Enrico Coiera, Kenneth D. Mandl and Adam G. Dunn. 2019. "Automatically Appraising the Credibility of Vaccine-Related Web Pages Shared on Social Media: A Twitter Surveillance Study." *Journal of Medical Internet Research* 21(11):e14007.

Sunstein, Cass R. 2001. *Republic.com*. Princeton University Press.

Telford, Taylor. 2019. "Anti-vaxxers are spreading conspiracy theories on Facebook, and the company is struggling to stop them." *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/.

Wolfe, Robert M. and Lisa K. Sharp. 2005. "Vaccination or immunization? The impact of search terms on the internet." *Journal of Health Communication* 10(6):537–551.

World Health Organization. 2019. "Ten threats to global health in 2019." Downloaded January 14, 2020 from https://www.who.int/emergencies/ten-threats-to-global-health-in-2019.

Zadrozny, Brandy. 2019. "Anti-vaccination groups still crowdfunding on Facebook despite crackdown." 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.

Zimmerman, Richard K., Robert M. Wolfe, Dwight E. Fox, Jake R. Fox, Mary Patricia Nowalk, Judith A. Troy and Lisa K. Sharp. 2005. "Vaccine criticism on the world wide web." *Journal of Medical Internet Research* 7(2):e17.

Online Appendix

Vaccine attitude question battery

Item wording:

- Getting vaccines is a good way to protect children from disease.
- Generally, I do what my doctor recommends about vaccines.
- New vaccines are recommended only if they are safe.
- Children do not need vaccines for diseases that are not common anymore.
- Enough children are vaccinated that even unvaccinated children are safe from disease.
- I am concerned about serious side effects of vaccines.
- Some vaccines cause autism in healthy children.
- Parents should have the right to refuse vaccines required for schools for any reason.
- The risks that vaccinations pose to children outweigh their health benefits.
- Vaccinations are one of the most significant achievements in improving public health.

Response options:
-Strongly agree (5)
-Agree (4)
-Neither agree nor disagree (3)
-Disagree (2)
-Strongly disagree (1)
-Don't know (only provided as a response option in 2015 data; coded as missing if selected)

Keyword subsetting procedure

1. First, webpage HTML is filtered through the `boilerpipeR` package in R. Specifically, the command `DefaultExtractor` is used to extract text.
2. Text is converted to individual words by splitting the string apart at spaces using `strsplit(x, '[[:space:]]', perl = TRUE)`.
3. Using regular expressions, we count the number of times "anti-vax" and "vaccin" appear in the text with the `grepl` command. Case is ignored. The search term "vaccin" allows for a variety of matches, including, e.g., "Vaccination", "anti-vaccine", "vaccines", etc.
4. A webpage is considered relevant for inclusion if it contains at least three vaccine-related keywords.

Vaccine content coding protocol

We use a binary classification procedure for the text of each webpage participants visited in which at least three vaccine-related keywords were included. Specifically, each text is coded as being negative/skeptical toward vaccines or not.

Coding as negative/skeptical

We use the terms "negative" and "skeptical" rather than "anti-vaccine" because they better capture the full set of messages to which people are likely to be exposed that could potentially harm public confidence in vaccines. "Anti-vaccine" implies a straightforward and possibly total opposition to vaccination, which is a rare viewpoint that is not commonly expressed in public communication. By contrast, skeptical or negative messages could undermine faith in approved vaccines even if they do not reflect an explicitly anti-vaccine viewpoint. If an article is skewed towards negative information and portrays vaccines as unnecessary or dangerous for the average person, or otherwise questions vaccine schedules, it should be coded as negative/skeptical. For instance, a message that acknowledges the potential benefits of vaccines but heavily emphasizes the risks could promote skepticism and vaccine hesitancy despite not being "anti-vaccine" as such. This category should also include messages that promote conspiracy theories about government agencies or vaccine manufacturers or messages that emphasize ineffectiveness (even if the argued cost is minimal). Common vaccine-skeptical claims include: people or animals are "over-vaccinated" or vaccinated too early; natural immunity is preferable to vaccination; chemicals or other ingredients in vaccines are dangerous (e.g., cause autism or death); vaccines are simply ineffective and unnecessary; vaccines are used for population control or part of some other nefarious plot.

Coding as NOT negative/skeptical

By contrast, we expect that the majority of vaccine-related content that participants encountered will be either neutral or supportive of vaccines and vaccination, reflecting the scientific and policy consensus that the vaccines available to the public are largely safe and effective, albeit imperfect. These articles should be coded as NOT negative/skeptical. Similarly, if vaccines are mentioned uncritically (if not positively) in passing or are characterized as being a medical necessity with some potential downsides, the page should be coded as NOT negative/skeptical as it is at least neutral. (We do not have a "mixed" category; a message must be coded as being, on balance, negative/skeptical toward vaccines or not.) Many of these non-skeptical pages will simply contain information about vaccines in passing, possibly as part of a medical page (e.g., vaccine histories, school or travel requirements, or recommendations while taking certain medications) or animal-related page (e.g., adoption information). Others may be news articles documenting, e.g., falling rates of vaccination, or scientific articles about a particular vaccine. Still other pages may include straightforward endorsements that urge readers to get vaccinations.

Complications and specific cases

1. Occasionally, a page will contain arguments from both pro- and anti-vaccine advocates (e.g., in a forum or comment section). Unless the overall orientation of the page is clearly positive,

these balanced or mixed formats should be coded as negative/skeptical given the prevalence of anti-vaccine viewpoints relative to the scientific and policy consensus.

2. Note that the coding rules above do not apply directly to simple descriptions of positions; describing a person or entity's position as anti-vaccine is different from putting forth a position that is anti-vaccine. It is therefore important to consider the speaker of the message as well as its contents.

3. It may be difficult to decide how to code seemingly neutral content. If the page portrays vaccines as potentially dangerous or not medically necessary (i.e., an agnostic position regarding vaccination), the page should be coded as negative/skeptical. If neutral simply means noting that there are rare complications and that not all vaccines are 100% effective (a more scientific and dispassionate view that reflects the scientific and policy consensus), the page should be coded as NOT vaccine skeptical. Similarly, if the page seems neutral because it lacks a recommendation—e.g., simply mentioning vaccines in passing without a position being stated—it should be coded as NOT vaccine skeptical.

4. It is important to flag an approach that is frequently used by anti-vaccine messages: cherry-picking. In some cases, the medical community acknowledges that a particular vaccine has severe side effects. If the article focuses on the danger, however, it could be considered scare-mongering and could potentially fuel fears of other vaccines. As such, even medically accurate content may be coded as negative/skeptical toward vaccines depending on how it is portrayed. Consideration of tone is key; reading the passages surrounding the actual mention carefully may be required to properly understand the context of the message.

5. It is critical to emphasize that this coding procedure does not consider the accuracy of the information in the articles, which is difficult to verify for thousands of articles. (Given the scientific and policy consensus, we expect misinformation to be more common in anti-vaccine content, but some pro-vaccine articles will include false claims and some anti-vaccine articles will contain accurate information that leaves a misleading impression or is taken out of context.)

6. It may be difficult to infer tone for certain messages. For example, a piece might be satirical. If anti-vaccine advocates are the subject of the joke, it should be coded as NOT negative/skeptical. If pro-vaccine advocates are being satirized, it should be coded as negative/skeptical.

7. Deciding how to code articles that describe the position of anti-vaccine advocates or individuals who express skepticism about the necessity of complete vaccination can be especially difficult. If the page is, e.g., a Wikipedia article stating a person's position, it should generally be coded as NOT negative/skeptical. However, if the article is more in the style of a news article, it should be coded as negative/skeptical if it shares a skeptical view without clearly stating the medical consensus that vaccines are not, in general, harmful. (Most pages with descriptions of vaccine skeptics should be coded as NOT negative/skeptical as they are usually accompanied by the dominant view that vaccination is good.)

8. If the page includes a call for people to report vaccine injuries (e.g., a legal advertisement), it should be coded as vaccine-skeptical. Cases like this emphasize that vaccines are dangerous.

9. Sometimes there is very little to go on to make the determination about whether the page should be considered vaccine-skeptical. For example, a page may simply include headlines or names of articles. If the relevant text hints at skepticism regarding vaccines (e.g., by positing a connection between them and some disease), the page should be coded as vaccine-skeptical.

10. If a page emphasizes the importance of individual and/or parental choice in getting vaccinated (i.e., treating it as an individual choice as opposed to recommending essentially full vaccination, which is necessary for effective herd immunity), it should be coded as vaccine-skeptical, even if it does not directly condemn vaccines wholesale. An argument need not be put forth about vaccines being dangerous necessarily. The individual liberty argument itself may serve as a cover for anti-vaccine skepticism, and at at any rate promotes vaccine hesitancy.

11. If a message laments some minor aspect about vaccines, e.g., the cost, it should not necessarily be coded as anti-vaccine unless vaccines themselves are questioned as efficacious or safe. For example, an author may express the belief that vaccines are too expensive. However, the message should only be coded as vaccine-skeptical if the vaccines themselves are considered unnecessary or dangerous.

12. Cases where page text was found to be unstable across time were dropped.

Negative/skeptical coding examples

Messages that are negative/skeptical toward vaccines are most easily identified by their tone, which is often angry, anxious, or condemnatory. Severe consequences are often emphasized, or the science or success of vaccines is questioned. The author may recommend against vaccinating, possibly condemning vaccination in broad strokes based on individual cases. These messages often express deeper skepticism toward the scientific method or related institutions (e.g., government, pharmaceutical companies, academia, etc.). We provide four examples below.

> Correspondent Dan Rather reports that in evaluating the potential danger of smallpox, the Bush administration has faced a deadly dilemma: Do not vaccinate the population against small pox and leave millions of people vulnerable to one of the worst scourges known to man. Or treat people with a vaccine that is extremely effective at blocking the disease but can cause dangerous, sometimes fatal, reactions...
>
> The vaccine effectively immunizes against smallpox. But that protection has a price. Some people die from it; and others have serious reactions, some permanent. Scientists say it's the most dangerous vaccine known to man.
>
> It could protect Americans from the unthinkable destruction of a smallpox attack. But the vaccine has a dark side.
>
> "We know if we immunize a million people, that there will be 15 people that will suffer severe, permanent adverse outcomes and one person who may die from the vaccine," says Dr. Paul Offit, one of the country's top infectious disease specialists, and he knows all about vaccines that prevent those diseases. In his lab at Children's Hospital of Philadelphia, he studies and creates new vaccines. There's nothing new about the smallpox vaccine.

Flu Remedies are a very hot topic these days especially since the H1N1 or Swine Flu scare. Most people place their trust in the flu vaccine to keep them safe; Not very effective as I will discuss shortly...

As you know you can get vaccines for the flu although I personally am not sold on their effectiveness.

There are over 200 variations of flu and they are constantly changing or mutating.

This makes it very difficult if not impossible for scientists to achieve any success in their formulations.

The fact is current vaccines are developed based on the previous year's virus and often are not the protection we expect and are sold on. I much prefer natural flu remedies that help your body strengthen it's own natural immunity rather than chemicals that try to "trick" an immune response.

35 Facts and Reasons Why I Became An Avid Ex-Vaxxer and No Longer Vaccinate...

So what happens when we start to question those in authority, especially in a subject so controversial, such as vaccination? What happens when we begin to think for ourselves, instead of feeding into the lies that we've been told for years? What happens when we finally listen to our "motherly instinct" and go against the grain to find the truth? What happens when we finally do some independent research and discover the truth behind the vaccine industry and its corruption? WE BECOME EX-VAXXERS.

Home Sci/Environment Flu Vaccines Are Killing Senior Citizens, Study Warns Flu Vaccines Are Killing Senior Citizens, Study Warns

A JAMA study has found that the flu vaccine, taken by 60% of people over 65-years-old, may be killing a significant number of senior citizens.

Sharyl Attkisson, a former investigative journalist for CBS, says the study shows there is no improvement in mortality rates among senior citizens who get flu shots, and may actually contribute to increased ill-health and death.

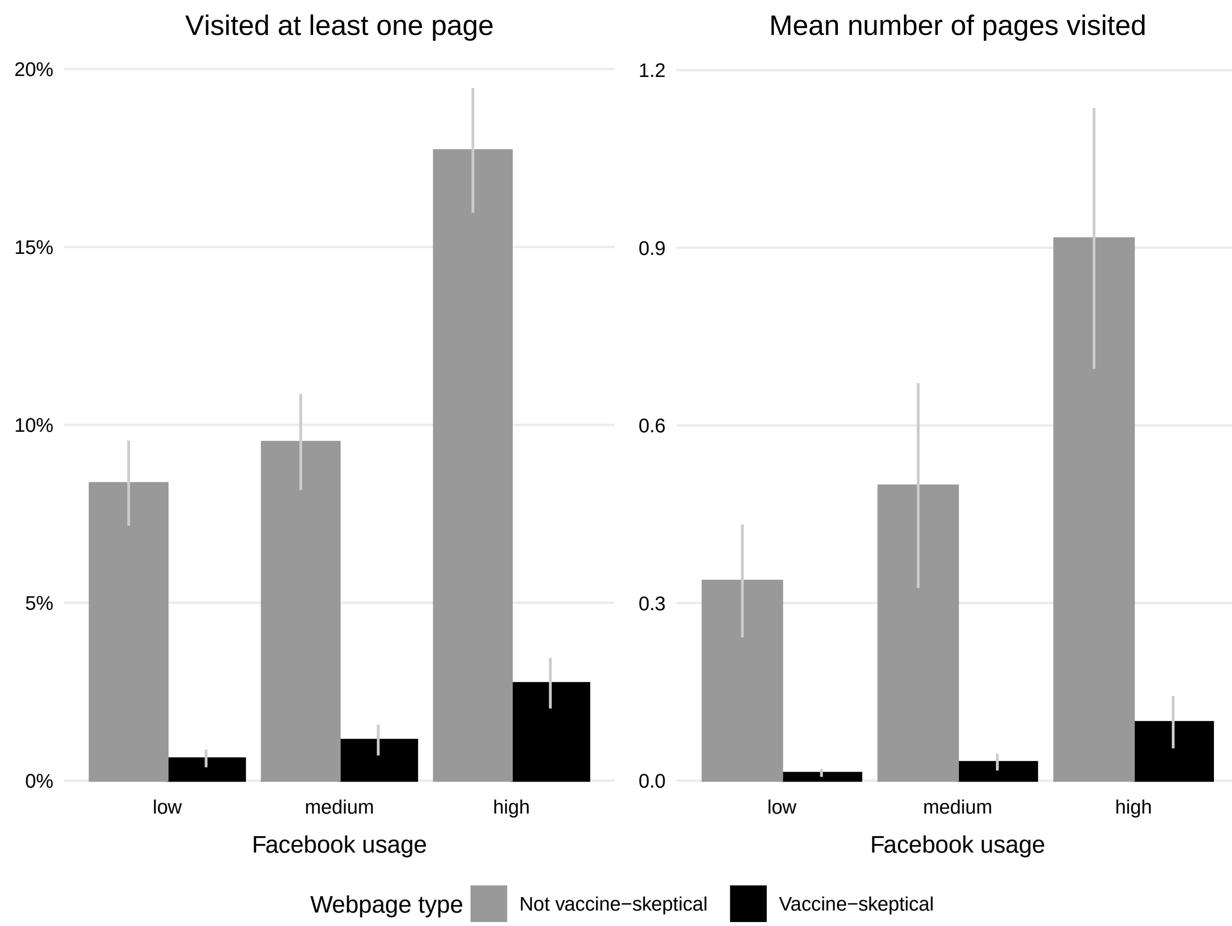
Inquisitr.com reports: The study "got little attention," she says, "because the science came down on the wrong side." Whereas the researchers had set out to prove that the push for massive flu vaccination would save the world, the researchers were "astonished" to find that the data did not support their presupposition at all. The data actually shows that deaths increased, not decreased, among seniors following vaccination.

Annualized rate calculation

In the main text, we convert the exposure rate to vaccine-related information in the period for which we have data to an estimated annualized rate. For instance, we find that 1.48% of respondents visited a vaccine-skeptical webpage, which we estimate is equivalent to an annualized rate of 18.5%. To calculate these quantities, we extrapolate the annual exposure rate e_{annual} from the sample rate of exposure in the period of data we observe, which we denote e_{observed}. Assuming this rate is uniform over time and respondents, we extrapolate the mean days of behavioral data we observe to a full year using the average number of days of behavioral data per respondent, d_{observed}, which is 26.59. These assumptions yield the following equation, which we use to calculate the estimated annualized rate of exposure:

$$e_{\text{annual}} = 1 - (1 - e_{\text{observed}})^{(365/d_{\text{observed}})}$$

Figure A1: Vaccine information exposure by Facebook usage



Means incorporating survey weights with 95% confidence intervals. 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 for details on coding protocol). Facebook usage measured by number of visits and split by terciles into low, medium, and high (cutpoints: 2, 85).

Table A1: Correlates of exposure to online vaccine-related information (unweighted)

	Non-skeptical webpages				Vaccine-skeptical webpages			
	Visited $\geq$ 1 page		Total pages visited		Visited $\geq$ 1 page		Total pages visited	
College graduate	0.020**	0.017	0.050	0.115	−0.001	−0.008	−0.004	−0.039
	(0.007)	(0.023)	(0.074)	(0.181)	(0.003)	(0.006)	(0.018)	(0.040)
Female	0.015*	−0.010	0.108	0.134	0.006*	0.014*	0.032	0.065*
	(0.007)	(0.021)	(0.067)	(0.158)	(0.003)	(0.007)	(0.018)	(0.031)
Nonwhite	−0.030***	−0.018	−0.173*	0.013	−0.008***	−0.015**	−0.032**	−0.074
	(0.007)	(0.020)	(0.073)	(0.232)	(0.002)	(0.005)	(0.010)	(0.045)
Age 30–44	−0.042***	−0.077	−0.657**	−0.896	−0.007	0.000	−0.021	0.031
	(0.012)	(0.053)	(0.241)	(0.815)	(0.004)	(0.005)	(0.021)	(0.037)
Age 45–59	−0.042***	−0.084	−0.738**	−1.061	−0.003	0.003	−0.030	−0.012
	(0.012)	(0.051)	(0.249)	(0.772)	(0.004)	(0.005)	(0.019)	(0.038)
Age 60+	−0.061***	−0.065	−0.834**	−1.113	0.001	0.014	0.013	0.018
	(0.012)	(0.052)	(0.257)	(0.765)	(0.005)	(0.007)	(0.025)	(0.028)
Parent	−0.008	−0.014	0.001	−0.223	0.002	0.009	0.014	0.073
	(0.009)	(0.029)	(0.088)	(0.192)	(0.003)	(0.009)	(0.013)	(0.083)
Days in panel	0.002***	0.001	0.010**	−0.002	0.000	0.000	0.000	0.001
	(0.000)	(0.001)	(0.003)	(0.005)	(0.000)	(0.000)	(0.000)	(0.001)
log(URLs visited)	0.025***	0.025***	0.142***	0.187**	0.005***	0.002	0.023***	0.025
	(0.002)	(0.006)	(0.016)	(0.063)	(0.001)	(0.001)	(0.005)	(0.017)
Vaccine favorability		0.047***		0.105		−0.004		−0.040
		(0.014)		(0.123)		(0.005)		(0.031)
Constant	−0.044**	−0.159	0.074	−0.403	−0.012	−0.014	−0.063	−0.098
	(0.017)	(0.082)	(0.296)	(0.475)	(0.007)	(0.029)	(0.048)	(0.081)
R^2	0.060	0.082	0.019	0.050	0.014	0.023	0.005	0.023
N	12,835	1,177	12,835	1,177	12,835	1,177	12,835	1,177

$^{*}p < 0.05$, $^{**}p < .01$, $^{***}p < .001$ (two-sided); OLS models with standard errors clustered by respondent. All models include panel fixed effects. Data from YouGov Pulse 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.

Table A2: Facebook use and online vaccine information exposure (behavioral data)

	Non-skeptical webpages		Vaccine-skeptical webpages	
	Visited $\geq$ 1 page	Total pages visited	Visited $\geq$ 1 page	Total pages visited
College graduate	0.021*	0.033	0.002	0.004
	(0.008)	(0.108)	(0.003)	(0.015)
Female	0.001	−0.001	0.002	0.020
	(0.009)	(0.106)	(0.003)	(0.016)
Nonwhite	−0.037***	−0.119	−0.008*	−0.025
	(0.010)	(0.126)	(0.003)	(0.014)
Age 30–44	−0.063***	−0.757**	−0.009	−0.033
	(0.016)	(0.236)	(0.005)	(0.021)
Age 45–59	−0.065***	−0.931***	−0.003	−0.039
	(0.017)	(0.234)	(0.006)	(0.022)
Age 60+	−0.082***	−1.039***	−0.005	−0.007
	(0.017)	(0.249)	(0.005)	(0.028)
Parent	−0.015	−0.039	−0.003	0.001
	(0.010)	(0.110)	(0.003)	(0.011)
Days in panel	0.002***	0.010	0.000	−0.001
	(0.000)	(0.006)	(0.000)	(0.000)
log(URLs visited)	0.020***	0.140***	0.004***	0.016***
	(0.003)	(0.034)	(0.001)	(0.004)
log(Facebook pages)	0.002	0.021	0.002*	0.009*
	(0.002)	(0.027)	(0.001)	(0.004)
Constant	0.000	0.074	−0.008	−0.051
	(0.026)	(0.299)	(0.008)	(0.039)
R^2	0.067	0.027	0.015	0.004
N	12,835	12,835	12,835	12,835

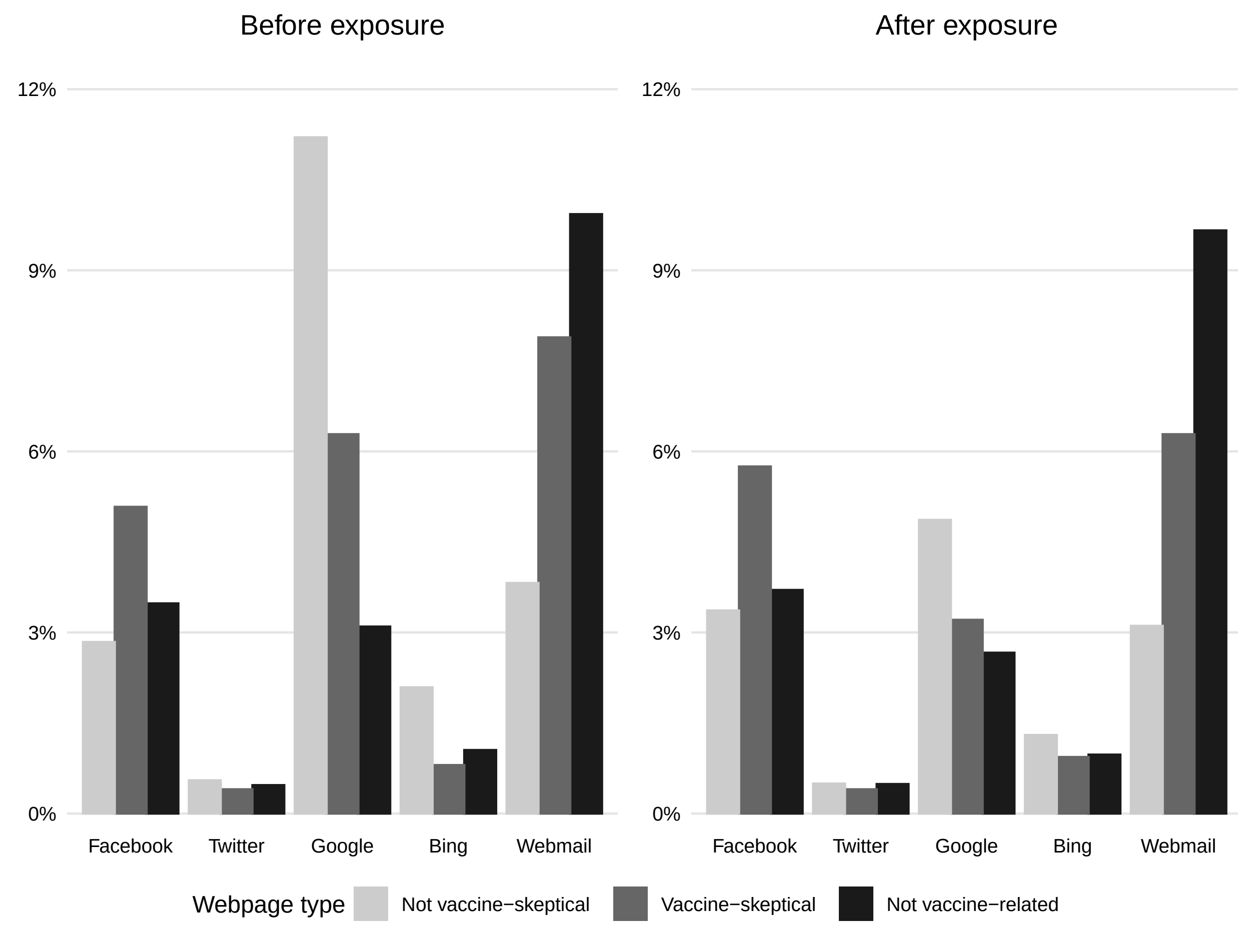
$^{*}p < 0.05$, $^{**}p < .01$, $^{***}p < .001$ (two-sided); OLS models with standard errors clustered by respondent and survey weights applied. All models include panel fixed effects. Data from YouGov Pulse panel participants with online traffic data available. Each observation represents behavioral web traffic data associated with responses to one of seven national surveys conducted from 2016–2019.

Table A3: Facebook use and online vaccine information exposure (unweighted)

	Non-skeptical webpages		Vaccine-skeptical webpages	
	Visited $\geq$ 1 page	Total pages visited	Visited $\geq$ 1 page	Total pages visited
College graduate	0.020**	0.056	0.000	−0.002
	(0.007)	(0.074)	(0.003)	(0.018)
Female	0.013*	0.094	0.005	0.027
	(0.007)	(0.066)	(0.003)	(0.018)
Nonwhite	−0.029***	−0.164*	−0.007**	−0.028**
	(0.007)	(0.072)	(0.002)	(0.010)
Age 30–44	−0.042***	−0.658**	−0.007	−0.021
	(0.012)	(0.241)	(0.004)	(0.021)
Age 45–59	−0.042***	−0.744**	−0.004	−0.033
	(0.012)	(0.250)	(0.004)	(0.019)
Age 60+	−0.062***	−0.840**	0.001	0.011
	(0.012)	(0.258)	(0.005)	(0.025)
Parent	−0.008	−0.003	0.002	0.012
	(0.009)	(0.088)	(0.003)	(0.013)
Days in panel	0.002***	0.009**	0.000	−0.001
	(0.000)	(0.003)	(0.000)	(0.000)
log(URLs visited)	0.023***	0.125***	0.004***	0.016***
	(0.002)	(0.016)	(0.001)	(0.004)
log(Facebook pages)	0.003*	0.029	0.003***	0.012**
	(0.002)	(0.015)	(0.001)	(0.004)
Constant	−0.041*	0.103	−0.010	−0.051
	(0.017)	(0.301)	(0.007)	(0.047)
R^2	0.061	0.019	0.016	0.005
N	12,835	12,835	12,835	12,835

$^{*}p < 0.05$, $^{**}p < .01$, $^{***}p < .001$ (two-sided); OLS models with standard errors clustered by respondent. All models include panel fixed effects. Data from YouGov Pulse 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.

Figure A2: 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 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 seconds of webpage exposure.

Anti-vaccine Facebook groups and pages

We identify prominent anti-vaccine Facebook groups and pages as those that were identified in data collected by Jamison et al. (2020) on the most frequent anti-vaccine advertisers on Facebook and those specifically identified by journalists as ranking among the most important anti-vaccine groups on Facebook by Broderick (2019); Madrigal (2019); Pilkington and Glenza (2019); Telford (2019); Zadrozny (2019). Some of the pages in question have been removed by Facebook or become defunct. We include personal pages when individuals from these groups have switched to using a personal account after an organization was banned. We then measure the prevalence of visits to these URLs or Facebook URLs containing the group ID list appears below.

- facebook.com/coloradoHealthChoiceAlliance
- facebook.com/MSParentsForVaccineRights
- facebook.com/GenerationRescue
- facebook.com/JB-Handley
- facebook.com/HealthRanger
- facebook.com/DrTenpennyOnVaccines
- facebook.com/HealthNutNews
- facebook.com/MarchAgainstMonstanto
- facebook.com/RevolutionForChoice
- facebook.com/pages/JB-Handley
- facebook.com/StopMandatoryVaccinationNow
- facebook.com/LarryCook333
- facebook.com/BriarandHPVinjury
- facebook.com/RhodeIslandersforVaccineChoice
- facebook.com/national.vaccine.information.center
- facebook.com/groups/VitaminCforoptimalhealth
- facebook.com/groups/VaccinationReEducationDiscussionForum
- facebook.com/groups/VaccineTruthMovement
- facebook.com/groups/VaccineResistanceMovement
- facebook.com/ICANdecide
- facebook.com/ChildrensHealthDefense
- facebook.com/First-Freedoms

Table A4: Top Google searches immediately prior to exposure to non-skeptical vaccine webpages

Search terms	N
plum island	8
flu vaccine	7
vaccine rate colorado by county	7
reducing number of pediatric vaccines	6
dallas county health department	6
mychart login	6
shingles	5
mychart	5
what rna interference and the pathways know the use of micrornas in stopping particular cancers	5
common variable immunodeficiency	4
tom price	4
102 6 fever in 9 year old girl	4
igra test near me	4
side effects of prednisone	4
grocery stores in douglas, wy	4
ur my chart	4
harmful ingredients in vaccines	4
petco	4
igra blood test	4
amanda kirchgessner	4
hep b vaccine newborn	4
daniel neides,m d	4
jill stein sons medical school	3
safeway near me	3
niaid gp120 lu	3
otitis media with effusion	3
valtrex commerical for people with shingles	3
polio history	3
symptons of cancer in cats	3
ragdoll kittens for sale in arkansas	3
hepatitis b	3
polio	3
staphylococcus aureus	3
permanently inadmissible to united states	3
fujifilm diosynth	3
w t norman trait model	3
igra blood test near me	3
albertsons in helena mt	3
jewel osco joliet	3
why are malaria vaccines so difficult to produce	3
side effects of rabies vaccine in cats	3
youll soon be free to fly	3
utd igra blood test	3
utd immunization	3
flu shot	3
twinkie	3
is control road open	3
university of cincinnati student health	3
los angeles aids 1980s	3
nerves jum tuberculosis	3
boxer breeders in texas	3
isidewith	3
city of garden grove	3

Google searches in the 30 seconds prior to exposure to a webpage classified as not skeptical about the safety and efficacy of vaccines (must be among the three most recently visited webpages).

Table A5: Google searches immediately prior to exposure to vaccine-skeptical webpages

Search terms	N
alabama issued immunization	4
daniel 2:43	2
pressure points for self defense chart	2
truth about vaccines	2
whats the best sleep cycle	1
robotic bees	1
watcher files ufos	1
george soros citizenship	1
malfeasance	1
now official fda announced vaccines causing autism	1
vaccines causing developmental delays	1
asthma cough homeopathic remedies	1
christians against vaccines	1
cosmetic ingredients that can cause brain damage in utero	1
lateral view of human skull	1
feces in vaccines	1
aluminum adjuvant autism	1
aluminum adjuvants autism	1
child vaccination debate	1
got flu shot arm hurts	1
dr wolfson on vaccines	1
anti-vaxxer tv show	1
christains against vaccines	1
cosmetic ingredients that can cause brain damge in utero	1
dr wolfson	1
vaccines cause disease jack wolfson	1
vaccines cause disease dr worlf	1

Google searches in the 30 seconds prior to exposure to a webpage classified as skeptical about the safety and efficacy of vaccines (must be among the three most recently visited webpages).