Nudging Recommendation Algorithms Increases News Consumption and Diversity on YouTube

Xudong Yu, Muhammad Haroon, Ericka Menchen-Trevino, Magdalena Wojcieszak

Magdalena Wojcieszak
E-mail: mwojcieszak@ucdavis.edu

This PDF file includes:

Supplementary text
Figs. S1 to S6
Tables S1 to S36
SI References

Supporting Information Text

A. Prompts Used and YouTube Interface in the User Nudge Treatment Condition.

In the user nudge treatment, we study two theoretically driven nudges—those reminding individuals of benefits to oneself (i.e., self-interest nudges) or benefits to the public (i.e., public interest nudges) of news and public affairs consumption. Reminding individuals of the personal benefits of news consumption (e.g., being able to better plan for one's future or make better financial decisions), should make news and politics more personally relevant. This reduces one of the barriers that prevent people from consuming news, namely the perceived disconnect between public affairs and one's daily life (1). In addition, priming self-interest heightens self-enhancement values that emphasize the pursuit of personal interests (2), which should motivate individuals to follow news for individual gains. Some work suggests that the enhanced relevance of news to one's self-interest promotes news consumption (3). In turn, public interest nudges emphasize the benefits of news to society and democracy at large (e.g., strengthening our democracy or ensuring an informed electorate). In addition, priming public interest stimulates self-transcendence values, which "emphasize concern for the welfare and interests of others" (2, p. 226). These values link oneself to other people to promote their well-being and prosocial behavior (2), which—in our case—could mean engaging with news on YouTube.

A total of seven nudges were developed and piloted to ensure that they were clear and worked as intended. Participants were asked to indicate what they read in the nudge, what the nudge was intended to mean, and how clear it was. The nudges, questions, response options, and descriptive statistics are as follows:

Self-Interest Nudges

1. Following news and politics is a good way to ensure that you are aware of the events that impact you.
2. In order to plan for your future, it is important to follow news and politics.
3. Following news and politics can help you make better financial decisions.
4. Impress your friends and co-workers with your knowledge by staying up to date with news and politics.

Public Interest Nudges

1. A healthy democracy requires citizens who stay up to date with news and politics.
2. In order to strengthen our democracy, it is important for citizens to follow politics and news.
3. To be a good citizen, it is important to follow what is going on in the country and the world.

Questions

1. Please indicate which statement is closer to what you read in the message above.
    1. Following news and politics can be beneficial for me personally
    2. Following news and politics is beneficial for democracy
2. In your opinion, what was the purpose of the message above? The purpose was to encourage you to:
    1. Follow news and political accounts
    2. Be more active on social media
    3. Vote in the upcoming midterm elections
3. How clear was the message above?
    1. Very unclear
    2. Unclear
    3. Somewhat unclear
    4. Neutral
    5. Somewhat clear
    6. Clear
    7. Very clear

Descriptive Statistics

The descriptive statistics for each of the questions is reported below.

Table S1. Descriptive Statistics of the Pilot Test

Nudge	Percent Correct		Mean (SD)
	Question 1	Question 2	Question 3
Self-Interest Nudge 1	86.1	72.6	5.17 (1.37)
Self-Interest Nudge 2	84.6	78.4	5.00 (1.47)
Self-Interest Nudge 3	88.5	82.3	5.09 (1.38)
Self-Interest Nudge 4	85.6	65.9	5.03 (1.38)
Public Interest Nudge 1	87.9	70.1	5.40 (1.26)
Public Interest Nudge 2	88.4	72.0	5.33 (1.34)
Public Interest Nudge 3	67.1	71.5	4.95 (1.33)

Although all nudges were clear to participants, we selected self-interest nudges 2 & 3 and public interest nudges 1 & 2 because they showed very high accuracy rates for both questions one and two. These four nudges were then slightly adapted to align with the research focus of this project—watching news and politics on YouTube—and were implemented in the user nudge treatment condition:

- *"In order to plan for your future, please watch news and politics videos on YouTube."*
- *"Please watch news and politics videos on YouTube, this can help you make better financial decisions."*
- *"Please watch news and politics videos on Youtube: A healthy democracy requires citizens who stay up to date with news and politics."*
- *"You can help strengthen our democracy, watch news and politics videos on YouTube today."*

Figures S1 and S2 show an example of the banner that appeared in the user nudge condition on the homepage and on any video page respectively. The specific text of the nudge was randomly chosen for any given user interaction with YouTube to ascertain that the participants were not exposed to one and the same prompt each time they opened the platform.

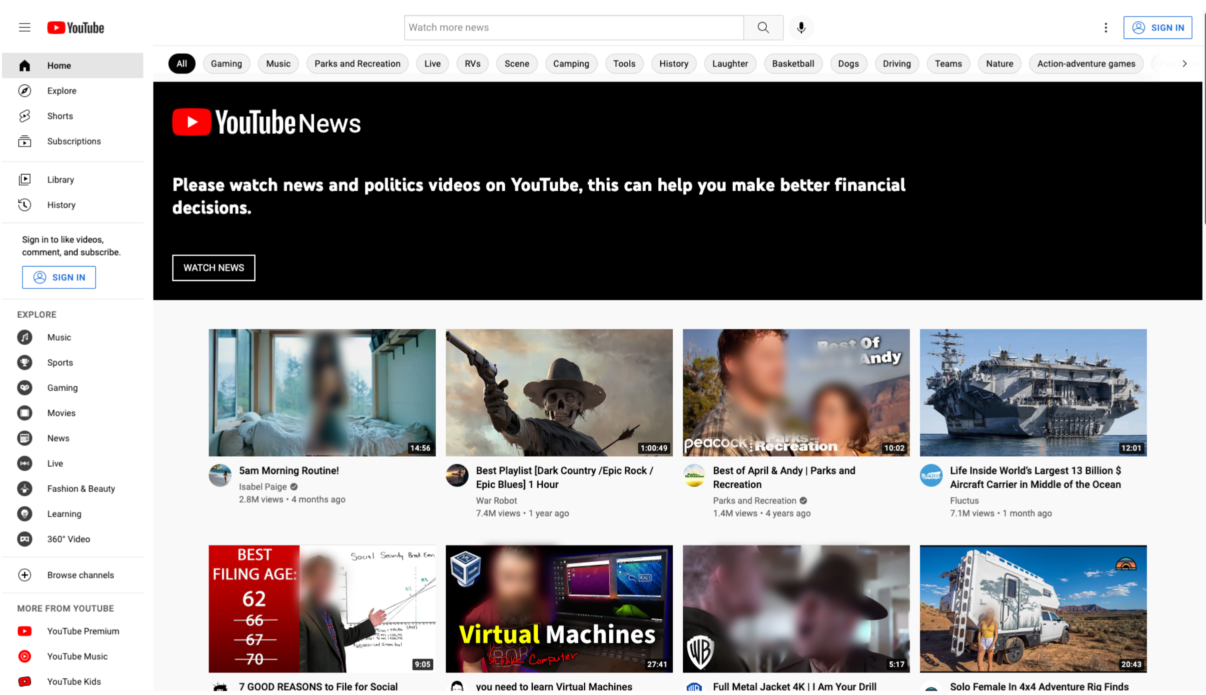


Fig. S1. Example of YouTube Homepage in the User Nudge Treatment Condition

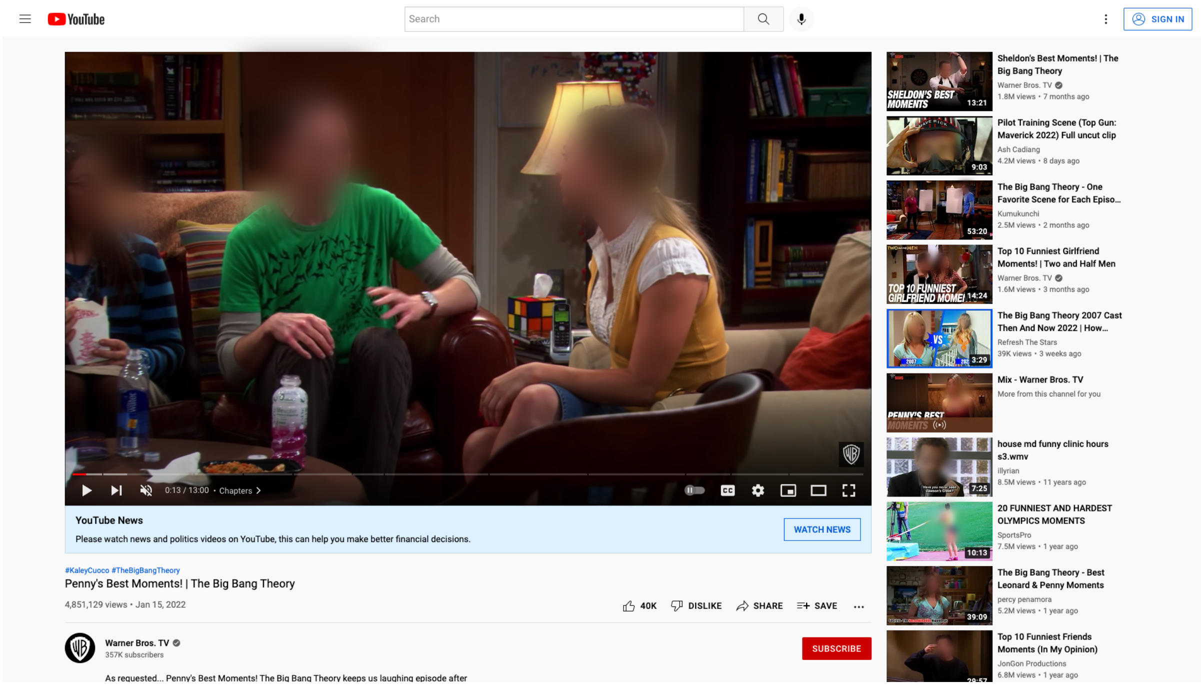


Fig. S2. Example of Individual Video Page in the User Nudge Treatment Condition

B. Channel Selection. We select news media outlets for the algorithmic intervention using Ad Fontes' Media Bias Chart (4)*.

Ad Fontes scores are based on the labeling of articles, radio, TV, and videos (called episodes) for numerous news sources by trained analysts. For each source, Ad Fontes selects a sample of episodes that are most prominently featured on that source's website over several news cycles (min. 15 articles/source, several dozen articles for top 100 sites/each, and over 100 articles for the largest sources, such as the New York Times and Washington Post). Ad Fontes uses a multi-person rating per episode system to minimize the impact of any one person's perspective on the published rating (5). Each episode is rated by at least three analysts from across the political spectrum (right, left, and center, based on self-reports). Once an analyst rates an episode, they compare their rating to the rating of (at least) two others and if there are large discrepancies more analysts review the episode. Ad Fontes has a team of 60 trained analysts who receive 30 hours of initial training and 40 hours of training annually to do this work. The analysts include academics, journalists, librarians, lawyers, military veterans, civil service professionals, and other professional groups that require rhetorical and analytical skills.

The scores are assigned along two dimensions. The first is reliability, which captures how much readers can rely on a particular source for actual news—the "who, what, when, and where" and which is rated on a scale from "contains inaccurate/fabricated information" to "original fact reporting." The second dimension is political bias, rated on a scale from "most extreme left" to "most extreme right". We select reliable and generally balanced sources, with a reliability score higher than 40 and a bias score between -18 and 18. For these sources, we also identify their corresponding YouTube channels and recommend news from these channels.

Different organizations evaluate news outlets on different, and sometimes non-overlapping, criteria and the operationalizations of bias, factuality, or transparency are not always clear. A recent study that compared six sets of expert ratings, AdFontes included, found that they correlate highly with one another (6). This suggests that experts generally agree on the relative placement of news domains. Because that study also generated a set of aggregate ratings, we compare the 39 sources selected from AdFontes to these aggregate ratings and ratings from NewsGuard and Media Bias/Fact Check. Table S2 shows that the news organizations we selected are ranked high in terms of reliability and credibility by other rating metrics. We used AdFontes because it is the only organization that labels outlets on both dimensions, i.e., reliability and political bias.

C. News Identification. To identify whether users were recommended or watched *news* videos, we analyzed if the recommended and watched videos came from a YouTube channel of a news organization from our extensive list of news domains (10). The overall list contains was composed from several sources: manually identified news domains from Alexa's Top 1,000 web domains list; the 1,000 most browsed domains in our trace data; and the 1,000 most shared domains by politicians on Twitter. We augmented the list with channels classified as mainstream by Ledwich et al. (11) or left, left-center, center, right-center, or right by Ribeiro et al. (12). Overall, we identified 1,625 YouTube channels from all news organizations on our lists.

D. Political Content Classification. To determine whether participants are recommended and watch videos about political issues, we develop a neural binary classifier. We conceptualize 'politics' rather broadly: videos considered as political include references

* https://adfontesmedia.com/interactive-media-bias-chart/

News Organization	Domain	Reliability Metrics			
		Ad Fontes (7)	Lin et al. (6)	NewsGuard (8)	MBFC (9)
AP	apnews.com	0.77	1.00	0.95	1.00
Reuters	reuters.com	0.76	1.00	1.00	1.00
PBS	pbs.org	0.76	0.87	—	0.83
Newsy	newsy.com	0.74	0.88	—	0.83
Voice of America	voanews.com	0.74	0.94	—	0.92
CBS News	cbsnews.com	0.73	0.88	0.90	0.83
CNBC	cnbc.com	0.73	0.85	—	0.75
BBC	bbc.com	0.72	0.88	—	0.83
NPR	npr.org	0.72	0.93	—	1.00
Wall Street Journal	wsj.com	0.72	0.80	—	0.67
USA Today	usatoday.com	0.71	0.90	—	0.92
NBC News	nbcnews.com	0.71	0.84	—	0.83
Axios	axios.com	0.71	0.90	—	0.92
Bloomberg News	bloomberg.com	0.71	0.84	—	0.75
Al Jazeera	aljazeera.com	0.71	0.78	—	0.67
LA Times	latimes.com	0.71	0.85	—	0.83
The New York Times	nytimes.com	0.70	0.86	—	0.83
The Hill	thehill.com	0.70	0.90	—	0.83
MarketWatch	marketwatch.com	0.69	0.85	—	0.83
The Economist	economist.com	0.69	0.93	—	0.92
Time Magazine	time.com	0.69	0.83	—	0.83
Washington Post	washingtonpost.com	0.68	0.82	—	0.75
The Guardian	theguardian.com	0.68	0.75	—	0.67
Forbes	forbes.com	0.68	0.83	—	0.75
FiveThirtyEight	fivethirtyeight.com	0.68	0.83	—	0.83
Fox Business	foxbusiness.com	0.68	0.79	0.69	0.75
Politico	politico.com	0.68	0.86	—	0.92
Insider	insider.com	0.68	0.85	—	0.83
Buzzfeed News	buzzfeednews.com	0.68	0.80	—	0.75
CNN	cnn.com	0.67	0.66	—	0.58
Foreign Policy	foreignpolicy.com	0.67	0.92	—	0.92
The Dispatch	thedispatch.com	0.66	0.84	0.88	0.83
Christianity Today	christianitytoday.com	0.66	0.83	—	0.83
NY Daily News	nydailynews.com	0.65	0.81	—	0.83
Vice	vice.com	0.65	0.75	—	0.75
The Intercept	theintercept.com	0.64	0.61	—	0.67
Vox	vox.com	0.64	0.65	—	0.67
The New Yorker	newyorker.com	0.63	0.66	—	0.75

Table S2. List of YouTube news channels that we sampled videos from on a daily basis along with reliability metrics from Ad Fontes (7), Lin et al. (6), NewsGuard (8), and Media Bias/Fact Check (9). The metrics have been scaled from 0 to 1 (0 - least reliable, 1 - most reliable). The lowest values for each metric in our list are defined as follows: Ad Fontes labels 0.63 as *"Mix of Fact Reporting and Analysis or Simple Fact Reporting"*, Lin et al. label 0.61 as *"medium-quality"*, NewsGuard labels 0.69 as *"Credible with Exceptions"*, and MBFC labels 0.58 as *"Mostly Factual"* and *"High Credibility"*. indicating that our news organizations are highly reliable and credible according to various reliability metrics. Several NewsGuard ratings have been redacted (—-) because NewsGuard permits the publication of only five example ratings.

to both political figures, policies, elections, news events, and specific political events *as well as* issues such as climate change, immigration, healthcare, gun control, sexual assault, racial, gender, sexual, ethnic, and religious minorities, the regulation of large tech companies, and crimes involving guns.

To train our political model, we fine-tune the RoBERTa NLP transformer-based model proposed in (13). RoBERTa builds upon BERT (14) and changes key hyperparameters (higher learning rates, mini batches, different pre-training procedure, etc.) for improved performance. To fine-tune, we first collect explicitly political and non-political comment data from Reddit. We collected a total of 9063 comments from the political subreddits: r/Socialism, r/conservatives, r/Conservatives, and from the non-political subreddits: r/soccer, r/nba, r/OnePiece. These were then annotated by trained undergraduate and graduate students and labeled as "0"("1") for "non-political"("political") comments. Manual annotation is required because there can be significant subtleties in deciding political ideology of comments. For fine-tuning the RoBERTa model, we train for 3 epochs, and employ a learning rate of 0.00002, batch sizes of 16, tokenization max length of 128, and gradient accumulation on each step.

Note that our original goal was to train a model that is highly performant irrespective of text type and social media platform. Therefore, we utilized a combined test set containing 3885 manually annotated comments, with 901 comments collected from YouTube, 594 comments from Facebook, and 2393 Reddit comments (similarly collected as for the training set). With our trained model, we obtain an accuracy of 91.47%, precision of 90.34%, and 90.77% recall on this platform-wise out-of-domain test set.

To ensure that the model performs well across platforms, we present results for the performance achieved for YouTube. We achieve an accuracy of 88.67%, precision of 87.67%, and recall of 87.87%. To further validate that our classifier works well with non-comment data, we collected 2044 video titles from YouTube which were then manually annotated for whether they were political or not. On this set, our model achieves an accuracy of 91.80% indicating that it is suitable for post/title-level data as well.

We note that the measure of political video consumption encompasses all videos that were not part of news channels (e.g., those from pundits, vloggers, or other political influencers, and also those by—say—celebrities, in which political issues are addressed).

E. Problematic Channels. For the analysis involving problematic channels, we relied on the manual channel classifications by Ribeiro et al. (12) and Ledwich et al. (11), in similar fashion to Hosseinmardi et al (15). We considered a video to be problematic if it belonged to a channel that was classified by Ribeiro et al. as either *intellectual dark web* (IDW), *Alt-lite*, or *Alt-right*. In the case of Ledwich et al.'s more granular classifications, we considered a video to be problematic if it belonged to a channel that was classified as one of *AntiSJW*, *Conspiracy*, *WhiteIdentitarian*, *Socialist*, *MRA*, *ReligiousConservative*, or *QAnon*. Thus, this analysis was carried out at a channel-level classification of either problematic or non-problematic. In total, we considered 4,150 channels as problematic, the breakdown of which can be seen in Table S3.

Label	Ribeiro et al. (12)	Ledwich et al. (16)	Number of channels
IDW	✓		91
Alt-lite	✓		113
Alt-right	✓		88
AntiSJW		✓	1,056
Conspiracy		✓	2,413
MRA		✓	111
Socialist		✓	148
QAnon		✓	366
WhiteIdentitarian		✓	70
ReligiousConservative		✓	818

Table S3. Number of problematic channels in each label

	Chronology	
	Random	Popular
News	1.91×	1.61×
Political	1.20×	1.24×

Number	0	1	3	5	6	7	8	9	15	20
News	0.97×	1.27×	1.51×	1.54×	1.71×	1.56×	1.83×	2.08×	2.75×	2.90×
Political	1.02×	1.13×	1.20×	1.21×	1.20×	1.26×	1.18×	1.22×	4.55×	1.51×

Percentage	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
News	0.94×	1.86×	2.38×	2.80×	3.67×	3.64×	4.18×	3.88×	5.30×	3.99×
Political	1.01×	1.18×	1.18×	1.26×	1.32×	1.33×	1.52×	1.55×	1.58×	1.55×

Table S4. Average increase in news and political recommendations as a result of different configurations. In the top table, we report the increase of injecting random and popular videos. In the middle table, we report the increase as we injected a fixed number of videos as in the top-row. In the bottom table, we report the increase as we inject a number of videos proportional to the user's watch history.

F. Sock Puppet Experiment. Here we present a more detailed description of the results from the sock puppet experiment (N= 8,400 sock puppets). With regards to chronology of injected videos, we see a marginal difference between injecting recent videos as opposed to the popular ones. Here, we normalize across all other parameters (i.e., number of videos injected) and report the average increase for just random and popular video injections. In the top of Table S4, we see that the increase in percentage of news videos is higher when injecting randomly selected videos (1.91×) as compared to popular videos (1.61×). In the case of political videos, both approaches are slightly more comparable at 1.24× and 1.20× with popular videos yielding the higher value. Thus, we conclude that the chronology of the injected videos is not a significant factor that yields higher news recommendations. Therefore, to promote diverse and fresh content, we choose to inject videos sampled from all the videos posted by the news channels in our list in the past 48 hours everyday.

Moving onto the proportion of injected videos, we try out various configurations involving injecting a fixed number of videos and injecting number of videos proportional to the number of videos already watched by the user. In both cases, we see that a higher values yield a higher increase in percentage. In the case of fixed injections, we see that the increase in percentage of news recommendations is not as pronounced as in the case of proportional injections. Furthermore, we see that the rate of increase is steady thus indicating that there is no "golden" number or proportion at which we see a sudden spike in news recommendations rather it grows ever so steadily as more videos are injected. This suggests a need for a balanced number of injections such that the number of news recommendations is neither too low nor too overwhelming for users.

G. Question wording, descriptives, and distributions of self-reported variables.

Table S5. Variable description

	Pre-test			Post-test		
	α	Mean	SD	α	Mean	SD
Intended political participation Are you likely to engage in the following behaviors in the next 3 months? Select all that apply. Attend a protest or rally Contribute money to a political candidate or organization Sign an online petition Try to convince someone how to vote (online or in-person) Write and post political messages online Talk about politics with someone you know None of the above		**26.69**	**22.81**		**27.62**	**21.82**
Perceived accuracy of true claims about current events The following is a list of events. Please indicate how certain you are about whether each event did or did not happen in the last few weeks. Version 1 Democratic Sen. Raphael Warnock wins re-election in Georgia's runoff election. WNBA star Brittney Griner is released from Russian detention in a prisoner swap. A landmark bill protecting same-sex marriage rights received bipartisan approval in the House and the Senate. Version 2 President Volodymyr Zelensky of Ukraine met with President Biden in Washington. The Supreme Court ruled that the border program to expel migrants must stay in place for now. The House Jan. 6 committee referred former President Donald Trump to the Department of Justice for criminal investigation. Version 3 Thousands of supporters of Brazil's former president stormed Brazil's Congress, Supreme Court and presidential offices. Kevin McCarthy (R) has been elected Speaker of the US House of Representatives. A 6-year-old child shot his teacher in a Virginia school.					**82.24**	**18.96**

Table S5. Variable description *(continued)*

	α	Mean	SD	α	Mean	SD
1 - Definitely didn't happen to 4 - Definitely did happen						
Perceived accuracy of false claims about current events The following is a list of events. Please indicate how certain you are about whether each event did or did not happen in the last few weeks. Version 1 Joe Biden formally announces he will seek reelection in 2024. Arizona Sen. Kyrsten Sinema (D) is switching parties and registering as a Republican. The Trump Organization is found innocent of criminal tax fraud. Version 2 House lawmakers voted 220-211 to reject a sweeping $1.7 trillion omnibus spending bill. Elon Musk says he will not resign as CEO of Twitter after Twitter users voted to oust him. Several busloads of migrants were dropped off near the residence of President Biden on Christmas Day. Version 3 Ukraine has accepted an offer from Russian President Vladimir Putin for a 36-hour ceasefire over Orthodox Christmas. The severe flooding in California was primarily caused by several levees breaking. Biden recently refused to visit the U.S.-Mexico border as president. 1 - Definitely didn't happen to 4 - Definitely did happen					**43.83**	**23.72**
Perceived polarization (politicians) How would you rate each of the following groups? Democratic Party candidates and elected officials Republican Party candidates and elected officials 1 - Very conservative to 5 - Very liberal		**55.67**	**24.77**		**53.69**	**24.66**
Perceived polarization (party supporters) How would you rate each of the following groups? People who support Democrats People who support Republicans 1 - Very conservative to 5 - Very liberal		**53.94**	**24.12**		**52.89**	**23.86**
Perceived polarization (4-item scale, 0-100) On the scale below, please indicate how strongly you agree or disagree with the following statements: Democrats and Republicans hate each other. The differences between Democrats and Republicans are too great to be reconciled. Americans are greatly divided when it comes to the most important values. Polarization in America is greater than ever before in my lifetime. 1 - Disagree strongly to 5 - Agree strongly	**.61**	**68.04**	**17.62**	**.63**	**66.53**	**18.44**
Affective polarization (presidents, 0-100) Please rate the person or group on a thermometer that runs from 0 to 100 degrees. Rating above 50 means that you feel favorable and warm toward the person or group. Rating below 50 means that you feel unfavorable and cool toward the person or group.		**79.94**	**17.84**		**80.12**	**18.56**

Table S5. Variable description *(continued)*

	α	Mean	SD	α	Mean	SD
Joe Biden Donald Trump						
0 - Very cold to 100 - Very warm						
Affective polarization (politicians, 0-100) Please rate the person or group on a thermometer that runs from 0 to 100 degrees. Rating above 50 means that you feel favorable and warm toward the person or group. Rating below 50 means that you feel unfavorable and cool toward the person or group.		**75.88**	**16.24**		**74.36**	**17.64**
Democratic Party candidates and elected officials Republican Party candidates and elected officials						
0 - Very cold to 100 - Very warm						
Prioritizing partisan ends over democratic means (0-100) Please indicate how strongly you agree or disagree with the following statements on a 5-point scale ranging from “disagree strongly” to “agree strongly”:	**.81**	**31.25**	**23.88**	**.81**	**32.11**	**23.91**
I think the [Democrats/Republicans] should do everything they can to hurt the [Republican/Democratic] party, even if it is at the short-term expense of the country. It’s OK to sacrifice US economic prosperity in the short term in order to hurt [Republicans’/Democrats’] chances in future elections. [Democrats/Republicans] should redraw districts to maximize their potential to win more seats in federal elections, even if it may be technically illegal. If [Democrats/Republicans] gain control of all branches of government, they should use the Federal Communications Commission to heavily restrict or shut down [Fox News/MSNBC] to stop the spread of fake news. I think the [Democrats/Republicans] should do everything in their power within the law to make it as difficult as possible for [Biden to run the government effectively/Republicans to take part in governing the country].						
1 - Disagree strongly to 5 - Agree strongly						

Note: All variables are re-scaled to 0-100.

In addition, we also show the distribution of the variables.

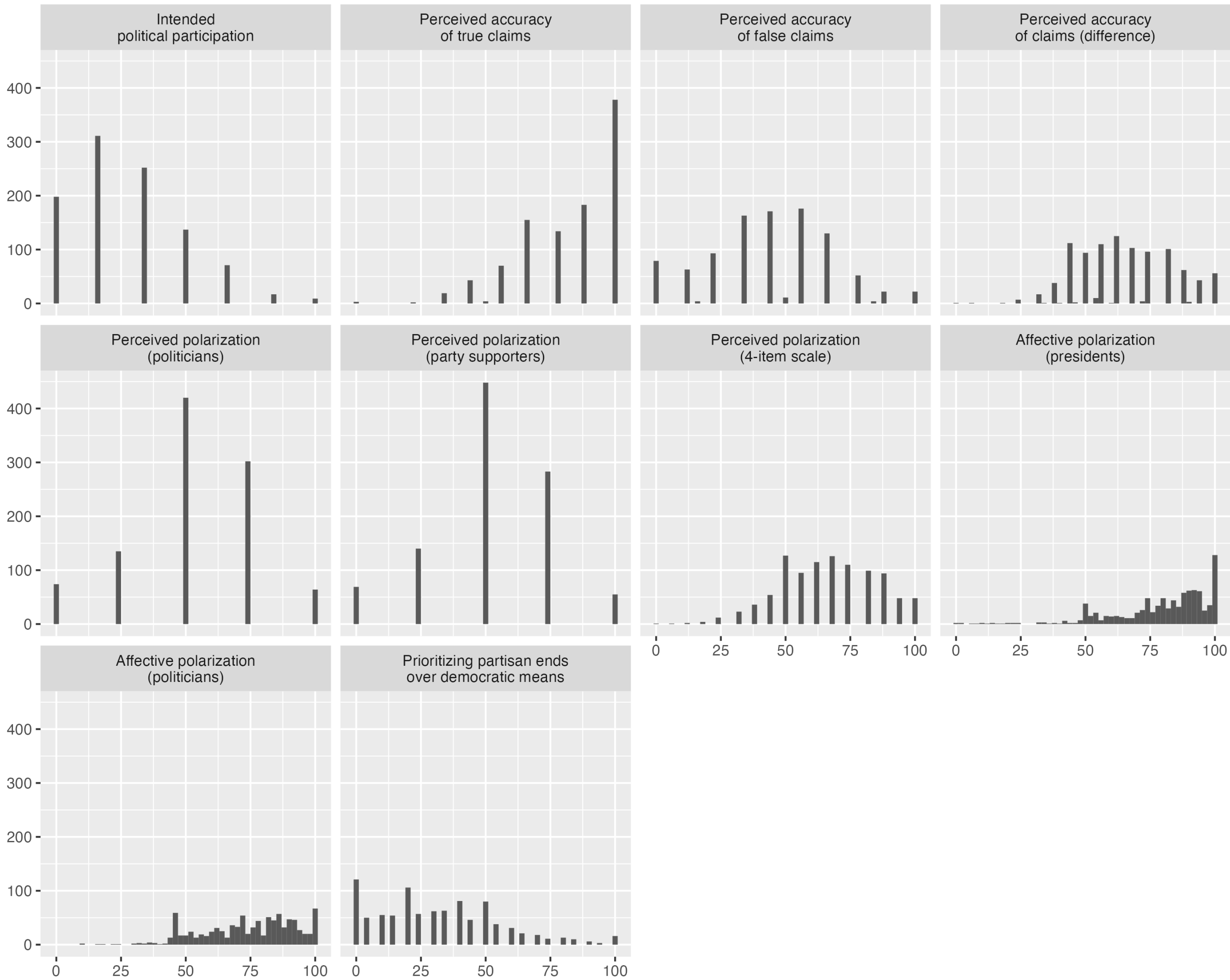


Fig. S3. The distribution of self-reported outcomes

H. Power Analysis. We used the software program G*Power to conduct a power analysis. For primary behavioral outcomes (i.e., effects during treatment weeks) and survey outcomes, our goal was to obtain .80 power to detect a very small effect size of 0.01 at the standard .05 alpha error probability with two tested predictors. The analysis showed that we would need to have at least 967 participants. For the over-time effects (i.e., effects in week four), our goal was to detect a small effect size of 0.02, and the required sample size was at least 485 participants.

I. Sample Description, Randomization, and Attrition. In this section of Supplementary Materials, we offer information on the socio-demographic characteristics of our final sample as well as on randomization and attrition. The analyses show that our sample did not substantially differ from the overall population of YouTube users, that randomization was successful, that overall exclusion rates do not differ between any two conditions, and that participants excluded do not differ from those included except for age.

First, we offer information on the demographics of the entire sample (N = 2,142) and the final sample (N = 1,188), as compared to the general YouTube population based on data from Pew Research (17).[†] As seen, our sample represents the YouTube population relatively well.

[†] This Pew dataset consists of a nationally representative, weighted panel of randomly selected U.S. adults. We applied the weight variable before excluding non-YouTube users. Subsequently, we calculated the proportions for each category of each demographic variable.

Table S6. Demographics of YouTube population/sample

	YouTube Population (Pew)	Final Sample (N = 1,188)	Entire Sample (N = 2,142)
Gender: Male	47%	46%	46%
Gender: Female	52%	54%	54%
Age: 18-29	22%	18%	20%
Age: 30-49	36%	33%	35%
Age: 50-64	25%	29%	28%
Age: 65+	17%	20%	18%
Education: Low (H.S. graduate or less)	33%	17%	18%
Education: Medium (4-year degree or some college)	51%	66%	64%
Education: High (Postgraduate)	15%	17%	17%
Ethnicity: White, non-Hispanic	62%	70%	69%
Ethnicity: Other	38%	30%	31%

Second, we show that randomization was successful. In the final sample used for analysis, age ($F(2, 1185) = 1.18$, $p = .31$), gender ($X^2(2) = .71$, $p = .70$), education ($X^2(4) = 2.71$, $p = .61$), ethnicity ($X^2(2) = 1.44$, $p = .49$), partisanship ($X^2(4) = 1.88$, $p = .76$) were evenly distributed among the three conditions. Third, we show the number of participants in total and in each condition when we use different inclusion criteria in Table S7. In addition, Table S8 reports the exclusion rates by condition for the final sample. A chi-square test indicated a significant difference in the number of participants excluded, $\chi^2(2) = 6.05, p = 0.049$, such that the algorithmic nudge condition had fewer participants excluded due to early uninstallation or lack of activity on YouTube than the user nudge or the control condition. Yet, pairwise comparisons using False Discovery Rate (FDR) correction to control for multiple comparisons showed no significant differences between the conditions: algorithmic nudge vs. user nudge ($p = .063$), algorithmic nudge vs. control ($p = .063$), and user nudge vs. control ($p = 1.00$), suggesting that the observed overall significance may be attributable to small and not individually robust differences between any specific pairs of conditions.

Table S7. Sample and sub-sample sizes when using different inclusion criteria

	Total	Algorithmic nudge	User nudge	Control
Participants who watched at least seven YouTube videos in weeks 1-3 and at least one video in weeks 2-3 (this sample was used for main analysis)	1,188	430	422	336
Participants watched at least one video during weeks 2-3	1,320	488	461	371
Participants watched at least five YouTube videos over weeks 1-3 and at least one video during weeks 2-3	1,228	448	435	345

Table S8. Exclusion Rates

Condition	Total	Excluded	Exclusion Rate (%)
Algo. Nudge	727	297	41%
User Nudge	787	365	46%
Control	628	292	46%

Importantly, we note that it is possible that the rates of uninstallation and/or YouTube activity —and the resulting data availability — may have been a result of the treatment (e.g., if participants were more likely to uninstall the extension or stop using YouTube in the treatment conditions, then excluding those participants introduces bias and prevents causal inference as the differences in observed behavior may be driven by who chooses to keep the extension installed or continues using YouTube rather than to causal effects of the treatments). In short, because we cannot do an Intent-To-Treat analysis (i.e., if someone uninstalled the extension in week 1 and/or did not use YouTube during the study, we do not have any online behavioral data for those individuals), we perform numerous tests for differences across conditions in each of the different kinds of missing data (i.e., early extension uninstallation and lack of YouTube visits).

In terms of the aggregate differences between those who were included in the analyses and those who were excluded, we show the following:

- There was no systematic bias across the conditions in terms of participant exclusion due to (1) early uninstallation of the extension and lack of activity and (2) also due to lack of activity without uninstallation (see details below). The participants excluded from the final sample did not vary across conditions in terms of age ($F(2, 951) = .94$, $p = .39$), gender ($X^2(2) = .15$, $p = .93$), education ($X^2(4) = 3.76$, $p = .44$), ethnicity ($X^2(2) = 3.89$, $p = .14$), partisanship ($X^2(4) = 3.88$, $p = .42$).
- A comparison between participants included in the final sample and those excluded reveals no statistically significant differences in gender ($X^2(1) = .00$, $p = 1.00$), education ($X^2(2) = 3.35$, $p = .19$), ethnicity ($X^2(1) = 1.53$, $p = .22$), or partisanship ($X^2(2) = 3.13$, $p = .21$). However, participants included in the final sample are, on average, 2.47 years older than those who were excluded ($se = .72$, $p < .001$). Because our multivariate models control for all those socio-demographic characteristics, age included, this small difference should not affect our conclusions.

Furthermore, we classify the exclusions into two distinct categories: (1) participants who uninstalled the extension before the study's conclusion (i.e., early uninstallers), and (2) participants who maintained the extension installed for the duration of the study but did not engage actively with YouTube. We provide detailed information regarding each category. In terms of uninstallations:

- Overall, as shown in Table S9, few participants uninstalled the extension before the study ended ($N = 340$, 16% of the initial sample). There is no difference in the number of early uninstallations across conditions ($X^2(2) = 2.21$, $p = .33$).
- Among early uninstallers, there are no statistically significant differences across the conditions in terms of age ($F(2, 337) = .03$, $p = .97$), gender ($X^2(2) = 1.74$, $p = .42$), education ($X^2(4) = 5.04$, $p = .28$), ethnicity ($X^2(2) = .48$, $p = .79$), or partisanship ($X^2(4) = 2.44$, $p = .65$).
- The early uninstallers are not significantly different from those who kept the extension throughout the study in terms of gender ($X^2(1) = .02$, $p = .89$), ethnicity ($X^2(1) = 2.27$, $p = .13$), or partisanship ($X^2(2) = 3.02$, $p = .22$). However, there are statistically significant differences on age and education. Early uninstallers are significantly younger by an average of 4.73 years ($se = 0.98$, $p < .001$), which may be because younger individuals have greater concerns about privacy (18) and are therefore less inclined to allow their online behavior to be tracked. Also, there is a significant association between education level and early uninstallation ($\chi^2(2) = 10.50$, $p = .005$). Individuals with lower education were overrepresented in the early uninstallation group (Residual = 3.22, $p = .004$), while no significant deviations were found for middle ($p = .23$) or high education levels ($p = .88$). We do not have an explanation for this difference, apart from the fact that online samples in general and those who install or keep extensions in particular are more likely to be more educated (19). Again, we control for all these demographics in our models to ensure that any observed effects are not confounded by demographic factors.
- Out of the early uninstallations, the majority occurred in week 1 ($N = 188$, 55%). Because this was before the treatments were activated, these uninstallations cannot be due to the treatment. We note that all those participants are excluded from the analyses because we do not have any on-platform data from them during the treatment weeks.
- There are no statistically significant differences in uninstallation rates in week 1 across conditions ($\chi^2(2) = 2.61$, $p = .27$).
- Only 152 (7% of the initial sample) participants uninstalled during weeks 2 to 4, and there are no significant differences in the number of uninstallations in weeks 2-4 across conditions ($\chi^2(2) = .96$, $p = .62$).
- Importantly, the majority of those who uninstalled during weeks 2 to 4 ($N = 109$, 72%) were included in the analysis because they met our criteria: watching at least seven YouTube videos during weeks 1–3 and at least one video during weeks 2–3. That is, because these participants actively used YouTube before uninstalling the extension, we could assess the impact of our interventions on their on-platform recommendations and behaviors during the treatment weeks. There are no statistically significant differences across conditions in the number of participants who uninstalled early but were still included in the final sample ($\chi^2(2) = .18$, $p = .91$).

Table S9. The number and percentage of early uninstallations

	Uninstalled during the study	Retained until the study ended
Algo. nudge	120 (16.5%)	607 (83.5%)
User nudge	113 (14.4%)	674 (85.6%)
Control	107 (17.0%)	521 (83.0%)
Total	340 (15.9%)	1802 (84.1%)

In terms of lack of YouTube activity:

- A total of 1,802 participants (84% of the initial sample) retained the extension until the end of the study (see Table S9). Of these, 723 participants (40% of 1,802) did not have sufficient YouTube activity during the treatment weeks and were therefore excluded from the final sample (see Table S10 for breakdown by condition).

- This exclusion differed significantly across conditions ($\chi^2(2) = 10.00$, $p = 0.007$). Post-hoc pairwise comparisons with FDR correction indicated that participants in the algorithmic nudge condition were significantly less likely to be inactive if they kept the extension installed for the entire four-week period compared to those in the user nudge condition ($p = .006$). However, no significant differences were found between the algorithmic nudge and control conditions ($p = .11$), nor between the user nudge and control conditions ($p = .27$). Therefore, our findings remain robust, as our primary focus is on the comparisons between the treatment conditions and the control condition.

- Among these inactive and excluded participants, there are no significant differences in age ($F(2, 720) = 0.76$, $p = 0.47$), gender ($\chi^2(2) = 1.64$, $p = 0.44$), education ($\chi^2(4) = 4.57$, $p = 0.33$), ethnicity ($\chi^2(2) = 5.33$, $p = 0.07$), or partisanship ($\chi^2(4) = 7.22$, $p = 0.12$) across conditions.

- Additionally, the group of inactive participants does not significantly differ from all other participants across age ($F(1, 2140) = .67$, $p = .41$), gender ($\chi^2(1) = .01$, $p = .92$), education ($\chi^2(2) = .55$, $p = .76$), ethnicity ($\chi^2(1) = .03$, $p = .87$), or partisanship ($\chi^2(2) = 2.24$, $p = .33$).

Table S10. The number and percentage of participants who ratained the extension until the study ended but were inactive

	Retained until the study ended but were inactive
Algo. nudge	215 (29.6%)
User nudge	293 (37.2%)
Control	215 (34.2%)
Total	723 (33.8%)

These analyses suggest that attrition bias is minimal or nonexistent, and should not affect the patterns observed in our findings. That is, uninstalling the extension or lack of on-platform activity is not systematically related to the treatment and there are minimal socio-demographic differences between those who uninstalled or were inactive and the participants included in the final analyses.

J. The Effects of Treatments on YouTube Engagement.

In this section, we present the results of a series of OLS models testing whether our interventions influence users' engagement on the platform. The models controlled for age, gender, education, race, partisanship, and the average number of videos watched per active day (or the number of active days, or the average time spent per active day) in week 1. As shown in tables below, we find that the intervention did *not* decrease YouTube usage among the sample. Neither the algorithmic nudge ($b = .78$, $se = .63$, $p = .22$) nor the user nudge ($b = .77$, $se = .63$, $p = .22$) affected the average number of videos viewed per active day during the treatment weeks relative to the control. In fact, participants in the algorithmic nudge condition had 0.57 *more* active days on YouTube during the treatment weeks ($se = .12$, $p < .001$) than the control, while the user nudge had no significant impact ($b = .08$, $se = .12$, $p = .51$). Also, neither the algorithmic ($b = -.05$, $se = .09$, $p = .61$) nor the user nudge ($b = .17$, $se = .09$, $p = .06$) conditions led to significant differences in average log(x+1) transformed time spent on YouTube per active day compared to the control.

Table S11. OLS Regressions Predicting the Effects of Treatments on YouTube Engagement

	Videos watched per active day (W2-3)	Number of active days per week (W2-3)	Time spent per active day (W2-3)
Intercept	3.55 (1.17) **	1.16 (0.23) ***	4.92 (0.30) ***
Age	-0.01 (0.02)	0.00 (0.00)	-0.00 (0.00)
Gender (non-male)	0.85 (0.51)	0.25 (0.10) *	0.00 (0.07)
Education (low)	-0.54 (0.87)	-0.01 (0.17)	-0.06 (0.12)
Education (middle)	-0.00 (0.68)	0.06 (0.13)	0.09 (0.10)
Ethnicity (White)	-0.52 (0.57)	0.12 (0.11)	0.00 (0.08)
Party (Independent)	1.26 (0.73)	0.17 (0.14)	0.13 (0.11)
Party (Republican)	0.36 (0.62)	-0.19 (0.12)	0.02 (0.09)
Videos watched per active day (W1)	0.44 (0.02) ***	-	-
Number of active days (W1)	-	0.25 (0.02) ***	-
Time spent per active day (W1)	-	-	0.40 (0.03) ***
Condition (algorithm)	0.78 (0.63)	0.57 (0.12) ***	-0.05 (0.09)
Condition (user)	0.77 (0.63)	0.08 (0.12)	0.17 (0.09)
N	1101	1188	1101
R^2	0.23	0.12	0.17

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. Time spent per active day (minutes) on YouTube (week 1 and weeks 2-3) were log(x+1) transformed. A total of 87 participants were excluded from models 1 and 3 because of missing data for "Videos watched per active day (W1)" or "Time spent per active day (W1)", which resulted from their non-use of YouTube in the first week.

K. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching.

Table S12. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Number Per Active Day)

	News watched per active day (W2-3)	News watched per active day (W4)	Pol. videos watched per active day (W2-3)	Pol. videos watched per active day (W4)
Intercept	-0.18 (0.22)	-0.20 (0.16)	0.11 (0.18)	0.03 (0.20)
Age	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Gender (non-male)	0.12 (0.10)	0.11 (0.07)	-0.08 (0.08)	0.09 (0.10)
Education (low)	0.13 (0.17)	-0.02 (0.13)	-0.16 (0.13)	0.05 (0.16)
Education (middle)	-0.01 (0.13)	-0.10 (0.09)	-0.06 (0.11)	0.04 (0.12)
Ethnicity (White)	-0.15 (0.11)	0.08 (0.08)	0.14 (0.09)	-0.22 (0.11) *
Party (Independent)	0.11 (0.14)	-0.09 (0.11)	0.22 (0.11)	0.03 (0.14)
Party (Republican)	0.08 (0.12)	0.13 (0.09)	-0.01 (0.10)	0.08 (0.12)
News watched per active day (W1)	0.41 (0.04) ***	0.37 (0.04) ***	-	-
Pol. videos watched per active day (W1)	-	-	0.58 (0.02) ***	0.61 (0.04) ***
Condition (algorithm)	1.01 (0.12) ***	0.25 (0.09) **	0.02 (0.10)	0.05 (0.12)
Condition (user)	0.14 (0.12)	0.10 (0.09)	0.03 (0.10)	0.13 (0.11)
N	1101	512	1101	512
R^2	0.16	0.15	0.44	0.29

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 87 participants were excluded from models 1 and 3 because of missing data for "News watched per active day (W1)" or "Pol. videos watched per active day (W1)", which resulted from their non-use of YouTube in the first week. An additional 589 participants were excluded from models 2 and 4 due to missing data for "News watched per active day (W4)" or "Pol. videos watched per active day (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

Table S13. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Percentage)

	Prop. of news watched (W2-3)	Prop. of news watched (W4)	Prop. of pol. videos watched (W2-3)	Prop. of pol. videos watched (W4)
Intercept	-1.38 (2.20)	-4.90 (3.29)	2.72 (1.69)	6.69 (3.50)
Age	0.05 (0.03)	0.07 (0.05)	0.09 (0.02) ***	0.06 (0.05)
Gender (non-male)	1.24 (1.00)	2.11 (1.55)	-0.23 (0.76)	-2.64 (1.65)
Education (low)	0.67 (1.69)	1.38 (2.61)	-1.69 (1.29)	3.29 (2.78)
Education (middle)	0.51 (1.32)	0.06 (1.96)	-1.05 (1.02)	2.14 (2.09)
Ethnicity (White)	-0.29 (1.11)	1.72 (1.74)	-0.33 (0.85)	-4.70 (1.86) *
Party (Independent)	-1.11 (1.43)	-1.75 (2.29)	0.45 (1.10)	-0.89 (2.44)
Party (Republican)	-1.54 (1.21)	2.02 (1.91)	-0.24 (0.93)	-1.21 (2.04)
Prop. of news watched (W1)	0.59 (0.04) ***	0.52 (0.06) ***	-	-
Prop. of pol. videos watched (W1)	-	-	0.40 (0.03) ***	0.43 (0.06) ***
Condition (algorithm)	16.29 (1.24) ***	6.90 (1.94) ***	-1.22 (0.95)	0.44 (2.07)
Condition (user)	0.80 (1.23)	1.87 (1.82)	0.19 (0.94)	-1.14 (1.94)
N	1101	512	1101	512
R^2	0.28	0.16	0.19	0.12

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 87 participants were excluded from models 1 and 3 because of missing data for "Prop. of news watched (W1)" or "Prop. of pol. videos watched (W1)", which resulted from their non-use of YouTube in the first week. An additional 589 participants were excluded from models 2 and 4 due to missing data for "Prop. of news watched (W4)" or "Prop. of pol. videos watched (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

Table S14. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Number Per Active Day, Algorithmic Nudge Condition As the Reference)

	News watched per active day (W2-3)	News watched per active day (W4)	Pol. videos watched per active day (W2-3)	Pol. videos watched per active day (W4)
Intercept	0.84 (0.22) ***	0.05 (0.16)	0.13 (0.18)	0.08 (0.21)
Age	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Gender (non-male)	0.12 (0.10)	0.11 (0.07)	-0.08 (0.08)	0.09 (0.10)
Education (low)	0.13 (0.17)	-0.02 (0.13)	-0.16 (0.13)	0.05 (0.16)
Education (middle)	-0.01 (0.13)	-0.10 (0.09)	-0.06 (0.11)	0.04 (0.12)
Ethnicity (White)	-0.15 (0.11)	0.08 (0.08)	0.14 (0.09)	-0.22 (0.11) *
Party (Independent)	0.11 (0.14)	-0.09 (0.11)	0.22 (0.11)	0.03 (0.14)
Party (Republican)	0.08 (0.12)	0.13 (0.09)	-0.01 (0.10)	0.08 (0.12)
News watched per active day (W1)	0.41 (0.04) ***	0.37 (0.04) ***	-	-
Pol. videos watched per active day (W1)	-	-	0.58 (0.02) ***	0.61 (0.04) ***
Condition (control)	-1.01 (0.12) ***	-0.25 (0.09) **	-0.02 (0.10)	-0.05 (0.12)
Condition (user)	-0.88 (0.11) ***	-0.14 (0.09)	0.01 (0.09)	0.07 (0.11)
N	1101	512	1101	512
R^2	0.16	0.15	0.44	0.29

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 87 participants were excluded from models 1 and 3 because of missing data for "News watched per active day (W1)" or "Pol. videos watched per active day (W1)", which resulted from their non-use of YouTube in the first week. An additional 589 participants were excluded from models 2 and 4 due to missing data for "News watched per active day (W4)" or "Pol. videos watched per active day (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

Table S15. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Percentage, Algorithmic Nudge Condition As the Reference)

	Prop. of news watched (W2-3)	Prop. of news watched (W4)	Prop. of pol. videos watched (W2-3)	Prop. of pol. videos watched (W4)
Intercept	14.91 (2.20) ***	2.00 (3.38)	1.49 (1.69)	7.13 (3.61) *
Age	0.05 (0.03)	0.07 (0.05)	0.09 (0.02) ***	0.06 (0.05)
Gender (non-male)	1.24 (1.00)	2.11 (1.55)	-0.23 (0.76)	-2.64 (1.65)
Education (low)	0.67 (1.69)	1.38 (2.61)	-1.69 (1.29)	3.29 (2.78)
Education (middle)	0.51 (1.32)	0.06 (1.96)	-1.05 (1.02)	2.14 (2.09)
Ethnicity (White)	-0.29 (1.11)	1.72 (1.74)	-0.33 (0.85)	-4.70 (1.86) *
Party (Independent)	-1.11 (1.43)	-1.75 (2.29)	0.45 (1.10)	-0.89 (2.44)
Party (Republican)	-1.54 (1.21)	2.02 (1.91)	-0.24 (0.93)	-1.21 (2.04)
Prop. of news watched (W1)	0.59 (0.04) ***	0.52 (0.06) ***	-	-
Prop. of pol. videos watched (W1)	-	-	0.40 (0.03) ***	0.43 (0.06) ***
Condition (control)	-16.29 (1.24) ***	-6.90 (1.94) ***	1.22 (0.95)	-0.44 (2.07)
Condition (user)	-15.49 (1.17) ***	-5.04 (1.84) **	1.42 (0.90)	-1.58 (1.97)
N	1101	512	1101	512
R^2	0.28	0.16	0.19	0.12

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 87 participants were excluded from models 1 and 3 because of missing data for "Prop. of news watched (W1)" or "Prop. of pol. videos watched (W1)", which resulted from their non-use of YouTube in the first week. An additional 589 participants were excluded from models 2 and 4 due to missing data for "Prop. of news watched (W4)" or "Prop. of pol. videos watched (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

L. OLS Regressions Predicting the Effects of Treatments on Consumption of Problematic Videos.

Table S16. OLS Regressions Predicting the Effects of Treatments on Consumption of Problematic Videos

	Problematic videos watched per active day (W2-3)	Problematic videos watched per active day (W4)	Prop. of problematic videos watched (W2-3)	Prop. of problematic videos watched (W4)
Intercept	0.00 (0.04)	-0.04 (0.05)	-0.42 (0.77)	0.71 (1.28)
Age	0.00 (0.00)	0.00 (0.00)	0.01 (0.01)	0.03 (0.02)
Gender (non-male)	-0.00 (0.02)	0.01 (0.02)	-0.23 (0.35)	0.06 (0.61)
Education (low)	0.03 (0.03)	-0.03 (0.04)	0.29 (0.59)	-2.59 (1.02) *
Education (middle)	-0.00 (0.02)	-0.01 (0.03)	0.62 (0.46)	-1.41 (0.77)
Ethnicity (White)	-0.01 (0.02)	-0.02 (0.03)	0.17 (0.39)	-1.11 (0.68)
Party (Independent)	0.02 (0.03)	0.04 (0.04)	0.35 (0.50)	0.15 (0.89)
Party (Republican)	0.01 (0.02)	0.09 (0.03) **	0.53 (0.43)	1.26 (0.75)
Problematic videos watched per active day (W1)	0.70 (0.02) ***	0.62 (0.03) ***	-	-
Prop. of problematic videos watched (W1)	-	-	0.49 (0.03) ***	0.78 (0.07) ***
Condition (algorithm)	-0.03 (0.02)	-0.00 (0.03)	-0.42 (0.43)	-0.52 (0.76)
Condition (user)	-0.00 (0.02)	0.04 (0.03)	-0.11 (0.43)	1.11 (0.71)
N	1101	512	1101	512
R^2	0.60	0.44	0.22	0.23

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 87 participants were excluded from models 1 and 3 because of missing data for "Problematic videos watched per active day (W1)" or "Prop. of problematic videos watched (W1)", which resulted from their non-use of YouTube in the first week. An additional 589 participants were excluded from models 2 and 4 due to missing data for "Problematic videos watched per active day (W4)" or "Prop. of problematic videos watched (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

M. Multilevel Regression Predicting the Effects of Background Playing of News Videos on News Recommendation. We tested whether background playing of news videos could increase news recommendation. Using hierarchical linear regression, we found that for every 1% increase in the percentage of news videos played by the extension out of all videos watched by the participant and played by the extension, there was a 0.07% increase in news videos recommended, factoring in the proportion of previous day's news recommended, watched, and demographics. In this model, the absence of an observable impact of news watching on recommendations could be attributed to two potential factors. Firstly, the relatively limited number of participants within the algorithmic condition might have hindered the statistical power required to detect the effect. Secondly, the lack of regular news video viewership amongst most participants in this condition could have induced a high variability (large standard deviation), further complicating the detection of any subtle influences.

Table S17. Multilevel Regression Predicting the Effects of Background Playing of News Videos on News Recommendation

	Prop. of news recommended
Intercept	2.93 (1.86)
Age	0.07 (0.03) **
Gender (non-male)	1.81 (0.83) *
Education (low)	1.02 (1.42)
Education (middle)	0.68 (1.09)
Ethnicity (White)	1.25 (0.93)
Party (Independent)	-1.30 (1.12)
Party (Republican)	-0.93 (1.04)
Prop. of news recommended (lagged)	0.36 (0.02) ***
Prop. of news watched (lagged)	0.01 (0.01)
Prop. of news played by the extension (lagged)	0.07 (0.01) ***
Random Effects	
σ^2	91.69
$\tau_{00\ \text{Participant}}$	43.21
ICC	0.32
$N_{\text{Participant}}$	392
Observations	2289
Marginal R^2 / Conditional R^2	0.230 / 0.476

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. Lagged denotes data sourced from the previous day. The Prop. of news played by the extension (lagged) represents the ratio of the number of news videos played by the extension on the previous day to the total number of videos - both played by the extension and watched by the participant - on the previous day. The algorithmic nudge condition has 430 participants, but 38 of them were excluded because they did not have enough activity data to perform lagged analysis during treatment weeks.

N. Multilevel Regressions Testing the Relationship Between News Watching and News Recommendation.

Table S18. Multilevel Regression Testing the Relationship Between News Watching and News Recommendation

	Prop. of news watched	Prop. of news recommended
Intercept	1.68 (1.76)	0.17 (0.55)
Age	0.04 (0.02)	0.05 (0.01) ***
Gender (non-male)	1.01 (0.80)	0.24 (0.25)
Education (low)	0.37 (1.36)	-0.12 (0.42)
Education (middle)	-0.51 (1.08)	-0.08 (0.33)
Ethnicity (White)	-0.99 (0.89)	0.15 (0.28)
Party (Independent)	0.13 (1.13)	-0.58 (0.35)
Party (Republican)	-0.34 (0.98)	-0.38 (0.30)
Prop. of news recommended (lagged)	0.16 (0.02) ***	0.59 (0.01) ***
Prop. of news watched (lagged)	0.21 (0.01) ***	0.02 (0.00) ***
Condition (algorithm)	11.14 (1.00) ***	4.30 (0.31) ***
Condition (user)	-0.23 (0.99)	0.14 (0.31)
Random Effects		
σ^2	382.85	54.87
$\tau_{00\ \text{Participant}}$	127.48	10.21
ICC	0.25	0.16
$N_{\text{Participant}}$	1173	1173
Observations	11196	11196
Marginal R^2 / Conditional R^2	0.155 / 0.366	0.537 / 0.609

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. Lagged denotes data sourced from the previous day. 15 participants were excluded because they did not have enough activities to perform lagged analysis.

O. 2SLS Regressions Testing the Relationship Between News Watching and News Recommendation.

Table S19. 2SLS Regressions Testing the Relationship Between News Watching and News Recommendation

	Prop. of news watched (W2-3)	Prop. of news recommended (W2-3)
Intercept	-1.09 (2.08)	-0.04 (0.97)
Age	0.00 (0.03)	0.03 (0.01) *
Gender (non-male)	0.73 (0.95)	1.31 (0.45) **
Education (low)	0.40 (1.60)	-0.54 (0.75)
Education (middle)	0.20 (1.25)	-0.27 (0.59)
Ethnicity (White)	-0.36 (1.05)	0.19 (0.49)
Party (Independent)	-0.46 (1.36)	-0.34 (0.63)
Party (Republican)	-0.85 (1.15)	-0.64 (0.54)
Condition (algorithm)	9.54 (1.73) ***	10.43 (1.12) ***
Condition (user)	0.46 (1.16)	0.22 (0.54)
Prop. of news watched (W1)	0.38 (0.06) ***	
Prop. of news recommended (W2-3)	0.55 (0.10) ***	
Prop. of news recommended (W1)		0.67 (0.05) ***
Prop. of news watched (W2-3)		0.12 (0.06) *
N	1100	1100
R^2	0.36	0.61

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. 87 participants were excluded because of missing data for "Prop. of news watched (W1)", which resulted from their non-use of YouTube in the first week. One additional participant was excluded because of missing data for "Prop. of news recommended (W1)", possibly due to a failure in collecting recommendation data in the first week.

P. OLS Regressions Predicting the Effects of News Exposure on Self-reported Outcomes.

Table S20. OLS Regressions Predicting the Effects of Overall News Exposure on Intended Political Participation, Perceived Accuracy of Political Claims, and Perceived Polarization

	Intended political participation	Perceived accuracy of true claims	Perceived accuracy of false claims	Perceived accuracy of claims (difference)	Perceived polarization (politicians)	Perceived polarization (party supporters)	Perceived polarization (4-item scale)
Intercept	11.29 (2.16) ***	71.43 (2.46) ***	52.76 (3.35) ***	54.17 (2.41) ***	13.58 (2.99) ***	24.15 (3.16) ***	21.27 (2.58) ***
Age	-0.02 (0.03)	0.38 (0.03) ***	-0.23 (0.05) ***	0.34 (0.03) ***	0.08 (0.04) *	0.03 (0.04)	-0.03 (0.03)
Gender (non-male)	0.66 (0.95)	-5.65 (1.10) ***	5.26 (1.49) ***	-6.08 (1.08) ***	-0.09 (1.24)	-0.38 (1.29)	0.63 (0.88)
Education (low)	-4.45 (1.64) **	-4.22 (1.89) *	4.61 (2.57)	-4.99 (1.85) **	-2.79 (2.12)	-3.51 (2.22)	-0.54 (1.51)
Education (middle)	-2.39 (1.27)	-1.73 (1.48)	2.51 (2.01)	-2.44 (1.44)	0.03 (1.66)	-0.85 (1.74)	-1.14 (1.18)
Ethnicity (White)	2.70 (1.07) *	1.16 (1.25)	-1.78 (1.69)	1.82 (1.22)	2.69 (1.40)	2.09 (1.46)	-1.27 (1.00)
Party (Independent)	-2.64 (1.38)	-10.95 (1.58) ***	2.95 (2.14)	-7.78 (1.54) ***	-3.79 (1.78) *	-4.87 (1.87) **	-0.03 (1.27)
Party (Republican)	-3.78 (1.16) **	-7.90 (1.34) ***	1.63 (1.82)	-5.33 (1.31) ***	2.88 (1.51)	4.34 (1.57) **	2.12 (1.07) *
Intended political participation (pre-test)	0.67 (0.02) ***	-	-	-	-	-	-
Perceived polarization (politicians, pre-test)	-	-	-	-	0.59 (0.03) ***	-	-
Perceived polarization (party supporters, pre-test)	-	-	-	-	-	0.50 (0.03) ***	-
Perceived polarization (4-item scale, pre-test)	-	-	-	-	-	-	0.71 (0.03) ***
Condition (algorithm)	1.17 (1.20)	1.10 (1.40)	-3.65 (1.90)	2.69 (1.37) *	2.51 (1.57)	0.38 (1.64)	0.10 (1.12)
Condition (user)	0.94 (1.18)	-1.10 (1.37)	-3.24 (1.86)	1.19 (1.34)	1.83 (1.54)	-0.06 (1.61)	-1.20 (1.10)
Total number of news videos watched (W2-3)	0.01 (0.04)	-0.03 (0.05)	-0.00 (0.06)	-0.02 (0.04)	-0.00 (0.05)	-0.04 (0.05)	0.03 (0.04)
N	995	991	990	989	995	995	995
R^2	0.55	0.20	0.05	0.17	0.40	0.30	0.45

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Among the remaining 995 participants, four had missing values for "Perceived accuracy of true claims", five for "Perceived accuracy of false claims", and six for "Perceived accuracy of claims (difference)", and were thus excluded from the respective analyses.

Table S21. OLS Regressions Predicting the Effects of Change in News Exposure on Intended Political Participation and Perceived Polarization

	Intended political participation	Perceived polarization (politicians)	Perceived polarization (party supporters)	Perceived polarization (4-item scale)
Intercept	11.31 (2.17) ***	13.54 (3.00) ***	24.10 (3.16) ***	21.21 (2.58) ***
Age	-0.02 (0.03)	0.08 (0.04) *	0.02 (0.04)	-0.03 (0.03)
Gender (non-male)	0.64 (0.95)	-0.05 (1.24)	-0.29 (1.30)	0.65 (0.89)
Education (low)	-4.43 (1.64) **	-2.78 (2.12)	-3.57 (2.21)	-0.46 (1.51)
Education (middle)	-2.39 (1.27)	0.04 (1.66)	-0.83 (1.74)	-1.13 (1.18)
Ethnicity (White)	2.69 (1.07) *	2.69 (1.39)	2.13 (1.46)	-1.31 (1.00)
Party (Independent)	-2.64 (1.38)	-3.78 (1.78) *	-4.87 (1.86) **	0.03 (1.26)
Party (Republican)	-3.79 (1.16) **	2.90 (1.51)	4.39 (1.57) **	2.13 (1.07) *
Intended political participation (pre-test)	0.67 (0.02) ***	-	-	-
Perceived polarization (politicians, pre-test)	-	0.59 (0.03) ***	-	-
Perceived polarization (party supporters, pre-test)	-	-	0.50 (0.03) ***	-
Perceived polarization (4-item scale, pre-test)	-	-	-	0.71 (0.03) ***
Condition (algorithm)	1.11 (1.20)	2.61 (1.56)	0.45 (1.63)	0.32 (1.12)
Condition (user)	0.93 (1.18)	1.84 (1.54)	-0.03 (1.61)	-1.18 (1.10)
Change in news consumption between W2-3 and W1	0.03 (0.06)	-0.03 (0.08)	-0.10 (0.09)	-0.00 (0.06)
N	995	995	995	995
R^2	0.55	0.40	0.30	0.45

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Change in news consumption between W2-3 and W1 refers to the difference in weekly news consumption between treatment weeks and the pre-treatment week.

Table S22. OLS Regressions Predicting the Heterogeneous Effects Between Partisan Identity Strength and Overall/Change in News Exposure on Affective Polarization and Prioritizing Partisan Ends over Democratic Means

	Model: Overall news exposure			Model: Change in news exposure		
	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means
Intercept	3.54 (1.67) *	3.63 (1.96)	11.94 (2.78) ***	3.85 (1.65) *	4.05 (1.93) *	12.34 (2.74) ***
Age	0.07 (0.02) ***	0.09 (0.02) ***	-0.07 (0.03) *	0.07 (0.02) ***	0.08 (0.02) ***	-0.07 (0.03) *
Gender (non-male)	0.10 (0.55)	0.93 (0.64)	1.66 (1.05)	0.04 (0.55)	1.00 (0.65)	1.84 (1.05)
Education (low)	-2.77 (0.94) **	0.06 (1.10)	-2.00 (1.79)	-2.73 (0.94) **	-0.00 (1.09)	-1.95 (1.78)
Education (middle)	-0.54 (0.72)	0.01 (0.85)	-1.46 (1.38)	-0.52 (0.72)	0.06 (0.85)	-1.45 (1.37)
Ethnicity (White)	-0.61 (0.62)	-0.28 (0.72)	0.01 (1.17)	-0.62 (0.62)	-0.20 (0.72)	-0.02 (1.17)
Partisanship (Republican)	-1.15 (0.62)	-1.36 (0.73)	-0.37 (1.17)	-1.19 (0.62)	-1.31 (0.73)	-0.30 (1.17)
Affective polarization (presidents, pre-test)	0.91 (0.02) ***	-	-	0.91 (0.02) ***	-	-
Affective polarization (politicians, pre-test)	-	0.82 (0.02) ***	-	-	0.82 (0.02) ***	-
Prioritizing partisan ends over democratic means(pre-test)	-	-	0.76 (0.02) ***	-	-	0.76 (0.02) ***
Condition (algorithm)	-0.15 (0.69)	-0.42 (0.81)	-0.82 (1.32)	-0.37 (0.69)	-0.59 (0.81)	-0.81 (1.31)
Condition (user)	-0.68 (0.68)	-0.72 (0.80)	1.82 (1.30)	-0.72 (0.68)	-0.70 (0.80)	1.83 (1.30)
Total number of news videos watched (W2-3)	0.09 (0.06)	0.04 (0.06)	0.05 (0.11)	-	-	-
Change in news consumption between W2-3 and W1	-	-	-	0.12 (0.09)	0.06 (0.11)	-0.23 (0.18)
Partisan identity strength	0.80 (0.29) **	1.64 (0.36) ***	-0.03 (0.54)	0.66 (0.27) *	1.51 (0.34) ***	-0.21 (0.50)
Total number of news videos watched (W2-3) × Partisan identity strength	-0.03 (0.02)	-0.04 (0.02)	-0.03 (0.04)	-	-	-
Change in news consumption between W2-3 and W1 × Partisan identity strength	-	-	-	-0.02 (0.03)	-0.05 (0.04)	0.07 (0.06)
N	941	940	942	941	940	942
R^2	0.81	0.70	0.57	0.81	0.70	0.58

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Among the remaining 995 participants, 53 were excluded because they were true partisan independents with no clear ideological leanings and did not vote for the presidential candidates of the two major parties (for independents, their partisanship was imputed based on their ideology and voting decisions in the 2020 presidential election). Additionally, one participant had a missing value for "Affective polarization (presidents, pre-test)," and two participants had missing values for "Affective polarization (politicians, pre-test)," leading to their exclusion from the respective analyses. Change in news consumption between W2-3 and W1 refers to the difference in weekly news consumption between treatment weeks and the pre-treatment week.

Table S23. OLS Regressions Predicting the Heterogeneous Effects Between Ideological Identity Strength and Overall/Change in News Exposure on Affective Polarization and Prioritizing Partisan Ends over Democratic Means

	Model: Overall news exposure			Model: Change in news exposure		
	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means
Intercept	4.08 (2.30)	4.14 (2.97)	9.81 (4.57) *	3.31 (2.24)	5.22 (2.88)	11.36 (4.46) *
Age	0.06 (0.02) ***	0.10 (0.02) ***	-0.10 (0.04) *	0.07 (0.02) ***	0.10 (0.02) ***	-0.10 (0.04) *
Gender (non-male)	-0.29 (0.58)	0.06 (0.75)	1.59 (1.28)	-0.33 (0.59)	0.06 (0.75)	1.55 (1.29)
Education (low)	-3.12 (1.02) **	0.62 (1.30)	-1.76 (2.23)	-3.06 (1.02) **	0.47 (1.30)	-1.95 (2.23)
Education (middle)	-0.69 (0.75)	0.21 (0.96)	-2.35 (1.65)	-0.68 (0.75)	0.21 (0.97)	-2.34 (1.66)
Ethnicity (White)	-0.61 (0.67)	-1.18 (0.86)	0.35 (1.48)	-0.68 (0.67)	-1.01 (0.86)	0.54 (1.47)
Ideology (Liberal)	1.42 (0.67) *	3.68 (0.86) ***	-0.99 (1.43)	1.40 (0.67) *	3.72 (0.87) ***	-1.00 (1.44)
Affective polarization (presidents, pre-test)	0.94 (0.02) ***	-	-	0.94 (0.02) ***	-	-
Affective polarization (politicians, pre-test)	-	0.83 (0.03) ***	-	-	0.82 (0.03) ***	-
Prioritizing partisan ends over democratic means(pre-test)	-	-	0.75 (0.03) ***	-	-	0.75 (0.03) ***
Condition (algorithm)	0.13 (0.72)	-0.90 (0.92)	-0.67 (1.59)	0.03 (0.71)	-1.10 (0.92)	-0.98 (1.57)
Condition (user)	-0.07 (0.72)	-0.62 (0.92)	2.06 (1.57)	-0.10 (0.72)	-0.60 (0.92)	2.10 (1.58)
Total number of news videos watched (W2-3)	-0.13 (0.13)	0.19 (0.16)	0.38 (0.28)	-	-	-
Change in news consumption between W2-3 and W1	-	-	-	0.12 (0.20)	0.00 (0.26)	0.24 (0.44)
Ideological identity strength	-0.48 (0.63)	0.92 (0.80)	2.07 (1.38)	-0.19 (0.59)	0.51 (0.76)	1.37 (1.31)
Total number of news videos watched (W2-3) × Ideological identity strength	0.06 (0.05)	-0.10 (0.07)	-0.18 (0.12)	-	-	-
Change in news consumption between W2-3 and W1 × Ideological identity strength	-	-	-	-0.03 (0.08)	-0.03 (0.10)	-0.11 (0.18)
N	648	647	649	648	647	649
R^2	0.82	0.68	0.58	0.82	0.68	0.57

Note: *** $< .001$ ** $< .01$ * $< .05$. A total of 193 participants were excluded because they did not complete the post-test. Among the remaining 995 participants, 346 were excluded because they were ideological independents. Additionally, one participant had a missing value for "Affective polarization (presidents, pre-test)," and two participants had missing values for "Affective polarization (politicians, pre-test)," leading to their exclusion from the respective analyses. Change in news consumption between W2-3 and W1 refers to the difference in weekly news consumption between treatment weeks and the pre-treatment week.

Table S24. OLS Regressions Predicting the Heterogeneous Effects Between Ideological Identity and Overall/Change in News Exposure on Affective Polarization and Prioritizing Partisan Ends over Democratic Means

	Model: Overall news exposure			Model: Change in news exposure		
	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means
Intercept	2.98 (1.80)	6.41 (2.35) **	14.53 (3.30) ***	2.95 (1.80)	6.24 (2.35) **	14.33 (3.30) ***
Age	0.06 (0.02) ***	0.10 (0.02) ***	-0.10 (0.04) **	0.06 (0.02) ***	0.10 (0.02) ***	-0.10 (0.04) **
Gender (non-male)	-0.28 (0.58)	0.06 (0.75)	1.59 (1.28)	-0.35 (0.58)	0.08 (0.75)	1.62 (1.29)
Education (low)	-3.05 (1.02) **	0.54 (1.30)	-1.92 (2.23)	-3.09 (1.01) **	0.47 (1.30)	-2.01 (2.23)
Education (middle)	-0.69 (0.75)	0.23 (0.97)	-2.33 (1.66)	-0.68 (0.75)	0.21 (0.96)	-2.34 (1.65)
Ethnicity (White)	-0.66 (0.67)	-1.08 (0.86)	0.54 (1.48)	-0.67 (0.67)	-1.00 (0.86)	0.60 (1.47)
Ideology (Liberal)	1.31 (0.69)	3.80 (0.90) ***	-0.76 (1.48)	1.45 (0.66) *	3.74 (0.86) ***	-0.76 (1.41)
Affective polarization (presidents, pre-test)	0.94 (0.02) ***	-	-	0.94 (0.02) ***	-	-
Affective polarization (politicians, pre-test)	-	0.83 (0.03) ***	-	-	0.83 (0.03) ***	-
Prioritizing partisan ends over democratic means(pre-test)	-	-	0.75 (0.03) ***	-	-	0.75 (0.03) ***
Condition (algorithm)	0.16 (0.72)	-0.94 (0.92)	-0.73 (1.59)	-0.03 (0.71)	-1.02 (0.91)	-0.82 (1.57)
Condition (user)	-0.09 (0.72)	-0.59 (0.92)	2.12 (1.58)	-0.11 (0.71)	-0.59 (0.92)	2.10 (1.58)
Total number of news videos watched (W2-3)	0.00 (0.04)	-0.06 (0.05)	-0.07 (0.09)	-	-	-
Change in news consumption between W2-3 and W1	-	-	-	0.18 (0.08) *	-0.18 (0.10)	-0.23 (0.17)
Total number of news videos watched (W2-3) × Ideology (Liberal)	0.01 (0.05)	0.01 (0.07)	0.03 (0.11)	-	-	-
Change in news consumption between W2-3 and W1 × Ideology (Liberal)	-	-	-	-0.16 (0.09)	0.17 (0.11)	0.26 (0.20)
N	648	647	649	648	647	649
R^2	0.82	0.68	0.57	0.83	0.68	0.57

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Among the remaining 995 participants, 346 were excluded because they were ideological independents. Additionally, one participant had a missing value for "Affective polarization (presidents, pre-test)," and two participants had missing values for "Affective polarization (politicians, pre-test)," leading to their exclusion from the respective analyses. Change in news consumption between W2-3 and W1 refers to the difference in weekly news consumption between treatment weeks and the pre-treatment week.

Q. OLS Regressions Predicting the Effects of News Exposure on Self-reported Outcomes (Not Controlling for Pre-test Values for the Outcomes).

Table S25. OLS Regressions Predicting the Effects of Overall News Exposure on Intended Political Participation, Perceived Accuracy of Political Claims, and Perceived Polarization (Not Controlling for Pre-test Values)

	Intended political participation	Perceived accuracy of true claims	Perceived accuracy of false claims	Perceived accuracy of claims (difference)	Perceived polarization (politicians)	Perceived polarization (party supporters)	Perceived polarization (4-item scale)
Intercept	25.51 (2.97) ***	71.43 (2.46) ***	52.76 (3.35) ***	54.17 (2.41) ***	40.34 (3.45) ***	47.60 (3.36) ***	67.94 (2.65) ***
Age	0.10 (0.04) *	0.38 (0.03) ***	-0.23 (0.05) ***	0.34 (0.03) ***	0.19 (0.05) ***	0.10 (0.05) *	-0.00 (0.04)
Gender (non-male)	-1.12 (1.33)	-5.65 (1.10) ***	5.26 (1.49) ***	-6.08 (1.08) ***	0.85 (1.54)	-0.07 (1.50)	-1.40 (1.18)
Education (low)	-12.60 (2.28) ***	-4.22 (1.89) *	4.61 (2.57)	-4.99 (1.85) **	-2.90 (2.65)	-3.45 (2.57)	-0.37 (2.03)
Education (middle)	-3.44 (1.78)	-1.73 (1.48)	2.51 (2.01)	-2.44 (1.44)	2.35 (2.07)	0.81 (2.01)	-1.20 (1.59)
Ethnicity (White)	5.82 (1.50) ***	1.16 (1.25)	-1.78 (1.69)	1.82 (1.22)	3.22 (1.74)	2.50 (1.69)	0.20 (1.33)
Party (Independent)	-11.10 (1.89) ***	-10.95 (1.58) ***	2.95 (2.14)	-7.78 (1.54) ***	-8.97 (2.20) ***	-10.05 (2.14) ***	-3.72 (1.69) *
Party (Republican)	-9.16 (1.62) ***	-7.90 (1.34) ***	1.63 (1.82)	-5.33 (1.31) ***	3.94 (1.88) *	4.68 (1.83) *	2.25 (1.44)
Condition (algorithm)	3.08 (1.69)	1.10 (1.40)	-3.65 (1.90)	2.69 (1.37) *	2.17 (1.96)	-0.05 (1.90)	1.24 (1.50)
Condition (user)	2.49 (1.65)	-1.10 (1.37)	-3.24 (1.86)	1.19 (1.34)	0.13 (1.92)	-1.97 (1.87)	-1.16 (1.47)
Total number of news videos watched (W2-3)	-0.03 (0.05)	-0.03 (0.05)	-0.00 (0.06)	-0.02 (0.04)	-0.04 (0.06)	-0.03 (0.06)	0.05 (0.05)
N	995	991	990	989	995	995	995
R^2	0.12	0.20	0.05	0.17	0.06	0.06	0.02

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Among the remaining 995 participants, four had missing values for "Perceived accuracy of true claims", five for "Perceived accuracy of false claims", and six for "Perceived accuracy of claims (difference)", and were thus excluded from the respective analyses.

Table S26. OLS Regressions Predicting the Effects of Change in News Exposure on Intended Political Participation and Perceived Polarization (Not Controlling for Pre-test Values)

	Intended political participation	Perceived polarization (politicians)	Perceived polarization (party supporters)	Perceived polarization (4-item scale)
Intercept	25.28 (2.97) ***	40.25 (3.46) ***	47.42 (3.36) ***	67.86 (2.65) ***
Age	0.10 (0.04) *	0.19 (0.05) ***	0.10 (0.05) *	-0.00 (0.04)
Gender (non-male)	-0.94 (1.33)	0.94 (1.55)	0.08 (1.50)	-1.35 (1.19)
Education (low)	-12.58 (2.27) ***	-2.95 (2.64)	-3.44 (2.57)	-0.24 (2.03)
Education (middle)	-3.39 (1.78)	2.37 (2.07)	0.84 (2.01)	-1.18 (1.59)
Ethnicity (White)	5.83 (1.49) ***	3.25 (1.74)	2.51 (1.69)	0.14 (1.33)
Party (Independent)	-11.03 (1.89) ***	-8.97 (2.20) ***	-10.00 (2.14) ***	-3.63 (1.69) *
Party (Republican)	-9.06 (1.61) ***	3.99 (1.88) *	4.76 (1.83) **	2.27 (1.44)
Condition (algorithm)	3.50 (1.67) *	2.23 (1.95)	0.28 (1.89)	1.65 (1.49)
Condition (user)	2.56 (1.65)	0.16 (1.92)	-1.91 (1.87)	-1.13 (1.47)
Change in news consumption between W2-3 and W1	-0.16 (0.09)	-0.09 (0.10)	-0.13 (0.10)	-0.01 (0.08)
N	995	995	995	995
R^2	0.12	0.06	0.06	0.02

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Change in news consumption between W2-3 and W1 refers to the difference in weekly news consumption between treatment weeks and the pre-treatment week.

Table S27. OLS Regressions Predicting the Heterogeneous Effects Between Partisan Identity Strength and Overall/Change in News Exposure on Affective Polarization and Prioritizing Partisan Ends over Democratic Means (Not Controlling for Pre-test Values)

	Model: Overall news exposure			Model: Change in news exposure		
	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means	Affective polarization (presidents)	Affective polarization (politicians)	Prioritizing partisan ends over democratic means
Intercept	49.00 (2.88) ***	39.58 (2.59) ***	30.43 (4.13) ***	49.25 (2.83) ***	39.79 (2.54) ***	31.77 (4.07) ***
Age	0.30 (0.03) ***	0.28 (0.03) ***	-0.14 (0.05) **	0.29 (0.03) ***	0.27 (0.03) ***	-0.15 (0.05) **
Gender (non-male)	2.69 (1.10) *	2.22 (0.99) *	0.76 (1.58)	2.84 (1.11) *	2.37 (0.99) *	0.97 (1.59)
Education (low)	0.28 (1.88)	0.11 (1.69)	0.98 (2.70)	0.21 (1.88)	0.01 (1.68)	1.20 (2.69)
Education (middle)	1.55 (1.45)	1.72 (1.31)	-1.15 (2.08)	1.61 (1.45)	1.81 (1.30)	-1.04 (2.08)
Ethnicity (White)	1.98 (1.24)	1.55 (1.11)	-1.76 (1.77)	2.06 (1.24)	1.69 (1.11)	-1.80 (1.78)
Partisanship (Republican)	-7.78 (1.23) ***	-6.27 (1.10) ***	1.94 (1.76)	-7.66 (1.23) ***	-6.12 (1.10) ***	1.99 (1.76)
Condition (algorithm)	1.09 (1.39)	0.12 (1.25)	1.55 (2.00)	1.06 (1.39)	0.17 (1.24)	1.52 (1.99)
Condition (user)	-0.80 (1.37)	-1.66 (1.23)	1.38 (1.96)	-0.74 (1.37)	-1.56 (1.23)	1.37 (1.97)
Partisan identity strength	5.08 (0.56) ***	6.98 (0.50) ***	2.88 (0.80) ***	4.93 (0.52) ***	6.82 (0.47) ***	2.38 (0.75) **
Total number of news videos watched (W2-3)	0.03 (0.11)	0.04 (0.10)	0.26 (0.16)	-	-	-
Change in news consumption between W2-3 and W1	-	-	-	-0.02 (0.19)	0.09 (0.17)	-0.04 (0.27)
Total number of news videos watched (W2-3) × Partisan identity strength	-0.04 (0.04)	-0.05 (0.04)	-0.10 (0.06)	-	-	-
Change in news consumption between W2-3 and W1 × Partisan identity strength	-	-	-	-0.05 (0.06)	-0.11 (0.06)	0.01 (0.09)
N	942	942	942	942	942	942
R^2	0.21	0.29	0.03	0.21	0.30	0.02

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 193 participants were excluded because they did not complete the post-test. Among the remaining 995 participants, 53 were excluded because they were true partisan independents with no clear ideological leanings and did not vote for the presidential candidates of the two major parties (for independents, their partisanship was imputed based on their ideology and voting decisions in the 2020 presidential election). Change in news consumption between W2-3 and W1 refers to the difference in weekly news consumption between treatment weeks and the pre-treatment week.

R. ResearchTube Web Extension Details. To monitor participants' behavior on the platform and run the various nudges, we developed a web browser extension and had the participants install it as part of the survey. This extension – dubbed ResearchTube – was developed for the Chrome web browser and used Manifest V3 capabilities to implement the various functionalities. Below we provide technical details of how ResearchTube was developed and deployed for participants.

R.1. Overview. We opted to develop and test ResearchTube for the Chrome web browser as it is the most commonly used browser globally (20). As such, only participants who used Chrome were allowed to participate in the survey. The extension consisted of two main parts: 1) the web extension itself which was installed on the participants' browsers and 2) a back-end API (Application Programming Interface) which was deployed on a remote AWS (Amazon Web Services) server. The web extension was responsible for administering the various nudges and monitoring the participants' behavior and platform recommendations.

The collected data would first be stored by the extension locally and then transmitted to the to the API which would then store it in a connected database. In addition to sending collected data, the extension could also request the updated list of videos to be injected from the API. To keep track of individual participants, the extension assigned a unique identifier to each participant which could then also be used for recovery if necessary. The extension was published on the Chrome Web Store to increase trust and to ensure that it adhered to the applicable privacy practices laid out by Google For privacy purposes, the extension was restricted to the YouTube domain through the web extension manifest. Finally, the extension kept track of what week the user was in from the time of installation and triggered changes in behavior as time passed.

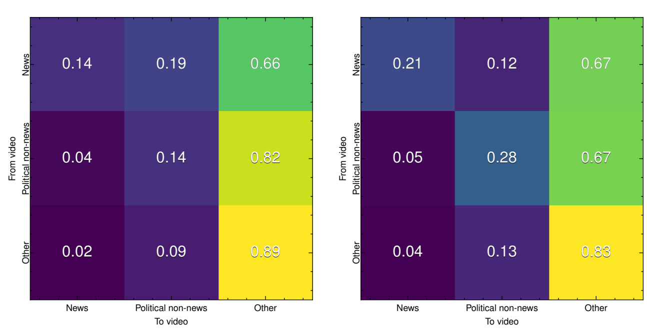
R.2. Assignment Process. Halfway through the initial screening survey, participants were redirected to the extension's installation page on the Chrome Web Store. Here, they followed Google's guidelines for installation and, post-installation, were redirected back to the rest of the survey by the extension. Through bounce-tracking, we were able to keep track of the participant's survey identifier which the extension would then link to its own internal identifier and pass to the API, allowing us to link participant data collected from the extension with their survey responses. The extension also uniformly randomly assigned the user a treatment (algorithmic or user nudge) or control group and passed that information to the API as well.

R.3. Activity Monitoring. Whenever a participant visited a YouTube webpage, the extension would start, retrieve, and parse the contents of the webpage's DOM (Document Object Model). The DOM of a webpage represents the layout and contents of the page in a standardized format (known as HTML) which can be easily parsed by the web extension. Through the URL of the webpage, the extension identified which type of webpage it was on. For example, if the URL was of the following format `...youtube.com/search?q=...`, the extension would know that it is on a search results page and extract the relevant information like what the user searched for. If the URL format was `...youtube.com/watch?v=...`, it indicated that the user was watching a video and the extension could extract the up-next recommendations. Finally, the extension would collect homepage recommendations if the user was on the homepage. To account for participants who did not frequent the homepage enough, the extension also loaded the homepage for the participant in a background tab in a fixed interval and sampled the recommendations from there as well.

R.4. Algorithmic Nudge. For users assigned the algorithmic nudge condition, the extension would retrieve the set of videos to be injected from the API and then play the videos in a background tab at set intervals in a manner that obscured it from the participant. The injected video would play for a few minutes in the muted tab so as not to draw the participant's attention. We further obscured the contents of the tab by overlaying it with a white banner so that an inquisitive user could not observe what was happening inside the tab.

R.5. User Nudge. For users assigned the user nudge condition, the extension manipulated the webpage DOM to add banners that motivated the user to consume more news. These banners appeared on the homepage at the top and under videos being played. See SM A for more details.

S. Transition Probabilities. To understand transition patterns of users, we compute transition probabilities between news, political, and all other videos. Figure S4 shows these probabilities for the algorithmic nudge participants pre and post-intervention. We observe a 1.51× increase between news → news consumption (from 14% to 21%, indicating higher retention on news videos post-intervention. We also see a 0.61× decrease in transitions from news → political (from 19% to 12%). The transition from news → other remains constant at 1.00× (from 66% to 77%). In other words, after the intervention, the users who were watching a news video were more likely to continue watching another news video, likely because they were less likely to watch a political video from outside news channels.



(a) Pre-intervention **(b)** Post-intervention

Fig. S4. Transition probabilities of participants between News, Political, and Other videos pre- and post-intervention. The rows sum to 1 and cell values indicate the percentage of transitions between from the state on the y-axis to the state on the x-axis.

For transitions from political videos, we see a marked increase in retention on political videos at 1.95× (from 14% to 28%). That is, after our intervention, users were almost twice as likely to continue watching political videos when already watching one. The drop of 0.82× in political → other videos (from 82% to 67%) indicates that users were less likely to move to another category videos post-intervention. There was also drop of 0.94× in the Other → Other (from 89% to 83%) transitions despite it still being the dominant watch category.

In short, although transitions away from news and political videos represent a larger pattern than vice versa, we observe sizeable decreases in this pattern suggesting that users' consumption sequences were influenced by the intervention.

T. Statistical Significance of Shifts in Political Leaning of Individual News Diets. In Table S28, we report the statistic values and their significance for the Kolmogorov-Smirnov test between the political slant of videos watched by and recommended to the participants of different ideologies pre, mid, and post-intervention stages. Table S29 shows statistic values for the same test but between recommended and watched videos.

Ideology	Stage	Type	Pre	Mid	Post
Very liberal	Pre	Watched	–	0.178*	0.260*
		Recommended	–	0.074*	0.105*
	Mid	Watched	0.178*	–	0.333*
		Recommended	0.074*	–	0.038*
	Post	Watched	0.260*	0.333*	–
		Recommended	0.105*	0.038*	–
Liberal	Pre	Watched	–	0.282*	0.296*
		Recommended	–	0.074*	0.087*
	Mid	Watched	0.282*	–	0.112
		Recommended	0.074*	–	0.027
	Post	Watched	0.296*	0.112	–
		Recommended	0.087*	0.027	–
Moderate	Pre	Watched	–	0.157*	0.362*
		Recommended	–	0.053*	0.078*
	Mid	Watched	0.157*	–	0.339*
		Recommended	0.053*	–	0.031
	Post	Watched	0.362*	0.339*	–
		Recommended	0.078*	0.031	–
Conservative	Pre	Watched	–	0.203*	0.246*
		Recommended	–	0.085*	0.106*
	Mid	Watched	0.203*	–	0.080
		Recommended	0.085*	–	0.030
	Post	Watched	0.246*	0.080	–
		Recommended	0.106*	0.030	–
Very conservative	Pre	Watched	–	0.349*	0.219
		Recommended	–	0.122*	0.091*
	Mid	Watched	0.349*	–	0.304*
		Recommended	0.122*	–	0.049*
	Post	Watched	0.219	0.304*	–
		Recommended	0.091*	0.049*	–

Table S28. Statistic values for the Kolmogorov-Smirnov test for significance between the ideologies of videos watched and recommended during the pre, mid, and post-intervention stages for different participant ideologies. The asterisk (*) is indicative of a statistically significant difference.

Ideology	Pre	Mid	Post
Very liberal	0.076*	0.052	0.319*
Liberal	0.186*	0.035	0.102
Moderate	0.146*	0.103*	0.248*
Conservative	0.248*	0.121*	0.190*
Very conservative	0.248*	0.035	0.248

Table S29. Statistic values for the Kolmogorov-Smirnov test for significance between the videos watched and recommended during the pre, mid, and post-intervention stages for different participant ideologies.

U. Breakdown of Daily Percentage Increases.

News						
	Algorithmic nudge		User nudge		Control	
Day	R	W	R	W	R	W
1	7.73%	4.01%	7.73%	4.02%	8.99%	6.08%
2	8.40%	5.30%	7.68%	4.41%	8.89%	4.29%
3	7.29%	4.15%	9.31%	4.86%	7.34%	3.14%
4	7.37%	4.62%	9.55%	4.07%	7.84%	3.39%
5	8.87%	4.66%	9.61%	3.87%	7.06%	2.97%
6	8.26%	4.13%	8.56%	3.42%	8.69%	4.40%
7	7.39%	4.65%	10.17%	4.97%	9.03%	5.25%
8	11.80%	16.50%	8.14%	5.10%	7.50%	2.95%
9	13.85%	18.25%	9.42%	4.95%	8.12%	3.28%
10	15.05%	15.62%	7.62%	4.68%	7.82%	3.20%
11	15.01%	16.49%	7.70%	3.76%	6.57%	3.09%
12	14.78%	19.57%	8.19%	3.76%	5.62%	2.20%
13	17.28%	24.94%	7.63%	4.47%	7.58%	3.88%
14	16.40%	15.20%	7.99%	2.72%	6.76%	3.57%
15	16.01%	14.74%	8.13%	3.99%	8.54%	3.98%
16	18.64%	23.66%	7.45%	3.79%	8.12%	4.34%
17	16.04%	15.63%	7.00%	4.11%	6.71%	2.58%
18	17.94%	18.26%	7.08%	3.80%	9.24%	2.96%
19	16.61%	20.31%	8.67%	6.24%	7.25%	3.36%
20	21.44%	21.41%	5.50%	3.83%	9.17%	5.05%
21	19.01%	20.86%	6.97%	4.49%	8.61%	5.88%
22	19.38%	11.38%	6.90%	3.76%	7.75%	4.05%
23	20.08%	8.58%	6.39%	3.06%	9.24%	3.53%
24	19.62%	8.92%	8.04%	5.46%	5.96%	1.72%
25	19.80%	10.02%	5.59%	1.22%	5.95%	3.60%
26	18.04%	6.84%	5.65%	5.22%	12.87%	8.36%
27	16.12%	7.99%	4.48%	1.57%	5.79%	1.79%
28	20.13%	7.27%	8.51%	3.83%	9.76%	3.69%

Political non-news						
	Algorithmic nudge		User nudge		Control	
Day	R	W	R	W	R	W
1	10.51%	10.50%	10.34%	10.18%	9.54%	11.64%
2	11.27%	12.52%	10.41%	10.42%	9.45%	11.03%
3	10.39%	10.02%	10.19%	8.91%	9.26%	12.39%
4	10.87%	10.74%	11.71%	10.51%	9.10%	10.41%
5	11.19%	12.08%	11.43%	9.26%	9.86%	11.02%
6	11.07%	13.18%	10.03%	8.56%	9.79%	12.83%
7	12.15%	11.77%	11.90%	10.31%	9.89%	11.79%
8	10.45%	7.96%	10.83%	10.24%	9.31%	12.12%
9	11.40%	8.57%	10.68%	8.58%	9.23%	9.40%
10	12.46%	10.32%	8.63%	8.31%	11.36%	15.60%
11	11.52%	7.31%	10.23%	8.20%	10.52%	14.79%
12	11.77%	9.85%	13.82%	12.13%	9.89%	10.33%
13	11.93%	9.12%	11.68%	6.74%	11.48%	9.90%
14	12.07%	14.41%	10.74%	8.47%	9.56%	7.53%
15	11.53%	10.62%	9.70%	7.29%	10.76%	10.07%
16	10.06%	6.49%	8.95%	9.74%	9.95%	7.20%
17	8.99%	7.54%	9.65%	9.58%	10.14%	9.83%
18	10.60%	9.13%	9.62%	7.98%	9.64%	9.05%
19	11.42%	11.96%	9.72%	10.94%	9.00%	8.03%
20	15.06%	13.96%	8.56%	5.99%	7.89%	8.25%
21	11.25%	9.83%	9.25%	7.03%	7.00%	5.88%
22	12.64%	12.14%	9.95%	13.47%	9.80%	9.12%
23	8.77%	8.58%	9.70%	8.91%	11.07%	10.83%
24	11.66%	11.89%	11.95%	10.75%	9.62%	8.11%
25	11.54%	10.24%	10.41%	7.14%	8.03%	11.83%
26	14.77%	13.46%	11.17%	7.83%	11.26%	10.80%
27	11.99%	11.38%	9.17%	6.79%	9.73%	9.52%
28	9.25%	11.82%	10.74%	6.77%	9.50%	12.55%

Table S30. Daily percentage of news videos watched by (W) and recommended to (R) participants. The separations at days 7 and 21 indicate the start and end of intervention respectively.

V. Participant Notices.

V.1. Invitation.

You are invited to participate in a research study on how you use YouTube. The study is sponsored by researchers at UC Davis and at the University of Amsterdam. It will involve two surveys over three weeks. Eligible respondents will earn 5,000 points for completing this survey. If you are willing to participate, YouGov will first re-direct you out of the YouGov survey platform to the Chrome Web Store, where you can learn more about the study and later install an open-source browser extension called ResearchTube. The open-source extension, designed by UC Davis and University of Amsterdam researchers, records the URL of the YouTube videos users watch. After installing ResearchTube, you will automatically return to the YouGov survey platform to complete a five-minute survey.

To be eligible to install ResearchTube, you must be a YouTube user and be completing the survey on a desktop using the Google Chrome browser. If you are using a mobile device but would like to participate, please click on the survey invitation again using a desktop computer. If you install the browser extension and return to the YouGov platform to complete the five-minute survey, you will receive 5,000 points.

Once you have completed the survey on YouGov's platform, the browser extension you have installed will record the URLs of any YouTube videos that you watch. After three weeks, YouGov will invite you to complete another 5-minute survey. You may also receive an invitation to this survey as a pop-up reminder from ResearchTube while browsing YouTube.

One week after the second survey, the study sponsors will deactivate ResearchTube, and YouGov will send you a confirmation email that the ResearchTube browser extension has been deactivated and is no longer collecting any data. We will also send you instructions about how you can remove the extension entirely from your desktop. You may remove the extension from your desktop at any time by following these https://support.google.com/chrome_webstore/answer/2664769?hl=en">instructions, but please note that if you uninstall ResearchTube during the three weeks of this study you will not be eligible for inclusion in the second survey.

IMPORTANT –YouGov will not have access to the data collected within ResearchTube. ResearchTube only collects your YouTube URLs with timestamps and does not collect any other information about you, such as your online behaviors or log-ins. Once the study has been completed, YouGov and the study sponsors will exchange an ID allowing the sponsors to link the results of the surveys to data collected through Research Tube. The results of this study may be published in professional and/or scientific journals. However, no individual subject will be identified and all information will be reported as group averages.

You will have an opportunity to read more about ResearchTube and its privacy practices before deciding whether to install it if you decide to follow the Chrome webstore where you can download ResearchTube. You can opt out of this study at any time by emailing youtubestudy@yougov.com. You may also contact the study sponsor with additional questions by emailing mwojcieszak@ucdavis.edu or magdalena.wojcieszak@uva.nl

V.2. Registration.

Welcome!

I'm Magdalena Wojcieszak, a Professor at UC Davis and an Associate Researcher at the University of Amsterdam. Together with my Team (see our photos below), we are inviting you to participate in a study that is part of an international research project (ERC EXPO). We want to understand people's opinions and how they use YouTube.

Why?

YouTube is a very popular and widely used platform. Knowing how people use it and what types of videos they watch is very important. We commit to protecting your confidentiality and privacy during our research.

What?

To understand how you use YouTube, we ask you to install a web browser extension on your computer. The extension will record the URL of the videos your visit on YouTube.

Important:

We do not collect any other information, only the YouTube videos you watch. Please note that your experience of YouTube may slightly change during these three weeks due to our study. It will not, however, interfere with your ability to use YouTube as usual. Our extension is not a black-box owned by a company or a corporation. We have created it ourselves at UC Davis and are committed to open-sourcing it at the end of our research. Our project is motivated by transparency, open access, and open research principles. Our research is in accordance with the highest ethical guidelines and standards and has been approved by the Ethical Review Board of the University of Amsterdam and also of the European Research Council. All the information about privacy can be found here. How?

The study will last 4 weeks. Today, you are asked to install our Chrome extension and complete a short survey (about 5 minutes). In three weeks, we will ask you to complete another short survey (about 5 minutes), and then leave the extension installed for one more week. If you participate, you will receive 5,000 YouGov polling points for installing the Chrome extension and taking the first survey. If you do not install the plug-in, you will not be able to take part in this study. You can only take each survey once.

Who?

Our team has many great members... among others:

Software development (browser extension) developed by Muhammad Haroon Informed consent and privacy policies developed by Ericka Menchen Trevino & Magdalena Wojcieszak Questionnaire and study development and design done by the entire Team Get in touch!

Drop an email to magdalena.wojcieszak@uva.nl or mwojcieszak@ucdavis.edu if you have any questions about the study Drop an email to mharoon@ucdavis.edu if you want more information about the plug-in We are also on Twitter @ERCEXPO and under our individual handles.

V.3. Debrief.

Thank you for your participation in the study. Your participation is very important to us. You were initially told that the purpose of this study is to analyze how people use YouTube. Now we can additionally tell you that the goal of the study was to investigate how to incentivize greater news consumption on the platform. Specifically, we were interested in whether encouraging you to search for news on YouTube, including a set of four news videos on the top of your YouTube homepage, and injecting into your watch history a set of videos from quality and balanced/not-strongly partisan news media organizations from this site https://adfontesmedia.com/ can encourage individuals to watch more news. You were randomly assigned to one of the following conditions: some of you were in the "search banner" condition, some of you were in the "panel with news" condition, some of you were in the "background news injection" condition, and some of you were in the control group (no change to your YouTube experience) In order to make the study realistic, we did not tell you any of this at the beginning of the experiment.

Please note that right now there is no need for you to do anything else related to our ResearchTube project. We will deactivate the plugin in 1 week after we have the chance to check the data.

Your responses will be included in statistical analyses and only in an aggregate form.

If you have any questions, please contact the project leader at any time: Magdalena Wojcieszak, mwojcieszak@ucdavis.edu or 530 304 XXXX.

Thank you again for your participation!

V.4. Privacy Policy.

Last updated: May 06, 2022

This Privacy Policy describes our policies and procedures on the collection, use and disclosure of Your information when You use the Service and tells You about Your privacy rights and how the law protects You.

We use Your Personal data to provide and improve the Service. By using the Service, You agree to the collection and use of information in accordance with this Privacy Policy.

Interpretation and Definitions

Interpretation

The words of which the initial letter is capitalized have meanings defined under the following conditions. The following definitions shall have the same meaning regardless of whether they appear in singular or in plural.

Definitions

For the purposes of this Privacy Policy:

- **Account** means a unique account created for You to access our Service or parts of our Service.
- **Application** means the software program provided by the ERC EXPO Project (PI Magdalena Wojcieszak) downloaded by You on any electronic device, named ResearchTube
- **Company** (referred to as either "the Company", "We", "Us" or "Our" in this Agreement) refers to ERC EXPO Project.
- **Country** refers to: California, United States
- **Device** means any device that can access the Service such as a computer, a cellphone or a digital tablet.
- **Personal Data** is any information that relates to an identified or identifiable individual.
- **Service** refers to the Application.
- **Service Provider** means any natural or legal person who processes the data on behalf of the Company. It refers to third-party companies or individuals employed by the Company to facilitate the Service, to provide the Service on behalf of the Company, to perform services related to the Service or to assist the Company in analyzing how the Service is used.
- **Usage Data** refers to data collected automatically, either generated by the use of the Service or from the Service infrastructure itself (for example, the duration of a page visit).
- **You** means the individual accessing or using the Service, or the company, or other legal entity on behalf of which such individual is accessing or using the Service, as applicable.

Collecting and Using Your Personal Data

Types of Data Collected

Personal Data

While using Our Service, We will not ask You to provide Us with any personally identifiable information.

Usage Data

Usage Data are collected automatically when using the Service.

Usage Data may include information such as Your Device's Internet Protocol address (e.g. IP address), browser type, browser version, the pages of our Service that You visit, the time and date of Your visit, the time spent on YouTube videos, unique device identifiers and other diagnostic data.

We may also collect information that Your browser sends whenever You visit our Service or when You access the Service by or through a mobile device.

- **For other purposes**: We may use Your information for other purposes, such as data analysis, identifying usage trends, determining the effectiveness of our promotional campaigns and to evaluate and improve our Service and your experience.

We will only share datasets that have been processed for scientific research with third scientist and researchers. That is, the data will be stripped of any identifiers (including unique respondent ID), and aggregated to the level that no re-identification of any individual respondent will not be possible.

Retention of Your Personal Data

The Project ERC EXPO will retain Your Data only for up to five years after the last publication from the project. The data will be stored at the secured SurfSara infrastructure.

In order to strike the right balance between data protection law's demand of deleting the data once they are no longer needed, and ethical research requirements to keep that data for verification, the retention period will be five years after the last research article has been published. At that point, all the files will be deleted.

Transfer of Your Personal Data

We collect and use personal data for scientific purposes only, and in accordance with the national and European law (including GDPR). Your information, including Personal Data, is stored safely at Surf Sara, the Dutch research computing infrastructure provider with extensive experience with storing and securing data from many privacy sensitive projects (https://www.surf.nl/en/about-surf/subsidiaries/surfsara). Your information will also be stored at the server of the network of the University of Amsterdam, a state-of-the-art network in terms of data protection. It is important to emphasize that the stored data are accessible only from within the University, through the university network, or from home through an advanced protected Virtual Private Network (VPN). On top of the secure infrastructure, we will use encryption technology every time data are stored or transferred. Your data will not be transferred to other countries with a lower level of data protection. Given the strict requirements about data security in the European data protection law, it is essential that the data be stored and processed in an EU member state. Your consent to this Privacy Policy followed by Your submission of such information represents Your agreement to that transfer.

The Company will take all steps necessary to ensure that Your data are treated securely and in accordance with this Privacy Policy and no transfer of Your Personal Data will take place to an organization or a country unless there are adequate controls in place including the security of Your data and other personal information.

Law enforcement

Under certain circumstances, the Project ERC EXPO may be required to disclose Your Personal Data if required to do so by law or in response to valid requests by public authorities (e.g. a court or a government agency). This situation is highly unlikely, however.

Other legal requirements

The Project ERC EXPO may disclose Your Personal Data in the good faith belief that such action is necessary to:

• Comply with a legal obligation • Prevent or investigate possible wrongdoing in connection with the Service • Protect the personal safety of Users of the Service or the public • Protect against legal liability

Security of Your Personal Data

The security of Your Personal Data is important to Us, but remember that no method of transmission over the Internet, or method of electronic storage is 100% secure. While We strive to use acceptable means to protect Your Personal Data, We cannot guarantee its absolute security.

Data protection law requires keeping User data safe and secure. Ensuring data security will be maximized by storing the raw data collected via the Service in the Dutch national research computing infrastructure provider, Surf Sara. Surf Sara provides data storage and processing services to several other privacy sensitive research projects, such as health and genomics research, and – as such - has the necessary expertise and security standards (https://www.surf.nl/en/about-surf/subsidiaries/surfsara). The employees of Surf Sara do not have access to the data. On top of the secure infrastructure, we will use encryption technology every time data are stored or transferred.

We will not grant access to the unprocessed personal data. We will only share data-sets that have been processed for scientific research with third scientist and researchers. That is, the data will be stripped of any identifiers (including unique respondent ID), and aggregated to the level that no re-identification of any individual respondent will not be possible.

Children's Privacy

Our Service does not address anyone under the age of 18. We do not knowingly collect personally identifiable information from anyone.

Links to Other Websites

Our Service may contain links to other websites that are not operated by Us. If You click on a third party link, You will be directed to that third party's site. We strongly advise You to review the Privacy Policy of every site You visit.

We have no control over and assume no responsibility for the content, privacy policies or practices of any third party sites or services.

Changes to this Privacy Policy

We may update Our Privacy Policy from time to time. We will notify You of any changes by posting the new Privacy Policy on this page.

We will let You know via email and/or a prominent notice on Our Service, prior to the change becoming effective and update the "Last updated" date at the top of this Privacy Policy.

You are advised to review this Privacy Policy periodically for any changes. Changes to this Privacy Policy are effective when they are posted on this page.

Contact Us

If you have any questions about this Privacy Policy, You can contact us:

- By email: mwojcieszak@ucdavis.edu

349 **W. Key Analysis Using Alternative Samples.**

Table S31. OLS Regressions Predicting the Effects of Treatments on YouTube Engagement (Participants Watched at Least One Video During Weeks 2-3)

	Videos watched per active day (W2-3)	Number of active days per week (W2-3)	Time spent per active day (W2-3)
Intercept	3.34 (1.11) **	0.87 (0.21) ***	4.45 (0.29) ***
Age	-0.01 (0.02)	0.00 (0.00)	-0.00 (0.00)
Gender (non-male)	0.84 (0.49)	0.23 (0.09) *	0.00 (0.07)
Education (low)	-0.57 (0.84)	0.00 (0.15)	-0.13 (0.13)
Education (middle)	-0.04 (0.65)	0.08 (0.12)	0.07 (0.10)
Ethnicity (White)	-0.54 (0.55)	0.10 (0.10)	0.02 (0.08)
Party (other)	1.29 (0.71)	0.15 (0.13)	0.16 (0.11)
Party (Republican)	0.25 (0.59)	-0.21 (0.11)	-0.01 (0.09)
Videos watched per active day (W1)	0.45 (0.02) ***	-	-
Number of active days (W1)	-	0.30 (0.02) ***	-
Time spent per active day (W1)	-	-	0.45 (0.03) ***
Condition (algorithm)	0.76 (0.61)	0.53 (0.11) ***	-0.03 (0.09)
Condition (user)	0.75 (0.60)	0.07 (0.11)	0.22 (0.09) *
N	1152	1320	1152
R^2	0.24	0.18	0.21

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. Time spent per active day (minutes) on YouTube (week 1 and weeks 2-3) were log(x+1) transformed. A total of 168 participants were excluded from models 1 and 3 because of missing data for "Videos watched per active day (W1)" or "Time spent per active day (W1)", which resulted from their non-use of YouTube in the first week.

Table S32. OLS Regressions Predicting the Effects of Treatments on YouTube Engagement (Participants watched at least five YouTube videos over weeks 1-3 and at least one video during weeks 2-3)

	Videos watched per active day (W2-3)	Number of active days per week (W2-3)	Time spent per active day (W2-3)
Intercept	3.45 (1.14) **	1.10 (0.22) ***	4.63 (0.30) ***
Age	-0.01 (0.02)	0.00 (0.00)	0.00 (0.00)
Gender (non-male)	0.84 (0.50)	0.25 (0.10) *	0.00 (0.07)
Education (low)	-0.54 (0.85)	-0.01 (0.16)	-0.09 (0.13)
Education (middle)	-0.03 (0.67)	0.05 (0.13)	0.09 (0.10)
Ethnicity (White)	-0.51 (0.56)	0.11 (0.11)	0.03 (0.08)
Party (other)	1.25 (0.72)	0.15 (0.14)	0.15 (0.11)
Party (Republican)	0.31 (0.61)	-0.19 (0.12)	-0.01 (0.09)
Videos watched per active day (W1)	0.44 (0.02) ***	-	-
Number of active days (W1)	-	0.26 (0.02) ***	-
Time spent per active day (W1)	-	-	0.43 (0.03) ***
Condition (algorithm)	0.76 (0.62)	0.55 (0.12) ***	-0.06 (0.09)
Condition (user)	0.73 (0.62)	0.06 (0.12)	0.19 (0.09) *
N	1125	1228	1125
R^2	0.24	0.14	0.18

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. Time spent per active day (minutes) on YouTube (week 1 and weeks 2-3) were log(x+1) transformed. A total of 103 participants were excluded from models 1 and 3 because of missing data for "Videos watched per active day (W1)" or "Time spent per active day (W1)", which resulted from their non-use of YouTube in the first week.

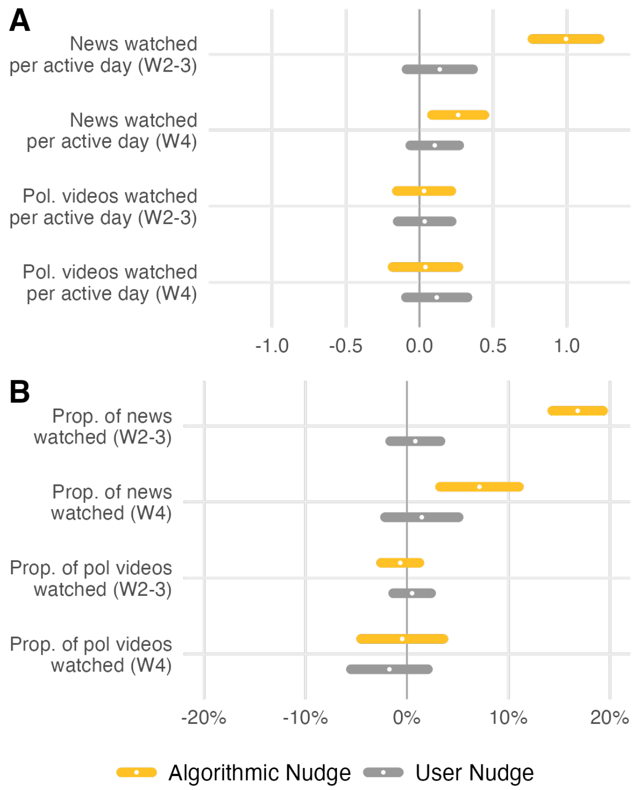


Fig. S5. Effects of treatments on the consumption of news videos and political videos. Dots represent coefficients and horizontal bars represent 95% confidence intervals. Participants included watched at least one video during weeks 2-3.

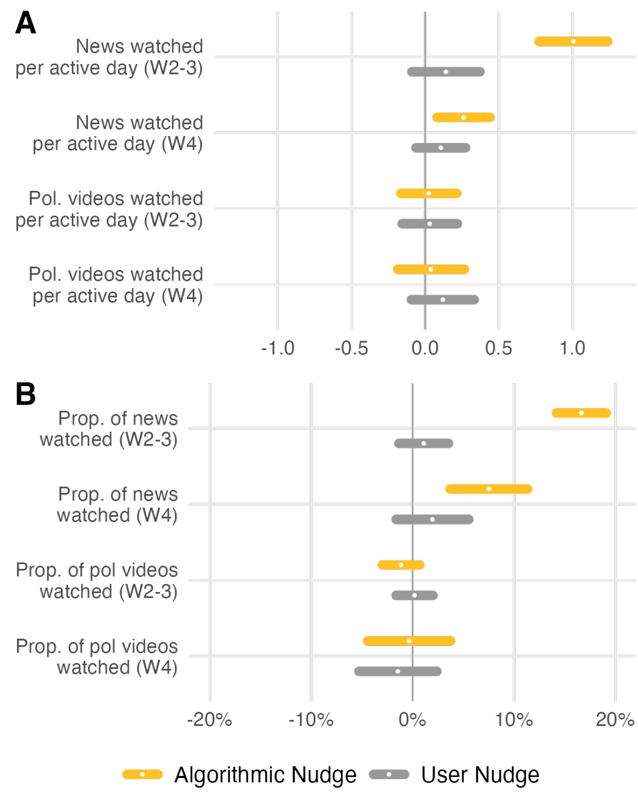


Fig. S6. Effects of treatments on the consumption of news videos and political videos. Dots represent coefficients and horizontal bars represent 95% confidence intervals. Participants included watched at least five YouTube videos over weeks 1-3 and at least one video during weeks 2-3.

Table S33. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Number Per Active Day, Participants Watched at Least One Video During Weeks 2-3)

	News watched per active day (W2-3)	News watched per active day (W4)	Pol. videos watched per active day (W2-3)	Pol. videos watched per active day (W4)
Intercept	-0.16 (0.21)	-0.18 (0.15)	0.10 (0.17)	0.04 (0.19)
Age	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Gender (non-male)	0.09 (0.09)	0.09 (0.07)	-0.08 (0.08)	0.08 (0.09)
Education (low)	0.12 (0.16)	-0.02 (0.12)	-0.15 (0.13)	0.04 (0.15)
Education (middle)	-0.00 (0.13)	-0.08 (0.09)	-0.06 (0.10)	0.04 (0.12)
Ethnicity (White)	-0.15 (0.11)	0.09 (0.08)	0.13 (0.09)	-0.21 (0.10) *
Party (other)	0.12 (0.14)	-0.09 (0.11)	0.21 (0.11)	0.02 (0.14)
Party (Republican)	0.06 (0.11)	0.13 (0.09)	-0.01 (0.09)	0.07 (0.11)
News watched per active day (W1)	0.41 (0.04) ***	0.37 (0.04) ***	-	-
Pol. videos watched per active day (W1)	-	-	0.58 (0.02) ***	0.61 (0.04) ***
Condition (algorithm)	0.99 (0.12) ***	0.26 (0.09) **	0.03 (0.09)	0.04 (0.11)
Condition (user)	0.14 (0.12)	0.10 (0.08)	0.03 (0.09)	0.12 (0.11)
N	1152	534	1152	534
R^2	0.16	0.14	0.45	0.29

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 168 participants were excluded from models 1 and 3 because of missing data for "News watched per active day (W1)" or "Pol. videos watched per active day (W1)", which resulted from their non-use of YouTube in the first week. An additional 618 participants were excluded from models 2 and 4 due to missing data for "News watched per active day (W4)" or "Pol. videos watched per active day (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

Table S34. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Percentage, Participants Watched at Least One Video During Weeks 2-3)

	Prop. of news watched (W2-3)	Prop. of news watched (W4)	Prop. of pol videos watched (W2-3)	Prop. of pol videos watched (W4)
Intercept	-0.14 (2.26)	-4.13 (3.35)	2.36 (1.71)	7.79 (3.50) *
Age	0.03 (0.03)	0.05 (0.05)	0.08 (0.02) ***	0.07 (0.05)
Gender (non-male)	0.38 (1.03)	1.17 (1.57)	-0.29 (0.78)	-3.12 (1.64)
Education (low)	0.51 (1.75)	1.26 (2.66)	-1.76 (1.32)	2.71 (2.78)
Education (middle)	1.19 (1.38)	0.90 (2.00)	-0.71 (1.04)	2.08 (2.10)
Ethnicity (White)	-0.21 (1.15)	2.22 (1.79)	-0.05 (0.87)	-4.59 (1.87) *
Party (other)	-0.91 (1.49)	-1.82 (2.35)	0.32 (1.13)	-1.28 (2.47)
Party (Republican)	-1.52 (1.24)	2.86 (1.92)	-0.08 (0.94)	-1.63 (2.01)
Prop. of news watched (W1)	0.58 (0.05) ***	0.50 (0.06) ***	-	-
Prop. of pol videos watched (W1)	-	-	0.39 (0.03) ***	0.44 (0.06) ***
Condition (algorithm)	16.83 (1.28) ***	7.15 (1.98) ***	-0.66 (0.96)	-0.46 (2.07)
Condition (user)	0.83 (1.27)	1.46 (1.85)	0.51 (0.96)	-1.73 (1.94)
N	1152	534	1152	534
R^2	0.26	0.14	0.17	0.12

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 168 participants were excluded from models 1 and 3 because of missing data for "Prop. of news watched (W1)" or "Prop. of pol. videos watched (W1)", which resulted from their non-use of YouTube in the first week. An additional 618 participants were excluded from models 2 and 4 due to missing data for "Prop. of news watched (W4)" or "Prop. of pol. videos watched (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

Table S35. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Number Per Active Day, Participants Watched at Least Five YouTube Videos Over Weeks 1-3 and at Least One Video During weeks 2-3)

	News watched per active day (W2-3)	News watched per active day (W4)	Pol. videos watched per active day (W2-3)	Pol. videos watched per active day (W4)
Intercept	-0.16 (0.21)	-0.18 (0.16)	0.11 (0.17)	0.04 (0.20)
Age	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Gender (non-male)	0.11 (0.10)	0.09 (0.07)	-0.08 (0.08)	0.08 (0.09)
Education (low)	0.12 (0.16)	-0.03 (0.12)	-0.15 (0.13)	0.04 (0.16)
Education (middle)	-0.01 (0.13)	-0.09 (0.09)	-0.06 (0.10)	0.04 (0.12)
Ethnicity (White)	-0.15 (0.11)	0.09 (0.08)	0.14 (0.09)	-0.21 (0.10) *
Party (other)	0.11 (0.14)	-0.09 (0.11)	0.21 (0.11)	0.02 (0.14)
Party (Republican)	0.07 (0.12)	0.13 (0.09)	-0.01 (0.09)	0.07 (0.11)
News watched per active day (W1)	0.41 (0.04) ***	0.37 (0.04) ***	-	-
Pol. videos watched per active day (W1)	-	-	0.58 (0.02) ***	0.61 (0.04) ***
Condition (algorithm)	1.01 (0.12) ***	0.26 (0.09) **	0.02 (0.10)	0.04 (0.12)
Condition (user)	0.14 (0.12)	0.11 (0.09)	0.03 (0.10)	0.12 (0.11)
N	1125	525	1125	525
R^2	0.16	0.14	0.45	0.29

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 103 participants were excluded from models 1 and 3 because of missing data for "News watched per active day (W1)" or "Pol. videos watched per active day (W1)", which resulted from their non-use of YouTube in the first week. An additional 600 participants were excluded from models 2 and 4 due to missing data for "News watched per active day (W4)" or "Pol. videos watched per active day (W4)", caused by either non-use of YouTube in week 4 or or the uninstallation of the extension before week 4.

Table S36. OLS Regressions Predicting the Effects of Treatments on News Watching and Political Watching (Percentage, Participants Watched at Least Five YouTube Videos Over Weeks 1-3 and at Least One Video During weeks 2-3)

	Prop. of news watched (W2-3)	Prop. of news watched (W4)	Prop. of pol videos watched (W2-3)	Prop. of pol videos watched (W4)
Intercept	-0.56 (2.23)	-4.09 (3.31)	2.86 (1.66)	6.97 (3.55)
Age	0.04 (0.03)	0.05 (0.05)	0.08 (0.02) ***	0.07 (0.05)
Gender (non-male)	0.73 (1.01)	1.54 (1.55)	-0.37 (0.75)	-3.09 (1.66)
Education (low)	0.33 (1.72)	1.07 (2.62)	-1.70 (1.28)	3.20 (2.82)
Education (middle)	0.64 (1.35)	0.37 (1.98)	-1.02 (1.00)	2.56 (2.13)
Ethnicity (White)	-0.33 (1.13)	1.97 (1.75)	-0.24 (0.84)	-4.15 (1.88) *
Party (other)	-0.66 (1.45)	-1.77 (2.31)	0.41 (1.08)	-1.32 (2.49)
Party (Republican)	-1.57 (1.23)	2.52 (1.90)	-0.27 (0.91)	-1.98 (2.05)
Prop. of news watched (W1)	0.59 (0.04) ***	0.52 (0.06) ***	-	-
Prop. of pol videos watched (W1)	-	-	0.40 (0.03) ***	0.43 (0.06) ***
Condition (algorithm)	16.65 (1.26) ***	7.53 (1.95) ***	-1.15 (0.93)	-0.34 (2.09)
Condition (user)	1.10 (1.24)	1.96 (1.83)	0.19 (0.93)	-1.46 (1.96)
N	1125	525	1125	525
R^2	0.27	0.15	0.19	0.12

Note: *** $<$ .001 ** $<$ .01 * $<$ .05. A total of 103 participants were excluded from models 1 and 3 because of missing data for "Prop. of news watched (W1)" or "Prop. of pol. videos watched (W1)", which resulted from their non-use of YouTube in the first week. An additional 600 participants were excluded from models 2 and 4 due to missing data for "Prop. of news watched (W4)" or "Prop. of pol. videos watched (W4)", caused by either non-use of YouTube in week 4 or the uninstallation of the extension before week 4.

References

1. Mikko Villi, Tali Aharoni, Keren Tenenboim-Weinblatt, Pablo J. Boczkowski, Kaori Hayashi, Eugenia Mitchelstein, Akira Tanaka, and Neta Kligler-Vilenchik. Taking a break from news: A five-nation study of news avoidance in the digital era. *Digital Journalism*, 10(1):148–164, Jan 2022. ISSN 2167-0811, 2167-082X. .
2. Shalom H Schwartz. Basic values: how they motivate and inhibit prosocial behavior. In M Mikulincer and P. R. Shaver, editors, *Prosocial motives, emotions, and behavior: The better angels of our nature*, pages 221–241. American Psychological Association, Washington, DC, 2010.
3. Toby Bolsen and Thomas J Leeper. Self-interest and attention to news among issue publics. *Political Communication*, 30 (3):329–348, 2013.
4. Vanessa Otero. Ad fontes media's first multi-analyst content analysis ratings project, 2019.
5. Vanessa Otero. Multi-analyst content analysis methodology, September 2021. URL https://adfontesmedia.com/white-paper-2021/.
6. Hause Lin, Jana Lasser, Stephan Lewandowsky, Rocky Cole, Andrew Gully, David G. Rand, and Gordon Pennycook. High level of correspondence across different news domain quality rating sets. *PNAS Nexus*, 2(9):pgad286, September 2023. ISSN 2752-6542. . URL https://doi.org/10.1093/pnasnexus/pgad286.
7. Ad Fontes Media. https://adfontesmedia.com/interactive-media-bias-chart/, 2022. Accessed: 2024-07-24.
8. NewsGuard. https://www.newsguardtech.com/solutions/news-reliability-ratings/, 2024. Accessed: 2024-07-24.
9. Media Bias/Fact Check. https://mediabiasfactcheck.com, 2024. Accessed: 2024-07-24.
10. Magdalena Wojcieszak, Sjifra de Leeuw, Ericka Menchen-Trevino, Seungsu Lee, Ke Maddie Huang-Isherwood, and Brian Weeks. No polarization from partisan news. *International Journal of Press and Politics*, Sep 2021. . URL https://osf.io/hqmuy/.
11. Mark Ledwich and Anna Zaitsev. Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. *First Monday*, 25(3), Feb. 2020. . URL https://journals.uic.edu/ojs/index.php/fm/article/view/10419.
12. Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgílio AF Almeida, and Wagner Meira Jr. Auditing Radicalization Pathways on YouTube. In *ACM Conference on Fairness, Accountability, and Transparency (FAT*)*, 2020.
13. Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Ro{bert}a: A robustly optimized {bert} pretraining approach, 2020. URL https://openreview.net/forum?id=SyxS0T4tvS.
14. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.
15. Homa Hosseinmardi, Amir Ghasemian, Aaron Clauset, Markus Mobius, David M Rothschild, and Duncan J Watts. Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences*, 118(32), 2021.
16. Mark Ledwich, Anna Zaitsev, and Anton Laukemper. Radical bubbles on YouTube? Revisiting algorithmic extremism with personalised recommendations. *First Monday*, 27(12), Dec. 2022. . URL https://journals.uic.edu/ojs/index.php/fm/article/view/12552.
17. Pew Research. American trends panel wave 112. https://www.pewresearch.org/dataset/american-trends-panel-wave-112/, 2024. Accessed: 2024-08-03.

18. Grant Blank, Gillian Bolsover, and Elizabeth Dubois. A new privacy paradox: Young people and privacy on social network sites. In *Prepared for the Annual Meeting of the American Sociological Association*, volume 17, 2014.
19. Magdalena Wojcieszak, Bernhard Clemm von Hohenberg, Andreu Casas, Ericka Menchen-Trevino, Sjifra de Leeuw, Alexandre Gonçalves, and Miriam Boon. Null effects of news exposure: a test of the (un) desirable effects of a 'news vacation'and 'news binging'. *Humanities and Social Sciences Communications*, 9(1):1–10, 2022.
20. U.S. desktop internet browser market share 2023. https://www.statista.com/statistics/272697/market-share-desktop-internet-browser-usa/, 2023. Accessed: 2023-7-19.