CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE 📊 🎓

Combating Hateful Attitudes and Online Browsing Behavior: The Case of Antisemitism

Catie Snow Bailard[1] 🆔, Matthew H. Graham[2] 🆔, Kimberly Gross[1], Ethan Porter[1] 🆔 and Rebekah Tromble[1] 🆔

[1]School of Media and Public Affairs, George Washington University, Washington, DC, USA and
[2]Department of Political Science, Temple University, Philadelphia, PA, USA
Corresponding author: Ethan Porter; Email: evporter@gwu.edu

Abstract
This study adds to the analogic perspective-taking literature by examining whether an online perspective-taking intervention affects both antisemitic attitudes and behaviors – in particular, engagement with antisemitic websites. Subjects who were randomly assigned to the treatment viewed a 90-s video of a college student describing an experience with antisemitism and reflected on its similarity to their own experiences. In a survey, treated subjects reported greater feelings of sympathy (+29 p.p.), more positive feelings toward Jews, a greater sense that Jews are discriminated against, and more support for policy solutions (+2–4 p.p.). However, these effects did not persist after 14 days. Examining our subjects' web browsing data, we find a 5% reduction in time spent viewing antisemitic content during the posttreatment period and some limited, suggestive evidence of effects on the number of site visits. These findings provide the first evidence that perspective-taking interventions may affect online browsing behavior.

What can be done to combat antisemitism online? Today, antisemitic attacks constitute the largest share of religion-based hate crimes in the United States (FBI, 2021). A majority of US Jews report feeling less safe than 5 years previously, with

anxiety especially acute among those who wear Jewish items of clothing (Alper and Cooperman, 2021). Rather than fading away as the population ages, antisemitic attitudes are most apparent among young people (Hersh and Royden, 2021).

In this paper, we test an online analogic perspective-taking intervention designed to combat antisemitism. Building on earlier work on the power of perspective-taking (Galinsky and Moskowitz, 2000; Galinsky and Ku, 2004), analogic perspective-taking exposes participants to a narrative about a target group and asks them to reflect on how it intersects with their own experiences (Kalla and Broockman, 2021). Analogic perspective-taking has been shown to reduce prejudiced attitudes toward the transgender community (Broockman and Kalla, 2016), refugees (Adida, Lo, and Platas, 2018), and undocumented immigrants (Kalla and Broockman, 2020). We find evidence that an online analogic perspective-taking intervention – a short video of a college student conveying his experiences with antisemitism and a prompt to relate one's own experiences to the student's – can reduce antisemitic attitudes. This intervention combines perhaps the most effective feature of analogic perspective-taking—hearing the perspective of the member of an out-group (Kalla and Broockman, 2021)—with an exercise meant to analogize the experience for the participant. However, we find no evidence that the attitudinal effects of our intervention persisted after 2 weeks.

Going beyond the typical focus on attitudes, we conduct the first test of a perspective-taking intervention's effect on web browsing behavior. To do so, we relied on YouGov to collect our subjects' Internet browsing histories. Our estimates suggest that the intervention caused about a 5% decline in the amount of time spent on antisemitic websites. Effects were concentrated among those who had visited at least one hateful or abusive (H/A) website in the pretreatment period. Among those participants, the treatment caused about a 35% decline in time spent on antisemitic websites. We also find suggestive but more limited evidence that the treatment reduced the proportion of subjects who visited at least one antisemitic website and the number of sites visited. This suggests that the primary effect of our intervention may have been to make antisemitic content less appealing once encountered, not to convince subjects to avoid it altogether.

Expectations

Though attitudes toward Jews have been well studied (e.g., Weil, 1985; Weisberg, 2019), the recent wave of research on prejudice reduction has largely overlooked this group. Based on efforts to reduce prejudice toward other groups, we expected the treatment to induce sympathetic feelings (H1) and increase warm feelings (H2) toward Jews (Burns and Granz, 2021; Kinder and Sanders, 1996; Baston et al., 1997, 2002; Broockman and Kalla, 2016). Also in line with previous findings (Kalla and Broockman, 2020), we expected it to increase the perception that Jews are discriminated against (H3) and increase support for policies to combat antisemitic discrimination (H4). Finally, as earlier work on perspective-taking linked it to a decreased reliance on stereotypes (Galinsky and Moskowitz, 2000), we anticipated that our intervention would reduce endorsements of antisemitic stereotypes (H5).

We did not set out with clear expectations regarding browsing behavior. On the one hand, our intervention could provoke the expected attitudinal changes, without

discernible effects on behavior. On the other hand, our intervention could change behavior even as beliefs remained stable, as has been observed elsewhere in the prejudice reduction literature (Paluck, 2009). Worse, the intervention could backfire, for example, by inducing curiosity about antisemitism. Given this, we articulated research questions pertaining to engagement with H/A content (RQ2) and other browsing behavior (RQ3). We also took advantage of our multiwave design to study whether analogic perspective-taking changes self-reported contact with the out-group (RQ1).

Research design

Identifying hateful or abusive content

We define H/A websites as those containing hateful, bigoted, discriminatory, and/or otherwise highly abusive content directed at groups of people, including the comments found on the website or in replies to posts. Here, we briefly describe our preregistered approach to identifying such sites. The appendix elaborates.

We seeded our list of URLs using reports on extremism from the Southern Poverty Law Center and the Anti-Defamation League. Next, we used Google searches conducted in "incognito" or "private" mode to identify relevant websites and social media accounts and channels associated with the individuals and organizations named in the reports. This included Bitchute, Dlive, Facebook, Gab, Parler, Telegram, Twitter, VK, YouTube, subreddits, and podcasts. We then used a snowball approach, examining each page for new names and evaluating these using the same criteria. We ended our collection when more than 70% of the links found on 20 consecutive pages were already part of our list.

The resulting URLs range from neo-Nazi and white supremacist sites to the so-called "alt lite" (ADL, 2017) and "intellectual dark web" (Weiss and Winter, 2018). We do not claim that any given website contains entirely, or even primarily, hateful content (although a number do), but rather, that someone spending much time on one of these sites would be likely to be exposed to hateful, discriminatory, bigoted, and/or highly abusive content and/or comments directed at groups of people. Thus, our list is meant to include individuals and organizations who do not directly espouse hateful or bigoted views themselves, but regularly provide space (e.g., in interviews, guest posts, etc.) for such views to be espoused. Given our interest in antisemitism, we placed the URLs in two categories, one for sites that promote or host antisemitic views and another for other H/A sites.

Perspective-taking experiment

The experiment proceeded as follows. First, YouGov provided us with two weeks of browsing history for panelists who had agreed to share their browsing data through YouGov Pulse. Pulse has been used in prior descriptive research (Guess, Nyhan, and Reifler, 2020; Chen et al., 2021), but to our knowledge has never been used to measure a behavioral outcome in an experiment. We examined this pretreatment data to identify participants who had previously visited H/A websites and directed YouGov to make an extra effort to recruit these individuals.
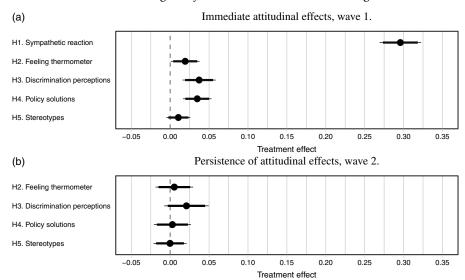
(a)

Immediate attitudinal effects, wave 1.

(b)

Persistence of attitudinal effects, wave 2.

Figure 1. Attitudinal effects.
Note: Figure displays treatment effect estimates for the attitudinal measures. Dots are point estimates. Thick (thin) error bars are 90 (95) percent confidence intervals.

Then, we conducted a two-wave survey experiment. Upon entering the first wave, participants answered pretreatment questions and were then randomly assigned to either treatment or control. Our decision to use a pure control was motivated by recent evidence showing that placebo content can inadvertently affect treatment effect estimates (Porter and Velez, 2021, Figure 1) and that demand effects in survey experiments are less frequent than commonly assumed (Mummolo and Peterson, 2019).

Emulating similar interventions (Broockman and Kalla, 2016), including those conducted in survey settings (Hertel-Fernandez and Porter, 2021), participants assigned to treatment were first asked if they knew someone who was Jewish and experienced antisemitism. If they answered yes, they were provided a text box to describe the experience; if not, they were asked to write about what it would be like.

Following this initial round of engagement on the topic of antisemitism, all treatment participants watched a video of a Jewish college student describing his experience with swastikas on campus. This student's experience echoes antisemitic incidents that occur regularly across US college campuses (Post, 2021; Fox, 2021). A video of the intervention can be found at https://youtu.be/IaCMgY2g13s. The complete script follows:

My name is Max, and I want to tell you about something that happened to me.

When I was eighteen years old, seven swastikas were found on my college campus.

I had just started college. I've been Jewish my whole life. But this was the first time I ever felt targeted for my identity. This was the first time I ever felt

uncertain about my place in a community because of my Judaism. And this was the first time I ever felt unsafe because of my religion.

Swastikas were drawn all over campus, in places I'd been familiar with. Places I'd walked by before. Places that were just a part of my everyday life.

The swastika is the ultimate symbol of the regime that killed millions of Jews. I can't see them without thinking about concentration camps.

The swastikas were going up all over campus during the holiest days of the Jewish religion. I remember getting a text message as I returned to my dorm from services that another one had been found.

Thinking about the swastikas now leaves a pit in my stomach. It was as if I was being told that I didn't belong because of my religion–that someone hated me just because of my religion.

Thanks for listening.

After watching the video, participants were asked to describe a time when they had been targeted for something that made them different from other people. If they could not recall a time, they were asked to describe when such an experience may have happened to a friend or family member. They then answered distractor questions about video quality.[1] In Appendix 2.5, we summarize subjects' responses to the open-ended items; the large majority of subjects appear to have taken the exercise seriously.

Following their assigned condition, all participants answered attitudinal outcome questions. To measure sympathy, participants answered how well a set of emotions (including sympathy) described how they were feeling on a 1–4 scale ("Not at all" to "A lot"). General warmth was measured with feeling thermometers for (in random order) 10 groups, including Jews. For perceptions of discrimination, participants rated how much discrimination 10 groups faced on a 1–4 scale ("A lot" to "None at all"). To evaluate support for policy solutions, participants were asked whether the government was doing enough for Jewish citizens on a 1–4 scale ("Definitely doing enough" to "Definitely not doing enough") and about support for hate crime legislation on a 1–5 scale ("Strongly favor" to "Strongly oppose"). Finally, to measure stereotype endorsement, participants were presented with thirteen stereotypes (in random order) and asked to evaluate the truthfulness of each one on a 1–5 scale ("Definitely true" to "Definitely false"). The complete text of all items can be found in the appendix.

For the next two weeks, we passively collected participants' Internet browsing history. After two weeks, we recontacted participants to measure the durability of attitudinal effects. Paluck and Green (2009) recommend that more work in the

---

[1]To be clear, treatment participants only watched the video once.

prejudice reduction literature measure the persistence of treatment effects. The present study joins a small but growing effort to take up this advice and investigate the durability of perspective-taking/getting efforts with panel designs (e.g., Adida, Lo, and Platas, 2018; Simonovitz, Kudzu, and Kardos, 2018; Kalla and Broockman, 2021).

Estimation

Our preregistered analytic plan included the following key features. (An anonymized version is available at https://bit.ly/3SwLdMG.) We estimated treatment effects using OLS regression with robust standard errors and covariate adjustment. We selected covariates by fitting a LASSO regression on the control group for each outcome variable, with the penalty term selected by cross-validation. All covariates selected by LASSO were used for covariate adjustment. The appendix includes unadjusted estimates of each treatment effect and corresponding regression tables.

Prior to analysis, we coded all categorical covariates and outcomes to range between 0 and 1, with equal space between intermediate levels (e.g., a three-level variable would be 0, 0.5, and 1). In the browsing data, our preregistered hypothesis tests transform all count variables (number of site visits and time spent) using the formula $\log(X + 1)$. This gives these outcomes a "percent change" interpretation and limits the influence of extreme values.

We carried out several quality checks. First, we found no evidence that treatment affected the probability of responding to the wave 2 survey or the types of subjects that responded (Appendix 1.1). Second, because Pulse is a dynamic panel that permits participants to drop out from being monitored, we investigated differential attrition based on Pulse participation; we found no evidence of this kind of attrition (Appendix 1.1). Third, we noticed that one of our treatment group subjects was an outlier in terms of their pretreatment level of browsing antisemitic sites. Our results are similar when dropping this respondent (Appendix 3.4). Finally, we conducted an exploratory analysis of differential breakoff within the wave 1 survey (i.e., within-wave attrition).[2] We found differential breakoff on average: among subjects who reached the final pretreatment question, 98.8% of control subjects completed the survey, compared with 95.0% of treated subjects (difference = 3.8, s.e. = 0.7; Table SI-5). Fortunately, pretreatment covariates were not collectively predictive of breakoff (Table SI-6), and estimates weighted using the procedure recommended by Gerber and Green (2012, Chapter 7) are substantively identical to our main estimates (Tables SI-15 and SI-30). In the concluding section, we discuss the implications for research design.

Results

Attitude change

Our first attitudinal hypothesis was that perspective-taking would make respondents feel more sympathetic, as measured by a single item in a battery of questions about the respondent's emotional state. The battery made no mention of

---

[2]We thank our reviewers for spurring this important analysis.

Jews. We find large effects: the treatment elevated sympathetic feelings from 0.48 to 0.77 (difference = 0.296, s.e. = 0.014; Figure 1, top row). This means that the effect amounted to more than half the distance between baseline levels of sympathy and the maximum amount that could be expressed using the scale.

These sympathetic feelings coincided with gains in explicit positive attitudes toward and regarding Jews. We find modest evidence of an increase in general warmth toward Jews (H2), equal to about two degrees on the feeling thermometer (difference = 0.019, s.e. = 0.009). We find a larger effect on perceptions that Jews are discriminated against (H3), with an estimated effect of 0.037 (s.e. = 0.011). We find similarly strong support for the policy solutions hypothesis (H4), with an effect of 0.035 (s.e. = 0.009). However, we find no evidence of a decrease in endorsements of antisemitic stereotypes (H5).[3]

As a measure of robustness to multiple testing, we preregistered a Bonferroni correction for our five attitudinal hypotheses, lowering the rejection threshold to $p < 0.01$. This changes our approach from individual testing to disjunctive testing: rather than assuring each null hypothesis has a 5% chance of being rejected due to sampling error, we ensure that we have a 5% chance of rejecting any hypothesis due to sampling error (Rubin, 2021).[4] The effects on sympathy discrimination perceptions and support for policy solutions retain statistical significance at this threshold, while the effect on the feeling thermometer does not. The fact that three of our five tests survived the correction suggests that sampling error is unlike to have created an illusion of positive attitudinal effects.

We find no evidence that the positive attitudinal effects persisted after two weeks.[5] Limiting the sample to only those who completed the wave 2 survey, Figure 1b displays our attitudinal effect estimates two weeks after initial exposure and measurement. Three of the four estimates are almost exactly zero. In the fourth case, regarding perceptions of discrimination (H3), the estimate remains positive but does not attain statistical significance.

We find minimal evidence of treatment effect heterogeneity in either wave of the survey. Among subjects who visited a H/A site in the pretreatment period, our estimates were effects similar in magnitude but only attained statistical significance in two cases (Appendix Table SI-12). There is no evidence that effects differ according to this factor or several other variables that are thought to predict susceptibility to perspective-taking interventions or online misinformation (cognitive reflection, conspiratorial beliefs, empathetic personality, personal contact with Jews, and political knowledge); see Appendix Tables SI-21 to SI-27.

---

[3]For the stereotype index, the control mean on the 0–1 scale was about 0.25, which raises some concern about floor effects. However, the control means for the feeling thermometer and discrimination perception indices were 0.73 and 0.69, which is only slightly further from the high end of the scale, and we were able to detect positive effects on these measures.

[4]The Bonferroni correction is known to be conservative. Strictly speaking, this means we have less than a 5% chance of rejecting any hypothesis due to sampling error.

[5]The second wave of our survey omitted the sympathy battery (as we had no reason to believe this outcome would be durable), but included all of the measures for H2–H5.

Browsing behavior

We examine three measures of browsing behavior: any visit, number of visits, and time spent. Before turning to treatment effect estimates, we begin by summarizing the data (Table 1). Among those for whom we were able to obtain browsing data ($n = 1,917$),[6] 14.7% ($n = 281$) visited an H/A site before treatment, with 1.9% ($n = 37$) of participants visiting antisemitic sites and 244 visiting other H/A sites. The average subject visited 4.1 H/A sites and spent a total of 162.91 seconds on such sites. Among these, 0.15 sites were categorized as explicitly antisemitic. These sites were viewed for an average of 6.38 seconds.

Past browsing behavior is a strong predictor of future browsing behavior (Table 1). Relative to those with no pretreatment visits to H/A sites, those with at least one pretreatment visit to a H/A site were more than 10 times more likely to visit a H/A site in the posttreatment period (71.4% vs. 6.1%); visit more than 50 times as many sites (25.3 vs. 0.47); and spend more than 50 times as much time on them (1,014 vs. 17 s). Those who visited antisemitic sites stood out even relative to those who had viewed only other H/A sites. Antisemitic site viewers spent about three times as much time on H/A sites in general (2,384 vs. 807 seconds) and about six times more on antisemitic sites in particular (185 vs. 29 seconds). This suggests that browsing histories can reliably identify subjects who may benefit from antiprejudice interventions.

Our treatment effect estimates suggest that the perspective-taking intervention reduced the appeal of antisemitic websites. We find suggestive evidence that the treatment reduced the percentage of respondents who visited at least one antisemitic site (difference = −1.0 p.p., s.e. = 0.6; Figure 2, top row) and weak evidence of a decline in the total number of H/A sites visited (difference = −0.8%, s.e. = 0.8). We find stronger evidence of an effect on time spent: treated subjects spent 5.2% less time viewing antisemitic sites (s.e. = 2.6).

The effects appear to be concentrated among subjects who visited H/A sites in the pretreatment period. Among this group, we estimate a 37.2% reduction in time spent on antisemitic sites due to treatment (s.e. = 13.3; Appendix Table SI-29). On the other two outcomes, our estimates are larger than the full sample but are again statistically insignificant. We estimate a 4.0 pp decline in the percentage who visited at least one antisemitic site (s.e. = 3.6) and a 3.2% decline in the average number of site visits (s.e. = 4.9).

In contrast to the effects on the appeal of antisemitic content, we find little evidence that the treatment reduced engagement with H/A websites overall (Figure 2, gray squares). Two of our three estimates are almost exactly zero, and none comes close to attaining statistical significance. This suggests that any effects on browsing behavior were limited to antisemitic websites in particular.

The appendix presents estimates of treatment effect heterogeneity according to several covariates (Appendix 3.6). We find some evidence that the treatment had larger effects on respondents with more conspiratorial beliefs and who are stronger Republicans. However, we conducted a large number of tests in this

---

[6]We could not obtain browsing data for respondents who left the Pulse panel between the time YouGov provided us with the pretreatment data and the time the survey was fielded.

Table 1. Summary of browsing data

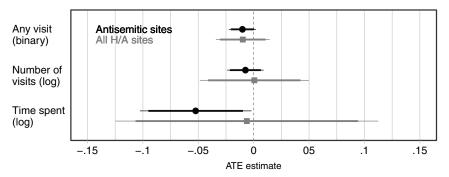| Pretreatment category | Condition | N | Posttreatment outcomes | | | | | |
| | | | All H/A sites | | | Antisemitic sites | | |
			Any Visit	Num. Visits	Time Spent	Any Visit	Num. Visits	Time Spent
All subjects	All respondents	1,917	0.156	4.10	162.97	0.026	0.15	6.38
	Control	953	0.164	4.02	188.17	0.030	0.18	8.85
	Treatment	964	0.149	4.18	138.05	0.021	0.13	3.94
No pretreatment H/A site visits	All respondents	1,636	0.061	0.47	16.72	0.007	0.02	2.44
	Control	810	0.067	0.50	21.68	0.009	0.02	4.16
	Treatment	826	0.054	0.44	11.86	0.006	0.01	0.75
Visited H/A site pretreatment	All respondents	281	0.715	25.25	1,014.42	0.132	0.96	29.32
	Control	143	0.713	23.96	1,131.22	0.154	1.08	35.41
	Treatment	138	0.717	26.59	893.38	0.109	0.83	23.01
Antisemitic sites	All respondents	37	0.730	42.76	2,383.70	0.405	6.00	185.03
	Control	18	0.722	30.06	2,703.11	0.444	6.89	216.67
	Treatment	19	0.737	54.79	2,081.11	0.368	5.16	155.05
Other H/A sites	All respondents	244	0.713	22.59	806.78	0.090	0.19	5.70
	Control	125	0.712	23.08	904.87	0.112	0.24	9.30
	Treatment	119	0.714	22.08	703.74	0.067	0.14	1.92

Figure 2. Effects on browsing behavior.
Note: Figure displays treatment effect estimates for browsing behavior. Dots are point estimates. Black dots describe only antisemitic sites, while gray dots describe all H/A sites. Thick (thin) error bars are 90 (95) percent confidence intervals.

section, and the partisanship tests are exploratory (added in response to a reviewer comment). We recommend that these results be replicated in future research.

Discussion

Our strongest interpretation of the results is that the perspective-taking intervention caused a short-term "sympathy shock" that reduced the appeal of antisemitic content online. The attitudinal results motivate the characterization as a sympathy shock: a large short-term effect on sympathy that heightened positive feelings, perceptions of discrimination, and support for policy solutions, but only in the short term. The minimal effects on visit frequency suggest that the intervention did not help our subjects avoid antisemitic content altogether. However, the presence of larger effects on visit length, concentrated among those who were most prone at baseline to view the content, suggests that the treatment caused subjects to find what antisemitic content they encountered to be less appealing.

Although we think our evidence should cause readers to update their beliefs in favor of thinking that perspective-taking interventions can reduce antisemitic attitudes and affect browsing behavior, our interpretation of the results is also informed by their weaknesses. In particular, our confidence that the treatment worked is tempered by the lack of evidence that the attitudinal effects persist after two weeks, the relative weakness of the effects on the number of site visits (as opposed to time spent), and the lack of evidence that our intervention reduced engagement with the larger universe H/A content. Given this, our view is that our results demonstrate our intervention's promise and suggest that it may have operated as sympathy shock.

Our results show that an intervention inspired by analogic perspective-taking not only affects attitudes toward outgroups (Broockman and Kalla, 2016; Adida, Lo, and Platas, 2018; Kalla and Broockman, 2020) but can also affect related behaviors. After exposure to our intervention, people who had previously visited H/A sites spent less time on antisemitic sites. However, even while browsing behavior changed, the effects on attitudes did not persist. This echoes Paluck (2009), who observed an

antiprejudice intervention changing behavior but not beliefs, as well as other work on the relationship between attitudes and behavior (Acharya, Blackwell, and Sen, 2018; Quintelier and Van Deth, 2014).

Our evidence also makes clear that analogic perspective-taking-inspired interventions can be effective in reducing prejudice toward a previously unexplored target group, Jewish people. Studying this group in the context of an online intervention is particularly important given the high levels of antisemitism observed among young people (Hersh and Royden, 2021). Given the frequency with which Jews are victims of hate crimes in the US and around the world, we hope that future research and practice will build on our evidence.

Future work could also address potential limitations with our data. Pulse does not reliably capture app data; if, say, participants navigated to hateful Instagram accounts on their smartphones, our data would not include those visits. Although this source of measurement error is orthogonal to our randomly assigned treatment, it could still skew our measurements of the prevalence and correlates of H/A site visits. It is also possible that the observed effects on browsing behavior are an artifact of social desirability bias: perhaps our results would have been different if our subjects did not know that they were being monitored. While we cannot rule this out entirely, we are struck by our finding that, in the pretreatment period, 14.6% of participants visited H/A sites although they had already consented to be monitored. Clearly, a large number of people are unafraid to visit otherwise socially unacceptable sites despite knowing their behavior is being observed. That being said, it remains possible that subjects change their behavior after they consent to monitoring. We encourage future research to find ways around this, perhaps by delivering treatments more naturalistically in settings that do not require informed consent. Researchers could partner with video or social media sites to vary exposure to prejudice-reduction videos and then passively observe subsequent video-watching behavior. But even that design would only rule out social desirability bias, as the treatment could still cause subjects to shift their consumption of hateful content to other platforms. Ultimately, there may be an inevitable trade-off between broad observation of individual-level online behavior (which requires consent) and the ability to completely rule out social desirability concerns.

Our experiences with differential attrition breakoff offer important lessons for survey researchers. By default, our vendor does not provide data on incomplete responses, which initially obscured the issue of differential breakoff. We are fortunate that the vendor was able to recover the data almost 2 years later, and that our preregistered strategy for dealing with other forms of attrition could be adapted to this case. However, we should not have assumed that survey vendors default to complete data disclosure. We recommend that researchers explicitly request data on incomplete responses during the planning stage, and that reviewers and editors require a full accounting of all potential forms of attrition and breakoff.

The behavioral effects' concentration among individuals who were prone to visit H/A sites at baseline suggests two further lessons for combating antisemitism and prejudice, as well as research into prejudice reduction. First, interventions such as ours do not appear to backfire by *increasing* engagement with antisemitic content (e.g., due to curiosity or a desire to reevaluate one's past browsing behavior). In contrast, some well-intentioned behavioral interventions appear to have the

opposite of their intended effect (Brinkman et al., 2016). Second, targeting interventions based on past browsing behavior can be fruitful for both researchers and practitioners. Engagement with antisemitic sites is a fairly stable tendency over time and can be predicted based on engagement with adjacent sites that do not traffic in explicitly antisemitic content.

Future research should build on our work by investigating intervention points other than audience tastes. For example, research targeting producers of online hate finds that the use of "sock puppet" accounts, which do not disclose the true identity of the person who controls them, can be effective in at least temporarily stemming the tide of such behavior (Munger, 2017; Siegel and Badaan, 2020). Strategies like this should be thought of as a complement to the approach we have taken. An all-hands-on-deck strategy for mitigating the effects of online hate speech would seek to reduce the amount of such content produced *and* shrink the audience for that content.

In pursuit of these goals, our findings suggest that online interventions can meaningfully reduce antisemitic attitudes and browsing behavior, especially if the interventions can be targeted at individuals whose prior behavior puts them at high risk of exposure to hateful content. This represents a meaningful step toward building healthier, less hateful communities, online and offline.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/XPS.2023.32.

Data availability. The data, code, and any additional materials required to replicate all analyses in this article are available in the *Journal of Experimental Political Science Dataverse* within the Harvard Dataverse Network, at https://doi.org/10.7910/DVN/U3FPBH (Graham, 2023).

Competing interests. The authors have no conflicts of interest.

Ethics statement. This research was approved by the George Washington University IRB (#NCR213417). This research adheres to APSA's Principles and Guidance for Human Subjects. For more details, see Appendix 4.1.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. "Explaining Preferences from Behavior: A Cognitive Dissonance Approach." *Journal of Politics* 80(2): 400–11.

Adida, Claire L., Adeline Lo, and Melina Platas. 2018. "Perspective Taking Can Promote Short-Term Inclusionary Behavior Toward Syrian Refugees." *Proceedings of the National Academy of Sciences* 115: 9521–26.

ADL. 2017. "From Alt Right to Alt Lite: Naming the Hate".

Alper, Becka A., and Alan Cooperman. 2021. "Jewish Americans in 2020".

Baston, C. Daniel, Johee Chang, Ryan Orr, and Jennifer Rowland. 2002. "Empathy, Attitudes and Action: Can Feelings for a Member of a Stigmatized Group Motivate One to Help the Group?" *Personality and Social Psychology Bulletin* 28: 1656–66.

Baston, Daniel C., Marina P. Polycarpou, Eddie Haron-Jones, Heidi J. Imhoff, Erin C. Mitchener, Lori L. Bednar, Tricia R. Klein, and Lori. Highberger. 1997. "Empathy and Attitudes: Can Feelings for a Member of a Stigmatized Group Improve Feelings Toward the Group?" *Journal of Personality and Social Psychology* 72: 105–18.

Brinkman, Sally A., Sarah E. Johnson, James P. Codde, Michael B. Hart, Judith A. Straton, Murthy M. Mittinty, and Sven R. Silburn. 2016. "Efficacy of Infant Simulator Programmes to Prevent Teenage

Pregnancy: A School-Based Cluster Randomised Controlled Trial in Western Australia." *The Lancet* 388(10057): 2264–71.

Broockman, David, and Joshua Kalla. 2016. "Durably Reducing Transphobia: A Field Experiment on Door-to-Door Canvassing." *Science* 352(6282): 220–24.

Burns, Mason D., and Erica L. Granz. 2021. ""Past Injustice and Present Prejudice": Reducing Racial Bias and Increasing Sympathy by Framing Historical Racism as Recent." *Group Processes & Intergroup Relations* 25(5): 1312–1332.

Chen, Annie Y., Brendan Nyhan, Jason Reifler, Ronald E. Robertson, and Christo Wilson. 2021. "Can Google Search Be Used to Counter White Supremacy?".

FBI. 2021. "FY 2019 Hate Crimes Statistics Report".

Fox. 2021. "Police Investigate Two Incidents of Anti-Semitic, Racist Vandalism at Yale Construction Site".

Galinsky, Adam D., and Gillian Ku. 2004. "The Effects of Perspective-Taking on Prejudice: The Moderating Role of Self-Evaluation." *Personality and Social Psychology Bulletin* 30(5): 594–604.

Galinsky, Adam D., and Gordon B. Moskowitz. 2000. "Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-Group Favoritism." *Journal of Personality and Social Psychology* 78: 708–24.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* W.W. Norton.

Graham, Matthew H. 2023. Replication Data for: Combatting Hateful Attitudes and Browsing Behavior: The Case of Antisemitism. Harvard Dataverse, v3.

Guess, A.M., B. Nyhan, and J. Reifler. 2020. "Exposure to Untrustworthy Websites in the 2016 US Election." *Nature Human Behaviour* 4(5): 472–480.

Hersh, Eitan, and Laura Royden. 2021. "Antisemitic Attitudes Across the Ideological Spectrum".

Hertel-Fernandez, Alexander, and Ethan Porter. 2021. "Analogic Perspective-Taking and Attitudes Toward Political Organizations: An Experiment with a Teachers' Union." *Journal of Experimental Political Science* 10: 100–111.

Kalla, Joshua L., and David E. Broockman. 2020. "Reducing Exclusionary Attitudes through Interpersonal Conversation: Evidence from Three Field Experiments." *American Political Science Review* 114(2): 410–25.

Kalla, Joshua L., and David E. Broockman. 2021. "Which Narrative Strategies Durably Reduce Prejudice? Evidence from Field and Survey Experiments Supporting the Efficacy of Perspective-Getting". *American Journal of Political Science* 67(1): 185–204.

Kinder, Donald R., and Lynn M. Sanders. 1996. *Divided by Color: Racial Politics and Democratic iDeals.* University of Chicago Press.

Mummolo, Jonathan, and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113(2): 517–29.

Munger, Kevin. 2017. "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment." *Political Behavior* 39(3): 629–49.

Paluck, Elizabeth Levy. 2009. "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda." *Journal of Personality and Social Psychology* 96(3): 574.

Paluck, Elizabeth Levy, and Donald P. Green. 2009. "Prejudice Reduction: What Works? A Review and Assessment of Research and Practice." *Annual Review of Psychology* 60(1): 339–67.

Porter, Ethan, and Yamil R. Velez. 2021. "Placebo Selection in Survey Experiments: An Agnostic Approach." *Political Analysis* 30(4): 481–94.

Quintelier, Ellen, and Jan W. Van Deth. 2014. "Supporting Democracy: Political Participation and Political Attitudes. Exploring Causality Using Panel Data." *Political Studies* 62(S1): 153–71.

Rubin, Mark. 2021. "When to Adjust Alpha During Multiple Testing: A Consideration of Disjunction, Conjunction, and Individual Testing." *Synthese* 199(3–4): 10969–1000.

Siegel, Alexandra A., and Vivienne Badaan. 2020. "#No2Sectarianism: Experimental Approaches to Reducing Sectarian Hate Speech Online." *American Political Science Review* 114(3): 837–55.

Simonovitz, Gabor, Gabor Kudzu, and Peter Kardos. 2018. "Seeing the World Through the Other's Eye: An Online Intervention Reducing Ethnic Prejudice." *American Political Science Review* 112(1): 186–93.

The Washington Post. 2021. "George Washington University Jewish Groups Discuss Antisemitism After Fraternity House Vandalized, Including Torah".

Weil, Frederick D. 1985. "The Variable Effects of Education on Liberal Attitudes: A Comparative-Historical Analysis of Anti-Semitism Using Public Opinion Survey Data." *American Sociological Review 50*(4): 458–74.

Weisberg, Herb. 2019. *The Politics of American Jews*. University of Michigan Press.

Weiss, Bari, and Damon Winter. 2018. "Meet the Renegades of the Intellectual Dark Web." *The New York Times*.