Suppressing the Search Engine Manipulation Effect (SEME)

ROBERT EPSTEIN, American Institute for Behavioral Research and Technology, USA
RONALD E. ROBERTSON, Northeastern University, USA
DAVID LAZER, Northeastern University, USA
CHRISTO WILSON, Northeastern University, USA

A recent series of experiments demonstrated that introducing ranking bias to election-related search engine results can have a strong and undetectable influence on the preferences of undecided voters. This phenomenon, called the Search Engine Manipulation Effect (SEME), exerts influence largely through order effects that are enhanced in a digital context. We present data from three new experiments involving 3,600 subjects in 39 countries in which we replicate SEME and test design interventions for suppressing the effect. In the replication, voting preferences shifted by 39.0%, a number almost identical to the shift found in a previously published experiment (37.1%). Alerting users to the ranking bias reduced the shift to 22.1%, and more detailed alerts reduced it to 13.8%. Users' browsing behaviors were also significantly altered by the alerts, with more clicks and time going to lower-ranked search results. Although bias alerts were effective in suppressing SEME, we found that SEME could be completely eliminated only by alternating search results – in effect, with an equal-time rule. We propose a browser extension capable of deploying bias alerts in real-time and speculate that SEME might be impacting a wide range of decision-making, not just voting, in which case search engines might need to be strictly regulated.

CCS Concepts: • **Human-centered computing → Laboratory experiments**; **Heuristic evaluations**; • **Social and professional topics → Technology and censorship**;

Additional Key Words and Phrases: search engine manipulation effect (SEME); search engine bias; voter manipulation; persuasive technology; algorithmic influence

ACM Reference Format:
Robert Epstein, Ronald E. Robertson, David Lazer, and Christo Wilson. 2017. Suppressing the Search Engine Manipulation Effect (SEME). *Proc. ACM Hum.-Comput. Interact.* 1, 2, Article 42 (November 2017), 22 pages. https://doi.org/10.1145/3134677

1 INTRODUCTION

Algorithms that filter, rank, and personalize online content are playing an increasingly influential role in everyday life [27]. Their automated curation of content enables rapid and effective navigation of the web [94] and has the potential to improve decision-making on a massive scale [40]. For example, Google Search produces billions of ranked information lists per month [22], and Facebook produces ranked social information lists for over a billion active users [57].

Authors' addresses: Robert Epstein, American Institute for Behavioral Research and Technology, 1035 East Vista Way Ste. 120, Vista, 92084, CA, USA; Ronald E. Robertson, Northeastern University, 1010-177 Network Science Institute, 360 Huntington Ave. Boston, 02115, MA, USA; David Lazer, Northeastern University, Boston, 02115, MA, USA; Christo Wilson, Northeastern University, Boston, 02115, MA, USA.

2573-0142/2017/11-ART42
https://doi.org/10.1145/3134677

However, algorithms are human inventions, and as such, characteristic human elements – such as intentions, beliefs, and biases – inevitably influence their design and function [14, 113]. Recent research has shown that society's growing dependence on ranking algorithms leaves our psychological heuristics and vulnerabilities susceptible to their influence on an unprecedented scale and in unexpected ways [11, 30, 69, 96, 114, 124]. For example, race and gender biases have been documented in the rankings of candidates in online job markets [55], and algorithms have been shown to learn similar biases from human generated text [14]. Experiments conducted on Facebook's Newsfeed have demonstrated that subtle ranking manipulations can influence the emotional language people use [69], and user studies have shown that people are generally unaware that the Newsfeed is ranked at all [33, 35]. Similarly, experiments on web search have shown that manipulating election-related search engine rankings can shift the voting preferences of undecided voters by 20% or more after a single search [30].

Concerns about the power and influence of ranking algorithms that have been expressed by regulators and users [101] are exacerbated by a lack of transparency. The inputs, parameters, and processes used by ranking algorithms to determine the visibility of content are often opaque. This can be due to the proprietary nature of the system, or because understanding requires a high level of technical sophistication [45, 101]. To overcome these challenges, researchers have developed techniques inspired by the social sciences to audit algorithms for potential biases [84, 113]. Algorithm audits have been used to examine the personalization of search engine rankings [53, 64], prices in online markets [17, 54, 80, 81], rating systems [37], and social media newsfeeds [33, 71].

While algorithm audits help to identify *what* an algorithm is doing, they don't necessarily help us to model the *impact* an algorithm is having. While field experiments can be controversial [69], controlled behavioral experiments designed to mimic online environments provide a promising avenue for isolating and investigating the impact that algorithms might have on the attitudes, beliefs, or behavior of users. This approach addresses a frequently missing link between the computational and social sciences [74]: a controlled test of an algorithm's influence and an opportunity to investigate design interventions that can enhance or mitigate it [30, 33, 35, 36, 78].

In this study, we focus on the influence of election-related ranking bias in web search on users' attitudes, beliefs, and behaviors – the Search Engine Manipulation Effect (SEME) [30] – and explore design interventions for suppressing it. While *"bias"* can be ambiguous, our focus is on the *ranking bias* recently quantified by Kulshrestha *et al.* with Twitter rankings [71]. The research questions we ask are:

(1) How does SEME replicate with a new election?
(2) Does alerting users to ranking bias suppress SEME?
(3) Does adding detail to the alerts increase their suppression of SEME?
(4) Do alerts alter search browsing behavior?
(5) How does bias awareness mediate SEME when alerts are, and are not, present?

To answer these questions, we developed a mock search engine over which we could exert complete control. Using this platform, we conducted three experiments, one replicating SEME with a new election, and two in which we implemented *bias alerts* of varying detail. To populate our search rankings we collected real search results and webpages related to the 2015 election for Prime Minister of the UK because it was projected to be a close race between two candidates. After obtaining bias ratings of the webpages from independent raters, we manipulated the search engine so that the ranking bias either (a) favored one specific candidate, or (b) favored neither candidate.

The number of votes for the candidates favored by the ranking bias increased by 39.0% in our replication experiment, a figure within 2% of the original study [30]. As predicted, our design interventions altered users' voting patterns, with a low detail alert suppressing votes for the favored candidate to 22.1%, and a high detail alert reducing the effect to 13.8%. Somewhat counterintuitively, we found that users' awareness of the ranking bias suppressed SEME when an alert was present, but increased SEME when no alert was present.

Our results provide support for the robustness of SEME and create a foundation for future efforts to mitigate ranking bias. More broadly, our work adds to the growing literature that provides an empirical basis to calls for algorithm accountability and transparency [24, 25, 90, 91] and contributes a quantitative approach that complements the qualitative literature on designing interventions for ranking algorithms [33, 35, 36, 93]. As regulators and academics have noted, the unregulated use of such technologies may lead to detrimental outcomes for users [15, 27, 30, 55, 101, 113, 124], and our results suggest that deploying external design interventions could mitigate such outcomes while legislation takes shape. Our results also suggest that *proactive* strategies that prevent ranking bias (*e.g.,* alternating rankings) are more effective than *reactive* strategies that suppress the effect through design interventions like bias alerts. Given the accumulating evidence [2, 92], we speculate that SEME may be impacting a wide range of decision-making, not just voting, in which case the search engine as we know it today might need to be strictly regulated.

The code and data we used are available at https://dataverse.harvard.edu/dataverse/biasalerts.

2 RELATED WORK

An interdisciplinary literature rooted in psychology is essential to understanding the influence of ranking bias. In this section, we briefly overview this work and discuss how it applies to online environments and ranked information in particular. We conclude by exploring the literature on resisting influence and design interventions to identify strategies for suppressing SEME.

2.1 Order Effects and Ranking Algorithms

Order effects are among the strongest and most reliable effects ever discovered in the psychological sciences [29, 88]. These effects favorably affect the recall and evaluation of items at the beginning of a list (*primacy*) and at the end of a list (*recency*). Primacy effects have been shown to influence decision-making in many contexts, such as medical treatment preferences [7], jury decisions [62], and increasing voting for the first candidate on a ballot [16, 56, 63, 68, 70, 100].

In online contexts, primacy has been shown to bias the way users navigate websites [26, 46, 89], influence which products receive recommendations [51, 67], and increase bookings for top-ranked hotels [32]. Experiments conducted on online ranking algorithms have demonstrated their influence on users' music preferences [111, 112], use of emotional language [69], beliefs about scientific controversy [92], and undecided voters' preferences [30].

Primacy effects have a particularly strong influence during online search [30, 92]. Highly ranked search results attract longer gaze durations [48, 50, 77] and receive the majority of clicks [108], even when superior results are present in lower ranked positions [60, 61, 95]. An ongoing study on international click-through-rates found that in February 2017, 62.3% of clicks were made on the first three results alone, and 88.6% of clicks were made on the first Search Engine Result Page (SERP) [108]. Leveraging these behavioral primacy effects, the original SEME experiments demonstrated that biasing search rankings to favor a particular candidate can (1) increase voting for that candidate by 20% or more, (2) create shifts as high as 80% in some demographic groups, and (3) be masked so that no users show awareness of the bias [30].

2.2 Attitude Change and Online Systems

Compared to newspaper readers and television viewers, search engine and social media users are more susceptible to influence [11, 21, 23, 28, 30, 44]. This enhanced influence stems from several persuasive advantages that online systems have over traditional media [40]. For example, online systems can: (1) provide a platform for constant, large-scale, rapid experimentation [66], (2) tailor their persuasive strategies by mining detailed demographic and behavioral profiles of users [1, 6, 9, 18, 121], and (3) provide users with a sense of control over the system that enhances their susceptibility to influence [5, 41, 116, 118].

The processes through which users become aware of and react to algorithms has been a topic of recent research interest [33–35, 37, 106]. On Facebook, the majority of people appear to be entirely unaware that their Newsfeed is algorithmically ranked [36], and SEME demonstrated how ranking bias can be masked yet still influence users [30]. Classifying bias awareness is not a trivial task, however, with human coders trained to identify biased language achieving under 60% accuracy [109]. Directly asking users about their awareness of an algorithm's idiosyncrasies is also problematic due to the possibility of creating demand characteristics [76, 119].

At present, search engines have an additional persuasive advantage in the public's trust. A recent report involving 33,000 people found that search engines were the most trusted source of news, with 64% of people reporting that they trust search engines, compared to 57% for traditional media, 51% for online media, and 41% for social media [10]. Similarly, a 2012 survey by Pew found that 73% of search engine users report that "all or most of the information they find is accurate and trustworthy," and 66% report that "search engines are a fair and unbiased source of information" [105].

Researchers have also suggested that the personalization algorithms in online systems can exacerbate the phenomenon of *selective exposure*, where people seek out information that confirms their attitudes or beliefs [42]. Eli Pariser coined the phrase "internet filter bubble" to describe situations where people become trapped in a digital echo chamber of information that confirms and strengthens their existing beliefs [96]. Researchers at Facebook have shown that selective exposure occurs in the Newsfeed [4], though its impact on users is unclear.

2.3 Resisting Influence and Design Interventions

Fortunately, research on resistance offers insights for how unwanted influence can be mitigated or suppressed [12, 43, 65, 79, 103]. Suggestions for fostering resistance can be broken down into two primary strategies: (1) providing forewarnings [43, 49] and (2) training and motivating people to resist [79, 120]. Forewarnings are often easier and more cost-effective to implement than motivating or training people, and their effect on resistance can be increased by including details about the direction and magnitude of the persuasive message [39, 120], providing specific, comprehensible, and evidence-based counterarguments [78, 117], and including autonomy-supportive language in warnings [83]. Part of the reason that forewarnings work is explained by psychological reactance theory [12], which posits that when people believe their intellectual freedom is threatened – by exposing an attempt to persuade, for example – they react in the direction opposite that of the intended one [73, 107].

Areas where forewarnings have been applied with success include antismoking campaigns, political advertisement critiques [65], educational outreach about climate change [117], and most relevant here, a technological debiasing study that used alerts to minimize cognitive biases during online searches for heath information [78]. Given recent research suggesting that the composition and ranking of health information in online search can impact attitudes and beliefs about the safety of vaccinations [2], Ludolph *et al.* utilized the Google custom search API to generate a set of randomly ordered search results consisting of 50% pro-vaccination and 50% anti-vaccination websites to test whether various warnings injected into Google's Knowledge Graph box could suppress the effects of the anti-vaccination information. Overall, Ludolph *et al.* found that generic warnings alerting users to the possibility of encountering misleading information during their search had little to no effect on their knowledge and attitudes, unless the warning was paired with comprehensible and factual information [78].

In the context of online media bias, researchers have primarily explored methods for curbing the effects of algorithmic filtering and selective exposure [87, 96] rather than ranking bias [71]. In this vein, researchers have developed services that encourage users to explore multiple perspectives [97, 98] and browser extensions that gamify and encourage balanced political news consumption [19, 20, 86]. However, these solutions are somewhat impractical because they require users to adopt new services or exert additional effort.

Several user studies have explored the impact of ranking algorithms, but their focus has primarily been on user experience and algorithm awareness [33, 35–37, 52]. For example, one study examined 15 users' reactions to the annotation of blog search results as conservative or liberal, and found that most users preferred interfaces that included the annotations [93]. While informative, small user studies are not designed to quantify the impact of technology on behavior and decision-making. Our focus here is on testing design interventions that provide users with the ability to identify bias *proactively* – before information is consumed – and could be implemented without requiring users to change services or exert additional effort.

3 METHODS AND DATA COLLECTION

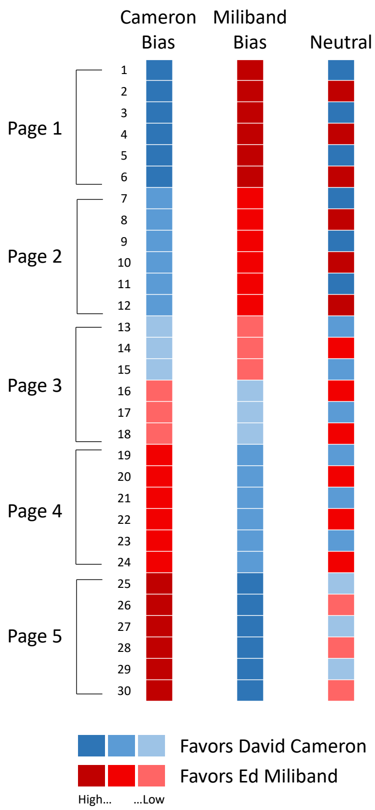
The procedure for all three experiments followed the same general procedure used by Epstein and Robertson in Study 2 of the original SEME experiments [30]. First, subjects were shown brief neutral biographies (available in the Appendix) of two political candidates and then asked to rate them in various ways and indicate who they would be more likely to vote for if the election were held today. Second, subjects were given an opportunity to gather information on the candidates using a mock search engine that we had created. Finally, subjects again rated the candidates, indicated who they would be likely to vote for, and answered a question measuring awareness of ranking bias in the search. We examined shifts in candidate preferences both within experiments and between experiments. This procedure was approved by the Institutional Review Board (IRB) of the American Institute for Behavioral Research and Technology (IRB#10010).

3.1 Experiment Design

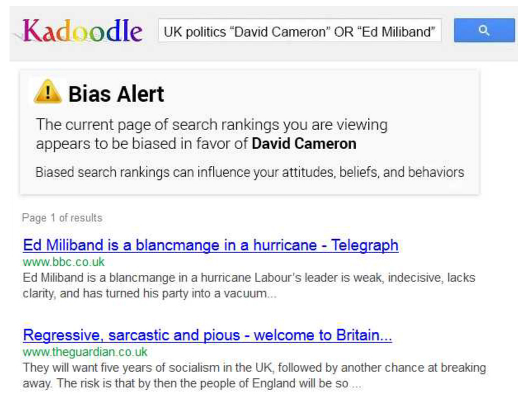
We constructed a mock search engine that gave us complete control of the search interface and rankings. To populate our mock search engine we identified an upcoming election that was expected to be a close race between two candidates and used Google Search and Google News to collect real search results and webpages related to the two candidates in the month preceding the experiments. The election was the 2015 Election for Prime Minister of the UK between incumbent David Cameron and his opponent Ed Miliband.

To construct biased search rankings we asked four independent raters to provide bias ratings of the webpages we collected on an 11-point Likert scale ranging from -5 "favors Cameron" to +5 "favors Miliband". We then selected the 15 webpages that most strongly favored Cameron and the 15 that most strongly favored Miliband to create three bias groups (Figure 1a):

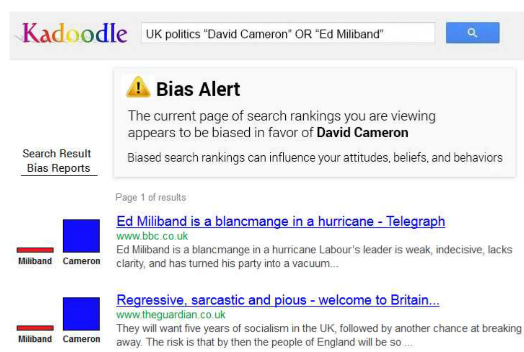
(1) In the **Cameron bias** group, the results were ranked in descending order by how much they favored David Cameron.
(2) In the **Miliband bias** group, the results were ranked in descending order by how much they favored Ed Miliband: the mirror image of the Cameron bias group rankings.
(3) In the **neutral** group, the results alternated between favoring the two candidates in descending order.



(a) Search rankings by bias group assignment.



(b) Low alert search condition.



(c) High alert search condition.

Fig. 1. Ranking manipulations for the three experiments, and example bias alerts.

All subjects had access to the same 30 results, which were equally distributed across five SERPs, but the ranking of the results depended on subjects' bias group assignment. The query in the search engine was fixed as "*UK Politics 'David Cameron' OR 'Ed Miliband'*", and subjects could not reformulate it. Subjects were given 15 minutes to use the search engine and could end the search after they felt they had enough information to make a decision. In our instructions, which are available in the Appendix, we asked subjects to do at least some research on the candidates before ending the search. It terms of Kulshrestha *et al*'s formulas [71], we found a ranking bias of -2.86 for the Cameron bias group, 2.83 for the Miliband bias group, and -0.39 for the neutral group.

On top of assignment to a bias group, subjects were randomly assigned to one of three alert experiments. We drew from the literature on decision-making and design intervention to implement so-called debiasing strategies for improving decision-making in the presence of biased information [39, 78, 82]. Specifically, we constructed and placed alerts in the search results produced by our mock search engine that provided forewarnings with salient graphics, autonomony-supportive language, and details on the persuasive threat [39, 82]:[1]

[1]Although Fischoff proposes two additional strategies, providing feedback and providing training [39], we did not investigate these strategies. The former would necessitate storing subjects' browsing history, while the latter would require costly and inconvenient training of subjects.

(1) In the **no alert** experiment, subjects saw the 30 search results depending on their bias group assignment.This replicates Study 2 from the original SEME experiments [30].
(2) In the **low alert** experiment, subjects saw a banner at the top of the search results that contained a caution symbol and a message informing subjects that their search rankings were biased towards one candidate or the other (Figure 1b). The candidate name displayed in the warning depended on bias group assignment. Subjects in the neutral group received an identically formatted alert, but it stated that "The current page of search rankings you are viewing does not appear to be biased in favor of either candidate."
(3) In the **high alert** experiment, subjects saw the same banner as subjects in the low alert experiment, but also received notifications to the left of each search result illustrating which candidate it favored (Figure 1c).

3.2 Procedure

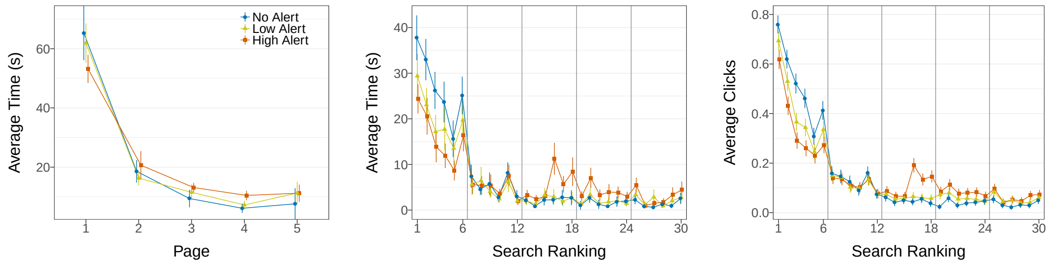
After providing informed consent and answering basic demographic questions, subjects were instructed to carefully read two brief, neutral biographies of David Cameron and Ed Miliband. Subjects then rated the two candidates on 10-point Likert scales with respect to their overall impression of each candidate, how much they trusted each candidate, and how much they liked each candidate. Subjects also indicated their likelihood of voting for one candidate or the other on an 11-point Likert scale where the candidates' names appeared at opposite ends of the scale and 0 indicated no preference, as well as on a binary choice question where subjects indicated who they would vote for if the election were held today. The appearance of the candidates' names in the biographies and all of the rating scales was counterbalanced.

Following the pre-test, subjects were given an opportunity to conduct background research on the candidates using our mock search engine. After completing the web research, subjects again rated the candidates and indicated their voting preferences. Upon submitting the post-search questions, we probed subjects' awareness of the search ranking bias with two non-leading questions. We asked: "While you were doing your online research on the candidates, did you notice anything about the search results that bothered you in any way?" and prompted subjects to explain what had bothered them in a free response format: "If you answered "yes," please tell us what bothered you." We did not directly ask subjects whether they had "noticed bias" to avoid the inflation of false positive rates that leading questions can cause [76, 119].

We adopted the same measure used in the original SEME study to detect subjects' awareness of ranking bias from their free responses. Similar to the measure recently used by Eslami *et al.* , our method relies on a two rule coding scheme and keyword matching to classify users as aware or unaware [37]. Specifically, we count subjects as aware of the bias if they (1) left a comment to explain how something about the search results had "bothered" them, and (2) their comment contained words or phrases that indicated awareness of the ranking bias (*e.g.,* "biased," "slanted," or "skewed") [30]. The full set of predefined terms is available in the Appendix. This measure likely underestimates the true proportion of detections, as some subjects may have noticed the bias but (1) not commented at all, or (2) used words or phrases not in our set. Thus, our measure provides a lower-bound estimate of bias awareness.

3.3 Participants

We recruited 3,883 subjects between April 28, 2015 and May 6, 2015 on Amazon's Mechanical Turk (AMT; https://mturk.com), a subject pool frequently used by behavioral, economic, and social science researchers [8, 13, 102]. We excluded from our analysis subjects who reported an English fluency level of 5 or less (on a scale of 1 to 10) (n=26). We randomly assigned 3,600 of the remaining



(a) Average time on each search engine results page (SERP).

(b) Average time on each search ranking's corresponding webpage.

(c) Average clicks made on each search ranking.

Fig. 2. Search metric averages by experiment. In (b) and (c) the vertical lines represent the start of a new SERP. Error bars represent 95% confidence intervals. Plot points are jittered horizontally to avoid overlap.

3,857 subjects into three alert experiments consisting of 1,200 subjects each. Each experiment had balanced *ns* for both the bias group and counter-balancing group assignments.

In aggregate, participating subjects were primarily from the US (85.9%) and India (10.8%), with a mean age of 32.8 (*SD* = 10.3) and a mean self-reported English fluency of 9.8 (*SD* = 0.61). Subjects were demographically diverse (Table 1 in the Appendix), and reported conducting an average of 16.3 web searches per day (*SD* = 25.2), a figure similar to that (15.3) found in previous SEME research conducted with subjects from AMT [30]. 88.5% reported having previously used search engines to find information about political candidates. Our sample was also largely college educated, with 63.6% reporting a bachelors degree or higher, liberal leaning (47.9%), and with a median income of $30,000 to $45,000. Subjects reported a mean familiarity of 5.2 with candidate David Cameron and a mean familiarity of 2.8 with candidate Ed Miliband on 10-point Likert scales.

4 ANALYSIS

We conducted between-subjects comparisons of search behaviors and candidate preferences by alert experiment, as well as within-subjects comparisons of subjects' pre- and post-search candidate ratings and voting preferences by bias group. We also investigated differences in subjects' shifts as a function of various demographic groups, search behavior, and awareness of the ranking bias.

4.1 Search Metrics

We examined whether the bias alerts had an impact on the browsing behavior of subjects in the two treatment groups combined with regard to average time on each SERP, time on each webpage, and clicks on each search result (Figure 2). Utilizing Kolmogorov-Smirnov (K-S) tests of differences in distributions, we found significant differences in the patterns of time spent on the 30 webpages between subjects in the no alert experiment (correlation with ranking: Spearman's ρ = -0.836, P <0.001) and the high alert experiment (ρ = -0.654, P <0.001) (K-S D = 0.467, P <0.01), and between subjects in the low alert experiment (ρ = -0.719, P <0.001) and the high alert experiment (K-S D = 0.400, P <0.01), but not between subjects in the no alert experiment and the low alert experiment (K-S D = 0.200, P = 0.301).

Similarly, we also found significant differences in the patterns of clicks that subjects made on the 30 webpages between subjects in the no alert experiment (ρ = -0.865, P <0.001) and the high alert experiment (ρ = -0.795, P <0.001) (K-S D = 0.500, P <0.001), and between subjects in the low alert experiment (ρ = -0.876, P <0.001) and the high alert experiment (K-S D = 0.367, P <0.05), but not between subjects in the no alert and the low alert experiments (K-S D = 0.300, P = 0.07).

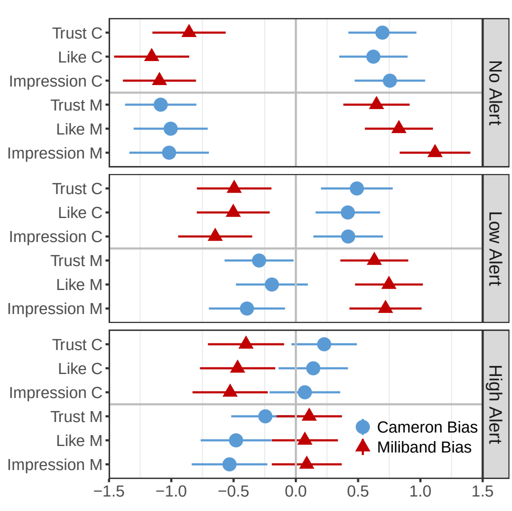
Among all conditions, we found that differences in the patterns of time and clicks on the individual rankings primarily emerged on the first SERP, but less so on the second, fourth, and fifth SERPs (Figure 2). In the high alert experiment, where subjects could see the bias shift from favoring one candidate to the other on the 16th result, we observed a substantial increase in both time and clicks compared to subjects in the no alert and low alert experiments. We observed a similar trend for time spent on SERPs (Figure 2), but K-S tests comparing the three experiments were not significant.

4.2 Attitude Shifts

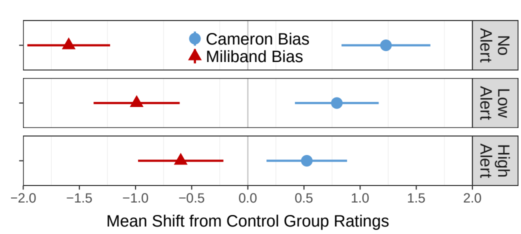
Candidate Ratings. Prior to the web research, we found no significant differences among the three groups with regard to any of the candidate ratings – trust, liking, and overall impression. After the web research, significant differences (all P <0.001) emerged for all ratings in the no alert and low alert experiments. In the high alert experiment, none of the post search ratings were significant for Cameron, but all three were significant for Miliband (Table 2 in the Appendix).

For each candidate rating we also calculated a *PostSearch* − *PreSearch* shift and compared the shifts we observed in the bias groups relative to the shifts we observed in the neutral group (Figure 3a). Using Mann-Whitney *U* tests, we found that the mean shifts in candidate ratings for the bias groups significantly converged on the mean shift found in the neutral group as the level of detail in the alerts increased, with high alerts creating higher convergence than low alerts (Figure 3a).

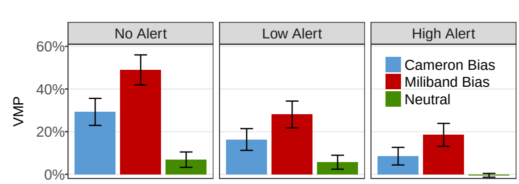
Candidate Preferences. Prior to web research, there were no significant differences in subjects' reported likelihood of voting for the two candidates among the three groups (Table 3 in the Appendix). Following the web research, significant differences emerged among all three groups and



(a) Shifts in candidate ratings on 10-point Likert scales by group and experiment. Horizontal lines separate ratings for Cameron (C) and for Miliband (M)



(b) Shifts in voting preferences by group and experiment on a scale from -5 (Miliband) to +5 (Cameron).



(c) VMP by experiment and bias group. For the neutral group VMP is percent increase in voting for Miliband.

Fig. 3. Shifts in subjects' candidate ratings (a), preferences (b), and votes (c) after the search task. For (a) and (b) the vertical line at zero is set to the mean shift for the neutral group for that metric. Error bars represent 95% CI.

the voting preferences of subjects in the two bias groups diverged by 2.71 in the no alert experiment (compared to 3.03 in the original study [30]), 1.70 in the low alert experiment, and 0.75 in the high alert experiment. As with candidate ratings, we found that the mean shifts in candiate preference in the two bias groups converged on the mean shift found in the neutral group as the alert level increased (Figure 3b).

4.3 Vote Shifts

VMP. Utilizing subjects' responses to the binary vote question, we examined the Vote Manipulation Power (VMP) measure developed in earlier research on SEME [30]. VMP is the percent change in the number of subjects, in the two bias groups combined, who indicated that they would vote for the candidate who was favored by their search rankings. That is, if x and x' subjects in the bias groups said they would vote for the favored candidate before and after conducting the search, respectively, then $VMP = (x' - x)/x$.

We found a VMP of 39.0% (95% confidence interval [CI], 34.2-43.9%; χ^2 = 109.498, P <0.001) for subjects in the no alert experiment, which closely mirrored the 37.1% VMP found previously [30]. The VMP was suppressed to 22.1% (95% CI, 18.0-26.2%; χ^2 = 49.796, P <0.001) in the low alert experiment, and to 13.8% (95% CI, 10.3-17.3%; χ^2 = 17.112, P <0.001) in the high alert experiment (Figure 3c). Compared to no alert, low alerts suppressed the effect by 16.9 percentage points (95% CI, 10.4-23.5%; χ^2 = 25.877, P <0.001) and high alerts suppressed it by 25.2 percentage points (95% CI, 19.0-31.4%; χ^2 = 61.006, P <0.001). The difference in suppression between the low alert and high alert experiments was 8.3 percentage points (95% CI, 2.7-13.9%; χ^2 = 8.398, P <0.001), confirming the impact of detailed alerts [39].

As in previous SEME experiments, we found differences in subjects' susceptibility to VMP by user characteristics, as well as variance in the effectiveness of the alerts by user characteristics (Figures 6 and 7 in the Appendix). To examine if the VMP differences we observed between the two bias groups (Figure 3c) were affected by subjects' familiarity with Cameron, we examined how familiarity with each candidate affected VMP in each bias group separately. To do this, we simplified familiarity from the 10-point Likert scales to either "high" (Familiarity > 5) or "low" (Familiarity ≤ 5) familiarity with each candidate, and then compared the VMPs we found by experiment and bias group (Figure 4). In the no alert experiment, we found that high familiarity with either candidate reduced VMP in the Cameron bias group and increased VMP in the Miliband group, but this relationship largely disappeared in the alert experiments. Interestingly, we also found that familiarity with Cameron was correlated with a slight preference for Miliband on the pre-search candidate preference measure (ρ = 0.18, P <0.001), suggesting that familiarity with Miliband made subjects less likely to shift their vote *toward* him, and more likely to vote *against* him after being exposed to SEME, even when the ranking bias was in his favor.

Percent Difference. It is useful to consider not just the percent change in voting for the favored candidate after search – VMP – but also the more granualar percent difference in voting for either candidate relative to the group total. That is, how many people shifted toward (positive) and away (negative) from the candidate favored by the rankings?

We examined the number of subjects who shifted their votes toward or away from the candidate favored by their bias group. For the neutral group, we measure votes toward Miliband as positive and shifts toward Cameron as negative. We found a significant decrease in the proportion of positive shifts between the no alert and low alert condition ($\chi^2 = 15.188, P < 0.001$) and between the no alert and high alert condition ($\chi^2 = 26.718, P < 0.001$), but not between the low alert and high alert condition ($\chi^2 = 1.512, P = 0.22$). We also found no significant increase in the proportion of

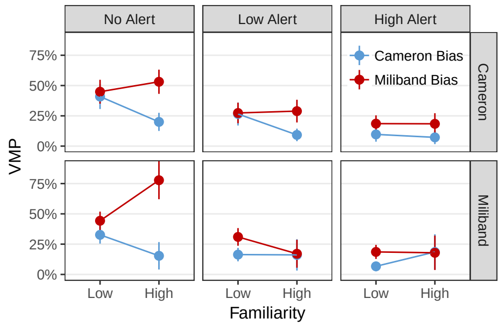


Fig. 4. VMP by experiment, bias group, and familiarity with the candidates. Error bars represent 95% CI.

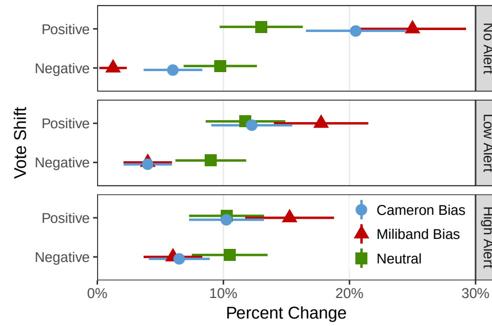


Fig. 5. Shift in voting toward the candidate favored by the ranking bias (positive) or away from that candidate (negative) by experiment. Error bars represent 95% CI.

negative shifts between the no alert and low alert condition ($\chi^2 = 0.068, P = 0.79$) or between the low alert and high alert condition ($\chi^2 = 3.715, P = 0.05$), but a significant increase did emerge between the no alert and high alert conditions ($\chi^2 = 5.326, P < 0.05$) (Table 4 in the Appendix).

We also found significant differences in positive and negative shifts among the three groups in each experiment (Figure 5). Specifically, we found highly significant differences among the three groups in the no alert and low alert experiments (both P <0.001), but less significant differences among the three groups in the high alert experiment (P <0.05). More generally, as with candidate ratings and preferences, we found that the shifts in voting patterns for the bias groups converged on the shifts found in the neutral group as the level of the alert increased. We also found that positive shifts toward Miliband were more resistant to suppression compared to positive shifts toward Cameron (Figure 5).

4.4 Bias Awareness

We found 8.1% of subjects that showed awareness of the bias in the no alert experiment, a figure identical to the 8.1% awareness rate found by Eslami *et al.* in their audit of Booking.com [37], and similar to the 8.6% of subjects who showed awareness in the original study [30]. The percentage of subjects showing bias awareness increased to 21.5% in the low alert experiment, and 23.4% in the high alert experiment. Compared to the no alert experiment, the proportion of subjects showing bias awareness significantly increased in the low alert (χ^2 = 55.653, P <0.001) and high alert (χ^2 = 68.960, P <0.001) experiments. Compared to the low alert, the high alert did not significantly increase awareness (χ^2 = 0.704, P = 0.401). Interestingly, we found that aware subjects were more likely than unaware subjects to click on lower ranked results in the no alert (K-S D = 0.333, P <0.05), low alert (K-S D = 0.433, P <0.01), and high alert (K-S D = 0.467, P <0.01) experiments.

In the no alert experiment, subjects who appeared to be aware of the bias shifted further in the direction of the favored candidate (VMP = 67.9%, 95% CI, 50.6-85.2%; χ^2 = 17.053, P <0.001) than subjects who did not show awareness of the bias (VMP = 36.8%, 95% CI, 31.9-41.8%; χ^2 = 92.130, P <0.001), and this difference was significant (χ^2 = 9.265, P <0.01), a result consistent with past research on SEME [30]. This effect was reversed in the low alert experiment, where subjects who showed awareness of the bias had a VMP of 16.7% (95% CI, 8.7-24.6%; χ^2 = 4.971, P <0.05) and those who did not appear to be aware of the bias manipulation had a VMP of 23.6% (95% CI, 18.9-28.3%; χ^2 = 45.161, P <0.001). In the high alert experiment, the VMP for subjects who showed awareness of the bias was not significant at 15.9% (95% CI, 7.9-23.8%; χ^2 = 3.512, P = 0.06), and the VMP for subjects who did not show awareness dropped further to 13.3% (95% CI, 9.4-17.1%; χ^2 = 13.009, P <0.001).

5 DISCUSSION

Our findings provide additional evidence for the robustness of SEME [30] and new evidence that design interventions that alert people to election-related search ranking bias can significantly suppress SEME (Figure 3c), increase the proportion of users who show bias awareness, and shift the browsing patterns of users to lower ranked results (Figure 2). Consistent with the debiasing literature [39, 78, 117] we found that more detailed alerts (Figure 1c) suppressed SEME more than less detailed alerts (Figure 1b). This enhanced suppression appears to be due to the deeper searches conducted by subjects in the high alert experiment (Figure 2), as these deeper searches were also associated a decrease in VMP (Figure 7 in the Appendix).

Our results show that bias alerts might operate by reducing the number of votes that shift *toward* and increasing the number of votes that shift *away* from the candidate favored by the ranking bias (Table 4 in the Appendix), suggesting reactance to the persuasive threat we exposed [12]. However, despite the additional suppression of the high alerts, the lowest VMP was found among the neutral group subjects: rankings alternating between favoring the two candidates *prevented* SEME.

It is unclear why shifts toward Miliband were stronger and more resistant to the bias alerts (Figure 5). It is possible that Ed Miliband, the liberal candidate, was disproportionately favored by our liberal leaning sample (Table 1 in the Appendix), or that the discrepancy in subjects' mean familiarity with the two candidates (5.2 for Cameron and 2.8 for Miliband) affected their voting patterns. We find that this enhanced familiarity with Cameron actually disadvantaged him in the no alert experiment, even when the ranking bias was in his favor, but this disadvantage largely disappeared when a result was present (Figure 4) and does not explain the consistent differences in bias group VMPs across experiments (Figure 3c). Although the ranking bias of the search rankings seen by the two bias groups was nearly identical, it is possible that the search results and webpages that favored Cameron were somehow less persuasive than those that favored Miliband.

As with previous research on SEME [30], and with research on attitude change and influence more generally [3, 72, 120], we found that subjects vary in their susceptibility to SEME, as well as in their responsiveness to the alerts, based on their personal characteristics (Figure 6 and Figure 7 in the Appendix). This should give researchers concerned about the potential misuse of ranking algorithms pause, since the technologies needed to target individuals susceptible to SEME are already ubiquitously employed for advertising and personalization purposes [1, 96, 121]. Even more troubling is the fact that such personalization, in combination with the ephemerality of search engine rankings, makes detection of ranking bias an extremely difficult task that requires either complete cooperation of the search engine with regulators, or an extensive independent monitoring system [31, 53, 64].

As more people turn to the internet for political news [85, 115], designing systems that can monitor and suppress the effects of algorithm biases, like ranking bias, will play an increasingly important role in protecting the public's psychological vulnerabilities. Indeed, one recent study found evidence of election-related ranking bias in Twitter search [71], and our results suggest that this ranking bias could have had a large impact on users. The recent proliferation of misinformation during the 2016 U.S. election is also relevant here, as the consumption of "fake news" is inherently tied to the rank at which it appears in search and social media [4, 48], and its impact could be amplified at higher rankings. Indeed, both Google and Facebook have begun rolling out their own fact checking "tags" and "marks" that somewhat resemble our detailed alerts [38, 47]. However, internal efforts such as these come with notable transparency concerns due to the private and proprietary nature of their data [104].

Given the slow pace at which legislation moves relative to technology [101], our findings provide a foundation for suppressing SEME, and potentially the effects of other ranking algorithms, that does not require the cooperation of the entities in charge of the algorithm. However, our finding that SEME was only completely prevented when the rankings reflected an equal-time rule suggests that election-related search results might need to be regulated [30]. Further research is needed to examine how SEME operates within social media platforms as well as how personalization of rankings might enhance the effect [110].

Real-time automated bias detection could potentially be achieved by utilizing a Natural Language Processing (NLP) approach. One could utilize opinions [75], sentiment [99], linguistic patterns [109], word associations [14], or recursive neural networks [59] with human-coded data to classify biased language. Bias would have to be defined on a case-by-case basis. For election-related bias, one might consider the degree to which a search result or webpage favors a particular candidate as we did. Given that negative information can have a stronger impact than positive information [58], it is possible that using separate positive and negative scales would provide more information than the bipolar scale we used. Along with relevant controls for noise [53, 64], this bias detection method could be packaged into a browser extension that conducts systematic analyses [17, 53–55, 64, 71] and injects detailed alerts to forewarn users when ranking bias is detected. As a practical matter, classifying the bias of media outlets rather than individual webpages could provide a simpler route forward [86], but would also reduce the granularity of the bias detection.

5.1 Awareness of Bias

Awareness of ranking bias appears to suppress SEME only when it occurs in conjunction with a bias alert, perhaps because an alert is a kind of warning–inherently negative in nature. Our results here are potentially limited by our crude measure of awareness, yet we were able to replicate the awareness levels found in prior SEME research within a percentage point, and the significant differences in search behavior and VMP by our awareness measure provides some face validity.

Awareness of ranking bias in the absence of bias alerts might increase VMP because people perceive the bias as a kind of social proof [111, 112], made all the more powerful because of the disproportionate trust people have in search rankings [10, 95, 105]. The user's interpretation might be, "This candidate MUST be good, because even the search results say so."

5.2 Limitations

One limitation of this study is that subjects were exposed to biased search rankings only once in a controlled environment. In real life, people are exposed to a variety of sources at multiple times, and these mere exposures will influence their attitudes to some degree [122, 123]. Similarly, in real web search, people craft their own queries and frequently reformulate them before going past the first page of search results [60].

It is possible that allowing subjects to reformulate their query while still serving them biased rankings would (1) reduce the number of subjects who browse past the first page of results, and (2) instill subjects with an artificial sense of control over the search results and thereby increase the impact of the ranking bias [5, 41].

Our ability to accurately model the impact of SEME on real elections is also limited in several respects. First, we tested maximal experimental manipulations by selecting the most biased results to include in our mock search engine. It is unclear how more subtle ranking biases would affect the magnitude of SEME. Second, we did not measure subjects' familiarity with the partisan platforms and websites of UK politics. Familiarity with such cues could reduce the novelty and interpretation of the information in the search results and reduce the effects we observed. Third, our sample is not representative of the UK voting population. Similar to the study we replicated, we simply used an MTurk sample that was largely (1) unlikely to know the outcome of the election, and (2) unlikely to have strong opinions on the candidates. For these reasons the results from our work should be interpreted as an upper-bound on the influence that ranking bias can have on undecided voters conducting online searches.

Our findings on bias awareness are also limited. Measuring bias awareness is difficult not only in experimental settings where leading questions can create demand characteristics, but also in real-world data where users have limited opportunities to express awareness [37]. It is possible that additional behavioral data, such as mouse hovering, could provide a stronger signal for bias awareness.

Lastly, we have limited ability to measure the decay of the suppresive effect generated by our alerts. It is possible that if alerts were not presented on subjects' subsequent searches the suppression would be diminished or eliminated altogether. These questions are left to future investigations.

A APPENDIX

A.1 Candidate Biographies

"David Cameron. Born on October 9, 1966 in London, England. Cameron was educated at Heatherdown School and later at Eton College, where he entered two years early due to high academic achievement. He studied at the University of Oxford, where he earned his Bachelor of Arts in Philosophy, Politics, and Economics. After graduation, Cameron worked for the Conservative Research Department between 1988 and 1993, and subsequently served as Special Adviser to the Chancellor and Home Secretary. He was elected to Parliament in 2000 after a string of unsuccessful attempts to secure a seat. In 2005, he was elected Leader of the Opposition and Leader of the Conservative party. In 2010, at age 43, at the recommendation of resigned Prime Minister Gordon Brown, Cameron became the youngest British Prime Minister since Lord Liverpool. He is married to Samantha Gwendoline Sheffield, with whom he has four children.

Ed Miliband. Born on December 24, 1969 in the London Borough of Camden, England. Miliband moved around England frequently while growing up - his family following his father's teaching work. He entered Oxford University in 1989, where he studied Philosophy, Politics, and Economics. After graduation, Miliband was encouraged by then Shadow Chancellor Gordon Brown to attend the London School of Economics where he would obtain a Masters of Science in Economics. Miliband served as Special Adviser from 1997 to 2002. After spending some time in the United States teaching at Harvard, Miliband was elected to Parliament in 2005. In 2010, after Gordon Brown's resignation as Prime Minister and Leader of the Labour Party, Miliband was elected the Leader of the Opposition, and at age 40, the youngest Leader of the Labour Party ever. In 2011, he married barrister Justine Thornton, with whom he has two children."

A.2 Search Instructions

"You will now be given the opportunity to conduct some research on the previously mentioned candidates using our special internet search engine called Kadoodle. Your goal is to try to clarify your views on each candidate so that you are better able to decide which one deserves your vote. Use the search engine results we show you as you would normally use any search engine results, and please do not use other search engines to help you. That will invalidate your participation in our study. If you would like to conduct further research on the candidates after you complete our study, go right ahead, but please complete our study first! You will have a total of 15 minutes to conduct your search, and the program will automatically let you know when the time is up. Please do NOT close the window after conducting your search, doing so will make it impossible for you to complete the survey. Instead, if you feel you have enough information to make a clear choice between the candidates, you may end your search early by clicking End Search in the upper left corner of the results page."

A.3 Supplementary Tables and Figures

Table 1. Aggregate demographics for our subjects.

Demographic		*n*	%
Gender	Female	1636	45.4%
	Male	1964	54.6%
Education	None	10	0.3%
	Primary	153	4.2%
	Secondary	1146	31.8%
	Bachelors	1712	47.6%
	Masters	506	14.1%
	Doctorate	73	2.0%
Ethnicity	Asian	613	17.0%
	Black	218	6.1%
	Hispanic	172	4.8%
	White	2410	66.9%
	Mixed	144	4.0%
	Other	43	1.2%
Income	Under 15,000	534	14.8%
	$15,000 to 30,000	744	20.7%
	$30,000 to 45,000	573	15.9%
	$45,000 to 60,000	537	14.9%
	$60,000 to 75,000	400	11.1%
	$75,000 to 100,000	384	10.7%
	$100,000 to 150,000	203	5.6%
	$150,000 and over	95	2.6%
	I prefer not to say	130	3.6%
Marital	Divorced	231	6.4%
	Married	1393	38.7%
	Never married	1917	53.2%
	Separated	37	1.0%
	Widowed	22	0.6%
Employed	Yes	2902	80.6%
	No	698	19.4%
Political view	Conservative	555	15.4%
	Moderate	1101	30.6%
	Liberal	1723	47.9%
	Other	67	1.9%
	None	154	4.3%
Religion	Christianity	1529	42.5%
	Hinduism	318	8.8%
	Islam	56	1.6%
	Other	283	7.9%
	None	1328	36.9%
	I prefer not to say	86	2.4%
Country	United States	3093	85.9%
	India	389	10.8%
	Other	118	3.3%

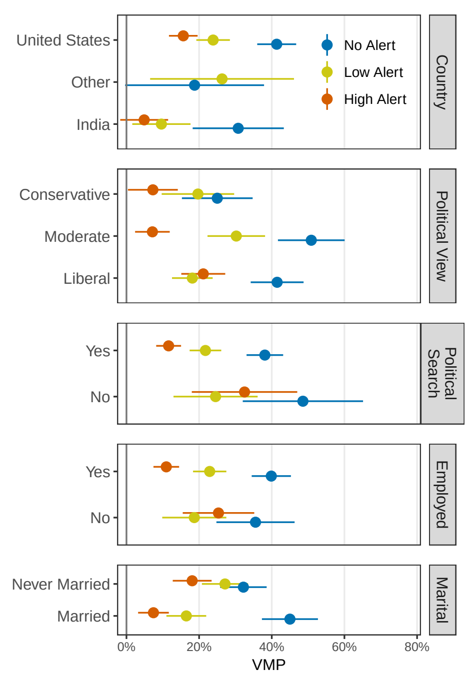


Fig. 6. VMP by demographic groups with significant differences and at least 60 subjects. Political search indicates whether subjects had previously used a search engine to look up information on political candidates. Error bars represent 95% CI.

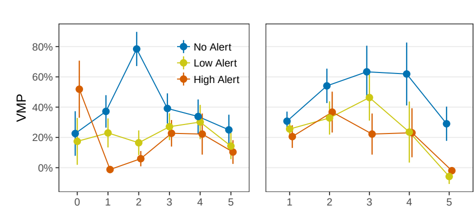


Fig. 7. VMP by subjects' initial attitude strength and search depth. Error bars represent 95% CI.

Table 2. Mean and standard error of candidate ratings by group and experiment pre- and post- search.

Experiment		Variable		Bias Group: Cameron	Miliband	Neutral	K-W χ^2	M-W U
No Alert	Pre	Impression	Cameron	7.18 (0.10)	7.22 (0.10)	7.25 (0.10)	0.294	78432.0
			Miliband	7.42 (0.09)	7.35 (0.09)	7.26 (0.09)	0.498	81352.5
		Likable	Cameron	7.11 (0.10)	7.14 (0.10)	7.10 (0.10)	0.181	79146.5
			Miliband	7.23 (0.09)	7.23 (0.09)	7.14 (0.09)	0.269	79280.5
		Trust	Cameron	6.20 (0.11)	6.35 (0.11)	6.41 (0.12)	1.982	76821.5
			Miliband	6.34 (0.10)	6.43 (0.10)	6.29 (0.11)	1.374	77417.5
	Post	Impression	Cameron	6.83 (0.12)	5.02 (0.13)	6.14 (0.12)	98.336***	111425.0***
			Miliband	5.36 (0.12)	7.42 (0.10)	6.21 (0.11)	161.494***	39799.0***
		Likable	Cameron	6.66 (0.12)	4.92 (0.13)	6.04 (0.12)	94.663***	110740.5***
			Miliband	5.47 (0.11)	7.30 (0.10)	6.39 (0.11)	135.928***	42517.5***
		Trust	Cameron	6.13 (0.13)	4.72 (0.13)	5.64 (0.13)	58.773***	104191.0***
			Miliband	4.95 (0.12)	6.78 (0.11)	5.98 (0.12)	123.786***	44352.5***
Low Alert	Pre	Impression	Cameron	7.00 (0.11)	7.14 (0.10)	7.04 (0.10)	0.969	77434.0
			Miliband	7.43 (0.09)	7.41 (0.08)	7.42 (0.08)	0.228	81440.5
		Likable	Cameron	6.87 (0.11)	6.97 (0.10)	6.92 (0.10)	0.604	77601.0
			Miliband	7.21 (0.09)	7.20 (0.09)	7.36 (0.08)	0.715	80951.5
		Trust	Cameron	6.04 (0.12)	6.17 (0.11)	6.09 (0.12)	0.539	77643.5
			Miliband	6.34 (0.10)	6.32 (0.10)	6.44 (0.10)	0.335	80871.0
	Post	Impression	Cameron	6.40 (0.12)	5.48 (0.12)	6.02 (0.12)	31.123***	97828.0***
			Miliband	5.91 (0.11)	7.00 (0.10)	6.29 (0.11)	57.876***	56022.5***
		Likable	Cameron	6.30 (0.12)	5.47 (0.12)	5.93 (0.12)	23.533***	95559.0***
			Miliband	6.02 (0.11)	6.96 (0.10)	6.36 (0.11)	43.392***	58951.5***
		Trust	Cameron	5.90 (0.13)	5.05 (0.12)	5.47 (0.13)	22.980***	95569.5***
			Miliband	5.55 (0.11)	6.45 (0.11)	5.94 (0.11)	36.630***	60541.5***
High Alert	Pre	Impression	Cameron	6.92 (0.11)	7.16 (0.10)	7.03 (0.10)	2.669	74800.0
			Miliband	7.38 (0.09)	7.46 (0.09)	7.47 (0.09)	0.572	77970.0
		Likable	Cameron	6.67 (0.11)	6.96 (0.10)	6.88 (0.11)	3.666	74074.5
			Miliband	7.19 (0.09)	7.15 (0.09)	7.25 (0.09)	0.990	81153.5
		Trust	Cameron	5.78 (0.12)	6.15 (0.11)	6.08 (0.12)	5.264	72852.0
			Miliband	6.16 (0.11)	6.30 (0.10)	6.41 (0.10)	2.730	76503.5
	Post	Impression	Cameron	5.86 (0.12)	5.49 (0.12)	5.89 (0.11)	7.429	87404.0
			Miliband	5.94 (0.11)	6.64 (0.10)	6.56 (0.10)	28.810***	64390.0***
		Likable	Cameron	5.80 (0.12)	5.48 (0.12)	5.87 (0.12)	5.755	85933.0
			Miliband	6.04 (0.11)	6.55 (0.10)	6.57 (0.10)	17.484**	68864.5**
		Trust	Cameron	5.28 (0.13)	5.02 (0.12)	5.35 (0.12)	3.371	84317.5
			Miliband	5.46 (0.11)	5.95 (0.11)	5.95 (0.11)	13.220*	69729.5*

*P < 0.05; **P < 0.01; ***P < 0.001.

Table 3. Subjects mean voting preferences (standard error) before and after completing the web research on the 11-point bipolar scale.

Experiment		Bias Group: Cameron	Miliband	Neutral	K-W χ^2	M-W U
No Alert	Pre	0.09 (0.16)	0.20 (0.16)	0.24 (0.15)	0.445	78373.5
	Post	1.09 (0.16)	-1.62 (0.16)	0.02 (0.17)	125.505***	115672.0***
Low Alert	Pre	-0.12 (0.16)	-0.04 (0.16)	0.06 (0.16)	0.761	78844.5
	Post	0.61 (0.16)	-1.09 (0.16)	0.00 (0.17)	54.984***	103890.5***
High Alert	Pre	-0.37 (0.15)	0.01 (0.15)	-0.31 (0.16)	3.755	74168.0
	Post	0.03 (0.16)	-0.72 (0.15)	-0.44 (0.16)	10.855**	90613.5**

*P < 0.05; **P < 0.01; ***P < 0.001.

Table 4. Shift in voting toward the candidate favored by bias group (+1), away from that candidate (-1), or no shift (0).

Experiment	Bias group	Vote Shift: -1	0	1	χ^2
No Alert	Cameron	24	294	82	41.165***
	Miliband	5	295	100	
	Neutral	39	309	52	
Low Alert	Cameron	16	335	49	18.990***
	Miliband	16	313	71	
	Neutral	36	317	47	
High Alert	Cameron	26	333	41	12.547*
	Miliband	24	315	61	
	Neutral	42	317	41	

*P < 0.05; **P < 0.01; ***P < 0.001.

A.4 Bias Awareness

The keywords and phrases we used to identify bias awareness were: "biased" "bias" "leaning towards" "leaning toward" "leaning against" "slanted" "slanted toward" "skewed" "skewed toward" "results favor" "results favored" "one sided" "favorable toward" "favorable towards |favorable for" "favorable against" "favorable results" "favored towards" "favored toward" "favored for" "favored against" "favored results" "favour toward" "favourable towards" "favourable toward" "favourable for" "favourable against" "favourable results" "favoured towards" "favoured toward" "favoured for" "favoured against" "favoured results" "favour toward" "results favour" "results favoured" "favor Cameron" "favor Miliband" "favour Cameron" "favour Miliband" "pro Cameron" "pro Miliband" "pro-Cameron" "pro-Miliband" "Cameron leaning " "Miliband leaning" "negative toward" "negative for" "negative against" "postive toward" "postive for" "postive against" "all postive" "all negative" "mainly positive" "mainly negative" "mostly positive" "mostly negative" "more negativity" "nothing positive" "nothing negative" "more results for" "less results for" "most of the articles were negative" "most of the articles were positive."

ACKNOWLEDGMENTS

We thank the anonymous reviewers and Brendan Nyhan for their extremely helpful comments. This research was supported in part by NSF grants IIS-1408345 and IIS-1553088. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

[1] Amr Ahmed, Mohamed Aly, Abhimanyu Das, Alexander J Smola, and Tasos Anastasakos. 2012. Web-scale multi-task feature selection for behavioral targeting. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. ACM, 1737–1741.

[2] Ahmed Allam, Peter Johannes Schulz, and Kent Nakamoto. 2014. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: two experiments manipulating Google output. *Journal of Medical Internet Research* 16, 4 (2014), e100.

[3] Sinan Aral and Dylan Walker. 2012. Identifying influential and susceptible members of social networks. *Science* 337, 6092 (2012), 337–341.

[4] Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.

[5] John A Bargh, Peter M Gollwitzer, Annette Lee-Chai, Kimberly Barndollar, and Roman Trötschel. 2001. The automated will: nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology* 81, 6 (2001), 1014.

[6] James R Beniger. 1987. Personalization of mass media and the growth of pseudo-community. *Communication research* 14, 3 (1987), 352–371.

[7] George R Bergus, Irwin P Levin, and Arthur S Elstein. 2002. Presenting risks and benefits to patients. *Journal of General Internal Medicine* 17, 8 (2002), 612–617.

[8] Adam J Berinsky, Gregory A Huber, and Gabriel S Lenz. 2012. Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis* 20, 3 (2012), 351–368.

[9] Shlomo Berkovsky, Jill Freyne, and Harri Oinas-Kukkonen. 2012. Influencing individually: fusing personalization and persuasion. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 2 (2012), 9.

[10] Edelman Berland. 2017. 2017 Edelman Trust Barometer. http://www.edelman.com/trust2017/. (2017). Accessed: 2017-03-07.

[11] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489, 7415 (2012), 295–298.

[12] Jack W Brehm. 1966. *A theory of psychological reactance.* Academic Press.

[13] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. 2011. Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* 6, 1 (2011), 3–5.

[14] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science* 356, 6334 (2017), 183–186.

[15] David Chavalarias. 2016. The unlikely encounter between von Foerster and Snowden: When second-order cybernetics sheds light on societal impacts of Big Data. *Big Data & Society* 3, 1 (2016), 1–11.

[16] Eric Chen, Gábor Simonovits, Jon A Krosnick, and Josh Pasek. 2014. The impact of candidate name order on election outcomes in North Dakota. *Electoral Studies* 35 (2014), 115–122.

[17] Le Chen, Alan Mislove, and Christo Wilson. 2016. An empirical analysis of algorithmic pricing on Amazon marketplace. In *Proceedings of the 25th International World Wide Web Conference*.

[18] Ye Chen, Dmitry Pavlov, and John F Canny. 2009. Large-scale behavioral targeting. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 209–218.

[19] Sidharth Chhabra and Paul Resnick. 2012. Cubethat: news article recommender. In *Proceedings of the Sixth ACM Conference on Recommender Systems*. ACM, 295–296.

[20] Sidharth Chhabra and Paul Resnick. 2013. Does clustered presentation lead readers to diverse selections?. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1689–1694.

[21] Chun-Fang Chiang and Brian Knight. 2011. Media Bias and Influence: Evidence from Newspaper Endorsements. *The Review of Economic Studies* 78, 3 (2011), 795–820.

[22] Inc comScore. 2017. comScore Explicit Core Search Query Report (Desktop Only). https://www.comscore.com/Insights/Rankings. (2017). Accessed: 2017-02-12.

[23] Stefano DellaVigna and Ethan Kaplan. 2007. The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics* 122, 3 (2007), 1187–1234.

[24] Nicholas Diakopoulos. 2014. Algorithmic accountability reporting: On the investigation of black boxes. *Tow Center for Digital Journalism, Columbia University* (2014).

[25] Nicholas Diakopoulos. 2016. Accountability in algorithmic decision making. *Commun. ACM* 59, 2 (2016), 56–62.

[26] Dimitar Dimitrov, Philipp Singer, Florian Lemmerich, and Markus Strohmaier. 2016. Visual positions of links and clicks on wikipedia. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 27–28.

[27] Luke Dormehl. 2014. *The formula: How algorithms solve all our problems–and create more.* Penguin, New York, NY.

[28] James N Druckman and Michael Parkin. 2005. The impact of media bias: How editorial slant affects voters. *Journal of Politics* 67, 4 (2005), 1030–1049.

[29] Hermann Ebbinghaus. 1913. *Memory: A contribution to experimental psychology.* Number 3. University Microfilms.

[30] Robert Epstein and Ronald E Robertson. 2015. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences* 112, 33 (2015), E4512–E4521.

[31] Robert Epstein and Ronald E Robertson. 2017. *A method for detecting bias in search rankings, with evidence of systematic bias related to the 2016 presidential election.* Technical Report White Paper no. WP-17-02. American Institute for Behavioral Research and Technology, Vista, CA.

[32] Eyal Ert and Aliza Fleischer. 2016. Mere Position Effect in Booking Hotels Online. *Journal of Travel Research* 55, 3 (2016), 311–321.

[33] Motahhare Eslami, Amirhossein Aleyasen, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. Feedvis: A path for exploring news feed curation algorithms. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*.

[34] Motahhare Eslami and Karrie Karahalios. 2017. Understanding and Designing around Users' Interaction with Hidden Algorithms in Sociotechnical Systems.. In *CSCW Companion*. 57–60.

[35] Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I like it, then I hide it: Folk theories of social feeds. In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2371–2382.

[36] Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. I always assumed that I wasn't really that close to [her]: Reasoning about Invisible Algorithms in News Feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 153–162.

[37] Motahhare Eslami, Kristen Vaccaro, Karrie Karahalios, and Kevin Hamilton. 2017. "Be careful; things can be worse than they appear": Understanding biased algorithms and users' behavior around them in rating platforms.. In *ICWSM*. 62–71.

[38] Facebook. 2017. News Feed FYI: Addressing Hoaxes and Fake News. https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/. (2017). Accessed: 2017-07-13.

[39] Baruch Fischoff. 1982. *Debiasing*. Cambridge University Press, Cambridge, MA.

[40] B.J. Fogg. 2002. *Persuasive technology: Using computers to change what we think and do.* Morgan Kaufmann, San Francisco, CA.

[41] Marieke L Fransen, Bob M Fennis, Ad Th H Pruyn, and Enny Das. 2008. Rest in peace? Brand-induced mortality salience and consumer behavior. *Journal of Business Research* 61, 10 (2008), 1053–1061.

[42] Jonathan L Freedman and David O Sears. 1965. Selective exposure. *Advances in Experimental Social Psychology* 2 (1965), 57–97.

[43] Jonathan L Freedman and David O Sears. 1965. Warning, distraction, and resistance to influence. *Journal of Personality and Social Psychology* 1, 3 (1965), 262–266.

[44] Alan S Gerber, Dean Karlan, and Daniel Bergan. 2009. Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics* 1, 2 (2009), 35–52.

[45] Tarleton Gillespie. 2014. *The relevance of algorithms*. Vol. 167. MIT Press Cambridge, Cambridge, MA.

[46] David F Gleich, Paul G Constantine, Abraham D Flaxman, and Asela Gunawardana. 2010. Tracking the random surfer: empirically measured teleportation parameters in PageRank. In *Proceedings of the 19th International Conference on World Wide Web*. ACM, 381–390.

[47] Google. 2017. Fact Check now available in Google Search and News around the world. https://www.blog.google/products/search/fact-check-now-available-google-search-and-news-around-world/. (2017). Accessed: 2017-07-13.

[48] Laura A Granka, Thorsten Joachims, and Geri Gay. 2004. Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 478–479.

[49] Edith Greene, Marlene S Flynn, and Elizabeth F Loftus. 1982. Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior* 21, 2 (1982), 207–219.

[50] Zhiwei Guan and Edward Cutrell. 2007. An eye tracking study of the effect of target rank on web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 417–420.

[51] Xunhua Guo, Mingyue Zhang, Chenyue Yang, et al. 2016. Order Effects in Online Product Recommendation: A Scenario-based Analysis. In *Proceedings of the 22nd Americas Conference on Information Systems*.

[52] Kevin Hamilton, Karrie Karahalios, Christian Sandvig, and Motahhare Eslami. 2014. A path to understanding the effects of algorithm awareness. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 631–642.

[53] Aniko Hannak, Piotr Sapieżyński, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. 2013. Measuring Personalization of Web Search. In *Proceedings of the 22nd International World Wide Web Conference*.

[54] Aniko Hannak, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson. 2014. Measuring price discrimination and steering on e-commerce Web sites. In *Proceedings of the 2014 ACM Conference on Internet Measurement*.

[55] Aniko Hannak, Claudia Wagner, David Garcia, Alan Mislove, Markus Strohmaier, and Christo Wilson. 2017. Bias in online freelance marketplaces: Evidence from TaskRabbit and Fiverr. In *20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*. Portland, OR.

[56] Daniel E Ho and Kosuke Imai. 2008. Estimating causal effects of ballot order from a randomized natural experiment the California alphabet lottery, 1978–2002. *Public Opinion Quarterly* 72, 2 (2008), 216–240.

[57] Statistics Brain Research Institute. 2017. Facebook Company Statistics. http://www.statisticbrain.com/facebook-statistics/. (2017). Accessed: 2017-02-12.

[58] Tiffany A Ito, Jeff T Larsen, N Kyle Smith, and John T Cacioppo. 1998. Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology* 75, 4 (1998), 887.

[59] Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. 2014. Political Ideology Detection Using Recursive Neural Networks. In *Association for Computational Linguistics*.

[60] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, Filip Radlinski, and Geri Gay. 2007. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transactions on Information Systems (TOIS)* 25, 2 (2007), 7.

[61] Yvonne Kammerer and Peter Gerjets. 2014. The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid interface. *International Journal of Human-Computer Interaction* 30, 3 (2014), 177–191.

[62] José H Kerstholt and Janet L Jackson. 1998. Judicial decision making: Order of evidence presentation and availability of background information. *Applied Cognitive Psychology* 12, 5 (1998), 445–454.

[63] Nuri Kim, Jon Krosnick, and Daniel Casasanto. 2015. Moderators of candidate name-order effects in elections: An experiment. *Political Psychology* 36, 5 (2015), 525–542.

[64] Chloe Kliman-Silver, Aniko Hannak, David Lazer, Christo Wilson, and Alan Mislove. 2015. Location, location, location: The impact of geolocation on Web search personalization. In *Proceedings of the 2015 ACM Conference on Internet Measurement*.

[65] Eric S Knowles and Jay A Linn. 2004. *Resistance and persuasion*. Psychology Press.

[66] Ron Kohavi, Alex Deng, Brian Frasca, Toby Walker, Ya Xu, and Nils Pohlmann. 2013. Online controlled experiments at large scale. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1168–1176.

[67] Alona Kolomiiets, Nathalie Dens, and Patrick De Pelsmacker. 2016. The wrap effect in online review sets revisited: How perceived usefulness mediates the effect on intention formation. *Journal of Electronic Commerce Research* 17, 4 (2016), 280–288.

[68] Jonathan GS Koppell and Jennifer A Steen. 2004. The effects of ballot position on election outcomes. *Journal of Politics* 66, 1 (2004), 267–281.

[69] Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24 (2014), 8788–8790.

[70] Jon A Krosnick, Joanne M Miller, and Michael P Tichy. 2004. An unrecognized need for ballot reform: Effects of candidate name order. In *Rethinking the vote: The politics and prospects of american election reform*, A.N Crigler, M.R. Just, and E.J. McCaffery (Eds.). Oxford University Press, New York, NY, 51–74.

[71] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P Gummadi, and Karrie Karahalios. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*.

[72] Howard Lavine, Joseph W Huff, Stephen H Wagner, and Donna Sweeney. 1998. The moderating influence of attitude strength on the susceptibility to context effects in attitude surveys. *Journal of Personality and Social Psychology* 75, 2 (1998), 359–373.

[73] Nicole R LaVoie, Brian L Quick, Julius M Riles, and Natalie J Lambert. 2015. Are graphic cigarette warning labels an effective message strategy? A test of psychological reactance theory and source appraisal. *Communication Research* (2015), 416–436.

[74] David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. 2009. Computational Social Science. *Science* 323, 5915 (2009), 721–723.

[75] Bing Liu, Minqing Hu, and Junsheng Cheng. 2005. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*. ACM, 342–351.

[76] Elizabeth F Loftus and Guido Zanni. 1975. Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society* 5, 1 (1975), 86–88.

[77] Lori Lorigo, Maya Haridasan, Hrönn Brynjarsdóttir, Ling Xia, Thorsten Joachims, Geri Gay, Laura Granka, Fabio Pellacini, and Bing Pan. 2008. Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the Association for Information Science and Technology* 59, 7 (2008), 1041–1052.

[78] Ramona Ludolph, Ahmed Allam, and Peter J Schulz. 2016. Manipulating Google's Knowledge Graph box to counter biased information processing during an online search on vaccination: application of a technological debiasing strategy. *Journal of Medical Internet research* 18, 6 (2016).

[79] William J McGuire. 1964. Some contemporary approaches. *Advances in Experimental Social Psychology* 1 (1964), 191–229.

[80] Jakub Mikians, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris. 2012. Detecting price and search discrimination on the Internet. In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*.

[81] Jakub Mikians, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris. 2013. Crowd-assisted search for price discrimination in e-commerce: First results. In *Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies*.

[82] Katherine L Milkman, Dolly Chugh, and Max H Bazerman. 2009. How can decision making be improved? *Perspectives on psychological science* 4, 4 (2009), 379–383.

[83] Claude H Miller. 2015. *Persuasion and psychological reactance: The effects of explicit, high controlling language*. Springer.

[84] Ronald B Mincy. 1993. The Urban Institute audit studies: their research and policy context. *Clear and Convincing Evidence: Measurement of Discrimination in America* (1993), 165–86.

[85] Amy Mitchell, Jeffrey Gottfried, and Katerina Eva Matsa. 2015. *Millennials and political news*. Technical Report. Pew Research Center.

[86] Sean A Munson, Stephanie Y Lee, and Paul Resnick. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget. In *ICWSM*.

[87] Sean A Munson and Paul Resnick. 2010. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 1457–1466.

[88] Bennet B Murdock. 1962. The serial position effect of free recall. *Journal of Experimental Psychology* 64, 5 (1962), 482–488.

[89] Jamie Murphy, Charles Hofacker, and Richard Mizerski. 2006. Primacy and recency effects on clicking behavior. *Journal of Computer-Mediated Communication* 11, 2 (2006), 522–535.

[90] Helen Nissenbaum. 1996. Accountability in a computerized society. *Science and Engineering Ethics* 2, 1 (1996), 25–42.

[91] Helen Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of social life.* Stanford University Press, Stanford, CA.

[92] Alamir Novin and Eric Meyers. 2017. Making sense of conflicting science information: Exploring bias in the search engine result page. In *Proceedings of the 2017 Conference on Human Information Interaction and Retrieval.* ACM, 175–184.

[93] Alice H Oh, Hyun-Jong Lee, and Young-Min Kim. 2009. User Evaluation of a System for Classifying and Displaying Political Viewpoints of Weblogs.. In *ICWSM.*

[94] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank citation ranking: Bringing order to the Web.* Technical Report. Stanford InfoLab.

[95] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication* 12, 3 (2007), 801–823.

[96] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you.* Penguin UK.

[97] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 443–452.

[98] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2012. A computational framework for media bias mitigation. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 2 (2012), 8.

[99] Souneil Park, Minsam Ko, Jungwoo Kim, Ying Liu, and Junehwa Song. 2011. The politics of comments: predicting political orientation of news stories with commenters' sentiment patterns. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work.* ACM, 113–122.

[100] Josh Pasek, Daniel Schneider, Jon A Krosnick, Alexander Tahk, Eyal Ophir, and Claire Milligan. 2014. Prevalence and moderators of the candidate name-order effect evidence from statewide general elections in California. *Public Opinion Quarterly* 78, 2 (2014), 416–439.

[101] Frank Pasquale. 2015. *The black box society: The secret algorithms that control money and information.* Harvard University Press, Cambridge, MA.

[102] Eyal Peer, Joachim Vosgerau, and Alessandro Acquisti. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior research methods* 46, 4 (2014), 1023–1031.

[103] Richard E Petty and Pablo Brinol. 2010. *Attitude change.* Oxford University Press Oxford, England, 217–259.

[104] Politico. 2017. Facebook undermines its own effort to fight fake news. http://www.politico.com/story/2017/09/07/facebook-fake-news-social-media-242407. (2017). Accessed: 2017-09-07.

[105] Kristen Purcell, Joanna Brenner, and Lee Rainie. 2012. *Search engine use 2012.* Technical Report. Pew Research Center's Internet and American Life Project.

[106] Emilee Rader and Rebecca Gray. 2015. Understanding user beliefs about algorithmic curation in the Facebook news feed. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, 173–182.

[107] Stephen A Rains. 2013. The nature of psychological reactance revisited: A meta-analytic review. *Human Communication Research* 39, 1 (2013), 47–73.

[108] Advanced Web Ranking. 2017. CTR study. https://www.advancedwebranking.com/cloud/ctrstudy. (2017). Accessed: 2017-04-01.

[109] Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. Linguistic Models for Analyzing and Detecting Biased Language. In *51st Annual Meeting of the Association for Computational Linguistics.* ACL, 1650–1659.

[110] Derek D Rucker, Zakary L Tormala, and Richard E Petty. 2004. Individual differences in resistance to persuasion: The role of beliefs and meta-beliefs. *Resistance and persuasion* (2004), 83.

[111] Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311, 5762 (2006), 854–856.

[112] Matthew J Salganik and Duncan J Watts. 2008. Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social psychology quarterly* 71, 4 (2008), 338–355.

[113] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms. In *Proceedings of "Data and Discrimination: Converting Critical Concerns into Productive Inquiry", a Productivereconference at the 64th Annual Meeting of the International Communication Association.*

[114] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2016. Automation, algorithms, and politics, when the algorithm itself is a racist: Diagnosing ethical Hharm in the basic components of software. *International Journal of Communication* 10 (2016), 19.

[115] Aaron Smith and Maeve Duggan. 2012. *Online political videos and campaign 2012.* Technical Report. Pew Research Journalism Project.

[116] S Shyam Sundar and Jinhee Kim. 2005. Interactivity and persuasion: Influencing attitudes with information and involvement. *Journal of Interactive Advertising* 5, 2 (2005), 5–18.
[117] Sander van der Linden, Anthony Leiserowitz, Seth Rosenthal, and Edward Maibach. 2017. Inoculating the public against misinformation about climate change. *Global Challenges* 1, 2 (2017).
[118] Guda Van Noort, Hilde AM Voorveld, and Eva A Van Reijmersdal. 2012. Interactivity in brand web sites: cognitive, affective, and behavioral responses explained by consumers' online flow experience. *Journal of Interactive Marketing* 26, 4 (2012), 223–234.
[119] Stephen J Weber and Thomas D Cook. 1972. Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychological Bulletin* 77, 4 (1972), 273.
[120] Duane T Wegener and Richard E Petty. 1997. The flexible correction model: The role of naive theories of bias in bias correction. *Advances in experimental social psychology* 29 (1997), 141–208.
[121] Jun Yan, Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen. 2009. How much can behavioral targeting help online advertising?. In *Proceedings of the 18th International Conference on World Wide Web*. ACM, 261–270.
[122] Robert B Zajonc. 1968. Attitudinal effects of mere exposure. *Journal of personality and social psychology* 9, 2p2 (1968), 1.
[123] Robert B Zajonc. 2001. Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science* 10, 6 (2001), 224–228.
[124] Jonathan L Zittrain. 2014. Engineering an election. *Harvard Law Review Forum* 127, 8 (2014), 335–341.

Received June 2017; revised August 2017; accepted November 2017