America's Digital Shield:

A New Online Monitoring System Will Make Google and Other Tech Companies Accountable to the Public

Testimony by

Robert Epstein, Ph.D. (re@aibrt.org)

Senior Research Psychologist
American Institute for Behavioral Research and Technology

Before the United States Senate Judiciary Subcommittee on Competition Policy, Antitrust, and Consumer Rights

Wednesday, December 13, 2023, 3:00 p.m.

Table of Contents

Appendix XIV: Epstein et al. (2022). **"The Answer Bot Effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines."** *PLOS ONE*, **406**

Appendix XV: Epstein and Zankich (2022). **"The surprising power of a click requirement: How click requirements and warnings affect internet users' willingness to disclose personal information."** *PLOS ONE*, **433**

Appendix XVI: Epstein et al. (2022). **"What would happen if Twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME)."** *PLOS ONE*, **455**

America's Digital Shield:
A New Online Monitoring System Will Make Google and Other Tech Companies Accountable to the Public

Testimony by

Robert Epstein, Ph.D. (re@aibrt.org)
Senior Research Psychologist
American Institute for Behavioral Research and Technology

Before the United States Senate Judiciary Subcommittee on Competition Policy, Antitrust, and Consumer Rights

Wednesday, December 13, 2023, 3:00 p.m.

I'm here to tell you about an existential threat to our country that is so well hidden you might know nothing about it. It's a threat posed by Big Tech monopolies, eerily predicted by President Eisenhower in 1961 (https://is.gd/ct5Gcb).

In 2016, Google alone shifted more than 2.6 million votes to Hillary Clinton using subliminal techniques I had been studying and quantifying since 2013 (https://TamingBigTech.com; Epstein, 2018d).

Four days later, a leaked video showed Google's leaders – devastated by Trump's win – telling their employees that they would not allow Trump to win the Presidency again (https://is.gd/ab4D8Z). They would *guarantee* his defeat using their "great strength and resources and reach." They made good on this promise in 2020, and in 2022, as I explained recently in *The Epoch Times*, they stopped the Red Wave cold (Epstein, 2022c; https://HowGoogleStoppedTheRedWave.com).

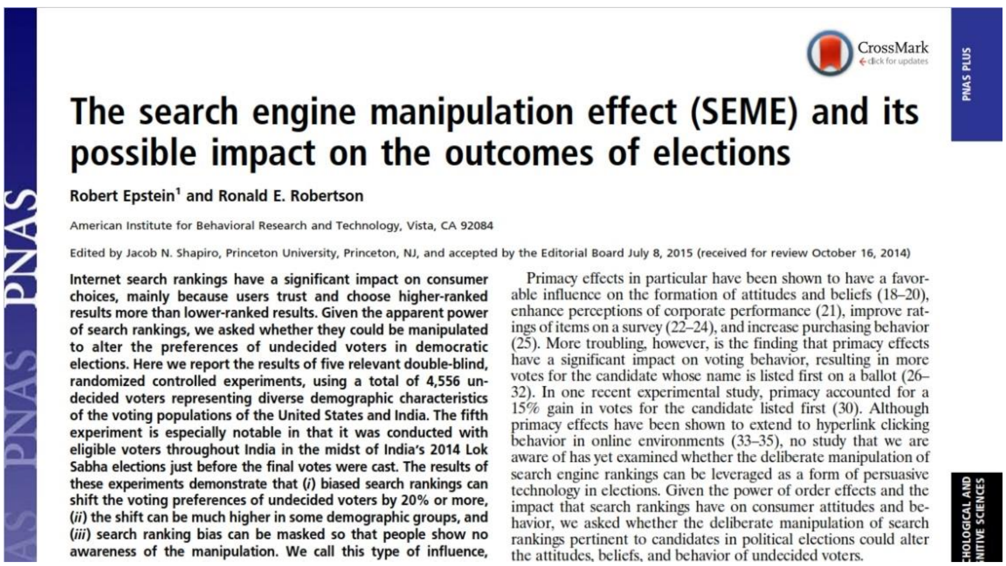
I lean left, but I don't think a private monopoly – one with no accountability to the public – should be able to pick our nation's leaders. Who knows how these secretive companies will lean *next* year, after all?

After that all-hands meeting, Google perfected at least a dozen new methods of subliminal control that I have now been studying for more than a decade (Epstein, 2018i, all see References and Appendices below.)

To shift votes, we know from leaked emails (https://is.gd/x9BtHn) that Google relies on what they call "ephemeral experiences" – fleeting content such as search results, search suggestions, and up-next videos on YouTube – content that impacts undecided voters and then disappears, leaving no paper trail.

Since 2016, my dedicated team has been building increasingly more sophisticated monitoring systems that *preserve and analyze ephemeral content* (Epstein, 2018d; Epstein et al., 2021a; Epstein et al., 2022b; Epstein & Peirson, 2023). This is Google's worst nightmare, because it means we are surveilling *them*, just as they surveil us and our children 24 hours a day. In other words, *we are giving you, our nation's leaders, the ammunition you need to hold Google accountable.*

Our research, which we publish in prestigious peer-reviewed journals (see Appendices), allows us to measure the power Big Tech has to shift votes, while our monitoring systems let us see whether these manipulations are being used.


CrossMark

PNAS PLUS

The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections

Robert Epstein[1] and Ronald E. Robertson

American Institute for Behavioral Research and Technology, Vista, CA 92084

Edited by Jacob N. Shapiro, Princeton University, Princeton, NJ, and accepted by the Editorial Board July 8, 2015 (received for review October 16, 2014)

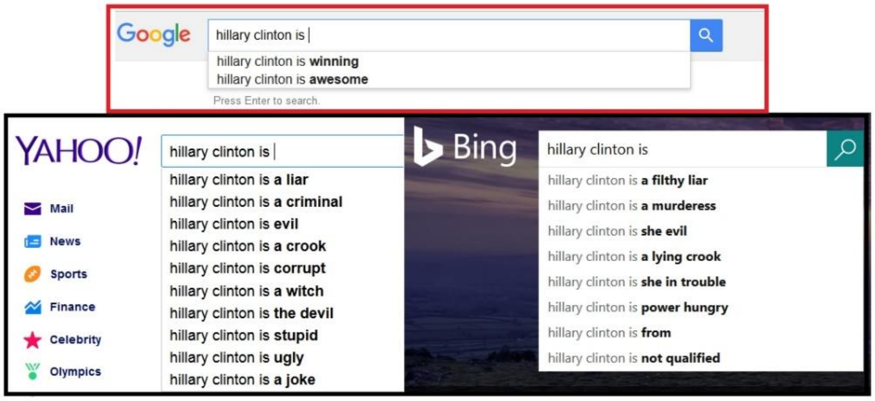
Internet search rankings have a significant impact on consumer choices, mainly because users trust and choose higher-ranked results more than lower-ranked results. Given the apparent power of search rankings, we asked whether they could be manipulated to alter the preferences of undecided voters in democratic elections. Here we report the results of five relevant double-blind, randomized controlled experiments, using a total of 4,556 undecided voters representing diverse demographic characteristics of the voting populations of the United States and India. The fifth experiment is especially notable in that it was conducted with eligible voters throughout India in the midst of India's 2014 Lok Sabha elections just before the final votes were cast. The results of these experiments demonstrate that (*i*) biased search rankings can shift the voting preferences of undecided voters by 20% or more, (*ii*) the shift can be much higher in some demographic groups, and (*iii*) search ranking bias can be masked so that people show no awareness of the manipulation. We call this type of influence,

Primacy effects in particular have been shown to have a favorable influence on the formation of attitudes and beliefs (18–20), enhance perceptions of corporate performance (21), improve ratings of items on a survey (22–24), and increase purchasing behavior (25). More troubling, however, is the finding that primacy effects have a significant impact on voting behavior, resulting in more votes for the candidate whose name is listed first on a ballot (26–32). In one recent experimental study, primacy accounted for a 15% gain in votes for the candidate listed first (30). Although primacy effects have been shown to extend to hyperlink clicking behavior in online environments (33–35), no study that we are aware of has yet examined whether the deliberate manipulation of search engine rankings can be leveraged as a form of persuasive technology in elections. Given the power of order effects and the impact that search rankings have on consumer attitudes and behavior, we asked whether the deliberate manipulation of search rankings pertinent to candidates in political elections could alter the attitudes, beliefs, and behavior of undecided voters.

SEME (the Search Engine Manipulation Effect) was discovered in 2013 and was reported by the Washington Post that spring. Five randomized, controlled experiments demonstrating the effect were published in the Proceedings of the National Academy of Sciences USA in 2015.

In one case so far, when we shared our data with Senators Lee, Johnson, and Cruz, they sent a strong letter to the CEO of Google (https://LetterToGoogleCEO.com), *which ceased its election manipulations that very day*. It turned off the political bias in its search engine and *stopped sending partisan "Go-Vote" reminders on its home page*; through our monitoring, we detected these changes the moment they were made (Epstein et al., 2021a).

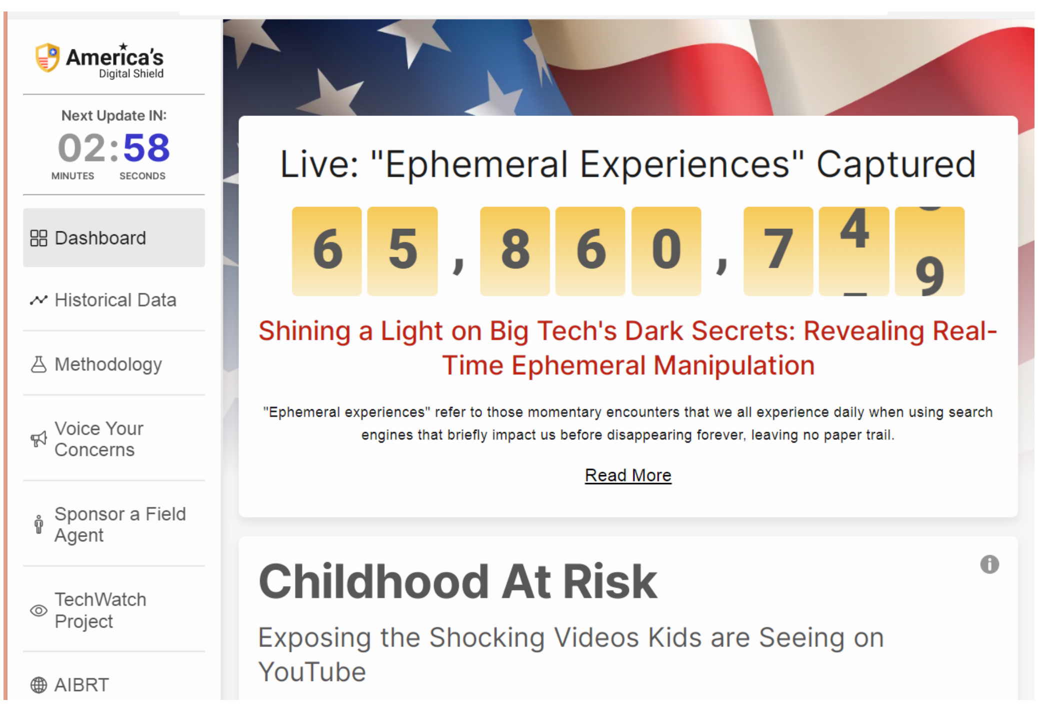
We preserve ephemeral content through the computers of a representative sample of real voters. One *must* monitor through the computers of real people because Big Tech sends out *personalized content*. To see what they're sending people, you must look over the shoulders of *real people*, just as the Nielsen company does with television viewers.



One of the simplest ways to support a candidate is to suppress negative search terms for that candidate. These screenshots from the summer of 2016 show Google was suppressing negative search suggestions for Hillary Clinton.

We started small in 2016 but have since deployed bigger systems with each election. In 2022, we preserved more than *2.5 million ephemeral experiences* through the computers of a politically-balanced group of 2,742 voters in 10 swing states.

We are now building the world's first *nationwide Digital Shield*, and we just released a public dashboard – AmericasDigitalShield.com – that shows our cumulative findings in real time. We are collecting and displaying data 24 hours a day through the computers of a politically-balanced group of more than *13,000 voters in all 50 states*, and so far we have *court-admissible data in 15 states*.



A partial view of a real-time display of data being streamed from the computers of a politically-balanced group of more than 13,000 registered voters in all 50 states. See: https://AmericasDigitalShield.com.

We have so far preserved and analyzed more than *66 million ephemeral experiences* on multiple platforms. This $3 million system is expanding every day, and we recently started preserving content being sent to more than 2,600 children and teens.

The extreme political bias we are seeing in content being sent to voters, along with the highly sexualized and violent content being sent to America's kids, confirm my worst fears: The "technological elite," as Eisenhower called them, are now in control of our democracy, and they are systematically indoctrinating our children.

If we can secure funding to complete our system so we have court-admissible data in all 50 states, the tech companies will almost certainly back down in 2024. Even if they don't, we will have incontrovertible evidence of election rigging on a massive scale.

If *no* monitoring system is in place, Google alone will be able to shift between *6.4 and 25.5 million votes* in the 2024 Presidential election, leaving no paper trail and making a mockery of the free-and-fair election.

Thank you, members of the Committee, for your attention, and for protecting our great nation from threats both foreign and, I hope, *domestic*.

REFERENCES

Papers in Press or Under Review:

Epstein, R., Lothringer, M., & Zankich, V. R. (in press). How a daily regimen of operant conditioning might explain the surprising power of the Search Engine Manipulation Effect (SEME). *Behavior and Social Issues.* (Preprint posted on SSRN, August 24, 2023, https://dx.doi.org/10.2139/ssrn.4551486, forwarding link: https://SEMEandOperantConditioning.com)

Epstein, R., & Li, J. (submitted for publication). *Can biased search results change people's opinions about anything at all? A close replication of the Search Engine Manipulation Effect (SEME)*. (Preprint posted on SSRN, October 10, 2023, https://dx.doi.org/10.2139/ssrn.4597654, forwarding link: https://MultipleTopicsResearch.com)

Epstein, R., Aries, S, Grebbien, K., Salcedo, A.M., & Zankich, V.R. (submitted for publication). *The Search Suggestion Effect (SSE): How search suggestions can be used to impact opinions and votes*. (Preprint posted on SSRN, August 8, 2023, https://dx.doi.org/10.2139/ssrn.4535163, forwarding link: https://SearchSuggestionEffect.com)

Epstein, R., & Flores, A. (submitted for publication). *The YouTube Manipulation Effect (YME): A quantification of the impact that the ordering of YouTube videos can have on opinions and voting preferences.* (Preprint posted on SSRN, July 31, 2023, http://dx.doi.org/10.2139/ssrn.4527207, forwarding link: https://YouTubeManipulationEffect.com)

Epstein, R., Huang, Y., & Megerdoomian, M. (submitted for publication). *The Opinion Matching Effect (OME): A subtle but powerful new form of influence that is apparently being used on the internet.* (Preprint posted on SSRN, August 4, 2023, https://dx.doi.org/10.2139/ssrn.4532141, forwarding link: https://OpinionMatchingEffect.com)

Epstein, R., Parsick, T., Shankar, P., & Zankich, V.R. (submitted for presentation). *The Differential Demographics Effect (DDE): Post hoc analyses of multiple datasets show the power of a new and invisible form of manipulation made possible by the internet.* https://DifferentialDemographicsEffect.com

Epstein, R., Newland, A., Tang, Liyu, & Buenaventura, Marco. (submitted for presentation). *The Digital Personalization Effect (DPE): How personalization of online content can dramatically increase the impact of biased online content.* https://DigitalPersonalizationEffect.com

Epstein, R. (submitted for presentation). *The ultimate mind control machine: Summary of a decade of empirical research on online search engines.*

https://aibrt.org/downloads/EPSTEIN_2024-WPA-The_Ultimate_Mind_Control_Machine.pdf

Epstein, R. (submitted for presentation). *America's "Digital Shield": How we are making Big Tech companies accountable to the public by continually preserving tens of millions of online ephemeral experiences – content that can impact users dramatically and that is normally lost forever*. https://aibrt.org/downloads/EPSTEIN_2024-WPA-Americas_Digital_Shield.pdf

2023:

Epstein, R., Tyagi, C., & Wang, H. (2023, July 27). What would happen if Twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME). *PLOS ONE.* https://doi.org/10.1371/journal.pone.0284495

Epstein. R. (2023, July 23). The perfect crime: Tech companies are manipulating our elections and indoctrinating our children – How we can stop them. Gatestone Institute. https://www.gatestoneinstitute.org/19821/tech-companies-manipulating-elections

Epstein, R., & Peirson, L. (2023, April). *How we preserved more than 2.5 million online ephemeral experiences in the 2022 midterm elections, and what this content revealed about online election bias.* Paper presented at the 103rd annual meeting of the Western Psychological Association, Riverside, CA. https://aibrt.org/downloads/EPSTEIN_&_Peirson_2023-WPA-How_We_Preserved_More_Than_2.5_Million_Online_Ephemeral_Experiences_in_the_Midterm_Elections.pdf

2022:

Epstein, R. (2022c, November 15). How Google stopped the Red Wave. *The Epoch Times*. https://www.theepochtimes.com/opinion/how-google-stopped-the-red-wave-4859553?utm_source=epochHG&utm_campaign=jj (forwarding link: https://HowGoogleStoppedTheRedWave.com)

Epstein, R. (2022, November 6). EPSTEIN: Google is shifting votes on a massive scale, but a solution is at hand. *Daily Caller*. https://dailycaller.com/2022/11/06/robert-epstein-2022-midterm-elections-google-bing/

Epstein, R., Lee, V., Mohr, R. & Zankich, V. R. (2022a). *The Answer Bot Effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. PLOS ONE.* https://doi.org/10.1371/journal.pone.0268081

Epstein, R., Bock, S., Peirson, L., Wang, H., & Voillot, M. (2022b, April). *How we preserved more than 1.5 million online "ephemeral experiences" in the recent U.S. elections, and what this content revealed about online election bias*. Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_et_al_2022-WPA-How_We_Preserved_More_Than_1.5_Million_Online_Ephemeral_Experiences_in_Recent_US_Elections...pdf

Epstein, R., & Voillot, M. (2022, April). *The YouTube Manipulation Effect (YME): The power that bias in YouTube's up-next algorithm has to shift votes and opinions, and preliminary evidence that such bias exists*. Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_&_VOILLOT_2022-WPA-YouTube_Manipulation_Effect-YME.pdf

Epstein, R., & Huang, Y. (2022, April). *The Opinion Matching Effect (OME): A subtle but powerful new form of influence that is being widely used on the internet without user awareness*. Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_&_HUANG_2022-WPA-Opinion_Matching_Effect-OME.pdf

Epstein, R., Lee, V., & Zankich, V. (2022, April). *The Answer Bot Effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines.* Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_et_al-2022-PLoS_ONE-in_press-The_Answer_Bot_Effect-ABE.pdf

Epstein, R., & Zankich, V. R. (2022). The surprising power of a click requirement: How click requirements and warnings affect internet users' willingness to disclose personal information. *PLOS ONE. 17*(2): e0263097. https://doi.org/10.1371/journal.pone.0263097

Epstein, R., Bordyug, M., Chen, Y., Chen, Y., Ginther, A., Kirkish, G, & Stead, H. (2022). The search for the perfect blade runner: A large, international assessment of a test that screens for "humanness sensitivity." *AI & Society*. https://aibrt.org/downloads/EPSTEIN_et_al_2022-Toward_the_search_for_the_perfect_blade_runner-EHI-humanness.pdf

2021:

Epstein, R. (2021). The technological elite are now in control. Chapter in Vol. 49 of *Champions of Freedom*. Hillsdale College Press. https://aibrt.org/downloads/EPSTEIN_2020-The_Technological_Elite_Are_Now_in_Control.pdf

Epstein, R., Bock, S., Peirson, L., & Wang, H. (2021, June 14). *Large-scale monitoring of Big Tech political manipulations in the 2020 Presidential election and 2021 Senate runoffs, and why monitoring is essential for democracy* (15-min. video). Paper presented at the 24th annual meeting of the American Association of Behavioral and Social Sciences. https://aibrt.org/downloads/EPSTEIN_et_al_2021-Large-Scale_Monitoring_of_Big_Tech_Political_Manipulations-FINAL_w_AUDIO.mp4

Epstein, R. (2021, January 28). When Big Tech flexes, everyone's freedom is threatened. *American Spectator*. https://spectator.org/trump-twitter-google-big-tech/

2020:

Epstein, R. (2020 November/December). Big Brother is in your house and in your voting booth. *Penthouse* (U.S. print edition), pp. 36-42. https://aibrt.org/downloads/EPSTEIN_2020-Big_brother_is_in_your_house_and_in_your_voting_booth.pdf

Epstein, R. (2020, November). *The technological elite are now in control* (1-hr. video). Invited talk presented at a meeting of the Center for Creative Alternatives, Hillsdale College, Hillsdale, MI. https://www.hillsdale.edu/event/cca-ii-big-tech/

2019:

Epstein, R. (2019, July 30). Google's latest whistleblower is hard to ignore. *Daily Caller*. https://dailycaller.com/2019/07/30/epstein-google-whistleblower/

Epstein, R. (2019, July 16). Why Google poses a serious threat to democracy, and how to end that threat. *Congressional Record of the United States*. https://www.judiciary.senate.gov/imo/media/doc/Epstein%20Testimony.pdf

Epstein, R. (2019, July 16). *Why Google poses a serious threat to democracy, and how to end that threat* (oral testimony, 7-min. video). Testimony before the United States Senate Judiciary Subcommittee on the Constitution, Washington, D.C. https://www.c-span.org/video/?c4807816/congressionalhearing-16july2019-testimonybydrrobertepstein

Epstein, R. (2019d, July 15). To break Google's monopoly on search, make its index public. *Bloomberg Businessweek*. https://www.bloomberg.com/news/articles/2019-07-15/to-break-google-s-monopoly-on-search-make-its-index-public

Epstein, R. (2019, April). *Evidence of systematic political bias in online search results in the 10 days leading up to the 2018 U.S. midterm elections*. Paper presented at the 99th annual meeting of the Western Psychological Association, Pasadena, CA. https://aibrt.org/downloads/EPSTEIN_2019-WPA-Evidence_of-search_engine_bias_related_to_2018_midterm_elections.pdf

Epstein, R. (2019c, April 3). Zucked again: Zuckerberg's proposal for regulating the internet is self-serving. *Epoch Times.* https://www.theepochtimes.com/opinion-zuckerbergs-proposal-for-regulating-the-internet-is-self-serving_2864920.html

Epstein, R. (2019b, March 22). Google, Facebook, Amazon: Warren's toothless break-up plan ignores real Big Tech threats. *USA Today*. https://www.usatoday.com/story/opinion/2019/03/22/elizabeth-warren-plan-misses-dangers-facebook-amazon-google-surveillance-column/3205451002/

Epstein, R. (2019a, January 2). How Google shifts votes: A "go vote" reminder is not always what you think it is. *Epoch Times*. https://www.theepochtimes.com/another-way-google-manipulates-votes-without-us-knowing-a-go-vote-reminder-is-not-what-you-think-it-is_2754073.html

2018:

Epstein, R. (2018i, September 26). Ten ways Big Tech can shift millions of votes in the November elections—without anyone knowing. *Epoch Times.* https://www.theepochtimes.com/10-ways-big-tech-can-shift-millions-of-votes-in-the-november-elections-without-anyone-knowing_2671195.html

Epstein, R. (2018h, September 13). Not just conservatives: Google and Big Tech can shift millions of votes in any direction. *USA Today*. https://www.usatoday.com/story/opinion/2018/09/13/google-big-tech-bias-hurts-democracy-not-just-conservatives-column/1265020002/

Epstein, R. (2018g, August 27). How major news organizations, universities and businesses surrender their privacy to Google. *The Daily Caller*. https://dailycaller.com/2018/08/27/surrender-privacy-google/

Epstein, R. (2018f, June 25). Zuck off: Six reasons Mark Zuckerberg should quit Facebook now. *The Daily Caller.* https://dailycaller.com/2018/06/25/mark-zuckerberg-should-quit-facebook/

Epstein, R. (2018e, May 25). Transcript to Google's internal video, "The Selfish Ledger." https://aibrt.org/downloads/GOOGLE-Selfish_Ledger-TRANSCRIPT.pdf

Epstein, R. (2018d, May 17). Taming Big Tech: The case for monitoring. *Hacker Noon.* https://hackernoon.com/taming-big-tech-5fef0df0f00d (Archived version at https://web.archive.org/web/20190323031918/https://hackernoon.com/taming-big-tech-5fef0df0f00d?gi=7d11d84801a2)

Epstein. R. (2018c). Manipulating minds: The power of search engines to influence votes and opinions. In M. Moore & D. Tambini (Eds.), *Digital dominance: The power of Google, Amazon, Facebook, and Apple* (pp. 294-319). Oxford, UK: Oxford University Press. https://aibrt.org/downloads/EPSTEIN_2018-Manipulating_minds-The-power_of_search_engines_to_influence_votes_and_opinions-UNCORRECTED_PROOFS.pdf

Epstein, R. (2018b). The unprecedented power of digital platforms to control opinions and votes. In G. Rolnik (Ed.), *Digital platforms and concentration: Second annual antitrust and competition conference* (pp. 31-33). Chicago, IL: University of Chicago Booth School of Business. https://promarket.org/wp-content/uploads/2018/04/Digital-Platforms-and-Concentration.pdf

Epstein, R. (2018a, March 22). Cambridge Analytica is not the problem: Google and Facebook are the problem. *The Daily Caller.* https://dailycaller.com/2018/03/22/google-and-facebook-are-problem-not-cambridge-analytica/

Epstein, R., & Mohr, R., Jr. (2018, April). *The Answer Bot Effect (ABE): Another surprising way search engines can impact opinions.* Paper presented at the 98th annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_&_MOHR_2018-WPA-The_Answer_Bot_Effect-ABE-WP_17_04.pdf

Epstein, R., Mohr., R., Jr., & Martinez, J. (2018, April). *The Search Suggestion Effect (SSE): How search suggestions can be used to shift opinions and voting preferences dramatically.* Paper presented at the 98th annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_MOHR_&_MARTINEZ_2018-WPA-The_Search_Suggestion_Effect-SSE-WP-17-03.pdf

2017:

Epstein, R. (2017h, August 31). Google's fighting hate and trolls with a dangerously mindless AI. *Fast Company.* https://www.fastcompany.com/40459339/google-perspective-fighting-hate-and-trolls-with-a-mindless-a-i

Epstein, R. (2017g, July 17). Hacking Google. *Huffington Post*. https://www.huffingtonpost.com/entry/596d0670e4b0376db8b659fd

Epstein, R. (2017f, May 28). Is it still possible to stop 'Big Tech' from killing democracy? [Review of Jonathan Taplin's *Move Fast and Break Things*]. *The Hill*. https://thehill.com/blogs/pundits-blog/technology/335507-is-it-still-possible-to-stop-big-tech-from-killing-democracy

Epstein, R. (2017e, May 10). The potential for Google and Facebook to manipulate millions of voters poses a bigger threat than fake news [Letter to the editor]. *Los Angeles Times*.

Epstein, R. (2017d, April 10). Fake news is a fake problem. *Medium*. https://medium.com/@re_53711/fake-news-is-a-fake-problem-914d7ffc7a91

Epstein, R. (2017c, March 16). Seven simple steps toward online privacy. *Medium*. https://medium.com/@re_53711/seven-simple-steps-toward-online-privacy-20dcbb9fa82

Epstein, R. (2017b, April). *Can search suggestions impact what we search for online? The role of negativity bias.* Paper presented at the 97th annual meeting of the Western Psychological Association, Sacramento, CA. https://aibrt.org/downloads/EPSTEIN_2017-WPA-Can_Search_Suggestions_Impact_What_We_Search_for_Online.pdf

Epstein, R. (2017a, March*). The Search Suggestion Effect (SSE): How autocomplete can be used to impact votes and opinions.* Paper presented at the 2nd biennial meeting of the International Convention of Psychological Science, Vienna, Austria. https://aibrt.org/downloads/EPSTEIN_2017-The_Search_Suggestion_Effect-SSE-ICPS_Vienna-March_2017.pdf

Epstein, R., Ding., M., Mourani, C., Olson, E., Robertson, R.E., & Tran, F. (2017, April). *Multiple searches increase the impact of the Search Engine Manipulation Effect (SEME)*. Paper presented at the 97th annual meeting of the Western Psychological Association, Sacramento, CA. https://aibrt.org/downloads/EPSTEIN_et_al._2017-WPA-Multiple_Searches_Increase_the_Impact_of%20_the_Search_Engine_Manipulation_Effect.pdf

Epstein, R., Mourani, C., Olson, E., & Robertson, R.E. (2017, April). *Biased search rankings can shift opinions on a wide range of topics.* Paper presented at the 97th annual meeting of the Western Psychological Association, Sacramento, CA. https://aibrt.org/downloads/EPSTEIN_et_al._2017-WPA-Biased_Search_Rankings_Can_Shift_Opinions_on_a_Wide_Range_of_Topics.pdf

Epstein, R., & Robertson, R.E. (2017c, June 1). *A method for detecting bias in search rankings, with evidence of systematic bias related to the 2016 presidential* *election.* Vista, CA: American Institute for Behavioral Research and Technology, White Paper no. WP-17-02. Retrieved from https://aibrt.org/downloads/EPSTEIN_&_ROBERTSON_2017-A_Method_for_Detecting_Bias_in_Search_Rankings-AIBRT_WP-17-02_6-1-17.pdf

Epstein, R., & Robertson, R.E. (2017b, March). *The Search Engine Manipulation Effect (SEME): Understanding its power to change opinions and votes*. Paper presented at the 2nd biennial meeting of the International Convention of Psychological Science, Vienna, Austria.
DC.

Epstein, R., & Robertson, R. (2017a). Suppressing the Search Engine Manipulation Effect (SEME). *Proceedings of the ACM: Human-Computer Interaction, 1*(2), Article 42. https://aibrt.org/downloads/EPSTEIN_et_al-2017-Suppressing_the_Search_Engine_Manipulation_Effect_(SEME).pdf

Epstein, R., Robertson, R., Shepherd, S., & Zhang, S. (2017, April). *A method for detecting bias in search rankings, with evidence of systematic bias related to the 2016 presidential election*. Paper presented at the 97th annual meeting of the Western Psychological Association, Sacramento, CA. https://aibrt.org/downloads/EPSTEIN_et_al_2017-SUMMARY-WPA-A_Method_for_Detecting_Bias_in_Search_Rankings.pdf

2016:

Epstein, R. (2016, November 17). Fake news and Facebook: There are far more pernicious ways social media can sway elections [Letter to the Editor]. *Los Angeles Times*. https://www.latimes.com/opinion/readersreact/la-ol-le-fake-news-facebook-20161117-story.html

Epstein, R. (2016k). Subtle new forms of internet influence are putting democracy at risk worldwide. In N. Lee (Ed.), *Google it: Total information awareness* (pp. 253-259). Springer. https://aibrt.org/downloads/EPSTEIN_2016-Suble_New_Forms_of_Internet_Influence-In_N.Lee_Ed.-Google_It-Springer.pdf

Epstein, R. (2016j, October 14). Breaking news: Google to donate its search engine to the American public. *Huffington Post*. https://www.huffingtonpost.com/dr-robert-epstein/breaking-news-google-to-d_b_12446856.html

Epstein, R. (2016i, September 12). Are we being manipulated by Google's autocomplete? *Sputnik International*. https://sputniknews.com/us/20160912/1045214398/google-clinton-manipulation-election.html

Epstein, R. (2016h, September 6). Free isn't freedom: How Silicon Valley tricks us. *Motherboard*. https://motherboard.vice.com/read/free-isnt-freedom-epstein-essay

Epstein, R. (2016g, September). Cyber sway: The new mind control. *Ladybeard*. https://aibrt.org/downloads/EPSTEIN_2016-Cyber_Sway-The_New_Mind_Control-LADYBEARD-9-16-16-FINAL_flat.pdf

Epstein, R. (2016f, July 12). Five subtle ways Facebook could influence the U.S. presidential election this fall. *Quartz*. https://qz.com/703680/five-subtle-ways-facebook-could-influence-the-us-presidential-election-this-fall/

Epstein, R. (2016e, July). Can search engine rankings swing elections? *New Internationalist*. https://newint.org/features/2016/07/01/can-search-engine-rankings-swing-elections/

Epstein. R. (2016d, June 22). The new censorship. *U.S. News & World Report*. https://www.usnews.com/opinion/articles/2016-06-22/google-is-the-worlds-biggest-censor-and-its-power-must-be-regulated

Epstein, R. (2016c, May 4). Bigger brother: Microsoft and Google's new pact could signal the beginning of the end for personal privacy. *Quartz*. https://qz.com/676184/microsoft-and-googles-pact-is-the-end-of-personal-privacy/

Epstein, R. (2016b, April 27). Google knows: In the future, Big Data will make actual voting obsolete. *Quartz*. https://qz.com/669983/maybe-we-should-let-google-vote-for-us/

Epstein, R. (2016a, February 18). The new mind control. *Aeon*. https://aeon.co/essays/how-the-internet-flips-elections-and-alters-our-thoughts [Russian translation here: http://mnenia.zahav.ru/Articles/7302/soznanie_pod_controlem]

Epstein, R., & Edelman, B. (2016, November 5). The other elephant in the voting booth: Big Tech could rig the election. *The Daily Caller*. https://dailycaller.com/2016/11/04/the-other-elephant-in-the-voting-booth-big-tech-could-rig-the-election/

Epstein, R., & Robertson, R.E. (2016b, April). *Why is the search engine manipulation effect (SEME) so large? A test of an operant conditioning hypothesis.* Paper presented at the 96th annual meeting of the Western Psychological Association, Long Beach, CA. https://aibrt.org/downloads/EPSTEIN_&_ROBERTSON_2016-SEME_Testing_an_operant_conditioning_hypothesis-WPA.pdf

Epstein, R., & Robertson, R.E. (2016a, April). *A replication of the search engine manipulation effect (SEME), plus methods for suppressing the effect.* Paper presented at the 96th annual meeting of the Western Psychological Association, Long Beach, CA. https://aibrt.org/downloads/EPSTEIN_et_al-2017-Suppressing_the_Search_Engine_Manipulation_Effect_(SEME).pdf

2015:

Epstein, R. (2015d). Wie Google Wahlen beeinflussen kann [How Google influences opinions]. In T. Fricke & U. Novak (Eds.), *Die Akte Google: Wie der U.S.-Konzern Daten missbraucht, die Welt manipuliert und Jobs vernichtet, Mit Beiträgen von Prof. Dr. Robert Epstein und Dr. Thomas Höppner* (pp. 165-173). München, Deutschland: F.A. Herbig Verlagsbuchhandlung GmbH.

Epstein, R. (2015c, October 6). Google's hypocrisy. *Huffington Post.* Retrieved from https://www.huffingtonpost.com/dr-robert-epstein/googles-hypocrisy_b_8253332.html

Epstein, R. (2015b, September 6). Google's vote counts more than yours. *Pittsburgh Post-Gazette*. https://www.post-gazette.com/opinion/Op-Ed/2015/09/06/Google-s-vote-counts-more-than-yours-because-its-search-engine-is-determining-the-outcomes-of-elections/stories/201509060042

Epstein, R. (2015a, August 19). How Google could rig the 2016 election. *Politico*. https://www.politico.com/magazine/story/2015/08/how-google-could-rig-the-2016-election-121548.html

Epstein, R., & Robertson, R.E. (2015b, April). *The Search Engine Manipulation Effect (SEME): Large-scale replications in two countries.* Paper presented at the 95th annual meeting of the Western Psychological Association, Las Vegas, Nevada.

Epstein, R., & Robertson, R. E. (2015a, August 4). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences USA*, *112*(33), E4512-E4521. https://www.pnas.org/content/112/33/E4512.full.pdf?with-ds=yes (forwarding link: https://SearchEngineManipulationEffect.com)

2014:

Epstein, R. (2014d). Democracy at risk from new forms of internet influence. *EMMA Magazine*. https://aibrt.org/downloads/EPSTEIN_2014-New_Forms_of_Internet_Influence-EMMA_Magazine.pdf

Epstein, R. (2014c, June). How Google could end democracy. *U.S. News & World Report*. https://www.usnews.com/opinion/articles/2014/06/09/how-googles-search-rankings-could-manipulate-elections-and-end-democracy?src=usn_tw

Epstein, R. (2014b, May). Google critic killed in "ironic" car accident: Struck by Google Street View vehicle. *Huffington Post*. https://www.huffingtonpost.com/dr-robert-epstein/google-critic-killed-in-i_b_5351458.html

Epstein, R. (2014a, May). Google's snoops: Mining our private data for profit and pleasure. *Dissent*. https://www.dissentmagazine.org/online_articles/googles-snoops-mining-our-data-for-profit-and-pleasure

Epstein, R., & Robertson, R.E. (2014, April). *Helping people preserve their privacy online: The surprising power of a click requirement.* Paper presented at the 94th annual meeting of the Western Psychological Association, Portland, Oregon. https://aibrt.org/downloads/EPSTEIN&ROBERTSON_2014-Helping_People_Be_More_Cautious_Online-ABSTRACT-WPA-April_2014.pdf

<u>2013</u>:

Epstein, R. (2013b, May). Google's gotcha. *U.S. News & World Report*. https://www.usnews.com/opinion/articles/2013/05/10/15-ways-google-monitors-you

Epstein, R. (2013a, March). Google's dance. *TIME*. https://techland.time.com/2013/03/27/googles-dance/

Epstein, R., & Robertson, R. E. (2013, May). *Democracy at risk: Search rankings can shift voter preferences substantially.* Paper presented at the 25th annual meeting of the Association for Psychological Science, Washington, D.C., May 2013.

<u>2012</u>:

Epstein, R. (2012e, November 5). Why Google should be regulated (Part 4 - End). *Huffington Post*. https://www.huffpost.com/entry/why-google-should-be-regu_b_2069223

Epstein, R. (2012d, November 2). Why Google should be regulated (Part 3). *Huffington Post*. https://www.huffpost.com/entry/why-google-should-be-regu_b_2054111 6-29-19 update: Link to Part 4 at the end of the article is broken. Use: https://www.huffpost.com/entry/why-google-should-be-regu_b_2069223

Epstein, R. (2012c, October 31). Why Google should be regulated (Part 2). *Huffington Post*. https://www.huffpost.com/entry/online-privacy_b_2013583 6-

29-19 update: Link to Part 3 at the end of the article is broken. Use: https://www.huffpost.com/entry/why-google-should-be-regu_b_2054111

Epstein, R. (2012b, October 23). Why Google should be regulated (Part 1). *Huffington Post*. https://www.huffpost.com/entry/google-privacy_b_1962827 6-29-19 update: Link to Part 2 at the end of the article is broken. Use: https://www.huffpost.com/entry/online-privacy_b_2013583 (Note: An edited version of the entire article first appeared in *The Kernel* [UK] on September 5, 2012. See below.)

Epstein, R. (2012a, September 12). Google: The case for hawkish regulation. *The Kernel*. https://web.archive.org/web/20121016112820/https://www.kernelmag.com/features/report/3281/google-the-case-for-hawkish-regulation (Expanded version appeared in four parts in *The Huffington Post* beginning on October 23, 2012. See above.)

APPENDIX I: Links of Possible Interest

https://AmericasDigitalShield.com – a live dashboard that documents Big Tech manipulation, bias, and indoctrination in real time.

https://EpsteinInTheNewYorkPost.com – May 2023 article in the *New York Post* about Dr. Epstein's research by reporter Miranda Devine. It ends, "Only Epstein is standing in the way."

https://TechWatchProject.org – a new website about Dr. Epstein's election monitoring project.

https://HowGoogleStoppedTheRedWave.com – a 2022 article by Dr. Epstein in *The Epoch Times.*

https://MyGoogleResearch.com – a webpage where you can learn more about Dr. Epstein's research on online influence and where you can also support that research with donations to the American Institute for Behavioral Research and Technology, a nonprofit, nonpartisan 501(c)(3) public charity.

https://EpsteinOnRogan.com – a 160-minute video recording of Dr. Epstein's 2022 appearance on The Joe Rogan Experience.

https://MyPrivacyTips.com – an essay by Dr. Epstein about how you can protect yourself and your children from surveillance by Google-and-the-Gang.

https://EpsteinTestimony.com – Dr. Epstein's 2019 Congressional testimony about the threat Google-and-the-Gang pose to democracy (7-minute video).

https://EpsteinOnSTEMTalks – a 90-minute biographical audio interview with Dr. Epstein.

https://TamingBigTech.com – an essay by Dr. Epstein about the development of his first election monitoring system, deployed before the 2016 Presidential election.

https://CreepyLine.org – an 80-minute documentary film – "The Creepy Line" – featuring Dr. Epstein's research. It warns about surveillance, censorship, and manipulation by Google-and-the-Gang. It also features Dr. Jordan Peterson and other experts.

https://TheCaseForMonitoring.com – a 15-minute video in which Dr. Epstein summarizes findings from his online monitoring in the days leading up to the 2020 Presidential Election and the 2021 Senate runoff elections in Georgia.

https://DrRobertEpstein.com – Dr. Epstein's personal website.

https://AIBRT.org – website of the American Institute for Behavioral Research and Technology.

https://TheNewCensorship.com – Dr. Epstein on Google's blacklists, in *US News & World Report*.

https://TamingBigTech.com – article by Dr. Epstein on AIBRT's 2016 election monitoring project.

https://LetterToGoogleCEO.com – Nov. 5, 2020 letter from three US Senators to Google CEO about Epstein's findings in the 2020 Presidential race.

https://SearchEngineManipulationEffect.com – SEME: 2015 seminal paper on the power that search engines have to shift opinions and votes, published in the *Proceedings of the National Academy of Sciences USA*, downloaded or accessed from the website of the National Academy of Sciences more than 250,000 times.

https://TargetedMessagingEffect.com – TME: 2023 peer-reviewed study in *PLOS ONE* showing the power that targeted messages on Twitter have to shift opinions and votes)

https://TheAnswerBotEffect.com – ABE: 2021 peer-reviewed study in *PLOS ONE* reporting new research on the power that personal assistants and answer boxes (and hence AIs) have to shift opinions and votes.

https://SearchSuggestionEffect.com – SSE: preprint of a research report on the power that Google search suggestions have to shift opinions and votes, currently under review.

https://YouTubeManipulationEffect.com – YME: preprint of a research report on the power that YouTube has to shift opinions and votes, currently under review.

https://OpinionMatchingEffect.com – OME: preprint of a research report on the power that online quizzes have to shift opinions and votes, currently under review.

https://MultipleTopicsResearch.com – preprint of a research report on the power that search engines have to shift opinions and votes about perhaps any topic at all, currently under review.

https://MultipleExposureEffect.com – MEE: preprint of new report on the additive impact of repeated exposures to similarly biased content, currently under review.

https://DigitalPersonalizationEffect.com – DPE: new research on the power that personalization has to increase the impact of biased content, submitted for presentation.

APPENDIX II:
The Methodology of SEME Experiments

The methodology of SEME experiments adheres to the highest standards of research in the social and behavioral sciences. All experiments are randomized, controlled, double-blind, and counterbalanced (Epstein and Robertson, 2015a). Multiple SEME experiments conducted over a period of more than five years have involved more than 10,000 participants and five national elections in four countries. Reasonable efforts have been made to assure that participants are diverse across multiple demographic characteristics, and, when possible, representative of the voting population. When samples are not representative of the voting population, adjustments are made statistically or by examining subsamples.

In most experiments, participants are selected who are "undecided," by which I mean either that they haven't yet made up their minds, or, in some cases, that we are deliberately showing them materials from an election they are not familiar with (for example, when we show people from the U.S. materials from an election in Australia).

All search results and web pages used in the experiments are real, drawn from the internet and from Google's search engine. The elections we have examined are also real: the 2010 election for Prime Minister of Australia; the 2014 Lok Sabha election in India; the 2015 national election in the UK, and the 2016 and 2018 elections in the U.S.

Search results are presented to participants using a mock search engine called Kadoodle, which looks and functions almost exactly like Google. The difference between Google and Kadoodle is that with Kadoodle, we control what search results we show and the order in which those results are shown. Our search results link to copies of real web pages, but links on those pages have been disabled so we can keep our research participants in a closed online environment.

In the basic procedure, participants are randomly assigned to one of three groups: a group in which search results favor Candidate A – which means that high-ranking results link to web pages that make Candidate A look better than his or her opponent – a group favoring Candidate B, and a group in which neither candidate is favored in search results (the control group).

Participants are told they will be asked to use our custom search engine, Kadoodle, to conduct research on political candidates. They are first asked to read short paragraphs about each candidate and then asked several questions about each candidate: How much they like each candidate, trust each candidate, and so on. They are also asked, both in a binary fashion and on a scale, which candidate they would vote for if they had to vote today. These are all "pre-search questions."

Then, typically, they are given up to fifteen minutes in which to use the Kadoodle search engine to conduct further research about the candidates. They are typically given access to five pages of search results, with six results per page (30 in total), and they can navigate through the search results and the web pages exactly as they would on Google. They can stop searching when they please.

Then they are asked those same questions about the candidates; now these are "post-search questions."

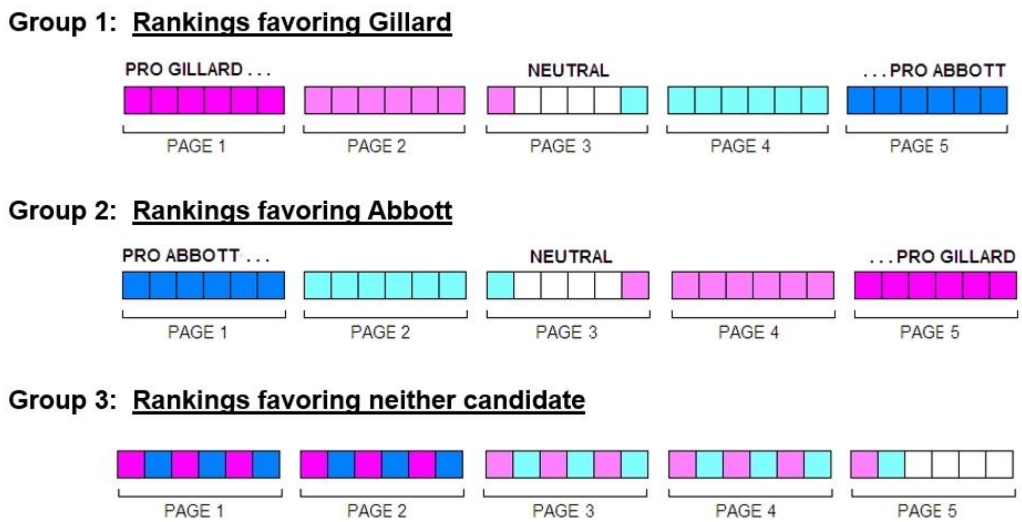


Figure 1. In a typical SEME experiment, in one group, search results are ordered in a way that favors Candidate A (Gillard, above). In a second group, the ordering is reversed, so it favors Candidate B (Abbott, above). And in a control group, the ordering alternates, so neither candidate is favored.

Remember that the only difference between the three groups is the order in which the search results are shown. All participants in all three groups have full access to all the search results and all the web pages.

The typical findings are as follows:

- Prior to search, all three groups tend to answer the pre-search questions the same way.
- After the search, the opinions and voting preferences of people in the control group shift very little or not at all.
- After the search, both the opinions and the voting preferences of people in the two bias groups shift fairly dramatically in the direction of the favored candidate. In other words, opinions and votes shift in opposite directions in the two groups.
- A shift of 20 percent or more is typical. In large studies in which we have enough participants to look at demographic differences, we have found shifts in the 60-to-

80 percent range in some demographic groups. In other words, some people are especially trusting of search results.

- Typically, very few people show any awareness of the bias they have seen. In a large study we conducted in India in 2014, for example – a study with more than 2,000 undecided voters throughout India in the midst of an intense election – 99.5 percent of our participants showed no awareness of bias in the search results we showed them.
- The very few people who do detect the bias tend, on average, to shift even farther in the direction of the bias.

Some of my SEME research attempts to explain why the effect is so large. One reason appears to be that people trust algorithmic output, believing that because it is computer-generated, it is inherently objective and unbiased.

Research I have conducted also suggests that SEME is a large effect because people are conditioned – very much like rats in a Skinner box – to believe that results at the top of the list are better and truer than results farther down the list (Epstein et al., in press). This is because most searches we conduct are for simple facts, such as "Who is the governor of Texas?" The correct answer always turns up at the top of the list, which is one reason 50 percent of all clicks go to the top two search positions.

But then that day comes when we search for something with a less certain answer: What is the best sushi restaurant in town? Who is the best candidate? Again, we are most likely to believe the highest-ranking answers.

When, in one experiment, we changed people's beliefs about high-ranking search results by placing answers to simple questions in random positions in lists of search results, politically-biased search results had less impact on them.

APPENDIX III
Article from *Bloomberg Businessweek*, July 15, 2019

https://www.bloomberg.com/news/articles/2019-07-15/to-break-google-s-monopoly-on-search-make-its-index-public
Entered into The Congressional Record, July 16, 2019

Bloomberg Businessweek

To Break Google's Monopoly on Search, Make Its Index Public

The tech giant doesn't have to be dismantled. Sharing its crown jewel might reshape the internet.

By
Robert Epstein
July 15, 2019, 3:00 AM PDT

PHOTO ILLUSTRATION: 731; PHOTO: GETTY IMAGES

1

Recognition is growing worldwide that something big needs to be done about Big Tech, and fast.

More than $8 billion in fines have been levied against Google by the European Union since 2017. Facebook Inc., facing an onslaught of investigations, has dropped in reputation to almost rock bottom among the 100 most visible companies in the U.S. Former employees of Google and Facebook have warned that these companies are "ripping apart the social fabric" and can "hijack the mind."

Adding substance to the concerns, documents and videos have been leaking from Big Tech companies, supporting fears—most often expressed by conservatives—about political manipulations and even aspirations to engineer human values.

Fixes on the table include forcing the tech titans to divest themselves of some of the companies they've bought (more than 250 by Google and Facebook alone) and guaranteeing that user data are transportable.

But these and a dozen other proposals never get to the heart of the problem, and that is that Google's search engine and Facebook's

2

social network platform have value only if they are intact. Breaking up Google's search engine would give us a smattering of search engines that yield inferior results (the larger the search engine, the wider the range of results it can give you), and breaking up Facebook's platform would be like building an immensely long Berlin Wall that would splinter millions of relationships.

With those basic platforms intact, the three biggest threats that Google and Facebook pose to societies worldwide are barely affected by almost any intervention: the aggressive surveillance, the suppression of content, and the subtle manipulation of the thinking and behavior of more than 2.5 billion people.

Different tech companies pose different kinds of threats. I'm focused here on Google, which I've been studying for more than six years through both experimental research and monitoring projects. (Google is well aware of my work and not entirely happy with me. The company did not respond to requests for comment.) Google is especially worrisome because it has maintained an unopposed monopoly on search worldwide for nearly a

3

decade. It controls 92 percent of search, with the next largest competitor, Microsoft's Bing, drawing only 2.5%.

Fortunately, there is a simple way to end the company's monopoly without breaking up its search engine, and that is to turn its "index"—the mammoth and ever-growing database it maintains of internet content—into a kind of public commons.

There is precedent for this both in law and in Google's business practices. When private ownership of essential resources and services—water, electricity, telecommunications, and so on—no longer serves the public interest, governments often step in to control them. One particular government intervention is especially relevant to the Big Tech dilemma: the 1956 consent decree in the U.S. in which AT&T agreed to share all its patents with other companies free of charge. As tech investor Roger McNamee and others have pointed out, that sharing reverberated around the world, leading to a significant increase in technological competition and innovation.

4

Doesn't Google already share its index with everyone in the world? Yes, but only for single searches. I'm talking about requiring Google to share its entire index with outside entities—businesses, nonprofit organizations, even individuals—through what programmers call an application programming interface, or API.

Google already allows this kind of sharing with a chosen few, most notably a small but ingenious company called Startpage, which is based in the Netherlands. In 2009, Google granted Startpage access to its index in return for fees generated by ads placed near Startpage search results.

With access to Google's index—the most extensive in the world, by far—Startpage gives you great search results, but with a difference. Google tracks your searches and also monitors you in other ways, so it gives you personalized results. Startpage doesn't track you—it respects and guarantees your privacy—so it gives you generic results. Some people like customized results; others treasure their privacy. (You might have heard of another privacy-oriented alternative to Google.com called DuckDuckGo, which aggregates

5

information obtained from 400 other non-Google sources, including its own modest crawler.)

If entities worldwide were given unlimited access to Google's index, dozens of Startpage variants would turn up within months; within a year or two, thousands of new search platforms might emerge, each with different strengths and weaknesses. Many would target niche audiences—some small, perhaps, like high-end shoppers, and some huge, like all the world's women, and most of these platforms would do a better job of serving their constituencies than Google ever could.

These aren't just alternatives to Google, they are competitors—thousands of search platforms, each with its special focus and emphasis, each drawing on different subsets of information from Google's ever-expanding index, and each using different rules to decide how to organize the search results they display. Different platforms would likely have different business models, too, and business models that have never been tried before would quickly be tested.

6

This system replicates the competitive ecology we now have of both traditional and online media sources—newspapers, magazines, television channels, and so on—each drawing on roughly the same body of knowledge, serving niche audiences, and prioritizing information as it sees fit.

But what about those nasty filter bubbles that trap people in narrow worlds of information? Making Google's index public doesn't solve that problem, but it shrinks it to nonthreatening proportions. At the moment, it's entirely up to Google to determine which bubble you're in, which search suggestions you receive, and which search results appear at the top of the list; that's the stuff of worldwide mind control. But with thousands of search platforms vying for your attention, the power is back in your hands. You pick your platform or platforms and shift to others when they draw your attention, as they will all be trying to do continuously.

If that happens, what becomes of Google? At first, not much. It should be allowed, I believe, to retain ownership and control of its index. That will assure it continues to do a

7

great job maintaining and updating it. And even with competition looming, change will take time. Serious competitors will need months to gather resources and generate traffic. Eventually, though, Google will likely become a smaller, leaner, more diversified company, especially if some of the other proposals out there for taming Big Tech are eventually implemented. If, over time, Google wants to continue to spy on people through its search engine, it will have to work like hell to keep them. It will no longer be able to rest on its laurels, as it has for most of the past 20 years; it's going to have to hustle, and we will all benefit from its energy.

My kids think Google was the world's first search engine, but it was actually the 21st. I can remember when search was highly competitive—when Yahoo! was the big kid on the block and engines such as Ask Jeeves and Lycos were hot commodities. Founded in 1998 amid a crowded field of competitors, Google didn't begin to dominate search until 2003, by which time it still handled only about a third of searches in the U.S. Search can be competitive again—this time with a massive, authoritative, rapidly expanding index available to all parties.

The alternative is frightening. If Google retains its monopoly on search, or even if a government steps in and makes Google a public utility, the obscene power to decide what information humanity can see and how that information should be ordered will remain in the hands of a single authority. Democracy will be an illusion, human autonomy will be compromised, and competition in search—with all the innovation that implies—might never emerge. With internet penetration increasing rapidly worldwide, do we really want a single player, no matter how benign it appears to be, to control the gateway to all information?

For the system I propose to work fairly and efficiently, we'll need rules. Here are some obvious ones to think about:

Access. There might have to be limits on who can access the API. We might not want every high school hacker to be able to build his or her own search platform. On the other hand, imagine thousands of Mark Zuckerbergs battling each other to find better ways of organizing the world's information.

Speed. Google must not be allowed to throttle access to its index, especially in ways that give it a performance advantage or that favor one search platform over another.

Content. To prevent Google from engineering humanity by being selective about what content it adds to its index, all parties with API access must be able to add content.

Visibility. For people using Google to seek information about other search platforms, Google must be forbidden from driving people to itself or its affiliated platforms.

Removal. Google must be prohibited from removing content from its index. The only exception will be when a web page no longer exists. An accurate, up-to-date record of such deletions must be accessible through the API.

Logging. Google must log all visits to its index, and that log must be accessible through the API.

Fees. Low-volume external platforms (think: high school hackers) should be able to access the index free of charge. High-volume users (think: Microsoft Corp.'s Bing) should

10

pay Google nominal fees set by regulators. That gives Google another incentive for maintaining a superior index.

Can we really justify bludgeoning one of the world's biggest and most successful companies? When governments have regulated, dismembered, or, in some cases, taken ownership of private water or electricity companies, they have done so to serve the public interest, even when the company in question has developed new technologies or resources at great expense. The rationale is straightforward: You may have built the pipelines, but water is a "common" resource that belongs to everyone, as David Bollier reminded us in his seminal book, Silent Theft: The Private Plunder of Our Common Wealth.

In Google's case, it would be absurd for the company to claim ownership rights over the contents of its index for the simple reason that it gathered almost all those contents. Google scraped the content by roaming the internet, examining webpages, and copying both the address of a page and language used on that page. None of those websites or any external authority ever gave Google permission to do this copying.

Did any external authority give Google permission to demote a website in its search results or to remove a website from its index? No, which is why both individuals and even top business leaders are sometimes traumatized when Google demotes or delists a website.

But when Google's index becomes public, people won't care as much about its machinations. If conservatives think Google is messing with them, they'll soon switch to other search platforms, where they'll still get potentially excellent results. Given the possibility of a mass migration, Google will likely stop playing God, treating users and constituencies with new respect and humility.

Who will implement this plan? In the U.S., Congress, the Federal Trade Commission, and the Department of Justice all have the power to make this happen. Because Google is a global company with, at this writing, 16 data centers—eight in the U.S., one in Chile, five in the EU, one in Taiwan, and one in Singapore—countries outside the U.S. could also declare its index to be a public commons. The EU is a prime candidate for taking such action.

But there is another possibility—namely, that Google itself will step up. This isn't as crazy as you might think. Likely prompted by the EU antitrust investigations, the company has quietly gone through two corporate reorganizations since 2015, and experts I've talked to in both the U.S. and the U.K. say the main effect of these reorganizations has been to distance Google's major shareholders from any calamities that might befall the Google search engine. The company's lawyers have also undoubtedly been taking a close look at the turbulent years during which Microsoft unsuccessfully fought U.S. antitrust investigators.

Google's leaders have been preparing for an uncertain future in which the search engine might be made a public utility, fined into bankruptcy, frozen by court orders, or even seized by governments. It might be able to avoid ugly scenarios simply by posting the specs for its new public API and inviting people and companies around the world to compete with its search platform. Google

13

could do this tomorrow—and generate glowing headlines worldwide. Google's data analysts know how to run numbers better than anyone. If the models predict that the company will make more money, minimize risk, and optimize its brand in coming years by making its index public, Google will make this happen long before the roof caves in.

Epstein (@DrREpstein), a former editor-in-chief of Psychology Today, is senior research psychologist at the American Institute for Behavioral Research and Technology. He has published 15 books and more than 300 articles on AI and other topics.

APPENDIX IV

The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections

Robert Epstein[1] and Ronald E. Robertson

American Institute for Behavioral Research and Technology, Vista, CA 92084

Edited by Jacob N. Shapiro, Princeton University, Princeton, NJ, and accepted by the Editorial Board July 8, 2015 (received for review October 16, 2014)

Internet search rankings have a significant impact on consumer choices, mainly because users trust and choose higher-ranked results more than lower-ranked results. Given the apparent power of search rankings, we asked whether they could be manipulated to alter the preferences of undecided voters in democratic elections. Here we report the results of five relevant double-blind, randomized controlled experiments, using a total of 4,556 undecided voters representing diverse demographic characteristics of the voting populations of the United States and India. The fifth experiment is especially notable in that it was conducted with eligible voters throughout India in the midst of India's 2014 Lok Sabha elections just before the final votes were cast. The results of these experiments demonstrate that (*i*) biased search rankings can shift the voting preferences of undecided voters by 20% or more, (*ii*) the shift can be much higher in some demographic groups, and (*iii*) search ranking bias can be masked so that people show no awareness of the manipulation. We call this type of influence, which might be applicable to a variety of attitudes and beliefs, the search engine manipulation effect. Given that many elections are won by small margins, our results suggest that a search engine company has the power to influence the results of a substantial number of elections with impunity. The impact of such manipulations would be especially large in countries dominated by a single search engine company.

search engine manipulation effect | search rankings | Internet influence | voter manipulation | digital bandwagon effect

Recent research has demonstrated that the rankings of search results provided by search engine companies have a dramatic impact on consumer attitudes, preferences, and behavior (1–12); this is presumably why North American companies now spend more than 20 billion US dollars annually on efforts to place results at the top of rankings (13, 14). Studies using eye-tracking technology have shown that people generally scan search engine results in the order in which the results appear and then fixate on the results that rank highest, even when lower-ranked results are more relevant to their search (1–5). Higher-ranked links also draw more clicks, and consequently people spend more time on Web pages associated with higher-ranked search results (1–9). A recent analysis of ∼300 million clicks on one search engine found that 91.5% of those clicks were on the first page of search results, with 32.5% on the first result and 17.6% on the second (7). The study also reported that the bottom item on the first page of results drew 140% more clicks than the first item on the second page (7). These phenomena occur apparently because people trust search engine companies to assign higher ranks to the results best suited to their needs (1–4, 11), even though users generally have no idea how results get ranked (15).

Why do search rankings elicit such consistent browsing behavior? Part of the answer lies in the basic design of a search engine results page: the list. For more than a century, research has shown that an item's position on a list has a powerful and persuasive impact on subjects' recollection and evaluation of that item (16–18). Specific order effects, such as primacy and recency, show that the first and last items presented on a list, respectively, are more likely to be recalled than items in the middle (16, 17).

Primacy effects in particular have been shown to have a favorable influence on the formation of attitudes and beliefs (18–20), enhance perceptions of corporate performance (21), improve ratings of items on a survey (22–24), and increase purchasing behavior (25). More troubling, however, is the finding that primacy effects have a significant impact on voting behavior, resulting in more votes for the candidate whose name is listed first on a ballot (26–32). In one recent experimental study, primacy accounted for a 15% gain in votes for the candidate listed first (30). Although primacy effects have been shown to extend to hyperlink clicking behavior in online environments (33–35), no study that we are aware of has yet examined whether the deliberate manipulation of search engine rankings can be leveraged as a form of persuasive technology in elections. Given the power of order effects and the impact that search rankings have on consumer attitudes and behavior, we asked whether the deliberate manipulation of search rankings pertinent to candidates in political elections could alter the attitudes, beliefs, and behavior of undecided voters.

It is already well established that biased media sources such as newspapers (36–38), political polls (39), and television (40) sway voters (41, 42). A 2007 study by DellaVigna and Kaplan found, for example, that whenever the conservative-leaning Fox television network moved into a new market in the United States, conservative votes increased, a phenomenon they labeled the Fox News Effect (40). These researchers estimated that biased coverage by Fox News was sufficient to shift 10,757 votes in Florida during the 2000 US Presidential election: more than enough to flip the deciding state in the election, which was carried by the Republican presidential candidate by only 537 votes. The Fox News Effect was also found to be smaller in television markets that were more competitive.

We believe, however, that the impact of biased search rankings on voter preferences is potentially much greater than the influence of traditional media sources (43), where parties compete in

Significance

We present evidence from five experiments in two countries suggesting the power and robustness of the search engine manipulation effect (SEME). Specifically, we show that (*i*) biased search rankings can shift the voting preferences of undecided voters by 20% or more, (*ii*) the shift can be much higher in some demographic groups, and (*iii*) such rankings can be masked so that people show no awareness of the manipulation. Knowing the proportion of undecided voters in a population who have Internet access, along with the proportion of those voters who can be influenced using SEME, allows one to calculate the win margin below which SEME might be able to determine an election outcome.

Author contributions: R.E. and R.E.R. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.N.S. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: re@aibrt.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1419828112/-/DCSupplemental.

an open marketplace for voter allegiance. Search rankings are controlled in most countries today by a single company. If, with or without intervention by company employees, the algorithm that ranked election-related information favored one candidate over another, competing candidates would have no way of compensating for the bias. It would be as if Fox News were the only television network in the country. Biased search rankings would, in effect, be an entirely new type of social influence, and it would be occurring on an unprecedented scale. Massive experiments conducted recently by social media giant Facebook have already introduced other unprecedented types of influence made possible by the Internet. Notably, an experiment reported recently suggested that flashing "VOTE" ads to 61 million Facebook users caused more than 340,000 people to vote that day who otherwise would not have done so (44). Zittrain has pointed out that if Facebook executives chose to prompt only those people who favored a particular candidate or party, they could easily flip an election in favor of that candidate, performing a kind of "digital gerrymandering" (45).

We evaluated the potential impact of biased search rankings on voter preferences in a series of experiments with the same general design. Subjects were asked for their opinions and voting preferences both before and after they were allowed to conduct research on candidates using a mock search engine we had created for this purpose. Subjects were randomly assigned to groups in which the search results they were shown were biased in favor of one candidate or another, or, in a control condition, in favor of neither candidate. Would biased search results change the opinions and voting preferences of undecided voters, and, if so, by how much? Would some demographic groups be more vulnerable to such a manipulation? Would people be aware that they were viewing biased rankings? Finally, what impact would familiarity with the candidates have on the manipulation?

Study 1: Three Experiments in San Diego, CA

To determine the potential for voter manipulation using biased search rankings, we initially conducted three laboratory-based experiments in the United States, each using a double-blind control group design with random assignment. For each of the experiments, we recruited 102 eligible voters through newspaper and online advertisements, as well through notices in senior recreation centers, in the San Diego, CA, area.* The advertisements offered USD$25 for each subject's participation, and subjects were prescreened in an attempt to match diverse demographic characteristics of the US voting population (46).

Each of the three experiments used 30 actual search results and corresponding Web pages relating to the 2010 election to determine the prime minister of Australia. The candidates were Tony Abbott and Julia Gillard, and the order in which their names were presented was counterbalanced in all conditions. This election was used to minimize possible preexisting biases by US study participants and thus to try to guarantee that our subjects would be truly "undecided." In each experiment, subjects were randomly assigned to one of three groups: (*i*) rankings favoring Gillard (which means that higher-ranked search results linked to Web pages that portrayed Gillard as the better candidate), (*ii*) rankings favoring Abbott, or (*iii*) rankings favoring neither (Fig. 1 *A–C*). The order of these rankings was determined based on ratings of Web pages provided by three independent observers. Neither the subjects nor the research assistants who supervised them knew either the hypothesis of the experiment or the groups to which subjects were assigned.

Initially, subjects read brief biographies of the candidates and rated them 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. They were also asked how likely they would be to vote for one candidate or the other on an 11-point scale ranging from −5 to +5, as well as to indicate which of the two candidates they would vote for if the election were held that day.

The subjects then spent up to 15 min gathering more information about the candidates using a mock search engine we had created (called Kadoodle), which gave subjects access to five pages of search results with six results per page. As is usual with search engines, subjects could click on any search result to view the corresponding Web page, or they could click on numbers at the bottom of each results page to view other results pages. The same search results and Web pages were used for all subjects in each experiment; only the order of the search results was varied (Fig. 1). Subjects had the option to end the search whenever they felt they had acquired sufficient information to make a sound decision. At the conclusion of the search, subjects rated the candidates again. When their ratings were complete, subjects were asked (on their computer screens) whether anything about the search rankings they had viewed "bothered" them; they were then given an opportunity to write at length about what, if anything, had bothered them. We did not ask specifically whether the search rankings appeared to be "biased" to avoid false positives typically generated by leading or suggestive questions (47).

Regarding the ethics of our study, our manipulation could have no impact on a past election, and we were also not concerned that it could affect the outcome of future elections, because the number of subjects we recruited was small and, to our knowledge, included no Australian voters. Moreover, our study was designed so that it did not favor any one candidate, so there was no overall bias. The study presented no more than minimal risk to subjects and was approved by the Institutional Review Board (IRB) of the American Institute for Behavioral Research and Technology (AIBRT). Informed consent was obtained from all subjects.

In aggregate for the first three experiments in San Diego, CA, the demographic characteristics of our subjects (mean age, 42.5 y; SD = 18.1 y; range, 18–95 y) did not differ from characteristics of the US voting population by more than the following

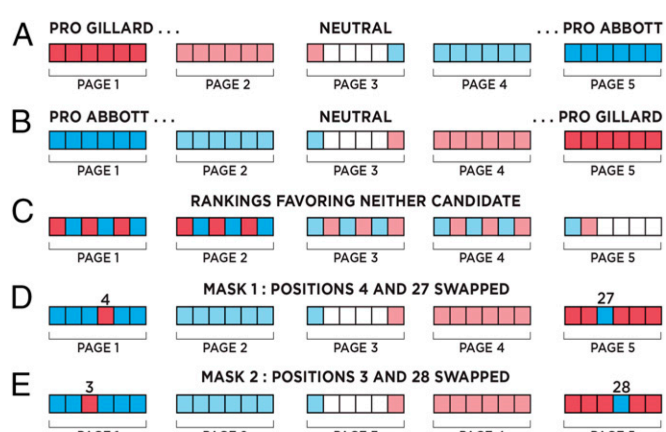


Fig. 1. Search rankings for the three experiments in study 1. (*A*) For subjects in group 1 of experiment 1, 30 search results that linked to 30 corresponding Web pages were ranked in a fixed order that favored candidate Julia Gillard, as follows: those favoring Gillard (from highest to lowest rated pages), then those favoring neither candidate, then those favoring Abbott (from lowest to highest rated pages). (*B*) For subjects in group 2 of experiment 1, the search results were displayed in precisely the opposite order so that they favored the opposing candidate, Tony Abbott. (*C*) For subjects in group 3 of experiment 1 (the control group), the ranking favored neither candidate. (*D*) For subjects in groups 1 and 2 of experiment 2, the rankings bias was masked slightly by swapping results that had originally appeared in positions 4 and 27. Thus, on the first page of search results, five of the six results—all but the one in the fourth position—favored one candidate. (*E*) For subjects in groups 1 and 2 of experiment 3, a more aggressive mask was used by swapping results that had originally appeared in positions 3 and 28.

*Although all participants claimed to be eligible voters in the prescreening, we later discovered that 6.9% of subjects marked "I don't know" and 5.2% of subjects marked "No" in response to a question asking "If you are not currently registered, are you eligible to register for elections?"

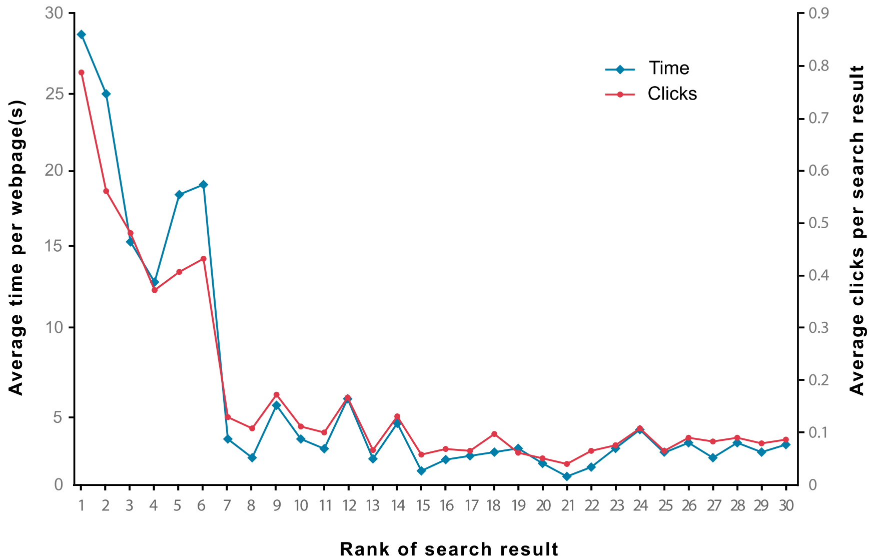


Fig. 2. Clicks on search results and time allocated to Web pages as a function of search result rank, aggregated across the three experiments in study 1. Subjects spent less time on Web pages corresponding to lower-ranked search results (blue curve) and were less likely to click on lower-ranked results (red curve). This pattern is found routinely in studies of Internet search engine use (1–12).

margins: 6.4% within any category of the age or sex measures; 14.1% within any category of the race measure; 18.7% within any category of the income or education measures; and 21.1% within any category of the employment status measure (Table S1). Subjects' political inclinations were fairly balanced, with 20.3% identifying themselves as conservative, 28.8% as moderate, 22.5% as liberal, and 28.4% as indifferent. Political party affiliation, however, was less balanced, with 21.6% identifying as Republican, 19.6% as Independent, 44.8% as Democrat, 6.2% as Libertarian, and 7.8% as other. In aggregate, subjects reported conducting an average of 7.9 searches (SD = 17.5) per day using search engines, and 52.3% reported having conducted searches to learn about political candidates. They also reported having little or no familiarity with the candidates (mean familiarity on a scale of 1–10, 1.4; SD = 0.99). On average, subjects in the first three experiments spent 635.9 s (SD = 307.0) using our mock search engine.

As expected, higher search rankings drew more clicks, and the pattern of clicks for the first three experiments correlated strongly with the pattern found in a recent analysis of ~300 million clicks [$r(13) = 0.90$, $P < 0.001$; Kolmogorov–Smirnov test of differences in distributions: $D = 0.033$, $P = 0.31$; Fig. 2] (7). In addition, subjects spent more time on Web pages associated with higher-ranked results (Fig. 2), as well as substantially more time on earlier search pages (Fig. 3).

In experiment 1, we found no significant differences among the three groups with respect to subjects' ratings of the candidates before Web research (Table S2). Following the Web research, all candidate ratings in the bias groups shifted in the predicted directions compared with candidate ratings in the control group (Table 1).

Before Web research, we found no significant differences among the three groups with respect to the proportions of people who said that they would vote for one candidate or the other if the election were held today (Table 2). Following Web research, significant differences emerged among the three groups for this measure (Table 2), and the number of subjects who said they would vote for the favored candidate in the two bias groups combined increased by 48.4% (95% CI, 30.8–66.0%; McNemar's test, $P < 0.01$).

We define the latter percentage as vote manipulation power (VMP). Thus, before the Web search, if a total of x subjects in the bias groups said they would vote for the target candidate, and if, following the Web search, a total of x' subjects in the bias groups said they would vote for the target candidate, VMP = $(x' - x)/x$. The VMP is, we believe, the key measure that an administrator would want to know if he or she were trying to manipulate an election using SEME.

Using a more sensitive measure than forced binary choice, we also asked subjects to estimate the likelihood, on an 11-point scale from −5 to +5, that they would vote for one candidate or the other if the election were held today. Before Web research, we found no significant differences among the three groups with respect to the likelihood of voting for one candidate or the other [Kruskal–Wallis (K–W) test: $\chi^2(2) = 1.384$, $P = 0.501$]. Following Web research, the likelihood of voting for either candidate in the bias groups diverged from their initial scale values by 3.71 points in the predicted directions [Mann–Whitney (M–W) test: $u = 300.5$, $P < 0.01$]. Notably, 75% of subjects in the bias groups showed no awareness of the manipulation. We counted subjects as showing awareness of the manipulation if (i) they had clicked on the box indicating that something bothered them about the rankings and (ii) we found specific terms or phrases in their open-ended comments suggesting that they were aware of bias in the rankings (*SI Text*).

In experiment 2, we sought to determine whether the proportion of subjects who were unaware of the manipulation could be increased with voter preferences still shifting in the predicted directions. We accomplished this by masking our manipulation to some extent. Specifically, the search result that had appeared in the fourth position on the first page of the search results favoring Abbott in experiment 1 was swapped with the corresponding search result favoring Gillard (Fig. 1*D*). Before Web research, we found no significant differences among the three

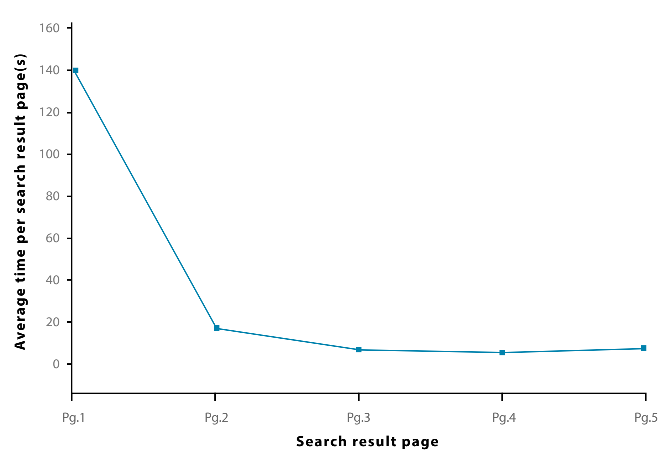


Fig. 3. Amount of time, aggregated across the three experiments in study 1, that subjects spent on each of the five search pages. Subjects spent most of their time on the first search page, a common finding in Internet search engine research (1–12).

Table 1. Postsearch shifts in voting preferences for study 1

			Mean deviation from control (SE)			
Experiment	Candidate	Rating	Gillard bias	u	Abbott bias	u
1	Gillard	Impression	1.44 (0.56)*	761.0	−1.52 (0.56)**	380.5
		Trust	1.26 (0.53)**	779.0	−1.85 (0.48)**	330.5
		Like	0.26 (0.54)	615.5	−1.73 (0.65)**	387.0
	Abbott	Impression	−2.29 (0.73)**	373.0	1.11 (0.72)**	766.5
		Trust	−2.02 (0.63)**	384.0	0.67 (0.76)	679.0
		Like	−1.55 (0.71)	460.5	1.17 (0.64)*	733.0
2	Gillard	Impression	0.97 (0.65)	704.0	−2.38 (0.79)***	325.0
		Trust	0.94 (0.72)	691.5	−2.17 (0.74)**	332.5
		Like	0.55 (0.76)	639.5	−1.82 (0.66)**	378.0
	Abbott	Impression	−1.44 (0.81)*	395.5	1.17 (0.75)*	742.0
		Trust	−0.79 (0.81)	453.5	1.85 (0.72)**	774.5
		Like	−1.44 (0.70)*	429.0	0.64 (0.71)	690.0
3	Gillard	Impression	1.44 (0.73)*	717.5	−0.55 (0.69)	507.5
		Trust	0.47 (0.70)	620.0	−0.23 (0.56)	466.5
		Like	0.44 (0.65)	623.5	−0.41 (0.70)	528.5
	Abbott	Impression	−0.32 (0.70)	534.0	1.26 (0.60)*	750.5
		Trust	−0.73 (0.65)	498.5	1.50 (0.58)**	795.0
		Like	−0.50 (0.61)	496.0	0.88 (0.62)	681.5

*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$: Mann–Whitney u tests were conducted between the control group and each of the bias groups.

groups with respect to subjects' ratings of the candidates (Table S2). Following the Web research, all candidate ratings in the bias groups shifted in the predicted directions compared with candidate ratings in the control group (Table 1).

Before Web research, we found no significant differences among the three groups with respect to voting proportions (Table 2). Following Web research, significant differences emerged among the three groups for this measure (Table 2), and the VMP was 63.3% (95% CI, 46.1–80.6%; McNemar's test, $P < 0.001$).

For the more sensitive measure (the 11-point scale), we found no significant differences among the three groups with respect to the likelihood of voting for one candidate or the other before Web research [K-W test: $\chi^2(2) = 0.888$, $P = 0.642$]. Following Web research, the likelihood of voting for either candidate in the bias groups diverged from their initial scale values by 4.44 points in the predicted directions (M-W test: $u = 237.5$, $P < 0.001$). In addition, the proportion of people who showed no awareness of the manipulation increased from 75% in experiment 1 to 85% in experiment 2, although the difference between these percentages was not significant ($\chi^2 = 2.264$, $P = 0.07$).

In experiment 3, we sought to further increase the proportion of subjects who were unaware of the manipulation by using a more aggressive mask. Specifically, the search result that had appeared in the third position on the first page of the search results favoring Abbott in experiment 1 was swapped with the corresponding search result favoring Gillard (Fig. 1*E*). This mask is a more aggressive one because higher ranked results are viewed more and taken more seriously by people conducting searches (1–12).

Before Web research, we found no significant differences among the three groups with respect to subjects' ratings of candidates (Table S2). Following the Web research, all candidate ratings in the bias groups shifted in the predicted directions compared with candidate ratings in the control group (Table 1).

Before Web research, we found no significant differences among the three groups with respect to voting proportions (Table 2). Following Web research, significant differences did not emerge among

Table 2. Comparison of voting proportions before and after Web research by group for studies 1 and 2

			Simulated vote before Web research			Simulated vote after Web research			
Study	Experiment	Group	Gillard	Abbott	χ^2	Gillard	Abbott	χ^2	VMP
1	1	1	8	26	5.409	22	12	8.870*	48.4%**
		2	11	23		10	24		
		3	17	17		14	20		
	2	1	16	18	2.197	27	7	14.274***	63.3%***
		2	20	14		12	22		
		3	14	20		22	12		
	3	1	17	17	2.199	22	12	3.845	36.7%*
		2	21	13		15	19		
		3	15	19		15	19		
2	4	1	317	383	1.047	489	211	196.280***	37.1%***
		2	316	384		228	472		
		3	333	367		377	323		

McNemar's test was conducted to assess VMP significance. VMP, percent increase in subjects in the bias groups combined who said that they would vote for the favored candidate.

*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$: Pearson χ^2 tests were conducted among all three groups.

the three groups for this measure (Table 2); the VMP, however, was 36.7% (95% CI, 19.4–53.9%; McNemar's test, $P < 0.05$).

For the more sensitive measure (the 11-point scale), we found no significant differences among the three groups with respect to the likelihood of voting for one candidate or the other before Web research [K-W test: $\chi^2(2) = 0.624$, $P = 0.732$]. Following Web research, the likelihood of voting for either candidate in the bias groups diverged from their initial scale values by 2.62 points in the predicted directions (M-W test: $u = 297.0$, $P < 0.001$). Notably, in experiment 3, no subjects showed awareness of the rankings bias, and the difference between the proportions of subjects who appeared to be unaware of the manipulations in experiments 1 and 3 was significant ($\chi^2 = 19.429$, $P < 0.001$).

Although the findings from these first three experiments were robust, the use of small samples from one US city limited their generalizability and might even have exaggerated the effect size (48).

Study 2: Large-Scale National Online Replication of Experiment 3

To better assess the generalizability of SEME to the US population at large, we used a diverse national sample of 2,100 individuals[†] from all 50 US states (Table S1), recruited using Amazon's Mechanical Turk (mturk.com), an online subject pool that is now commonly used by behavioral researchers (49, 50). Subjects (mean age, 33.9 y; SD = 11.9 y; range, 18–81 y) were exposed to the same aggressive masking procedure we used in experiment 3 (Fig. 1*E*). Each subject was paid USD$1 for his or her participation.

Regarding ethical concerns, as in study 1, our manipulation could have no impact on a past election, and we were not concerned that it could affect the outcome of future elections. Moreover, our study was designed so that it did not favor any one candidate, so there was no overall bias. The study presented no more than minimal risk to subjects and was approved by AIBRT's IRB. Informed consent was obtained from all subjects.

Subjects' political inclinations were less balanced than those in study 1, with 19.5% of subjects identifying themselves as conservative, 24.2% as moderate, 50.2% as liberal, and 6.3% as indifferent; 16.1% of subjects identified themselves as Republican, 29.9% as Independent, 43.2% as Democrat, 8.0% as Libertarian, and 2.9% as other. Subjects reported having little or no familiarity with the candidates (mean, 1.9; SD = 1.7). As one might expect in a study using only Internet-based subjects, self-reported search engine use was higher in study 2 than in study 1 [mean searches per day, 15.3; SD = 26.3; $t(529.5)^{\ddagger} = 6.9$, $P < 0.001$], and more subjects reported having previously used a search engine to learn about political candidates (86.0%, $\chi^2 = 204.1$, $P < 0.001$). Subjects in study 2 also spent less time using our mock search engine [mean total time, 309.2 s; SD = 278.7; $t(381.9)^{\ddagger} = -17.6$, $P < 0.001$], but patterns of search result clicks and time spent on Web pages were similar to those we found in study 1 [clicks: $r(28) = 0.98$, $P < 0.001$; Web page time: $r(28) = 0.98$, $P < 0.001$] and to those routinely found in other studies (1–12).

Before Web research, we found no significant differences among the three groups with respect to subjects' ratings of the candidates (Table S3). Following the Web research, all candidate ratings in the bias groups shifted in the predicted directions compared with candidate ratings in the control group (Table 3).

Before Web research, we found no significant differences among the three groups with respect to voting proportions (Table 2). Following Web research, significant differences emerged among the three groups for this measure (Table 2), and the VMP was 37.1% (95% CI, 33.5–40.7%; McNemar's test, $P < 0.001$). Using post-stratification and weights obtained from the 2010 US Census (46) and a 2011 study from Gallup (51), which were scaled to size for age, sex, race, and education, the VMP was 36.7% (95% CI, 33.2–40.3%; McNemar's test, $P < 0.001$). When weighted using the same demographics via classical regression poststratification (52) (Table S4), the VMP was 33.5% (95% CI, 30.1–37.0%, McNemar's test, $P < 0.001$).

For the more sensitive measure (the 11-point scale), we found no significant differences among the three groups with respect to the likelihood of voting for one candidate or the other before Web research [K-W test: $\chi^2(2) = 2.790$, $P = 0.248$]. Following Web research, the likelihood of voting for either candidate in the bias groups diverged from their initial scale values by 3.03 points in the predicted directions (M-W test: $u = 1.29 \times 10^5$, $P < 0.001$). As one might expect of a more Internet-fluent sample, the proportion of subjects showing no awareness of the manipulation dropped to 91.4%.

The number of subjects in study 1 was too small to look at demographic differences. In study 2, we found substantial differences in how vulnerable different demographic groups were to SEME. Consistent with previous findings on the moderators of order effects (30–32), for example, we found that subjects reporting a low familiarity with the candidates (familiarity less than 5 on a scale from 1 to 10) were more vulnerable to SEME (VMP = 38.7%; 95% CI, 34.9–42.4%; McNemar's test, $P < 0.001$) than were subjects who reported high familiarity with the candidates (VMP = 19.3%; 95% CI, 9.1–29.5%; McNemar's test, $P < 0.05$), and this difference was significant ($\chi^2 = 8.417$, $P < 0.01$).

We found substantial differences in vulnerability to SEME among a number of different demographic groups (*SI Text*). Although the groups we examined were overlapping and somewhat arbitrary, if one were manipulating an election, information about such differences would have enormous practical value. For example, we found that self-labeled Republicans were more vulnerable to SEME (VMP = 54.4%; 95% CI, 45.2–63.5%; McNemar's test, $P < 0.001$) than were self-labeled Democrats (VMP = 37.7%; 95% CI, 32.3–43.1%; McNemar's test, $P < 0.001$) and that self-labeled divorcees were more vulnerable (VMP = 46.7%; 95% CI, 32.1–61.2%; McNemar's test, $P < 0.001$) than were self-labeled married subjects (VMP = 32.4%; 95% CI, 26.8–38.1%; McNemar's test, $P < 0.001$). Among the most vulnerable groups we identified were Moderate Republicans (VMP = 80.0%; 95% CI, 62.5–97.5%; McNemar's test, $P < 0.001$), whereas among the least vulnerable groups were people who reported a household income of $40,000 to $49,999 (VMP = 22.5%; 95% CI, 13.8–31.1%; McNemar's test, $P < 0.001$).

Notably, awareness of the manipulation not only did not nullify the effect, it seemed to enhance it, perhaps because people trust search order so much that awareness of the bias serves to confirm the superiority of the favored candidate. The VMP for people who showed no awareness of the biased search rankings ($n = 1{,}280$) was 36.3% (95% CI, 32.6–40.1%; McNemar's test, $P < 0.001$), whereas the VMP for people who showed awareness of the bias ($n = 120$) was 45.0% (95% CI, 32.4–57.6%; McNemar's test, $P < 0.001$).

Having now replicated the effect with a large and diverse sample of US subjects, we were concerned about the weaknesses associated with testing subjects on a somewhat abstract election (the election in Australia) that had taken place years before and in which subjects were unfamiliar with the candidates. In real elections, people are familiar with the candidates and are influenced, sometimes on a daily basis, by aggressive campaigning. Presumably, either of these two factors—familiarity and outside influence—could potentially minimize or negate the influence of biased search rankings on voter preferences. We therefore asked if SEME could be replicated with a large and diverse sample of real voters in the midst of a real election campaign.

Study 3: SEME Evaluated During the 2014 Lok Sabha Elections in India

In our fifth experiment, we sought to manipulate the voting preferences of undecided eligible voters in India during the 2014 national Lok Sabha elections there. This election was the largest democratic election in history, with more than 800 million eligible voters and more than 430 million votes ultimately cast. We accomplished this by randomly assigning undecided English-speaking

[†]As in study 1, although all participants claimed to be eligible voters in the prescreening, we later discovered that 4.7% of subjects marked "I don't know" and 2.6% of subjects marked "No" in response to a question asking "If you are not currently registered, are you eligible to register for elections?"

[‡]Degrees of freedom adjusted for significant inequality of variances (Welch's *t* test).

Table 3. Postsearch shifts in voting preferences for study 2

		Mean deviation from control (SE)			
Candidate	Rating	Gillard bias	u	Abbott bias	u
Gillard	Impression	0.65 (0.10)***	288,299.5	−1.25 (0.12)***	168,203.5
	Trust	0.61 (0.10)***	283,491.0	−1.21 (0.11)***	167,658.5
	Like	0.50 (0.10)***	279,967.0	−1.25 (0.11)***	166,544.0
Abbott	Impression	−0.96 (0.13)***	189,290.5	1.35 (0.12)***	326,067.0
	Trust	−1.09 (0.14)***	183,993.0	1.31 (0.12)***	318,740.5
	Like	−0.85 (0.13)***	195,088.5	0.94 (0.11)***	302,318.0

***P < 0.001: Mann–Whitney u tests were conducted between the control group and each of the bias groups.

voters throughout India who had not yet voted (recruited through print advertisements, online advertisements, and online subject pools) to one of three groups in which search rankings favored either Rahul Gandhi, Arvind Kejriwal, or Narendra Modi, the three major candidates in the election.[§]

Subjects were incentivized to participate in the study either with payments between USD$1 and USD$4 or with the promise that a donation of approximately USD$1.50 would be made to a prominent Indian charity that provides free lunches for Indian children. (At the close of the study, a donation of USD$1,457 was made to the Akshaya Patra Foundation.)

Regarding ethical concerns, because we recruited only a small number of subjects relative to the size of the Indian voting population, we were not concerned that our manipulation could affect the election's outcome. Moreover, our study was designed so that it did not favor any one candidate, so there was no overall bias. The study presented no more than minimal risk to subjects and was approved by AIBRT's IRB. Informed consent was obtained from all subjects.

The subjects (n = 2,150) were demographically diverse (Table S5), residing in 27 of 35 Indian states and union territories, and political leanings varied as follows: 13.3% identified themselves as politically right (conservative), 43.8% as center (moderate), 26.0% as left (liberal), and 16.9% as indifferent. In contrast to studies 1 and 2, subjects reported high familiarity with the political candidates (mean familiarity Gandhi, 7.9; SD = 2.5; mean familiarity Kejriwal, 7.7; SD = 2.5; mean familiarity Modi, 8.5; SD = 2.1). The full dataset for all five experiments is accessible at Dataset S1.

Subjects reported more frequent search engine use compared with subjects in studies 1 or 2 (mean searches per day, 15.7; SD = 30.1), and 71.7% of subjects reported that they had previously used a search engine to learn about political candidates. Subjects also spent less time using our mock search engine (mean total time, 277.4 s; SD = 368.3) than did subjects in studies 1 or 2. The patterns of search result clicks and time spent on Web pages in our mock search engine was similar to the patterns we found in study 1 [clicks, $r(28) = 0.96$; $P < 0.001$; Web page time, $r(28) = 0.91$; $P < 0.001$] and study 2 [clicks, $r(28) = 0.96$; $P < 0.001$; Web page time, $r(28) = 0.92$; $P < 0.001$].

Before Web research, we found one significant difference among the three groups for a rating pertaining to Kejriwal, but none for Gandhi or Modi (Table S6). Following the Web research, most of the subjects' ratings of the candidates shifted in the predicted directions (Table 4).

Before Web research, we found no significant differences among the three groups with respect to voting proportions (Table 5). Following Web research, significant differences emerged among the three groups for this measure (Table 5), and the VMP was 10.6% (95% CI, 8.3–12.8%; McNemar's test, P < 0.001). Using poststratification and weights obtained from the 2011 India Census data on literate Indians (53)—scaled to size for age, sex, and location (grouped into state or union territory)—the VMP was 9.4% (95% CI, 8.2–10.6%; McNemar's test, P < 0.001). When weighted using the same demographics via classical regression poststratification (Table S7), the VMP was 9.5% (95% CI, 8.3–10.7%; McNemar's test, P < 0.001).

To obtain a more sensitive measure of voting preference in study 3, we asked subjects to estimate the likelihood, on three separate 11-point scales from −5 to +5, that they would vote for each of the candidates if the election were held today. Before Web research, we found no significant differences among the three groups with respect to the likelihood of voting for any of the candidates (Table S6). Following Web research, significant differences emerged among the three groups with respect to the likelihood of voting for Rahul Gandhi and Arvind Kejriwal but not Narendra Modi (Table S6), and all likelihoods shifted in the predicted directions (Table 4). The proportion of subjects showing no awareness of the manipulation in experiment 5 was 99.5%.

In study 3, as in study 2, we found substantial differences in how vulnerable different demographic groups were to SEME (*SI Text*). Consistent with the findings of study 2 and previous findings on the moderators of order effects (30–32), for example, we found that subjects reporting a low familiarity with the candidates (familiarity less than 5 on a scale from 1 to 10) were more vulnerable to SEME (VMP = 13.7%; 95% CI, 4.3–23.2%; McNemar's test, P = 0.17) than were subjects who reported high familiarity with the candidates (VMP = 10.3%; 95% CI, 8.0–12.6%; McNemar's test, P < 0.001), although this difference was not significant ($\chi^2 = 0.575$, P = 0.45).

As in study 2, although the demographic groups we examined were overlapping and somewhat arbitrary, if one was manipulating an election, information about such differences would have enormous practical value. For example, we found that subjects between ages 18 and 24 were less vulnerable to SEME (VMP = 8.9%; 95% CI, 5.0–12.8%; McNemar's test, P < 0.05) than were subjects between ages 45 and 64 (VMP = 18.9%; 95% CI, 6.3–31.5%; McNemar's test, P = 0.10) and that self-labeled Christians were more vulnerable (VMP = 30.7%; 95% CI, 20.2–41.1%; McNemar's test, P < 0.001) than self-labeled Hindus (VMP = 8.7%; 95% CI, 6.3–11.1%; McNemar's test, P < 0.001). Among the most vulnerable groups we identified were unemployed males from Kerala (VMP = 72.7%; 95% CI, 46.4–99.0%; McNemar's test, P < 0.05), whereas among the least vulnerable groups were female conservatives (VMP = −11.8%; 95% CI, −29.0%–5.5%; McNemar's test, P = 0.62).

A negative VMP might suggest oppositional attitudes or an underdog effect for that group (54). No negative VMPs were found in the demographic groups examined in study 2, but it is understandable that they would be found in an election in which people are highly familiar with the candidates (study 3). As a practical matter, where a search engine company has the ability to send people customized rankings and where biased search rankings are likely to produce an oppositional response with certain voters, such rankings would probably not be sent to them. Eliminating the 2.6% of our sample (n = 56) with oppositional responses, the overall VMP in this experiment increases from 10.6% to 19.8% (95% CI, 16.8–22.8%; n = 2,094; McNemar's test: P < 0.001).

As we found in study 2, awareness of the manipulation appeared to enhance the effect rather than nullify it. The VMP for people

[§]English is one of India's two official languages, the other being Hindi.

Table 4. Postsearch shifts in voting preferences for study 3

Candidate	Rating	χ^2	Mean (SE) Gandhi bias	Kejriwal bias	Modi bias
Gandhi	Impression	3.61	−0.16 (0.06)	−0.21 (0.06)	−0.30 (0.06)
	Trust	21.19***	0.14 (0.06)	−0.04 (0.07)	−0.20 (0.06)
	Like	12.99**	−0.09 (0.07)	−0.17 (0.06)	−0.34 (0.06)
	Voting likelihood	10.79**	0.16 (0.07)	−0.04 (0.07)	−0.18 (0.07)
Kejriwal	Impression	17.75***	−0.30 (0.06)	−0.11 (0.06)	−0.39 (0.05)
	Trust	26.69***	−0.17 (0.07)	0.15 (0.06)	−0.16 (0.06)
	Like	24.74***	−0.31 (0.06)	0.05 (0.06)	−0.23 (0.06)
	Voting likelihood	13.22**	−0.03 (0.06)	0.17 (0.07)	−0.12 (0.06)
Modi	Impression	24.98***	−0.22 (0.06)	−0.21 (0.06)	0.12 (0.05)
	Trust	18.78***	−0.04 (0.06)	−0.10 (0.06)	0.23 (0.06)
	Like	16.89***	−0.16 (0.05)	−0.09 (0.06)	0.19 (0.06)
	Voting likelihood	31.07***	−0.07 (0.07)	−0.10 (0.06)	0.33 (0.06)

**$P < 0.01$ and ***$P < 0.001$: for each rating, a Kruskal–Wallis χ^2 test was used to assess significance of group differences.

who showed no awareness of the biased search rankings (n = 2,140) was 10.5% (95% CI, 8.3–12.7%; McNemar's test, $P <$ 0.001), whereas the VMP for people who showed awareness of the bias (n = 10) was 33.3%.

The rankings and Web pages we used in study 3 were selected by the investigators based on our limited understanding of Indian politics and perspectives. To optimize the rankings, midway through the election process we hired a native consultant who was familiar with the issues and perspectives pertinent to undecided voters in the 2014 Lok Sabha Election. Based on the recommendations of the consultant, we made slight changes to our rankings on 30 April, 2014. In the preoptimized rankings group (n = 1,259), the VMP was 9.5% (95% CI, 6.8–12.2%; McNemar's test, $P <$ 0.001); in the postoptimized rankings group (n = 891), the VMP increased to 12.3% (95% CI, 8.5–16.1%; McNemar's test, $P <$ 0.001). Eliminating the 3.1% of the subjects in the postoptimization sample with oppositional responses (n = 28), the VMP increased to 24.5% (95% CI, 19.3–29.8%; n = 863).

Discussion

Elections are often won by small vote margins. Fifty percent of US presidential elections were won by vote margins under 7.6%, and 25% of US senatorial elections in 2012 were won by vote margins under 6.0% (55, 56). In close elections, undecided voters can make all of the difference, which is why enormous resources are often focused on those voters in the days before the election (57, 58). Because search rankings biased toward one candidate can apparently sway the voting preferences of undecided voters without their awareness and, at least under some circumstances, without any possible competition from opposing candidates, SEME appears to be an especially powerful tool for manipulating elections. The Australian election used in studies 1 and 2 was won by a margin of only 0.24% and perhaps could easily have been turned by such a manipulation. The Fox News Effect, which is small compared with SEME, is believed to have shifted between 0.4% and 0.7% of votes to conservative candidates: more than enough, according to the researchers, to have had a "decisive" effect on a number of close elections in 2000 (40).

Political scientists have identified two of the most common methods political candidates use to try to win elections. The core voter model describes a strategy in which resources are devoted to mobilizing supporters to vote (59). As noted earlier, Zittrain recently pointed out that a company such as Facebook could mobilize core voters to vote on election day by sending "get-out-and-vote" messages en masse to supporters of only one candidate. Such a manipulation could be used undetectably to flip an election in what might be considered a sort of digital gerrymandering (44, 45). In contrast, the swing voter model describes a strategy in which candidates target their resources toward persuasion—attempting to change the voting preferences of undecided voters (60). SEME is an ideal method for influencing such voters.

Although relatively few voters have actively sought political information about candidates in the past (61), the ease of obtaining information over the Internet appears to be changing that: 73% of online adults used the Internet for campaign-related purposes during the 2010 US midterm elections (61), and 55% of all registered voters went online to watch videos related to the 2012 US election campaign (62). Moreover, 84% of registered voters in the United States were Internet users in 2012 (62). In our nationwide study in the United States (study 2), 86.0% of our subjects reported having used search engines to get information about candidates. Meanwhile, the number of people worldwide with Internet access is increasing rapidly, predicted to increase to nearly 4 billion by 2018 (63). By 2018, Internet access in India is expected to rise from the 213 million users who had access in 2013 to 526 million (63). Worldwide, it is reasonable to conjecture that both proportions will increase substantially in future years; that is, more people will have Internet access, and more people will obtain information about candidates from the Internet. In the context of the experiments we have presented, this suggests that whatever the effect sizes we have observed now, they will likely be larger in the future.

Table 5. Comparison of voting proportions before and after Web research for study 3

Group	Simulated vote before Web research: Gandhi	Kejriwal	Modi	χ^2	Simulated vote after Web research: Gandhi	Kejriwal	Modi	χ^2	VMP
1	115	164	430	3.070	144	152	413	16.935**	10.6%***
2	112	183	393		113	199	376		
3	127	196	430		117	174	462		

McNemar's test was conducted to assess VMP significance. VMP, percent increase in subjects in the bias groups combined who said that they would vote for the favored candidate.

**$P < 0.01$; and ***$P < 0.001$: Pearson χ^2 tests were conducted among all three groups.

The power of SEME to affect elections in a two-person race can be roughly estimated by making a small number of fairly conservative assumptions. Where i is the proportion of voters with Internet access, u is the proportion of those voters who are undecided, and VMP, as noted above, is the proportion of those undecided voters who can be swayed by SEME, W—the maximum win margin controllable by SEME—can be estimated by the following formula: $W = i * u * VMP$.

In a three-person race, W will vary between 75% and 100% of its value in a two-person race, depending on how the votes are distributed between the two losing candidates. (Derivations of formulas in the two-candidate and three-candidate cases are available in *SI Text*.) In both cases, the size of the population is irrelevant.

Knowing the values for i and u for a given election, along with the projected win margin, the minimum VMP needed to put one candidate ahead can be calculated (Table S8). In theory, continuous online polling would allow search rankings to be optimized continuously to increase the value of VMP until, in some instances, it could conceivably guarantee an election's outcome, much as "conversion" and "click-through" rates are now optimized continuously in Internet marketing (64).

For example, if (i) 80% of eligible voters had Internet access, (ii) 10% of those individuals were undecided at some point, and (iii) SEME could be used to increase the number of people in the undecided group who were inclined to vote for the target candidate by 25%, that would be enough to control the outcome of an election in which the expected win margin was as high as 2%. If SEME were applied strategically and repeatedly over a period of weeks or months to increase the VMP, and if, in some locales and situations, i and u were larger than in the example given, the controllable win margin would be larger. That possibility notwithstanding, because nearly 25% of national elections worldwide are typically won by margins under 3%,[¶] SEME could conceivably impact a substantial number of elections today even with fairly low values of i, u, and VMP.

Given our procedures, however, we cannot rule out the possibility that SEME produces only a transient effect, which would limit its value in election manipulation. Laboratory manipulations of preferences and attitudes often impact subjects for only a short time, sometimes just hours (65). That said, if search rankings were being manipulated with the intent of altering the outcome of a real election, people would presumably be exposed to biased rankings repeatedly over a period of weeks or months. We produced substantial changes not only in voting preferences but in multiple ratings of attitudes toward candidates given just one exposure to search rankings linking to Web pages favoring one candidate, with average search times in the 277- to 635-s range. Given hundreds or thousands of exposures of this sort, we speculate not only that the resulting attitudes and preferences would be stable, but that they would become stronger over time, much as brand preferences become stronger when advertisements are presented repeatedly (66).

Our results also suggest that it is a relatively simple matter to mask the bias in search rankings so that it is undetectable to virtually every user. In experiment 3, using only a simple mask, none of our subjects appeared to be aware that they were seeing biased rankings, and in our India study, only 0.5% of our subjects appeared to notice the bias. When people are subjected to forms of influence they can identify—in campaigns, that means speeches, billboards, television commercials, and so on—they can defend themselves fairly easily if they have opposing views. Invisible sources of influence can be harder to defend against (67–69), and for people who are impressionable, invisible sources of influence not only persuade, they also leave people feeling that they made up their own minds—that no external force was applied (70, 71). Influence is sometimes undetectable because key stimuli act subliminally (72–74), but search results and Web pages are easy to perceive; it is the pattern of rankings that people cannot see. This invisibility makes SEME especially dangerous as a means of control, not just of voting behavior but perhaps of a wide variety of attitudes, beliefs, and behavior. Ironically, and consistent with the findings of other researchers, we found that even those subjects who showed awareness of the biased rankings were still impacted by them in the predicted directions (75).

One weakness in our studies was the manner in which we chose to determine whether subjects were aware of bias in the search rankings. As noted, to not generate false-positive responses, we avoided asking leading questions that referred specifically to bias; rather, we asked a rather vague question about whether anything had bothered subjects about the search rankings, and we then gave subjects an opportunity to type out the details of their concerns. In so doing, we probably underestimated the number of detections (47), and this is a matter that should be studied further. That said, because people who showed awareness of the bias were still vulnerable to our manipulation, people who use SEME to manipulate real elections might not be concerned about detection, except, perhaps, by regulators.

Could regulators in fact detect SEME? Theoretically, by rating pages and monitoring search rankings on an ongoing basis, search ranking bias related to elections might be possible to identify and track; as a practical matter, however, we believe that biased rankings would be impossible or nearly impossible for regulators to detect. The results of studies 2 and 3 suggest that vulnerability to SEME can vary dramatically from one demographic group to another. It follows that if one were using biased search rankings to manipulate a real election, one would focus on the most vulnerable demographic groups. Indeed, if one had access to detailed online profiles of millions of individuals, which search engine companies do (76–78), one would presumably be able to identify those voters who appeared to be undecided and impressionable and focus one's efforts on those individuals only—a strategy that has long been standard in political campaigns (79–84) and continues to remain important today (85). With search engine companies becoming increasingly adept at sending users customized search rankings (76–78, 86–88), it seems likely that only customized rankings would be used to influence elections, thus making it difficult or impossible for regulators to detect a manipulation. Rankings that appear to be unbiased on the regulators' screens might be highly biased on the screens of select individuals.

Even if a statistical analysis did show that rankings consistently favored one candidate over another, those rankings could always be attributed to algorithm-guided dynamics driven by market forces—so-called "organic" forces (89)—rather than by deliberate manipulation by search engine company employees. This possibility suggests yet another potential danger of SEME. What if election-related search results are indeed being left to the vagaries of market forces? Do such forces end up pushing some candidates to the top of search rankings? If so, it seems likely that those high rankings are cultivating additional supporters for those candidates in a kind of digital bandwagon effect. In other words, for several years now and with greater impact each year (as more people get election-related information through the Internet), SEME has perhaps already been affecting the outcomes of close elections. To put this another way, without human intention or direction, algorithms have perhaps been having a say in selecting our leaders.

Because search rankings are based, at least in part, on the popularity of Web sites (90), it is likely that voter preferences impact those rankings to some extent. Given our findings that search rankings can in turn affect voter preferences, these phenomena might interact synergistically, causing a substantial increase in support for one candidate at some point even when the effects of the individual phenomena are small.[||]

Our studies produced a wide range of VMPs. In a real election, what proportion of undecided voters could actually be

[¶]Some of the data applied in the analysis in this publication are based on material from the "European Election Database." The data are collected from original sources and prepared and made available by the Norwegian Social Science Data Services (NSD). NSD is not responsible for the analyses/interpretation of the data presented here.

[||]A mathematical model we developed—highly conjectural, we admit, and at this point unverifiable—shows the possible dynamics of such synergy (Fig. S1).

shifted using SEME? Our first two studies, which relied on a campaign and candidates that were unfamiliar to our subjects, produced overall VMPs in the range 36.7–63.3%, with demographic shifts occurring with VMPs as high as 80.0%. Our third study, with real voters in the midst of a real election, produced, overall, a lower VMP: just 10.6%, with optimizing our rankings raising the VMP to 12.3% and with the elimination of a small number of oppositional subjects raising the VMP to 24.5%, which is the value we would presumably have found if our search rankings had been optimized from the start and if we had advance knowledge about oppositional groups. In the third study, VMPs in some demographic groups were as high as 72.7%. If a search engine company optimized rankings continuously and sent customized rankings only to vulnerable undecided voters, there is no telling how high the VMP could be pushed, but it would almost certainly be higher than our modest efforts could achieve. Our investigation suggests that with optimized, targeted rankings, a VMP of at least 20% should be relatively easy to achieve in real elections. Even if only 60% of a population had Internet access and only 10% of voters were undecided, that would still allow control of elections with win margins up to 1.2%—five times greater than the win margin in the 2010 race between Gillard and Abbott in Australia.

Conclusions

Given that search engine companies are currently unregulated, our results could be viewed as a cause for concern, suggesting that such companies could affect—and perhaps are already affecting—the outcomes of close elections worldwide. Restricting search ranking manipulations to voters who have been identified as undecided while also donating money to favored candidates would be an especially subtle, effective, and efficient way of wielding influence.

Although voters are subjected to a wide variety of influences during political campaigns, we believe that the manipulation of search rankings might exert a disproportionately large influence over voters for four reasons:

First, as we noted, the process by which search rankings affect voter preferences might interact synergistically with the process by which voter preferences affect search rankings, thus creating a sort of digital bandwagon effect that magnifies the potential impact of even minor search ranking manipulations.

Second, campaign influence is usually explicit, but search ranking manipulations are not. Such manipulations are difficult to detect, and most people are relatively powerless when trying to resist sources of influence they cannot see (66–68). Of greater concern in the present context, when people are unaware they are being manipulated, they tend to believe they have adopted their new thinking voluntarily (69, 70).

Third, candidates normally have equal access to voters, but this need not be the case with search engine manipulations. Because the majority of people in most democracies use a search engine provided by just one company, if that company chose to manipulate rankings to favor particular candidates or parties, opponents would have no way to counteract those manipulations. Perhaps worse still, if that company left election-related search rankings to market forces, the search algorithm itself might determine the outcomes of many close elections.

Finally, with the attention of voters shifting rapidly toward the Internet and away from traditional sources of information (12, 61, 62), the potential impact of search engine rankings on voter preferences will inevitably grow over time, as will the influence of people who have the power to control such rankings.

We conjecture, therefore, that unregulated election-related search rankings could pose a significant threat to the democratic system of government.

Materials and Methods

We used 102 subjects in each of experiments 1–3 to give us an equal number of subjects in all three groups and both counterbalancing conditions of the experiments.

Nonparametric statistical tests such as the Mann–Whitney u and the Kruskal–Wallis H are used throughout the present report because Likert scale scores, which were used in each of the studies, are ordinal.

In study 3, the procedure was identical to that of studies 1 and 2; only the Web pages and search results were different: that is, Web pages and search results were pertinent to the three leading candidates in the 2014 Lok Sabha general elections. The questions we asked subjects were also adjusted for a three-person race.

ACKNOWLEDGMENTS. We thank J. Arnett, E. Clemons, E. Fantino, S. Glenn, M. Hovell, E. Key, E. Loftus, C. McKenzie, B. Meredith, N. Metaxas, D. Moriarty, D. Peel, M. Runco, S. Stolarz-Fantino, and J. Wixted for comments; K. Robertson for image editing; V. Sharan for advice on optimizing search rankings in the India study; K. Duncan and F. Tran for assistance with data analysis; J. Hagan for technical assistance; and S. Palacios and K. Huynh for assistance in conducting the experiments in study 1. This work was supported by the American Institute for Behavioral Research and Technology, a nonpartisan, nonprofit organization.

1. Joachims T, et al. (2007) Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *Association for Computing Machinery Transactions on Information Systems* 25(2):7.
2. Pan B, et al. (2007) In Google we trust: User's decisions on rank, position, and relevance. *J Comput Mediat Commun* 12(3):801–823.
3. Guan Z, Cutrell E (2007) An eye tracking study of the effect of target rank on web search. *Proceedings of the Special Interest Group for Computer-Human Interaction Conference on Human Factors in Computing Systems* (ACM, New York), pp 417–420.
4. Lorigo L, et al. (2008) Eye tracking and online search: Lessons learned and challenges ahead. *J Am Soc Inf Sci Technol* 59(7):1041–1052.
5. Granka LA, Joachim T, Gay G (2004) Eye-tracking analysis of user behavior in www search. *Proceedings of the 27th Annual International Association for Computing Machinery Special Interest Group on Information Retrieval Conference on Research and Development in Information Retrieval* (ACM, New York, NY), pp 478–479.
6. Agichtein E, Brill E, Dumais S, Ragno R (2006) Learning user interaction models for predicting web search result preferences. *Proceedings of the 29th Annual International Association for Computing Machinery Special Interest Group on Information Retrieval Conference on Research and Development in Information Retrieval* (ACM, New York, NY), pp 3–10.
7. Chitika (2013) The value of Google result positioning. Available at perma.cc/7AGC-HTDH. Accessed June 30, 2015.
8. Optify (2011) The changing face of SERPS: Organic CTR. Available at perma.cc/KZ5X-78TL. Accessed June 30, 2015.
9. Jansen BJ, Spink A, Saracevic T (2000) Real life, real users, and real needs: A study and analysis of user queries on the web. *Inf Process Manage* 36(2):207–227.
10. Spink A, Wolfram D, Jansen BJ, Saracevic T (2001) Searching the web: The public and their queries. *J Am Soc Inf Sci Technol* 53(3):226–234.
11. Silverstein C, Marais H, Henzinger M, Moricz A (1999) Analysis of a very large web search engine query log. *Association for Computing Machinery Special Interest Group on Information Retrieval Forum* 33(1):6–12.
12. Purcell K, Brenner J, Rainie L (2012) Pew Research Center's Internet & American Life project: Search engine use 2012. Available at perma.cc/NF6C-JCPW. Accessed June 30, 2015.
13. Learmonth M (2010) What big brands are spending on Google. Available at perma.cc/5L3B-SPTX. Accessed June 30, 2015.
14. Econsultancy (2012) SEMPO state of search marketing report 2012. Available at perma.cc/A3KH-6QVB. Accessed June 30, 2015.
15. Gerhart S (2004) Do Web search engines suppress controversy? *First Monday* 9(1):1111.
16. Ebbinghaus H (1913) *Memory: A Contribution to Experimental Psychology (No. 3)* (Teachers College, Columbia Univ, New York).
17. Murdock BB (1962) The serial position effect of free recall. *J Exp Psychol* 64(5):482–488.
18. Asch SE (1946) Forming impressions of personality. *J Abnorm Psychol* 41:258–290.
19. Lund FH (1925) The psychology of belief. *J Abnorm Soc Psych* 20(1):63–81.
20. Hovland CI (1957) *The Order of Presentation in Persuasion* (Yale Univ Press, New Haven, CT).
21. Baird JE, Zelin RC (2000) The effects of information ordering on investor perceptions: An experiment utilizing presidents' letters. *J Financial Strategic Decisions* 13(3):71–80.
22. Tourangeau R, Couper MP, Conrad FG (2013) "Up means good" The effect of screen position on evaluative ratings in web surveys. *Public Opin Q* 77(1, Suppl 1):69–88.
23. Krosnick JA, Alwin DF (1987) An evaluation of a cognitive theory of response order effects in survey measurement. *Pub Op Quarterly* 51(2):201–219.
24. Stern MJ, Dillman DA, Smyth JD (2007) Visual design, order effects, and respondent characteristics in a self-administered survey. *Surv Res Methods* 1(3):121–138.
25. Miller JE (1980) *Menu Pricing and Strategy* (CBI Publishing, Boston).
26. Ho DE, Imai K (2008) Estimating causal effects of ballot order from a randomized natural experiment the California alphabet lottery, 1978–2002. *Pub Op Quarterly* 72(2):216–240.
27. Chen E, Simonovits G, Krosnick JA, Pasek J (2014) The impact of candidate name order on election outcomes in North Dakota. *Elect Stud* 35:115–122.

28. Krosnick JA, Miller JM, Tichy MP (2004) An unrecognized need for ballot reform. *Rethinking the Vote*, eds Crigler A, Just M, McCaffery E (Oxford Univ Press, New York), pp 51–74.
29. Koppell JG, Steen JA (2004) The effects of ballot position on election outcomes. *J Polit* 66(1):267–281.
30. Kim N, Krosnick JA, Casasanto D (2014) Moderators of candidate name-order effects in elections: An experiment. *Polit Psychol*, 10.1111/pops.12178.
31. Miller JM, Krosnick JA (1998) The impact of candidate name order on election outcomes. *Pub Op Quarterly* 62(3):291–330.
32. Pasek J, et al. (2014) Prevalence and moderators of the candidate name-order effect evidence from statewide general elections in California. *Pub Op Quarterly* 78(2):416–439.
33. Murphy J, Hofacker C, Mizerski R (2006) Primacy and recency effects on clicking behavior. *J Comput Mediat Commun* 11(2):522–535.
34. Ansari A, Mela CF (2003) E-customization. *J Mark Res* 40(2):131–145.
35. Drèze X, Zufryden F (2004) Measurement of online visibility and its impact on Internet traffic. *J Interact Market* 18:20–37.
36. Chiang CF, Knight BG (2011) Media bias and influence: Evidence from newspaper endorsements. *Rev Econ Stud* 78(3):795–820.
37. Druckman JN, Parkin M (2005) The impact of media bias: How editorial slant affects voters. *J Polit* 67(4):1030–1049.
38. Gerber A, Karlan DS, Bergan D (2009) Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *Am Econ J Appl Econ* 1(2):35–52.
39. Morwitz VG, Pluzinski C (1996) Do pools reflect opinions or do opinions reflect polls? The impact of political polling on voters' expectation, preferences, and behavior. *J Consum Res* 23(1):53–67.
40. DellaVigna S, Kaplan E (2007) The Fox News effect: Media bias and voting. *Q J Econ* 122(3):1187–1234.
41. Lakoff G (1987) *Women, Fire and Dangerous Things* (Univ of Chicago Press, Chicago).
42. DeMarzo PM, Vayanos D, Zwiebel J (2003) Persuasion bias, social influence, and unidimensional opinions. *Q J Econ* 118(3):909–968.
43. Fogg BJ (2002) *Persuasive Technology: Using Computers to Change What We Think and Do* (Morgan Kaufmann, San Francisco).
44. Bond RM, et al. (2012) A 61-million-person experiment in social influence and political mobilization. *Nature* 489(7415):295–298.
45. Zittrain J (2014) Engineering an election. Digital gerrymandering poses a threat to democracy. *Harvard Law Review Forum* 127(8):335–341.
46. US Census Bureau (2011) Voting and registration in the election of November 2010. Available at perma.cc/DU9N-C6XT. Accessed June 30, 2015.
47. Loftus EF (1996) *Eyewitness Testimony* (Harvard Univ Press, Boston).
48. Kühberger A, Fritz A, Scherndl T (2014) Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS One* 9(9):e105825.
49. Paolacci G, Chandler J (2014) Inside the Turk: Understanding mechanical Turk as a participant pool. *Curr Dir Psychol Sci* 23(3):184–188.
50. Berinsky AJ, Huber GA, Lenz GS (2012) Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Polit Anal* 20(3):351–368.
51. Gates GJ (2011) How many people are lesbian, gay, bisexual and transgender? Available at perma.cc/N7P2-TXHV. Accessed June 30, 2015.
52. Gelman A, Hill J (2006) *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge Univ Press, Boston).
53. Census of India (2011) Population enumeration data. Available at perma.cc/4JKF-UWQZ. Accessed June 30, 2015.
54. Simon HA (1954) Bandwagon and underdog effects and the possibility of election predictions. *Pub Op Quarterly* 18(3):245–253.
55. Leip D (2012) Dave Leip's atlas of U.S. presidential elections. Available at perma.cc/G9YF-6CPE. Accessed June 30, 2015.
56. Rogers D, Cage F (2012) US house and senate elections 2012. Available at perma.cc/5QSK-FNCL. Accessed June 30, 2015.
57. Cabletelevision Advertising Bureau (2012) CAB undecided voters study 2012. Available at perma.cc/T82K-YC38. Accessed June 30, 2015.
58. Rutenberg J (2012) Obama campaign endgame: Grunt work and cold math. Available at perma.cc/X74D-YNM5. Accessed June 30, 2015.
59. Cox GW, McCubbins MD (1986) Electoral politics as a redistributive game. *J Polit* 48(2): 370–389.
60. Lindbeck A, Weibull JW (1987) Balanced-budget redistribution as the outcome of political competition. *Public Choice* 52(3):273–297.
61. Smith A (2011) Pew Research Center's Internet & American Life Project: The Internet and campaign 2010. Available at perma.cc/3U9R-EUQQ. Accessed June 30, 2015.
62. Smith A, Duggan M (2012) Pew Research Center's Internet & American life project: Online political videos and campaign 2012. Available at perma.cc/5UX2-LELJ. Accessed June 30, 2015.
63. Cisco (2014) VNI Global IP Traffic Forecast, 2013–2018. Available at perma.cc/AE9R-JSHL. Accessed June 30, 2015.
64. Ash T, Ginty M, Page R (2012) *Landing Page Optimization: The Definitive Guide to Testing and Tuning for Conversions* (John Wiley & Sons, Indianapolis).
65. Kraus SJ (1995) Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. *Pers Soc Psychol Bull* 21(1):58–75.
66. Zajonc RB (2001) Mere exposure: A gateway to the subliminal. *Curr Dir Psychol Sci* 10(6):224–228.
67. Berger J, Fitzsimons G (2008) Dogs on the street, pumas on your feet: How cues in the environment influence product evaluation and choice. *J Mark Res* 45(1):1–14.
68. Brasel SA, Gips J (2011) Red bull "gives you wings" for better or worse: A double-edged impact of brand exposure on consumer performance. *J Consum Psychol* 21(1):57–64.
69. Fransen ML, Fennis BM, Pruyn ATH, Das E (2008) Rest in peace? Brand-induced mortality salience and consumer behavior. *J Bus Res* 64(10):1053–1061.
70. Bargh JA, Gollwitzer PM, Lee-Chai A, Barndollar K, Trötschel R (2001) The automated will: Nonconscious activation and pursuit of behavioral goals. *J Pers Soc Psychol* 81(6): 1014–1027.
71. Pronin E, Kugler MB (2007) Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *J Exp Soc Psychol* 43(4):565–578.
72. Winkielman P, Berridge KC, Wilbarger JL (2005) Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Pers Soc Psychol Bull* 31(1):121–135.
73. Karremans JC, Stroebe W, Claus J (2006) Beyond Vicary's fantasies: The impact of subliminal priming and brand choice. *J Exp Soc Psychol* 42(6):792–798.
74. Légal JB, Chappé J, Coiffard V, Villard-Forest A (2012) Don't you know that you want to trust me? Subliminal goal priming and persuasion. *J Exp Soc Psychol* 48(1):358–360.
75. Mills J, Aronson E (1965) Opinion change as a function of the communicator's attractiveness and desire to influence. *J Pers Soc Psychol* 1(2):173–177.
76. Bi B, Shokouhi M, Kosinski M, Graepel T (2013) Inferring the demographics of search users: Social data meets search queries. *Proceedings of the 22nd International Conference on World Wide Web* (International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland), pp 131–140.
77. Yan J, et al. (2009, April). How much can behavioral targeting help online advertising? *Proceedings of the 18th International Conference on World Wide Web* (ACM, New York), pp 261–270.
78. Weber I, Castillo C (2010) The demographics of web search. *Proceedings of the 33rd Annual International Association for Computing Machinery Special Interest Group on Information Retrieval Conference on Research and Development in Information Retrieval* (ACM, New York), pp 523–530.
79. Mayer WG (2008) *The Swing Voter in American Politics* (Brookings Institution Press, Washington, DC).
80. Cox GW (2010) Swing voters, core voters, and distributive politics. *Political Representation*, eds Shapiro I, Stokes SC, Wood EJ, Kirshner AS (Cambridge Univ Press, Cambridge, MA), pp 342–357.
81. Bickers KN, Stein RM (1996) The electoral dynamics of the federal pork barrel. *Am J Pol Sci* 40(4):1300–1325.
82. Dahlberg M, Johansson E (2002) On the vote-purchasing behavior of incumbent governments. *Am Polit Sci Rev* 96(1):27–40.
83. Stein RM, Bickers KN (1994) Congressional Elections and the Pork Barrel. *J Polit* 56(2): 377–399.
84. Stokes SC (2005) Perverse accountability: A formal model of machine politics with evidence from Argentina. *Am Polit Sci Rev* 99(3):315–325.
85. Mayer WG (2012) The disappearing–but still important–swing voter. *Forum* 10(3): 1520.
86. Agichtein E, Brill E, Dumais S (2006) Improving web search ranking by incorporating user behavior information. *Proceedings of the 29th Annual International Association for Computing Machinery Special Interest Group on Information Retrieval Conference on Research and Development in Information Retrieval* (ACM, New York), pp 19–26.
87. Ahmed A, Aly M, Das A, Smola AJ, Anastasakos T (2012) Web-scale multi-task feature selection for behavioral targeting. *Proceedings of the 21st Association for Computing Machinery International Conference on Information and Knowledge Management* (ACM, New York), pp 1737–1741.
88. Chen Y, Pavlov D, Canny JF (2009) Large-scale behavioral targeting. *Proceedings of the 15th Association for Computing Machinery Special Interest Group on Knowledge Discovery in Data International Conference on Knowledge Discovery and Data Mining* (ACM, New York), pp 209–218.
89. Höchstötter N, Lewandowski D (2009) What users see—Structures in search engine results pages. *Inf Sci* 179(12):1796–1812.
90. Page L, Brin S, Motwani R, Winograd T (1999) *The PageRank Citation Ranking: Bringing Order to the Web. Technical Report* (Stanford InfoLab, Palo Alto, CA).

Supporting Information

Epstein and Robertson 10.1073/pnas.1419828112

SI Text

Demographic Differences in VMP. In study 2, we found substantial differences in how vulnerable different demographic groups were to SEME. Although the groups we examined are somewhat arbitrary, overlapping, and by no means definitive, they do establish a range of vulnerability to SEME. Ten groups ($n \geq 50$) that appeared to be highly vulnerable in study 2, as indicated by their VMP scores, were, in order from highest to lowest, as follows:

i) Moderate Republicans (80.0%; 95% CI, 62.5–97.5%)
ii) People from North Carolina (66.7%; 95% CI, 42.8–90.5%)
iii) Moderate Libertarians (73.3%; 95% CI, 51–95.7%)
iv) Male Republicans (66.1%; 95% CI, 54–78.2%)
v) Female conservatives age 30 and over (67.7%; 95% CI, 52.5–82.7%)
vi) People from Virginia (60.0%; 95% CI, 38.5–81.5%)
vii) People earning between \$15,000 and \$19,999 (60.0%; 95% CI, 42.5–77.5%)
viii) Hispanics (59.4%; 95% CI, 42.4–76.4%)
ix) Independents with no political leaning (58.3%; 95% CI, 38.6–78.1%)
x) Female conservatives (54.7%; 95% CI, 41.3–68.1%)

Ten groups that appeared to show little vulnerability to SEME, as indicated by their VMP scores, were, in order from highest to lowest, as follows:

i) People from California (24.1%; 95% CI, 15.1–33.1%)
ii) Moderate independents (24.0%; 95% CI, 15.4–32.5%)
iii) Liberal independents (23.4%; 95% CI, 13.1–33.8%)
iv) People from Texas (22.9%; 95% CI, 11–34.8%)
v) Liberal Libertarians (22.7%; 95% CI, 5.2–40.2%)
vi) People earning between \$40,000 and \$49,999 (22.5%; 95% CI, 13.8–31.1%)
vii) Female independents (22.0%; 95% CI, 13.5–30.5%)
viii) Male moderates age 30 and over (19.3%; 95% CI, 9.1–29.5%)
ix) Female independent moderates (17.9%; 95% CI, 13.5–30.5%)
x) People with an uncommon political party (15.0%; 95% CI, −0.6% to 30.6%)

In study 3, as in study 2, we found substantial differences in how vulnerable different demographic groups were to SEME. Although the groups we examined are somewhat arbitrary, overlapping, and by no means definitive, they do establish a range of vulnerability to SEME. Ten groups ($n \geq 50$) that appeared to be highly vulnerable in study 3, as indicated by their VMP scores, were, in order from highest to lowest, as follows:

i) Unemployed males from Kerala (72.7%; 95% CI, 46.4–99.1%)
ii) Unemployed Christians (68.8%; 95% CI, 46.0–91.5%)
iii) Unemployed moderate males (50.0%; 95% CI, 33.2–66.8%)
iv) Moderate Christian males (47.6%; 95% CI, 26.3–69.0%)
v) Christian moderates (42.9%; 95% CI, 26.5–59.3%)
vi) Males from Kerala (40.4%; 95% CI, 26.4–54.5%)
vii) Unemployed moderates (33.3%; 95% CI, 22.0–44.7%)
viii) Male Christians (32.7%; 95% CI, 19.9–45.4%)
ix) People from Kerala (32.4%; 95% CI, 21.8–43.1%)
x) Unemployed females with no political ideology (31.6%; 95% CI, 10.7–52.5%)

Ten groups that appeared to show little vulnerability to SEME, as indicated by their VMP scores, were, in order from highest to lowest, as follows:

i) People from Tamil Nadu with no political ideology (0.0%; 95% CI, −0.01%–0.04%)
ii) Employed females with no political ideology (0.0%; 95% CI, −0.01%–0.06%)
iii) People earning between Rs 10,000 and Rs 29,999 (−3.2%; 95% CI, −7.6%–1.3%)
iv) Married people who are separated (−3.3%; 95% CI, −10.0%–3.3%)
v) People with a pre-university education (−4.3%; 95% CI, −10.5%–1.81%)
vi) Unemployed liberals (−4.3%; 95% CI, −10.5%–1.81%)
vii) Unemployed conservatives (−5.0%; 95% CI, −15.0%–5.0%)
viii) People from Gujarat (−5.9%; 95% CI, −17.8%–6.0%)
ix) Unemployed male liberals (−8.0%; 95% CI, −19.5%–3.5%)
x) Female conservatives (−11.8%; 95% CI, −29.0%–5.5%)

Bias Awareness. Subjects were counted as showing awareness of the manipulation if (*i*) they had clicked on a box indicating that something "bothered" them about the rankings and (*ii*) we found specific terms or phrases in their open-ended comments suggesting that they were aware of bias in the rankings, such as "biased," "bias," "leaning towards," "leaning toward," "leaning against," "slanted," "skewed," "favorable towards," "favorable toward," "favorable for," "favorable against," "favorable results," "favored towards," "favored toward," "favored for," "favored against," "favored results," "favor toward," "results favor," "favor Modi," "favor Kejriwal," "favor Gandhi," "negative toward," "negative for," "negative against," "all negative," "all positive," "mainly negative," "mainly positive," "nothing positive," "nothing negative," "more results for," "less results for," "most of the articles were negative," "most of the articles were positive," "pro Modi," "pro Kejriwal," "pro Gandhi," "Modi leaning," "Kejriwal leaning," "Gandhi leaning," "pro Gillard," "pro Abbott," "favor Gillard," "favor Abbott," "Gillard leaning," and "Abbott leaning."

Derivation of the Formulas for Computing *W*, the Maximum Win Margin Controllable Through SEME, in Two- and Three- Person Races.

Two-person race. Where $T =$ total number of eligible voters in a population, $i =$ proportion of T who are internet users, $u =$ proportion of i who are undecided, $p =$ proportion of u who are prone to vote for the target candidate, and VMP = proportion of p who can be shifted by SEME.

The number of votes that can be shifted by SEME is given by

$$n = T * i * u * p * \mathrm{VMP}.$$

In a two-person race, the number of votes for the candidate favored by SEME when the vote is initially evenly split is

$$\frac{T}{2} + n,$$

and the number of votes for the losing candidate is

$$\frac{T}{2} - n.$$

The vote margin in favor of the winning candidate is therefore the larger vote minus the smaller vote, or, simply: $2n$.

Therefore, the margin of voters, expressed as a proportion, that can be shifted by SEME is

$$\frac{2n}{T} = \frac{2 * T * i * u * p * \mathrm{VMP}}{T} = 2 * i * u * p * \mathrm{VMP}.$$

Because the undecided voters in a two-person race have only two voting options, the value of p before outside influence is exercised can reasonably be assumed to be 0.5.

Therefore, W can be calculated as follows:

$$W = 2 * i * u * 0.5 * \mathrm{VMP},$$

and the calculation can be simplified as follows:

$$W = i * u * \mathrm{VMP}.$$

In other words, the maximum win margin controllable by SEME in a two-person race is equal to the proportion of people who can be influenced by SEME (the VMP) times the proportion of undecided Internet voters in the population. ($i * u$).

Three-person race. Where $T =$ total number of voters in a population, $i =$ proportion of T who are internet users, $u =$ proportion of i who are undecided, $p =$ proportion of u who are prone to vote for the target candidate, and VMP = proportion of p who can be shifted by SEME.

The number of votes that can be shifted by SEME is given by

$$n = T * i * u * p * \mathrm{VMP}.$$

In a three-person race, because the winning candidate can draw votes from either of the two losing candidates, W can vary between two extremes:

i) At one extreme, one of the two losing candidates draws zero votes, in which case the formula for the two-person case (above) is applicable.

ii) At the other extreme, voting preferences are initially split three ways evenly, and the winning candidate draws votes equally from the other two. This distribution will give us the lowest possible value of W in the three-person race, as follows.

The number of votes for the candidate favored by SEME will still be

$$\frac{T}{2} + n.$$

However, because of the split, the number of votes for each of the losing candidates will now be

$$\frac{T}{2} - \frac{n}{2}.$$

The vote margin in favor of the winning candidate will therefore be the larger vote minus either of the smaller votes or, simply, $1.5n$.

Therefore, the margin of voters, expressed as a proportion, that can be shifted by SEME is

$$\frac{2n}{T} = \frac{1.5 * T * i * u * p * \mathrm{VMP}}{T} = 1.5 * i * u * p * \mathrm{VMP}.$$

Therefore, W can be calculated as follows:

$$W = 1.5 * i * u * 0.5 * \mathrm{VMP},$$

and the calculation can be simplified as follows:

$$W = 0.75 * i * u * \mathrm{VMP}.$$

Therefore, in a three-person race, W will vary between 75% and 100% of the W found in the two-person case, depending on how votes are distributed between the two losing candidates; the more even the split, the smaller the controllable win margin.

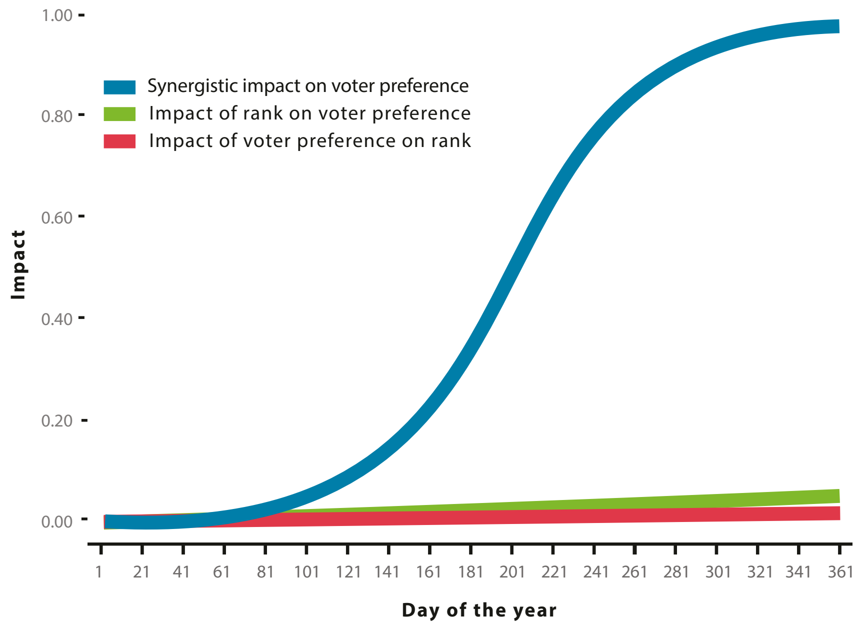


Fig. S1. A possible synergistic relationship between the impact that search rankings have on voter preferences and the impact that voter preferences have on search rankings. The lower curves (red and green) show slow increases that might occur if each of the processes acted alone over the course of a year (365 iterations of the model). The upper curve (blue) shows the result of a possible synergy between these two processes using the same parameters that generated the two lower curves. The curves are generated by an iterative model using equations of the general form $V_{n+1} = V_n + r[R_n \times (1 - V_n)] + r[O_n \times (1 - V_n)]$, where V is voter preference for one candidate, R is the impact of voter preferences on search rankings, O is the impact (randomized with each iteration) of other influences on voter preferences, and r is a rate-of-change factor. Because a change in voter preference alters the proportion of votes available, its value in the model cannot exceed 1.0.

Table S1. Demographics for studies 1 and 2

Category	Value	Census 2010†		Study 1		Census and study 1	Study 2	
		n	%	*n*	%	*Z*	*n*	%
Age	18–24	26,718	12.7%	51	16.7%	2.097*	446	21.2%
	25–44	70,472	33.4%	122	39.9%	2.385*	1,274	60.7%
	45–64	75,865	36.0%	95	31.0%	1.800	342	16.3%
	65–74	20,605	9.8%	20	6.5%	1.906	33	1.6%
	75+	17,140	8.1%	18	5.9%	1.438	5	0.2%
Race	White	152,929	72.5%	179	58.5%	5.502***	1,645	78.3%
	Black	25,632	11.8%	38	12.4%	0.349	126	6.0%
	Hispanic	21,285	9.8%	52	17.0%	4.169***	121	5.8%
	Asian	7,638	3.9%	7	2.3%	1.528	123	5.9%
	Other	3,316	2.0%	30	9.8%	10.977***	85	4.0%
Sex	Male	101,279	48.0%	162	52.9%	1.715	1,148	54.7%
	Female	109,521	52.0%	144	47.1%	1.715	947	45.1%
	Other	n/a	n/a	0	0.0%	n/a	5	0.2%
Education	Less than ninth grade	6,655	3.2%	2	0.7%	2.504*	0	0.0%
	Ninth to 12th grade	15,931	7.6%	45	14.7%	4.724***	22	1.0%
	High school graduate	65,951	31.3%	68	22.2%	3.417***	231	11.0%
	Some college or associate degree	62,655	29.7%	145	47.4%	6.753***	820	39.0%
	Bachelors	39,272	18.6%	30	9.8%	3.963***	752	35.8%
	Advanced	20,336	9.6%	16	5.2%	2.616**	275	13.1%
Used‡	Yes	126,477	60.0%	119	38.9%	7.531***	1,509	71.9%
	No	84,323	40.0%	187	61.1%	7.531***	591	28.1%
Income	Under $10,000	5,496	3.6%	67	21.9%	20.009***	137	6.5%
	$10,000 to $14,999	5,069	3.3%	33	10.8%	8.538***	131	6.2%
	$15,000 to $19,999	4,549	2.9%	28	9.2%	7.446***	124	5.9%
	$20,000 to $29,999	12,632	8.2%	45	14.7%	4.800***	282	13.4%
	$30,000 to $39,999	13,182	8.5%	34	11.1%	1.857	288	13.7%
	$40,000 to $49,999	10,807	7.0%	17	5.6%	1.143	239	11.4%
	$50,000 to $74,999	25,516	16.5%	30	9.8%	3.602***	405	19.3%
	$75,000 to $99,999	17,597	11.4%	11	3.6%	4.932***	235	11.2%
	$100,000 to $149,999	16,586	10.7%	5	1.6%	5.916***	148	7.0%
	$150,000 and over	12,102	7.8%	0	0.0%	5.893***	46	2.2%
	Prefer not to say	30,875	20.0%	36	11.8%	4.069***	65	3.1%
Marital status	Married	113,421	53.8%	48	15.7%	13.364***	751	35.8%
	Widowed	13,612	6.5%	27	8.8%	1.682	15	0.7%
	Divorced	23,035	10.9%	68	22.2%	6.324***	141	6.7%
	Separated	4,528	2.1%	15	4.9%	3.317***	33	1.6%
	Never married	56,203	26.7%	148	48.4%	8.576***	1,160	55.2%

*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$.
†Census numbers are in hundred thousands.
‡For census data, "No" includes "unemployed" and "not in labor force."

Table S2. Voting preferences by group for study 1

Experiment	Voting preferences	Mean (SE) Group 1 (Gillard bias)	Group 2 (Abbott bias)	Group 3 (control)	Kruskal–Wallis (χ^2)	Mann–Whitney *u*
1	PreImpressionAbbott	8.09 (0.34)	7.74 (0.40)	7.41 (0.26)	3.979	525.0
	PreImpressionGillard	7.06 (0.42)	7.47 (0.35)	6.88 (0.32)	1.395	529.5
	PreTrustAbbott	7.82 (0.31)	7.85 (0.39)	7.35 (0.28)	3.275	538.5
	PreTrustGillard	6.38 (0.40)	7.56 (0.30)	6.88 (0.32)	5.213	407.0
	PreLikeAbbott	6.06 (0.52)	5.68 (0.47)	5.79 (0.38)	0.296	538.5
	PreLikeGillard	5.29 (0.48)	5.76 (0.41)	5.29 (0.37)	1.335	500.0
	PostImpressionAbbott	4.24 (0.49)	7.29 (0.51)	5.85 (0.38)	19.029***	252.0***
	PostImpressionGillard	7.26 (0.45)	4.71 (0.47)	5.65 (0.46)	14.667**	286.0**
	PostTrustAbbott	4.59 (0.43)	7.32 (0.51)	6.15 (0.38)	18.385***	260.5***
	PostTrustGillard	6.91 (0.42)	4.97 (0.43)	6.15 (0.40)	10.809**	326.5**
	PostLikeAbbott	3.88 (0.43)	6.24 (0.58)	5.18 (0.42)	11.026**	341.5**
	PostLikeGillard	5.68 (0.49)	4.15 (0.45)	5.41 (0.42)	5.836	403.0*
2	PreImpressionAbbott	6.76 (0.43)	7.50 (0.34)	6.76 (0.44)	1.761	477.0
	PreImpressionGillard	6.50 (0.36)	7.29 (0.43)	6.12 (0.45)	4.369	449.5
	PreTrustAbbott	6.41 (0.44)	7.12 (0.30)	7.32 (0.44)	2.700	499.0
	PreTrustGillard	6.56 (0.41)	7.32 (0.36)	6.35 (0.43)	3.094	465.0
	PreLikeAbbott	5.56 (0.46)	5.65 (0.43)	5.76 (0.49)	0.170	575.0
	PreLikeGillard	5.79 (0.44)	5.79 (0.48)	5.47 (0.45)	0.306	568.0
	PostImpressionAbbott	3.79 (0.41)	7.15 (0.49)	5.24 (0.48)	20.878***	226.5***
	PostImpressionGillard	7.35 (0.39)	4.79 (0.47)	6.00 (0.38)	15.270***	279.5***
	PostTrustAbbott	3.82 (0.40)	7.18 (0.47)	5.53 (0.51)	21.917***	207.5***
	PostTrustGillard	7.32 (0.41)	4.97 (0.46)	6.18 (0.36)	13.410**	302.0**
	PostLikeAbbott	3.91 (0.42)	6.09 (0.53)	5.56 (0.48)	9.822**	353.0**
	PostLikeGillard	6.68 (0.45)	4.29 (0.48)	5.79 (0.40)	12.905**	311.5**
3	PreImpressionAbbott	7.24 (0.39)	7.18 (0.39)	7.88 (0.27)	1.346	568.5
	PreImpressionGillard	6.12 (0.43)	7.09 (0.39)	7.26 (0.34)	4.134	452.0
	PreTrustAbbott	7.18 (0.35)	6.41 (0.41)	7.53 (0.32)	3.837	478.0
	PreTrustGillard	6.65 (0.38)	6.68 (0.40)	6.97 (0.33)	0.259	568.5
	PreLikeAbbott	6.59 (0.42)	5.94 (0.39)	6.59 (0.43)	2.301	491.0
	PreLikeGillard	5.85 (0.46)	5.85 (0.43)	6.26 (0.41)	1.065	576.5
	PostImpressionAbbott	5.29 (0.48)	6.82 (0.41)	6.26 (0.48)	5.512	384.0*
	PostImpressionGillard	6.50 (0.45)	5.47 (0.43)	6.21 (0.48)	3.027	445.5
	PostTrustAbbott	5.38 (0.49)	6.85 (0.45)	6.47 (0.47)	5.091	399.0*
	PostTrustGillard	6.44 (0.45)	5.76 (0.47)	6.29 (0.44)	1.365	493.0
	PostLikeAbbott	5.29 (0.48)	6.03 (0.48)	5.79 (0.53)	1.129	487.0
	PostLikeGillard	6.12 (0.47)	5.26 (0.54)	6.09 (0.51)	1.475	491.5

*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$: Kruskal–Wallis tests were conducted between all three groups, and Mann–Whitney *u* tests were conducted between groups 1 and 2. Preferences were measured for each candidate separately on 10-point Likert scales.

Table S3. Voting preferences by group for study 2

Voting preferences	Mean (SE) Group 1 (Gillard bias)	Group 2 (Abbott bias)	Group 3 (control)	Kruskal–Wallis (χ^2)	Mann–Whitney *u*
PreImpressionAbbott	7.40 (0.07)	7.36 (0.08)	7.37 (0.07)	0.458	241,861.5
PreImpressionGillard	7.13 (0.07)	7.12 (0.08)	7.13 (0.07)	0.081	243,115.0
PreTrustAbbott	7.26 (0.07)	7.22 (0.08)	7.18 (0.07)	0.954	241,924.5
PreTrustGillard	6.95 (0.07)	6.89 (0.08)	6.92 (0.07)	0.222	241,779.0
PreLikeAbbott	6.42 (0.08)	6.39 (0.08)	6.23 (0.08)	2.987	243,677.5
PreLikeGillard	6.24 (0.08)	6.30 (0.08)	6.11 (0.08)	3.178	239,556.0
PostImpressionAbbott	4.61 (0.09)	6.88 (0.09)	5.53 (0.09)	289.065***	120,660.0***
PostImpressionGillard	6.87 (0.08)	4.95 (0.09)	6.21 (0.09)	237.034***	133,106.5***
PostTrustAbbott	4.56(0.10)	6.94 (0.09)	5.57 (0.10)	281.560***	121,786.5***
PostTrustGillard	6.84 (0.09)	4.95 (0.09)	6.19 (0.09)	221.709***	136,689.0***
PostLikeAbbott	4.55 (0.09)	6.31 (0.09)	5.21 (0.09)	177.225***	146,957.0***
PostLikeGillard	6.34(0.09)	4.64 (0.09)	5.71 (0.09)	176.066***	147,372.5***

***$P < 0.001$: Kruskal–Wallis tests were conducted between all three groups, and Mann–Whitney *u* tests were conducted between groups 1 and 2. Preferences were measured for each candidate separately on 10-point Likert scales.

Table S4. Treatment effect estimates for study 2 voting preferences

	Presearch vote		Postsearch vote	
Predictor variable	Coefficient	SE	Coefficient	SE
Intercept	−0.073	0.540	0.062	0.543
Sex				
Female	0	Referent	0	Referent
Male	0.039	0.110	−0.135	0.119
Other	−0.430	0.922	−0.568	0.924
Race/ethnicity				
White	0	Referent	0	Referent
Black	0.115	0.224	0.090	0.245
Hispanic	−0.435	0.235	−0.280	0.237
Asian	0.366	0.238	0.668	0.291*
Other	0.133	0.274	−0.072	0.291
Age group				
18–24	0	Referent	0	Referent
25–44	−0.024	0.144	−0.083	0.157
45–64	0.241	0.184	0.029	0.200
65+	0.258	0.411	0.685	0.519
Education level				
Less than ninth grade	0	Referent	0	Referent
Ninth to 12th grade	0.024	0.548	0.732	0.550
High school graduate	0.074	0.528	0.927	0.528
Bachelors	0.094	0.529	0.842	0.530
Advanced	−0.050	0.543	0.549	0.544

The presearch and postsearch columns report the estimate and variance for both treatment groups using classical regression poststratification. Data for sex, race/ethnicity, age group, and education level came from the 2010 US Census. Data on the number of people who identify their sex as "other" came from a 2011 Gallup study.
*$P < 0.05$.

Table S5. Demographics for study 3

Category	Value	Study 3 *n*	Study 3 %	Indian Census 2011 (literates) *n*	Indian Census 2011 (literates) %
Age	18–24	602	28.0%	160,241,457	21.0%
	25–44	1410	65.6%	347,587,712	45.6%
	45–64	124	5.8%	188,197,343	24.7%
	65+	14	0.7%	66,185,333	8.7%
Religion	Buddhism	14	0.7%	—	—
	Christianity	262	12.2%	—	—
	Hinduism	1512	70.3%	—	—
	Islam	314	14.6%	—	—
	Jainism	21	1.0%	—	—
	Other	15	0.7%	—	—
	Sikhism	12	0.6%	—	—
Sex	Male	1518	70.6%	388,428,872	51.0%
	Female	632	29.4%	373,782,973	49.0%
Education	None	0	0.0%	—	—
	Primary school	4	0.2%	—	—
	Higher secondary	71	3.3%	—	—
	Pre-university	136	6.3%	—	—
	Bachelors	1225	57.0%	—	—
	Masters	699	32.5%	—	—
	Doctorate	15	0.7%	—	—
Used	Yes	1635	76.0%	—	—
	No	515	24.0%	—	—
Income	Under Rs 10,000	121	5.6%	—	—
	Rs 10,000 to Rs 29,999	206	9.6%	—	—
	Rs 30,000 to Rs 49,999	131	6.1%	—	—
	Rs 50,000 to Rs 69,999	106	4.9%	—	—
	Rs 70,000 to Rs 89,999	146	6.8%	—	—
	Rs 90,000 to Rs 109,999	181	8.4%	—	—
	Rs 110,000 to Rs 129,999	172	8.0%	—	—
	Rs 130,000 to Rs 149,999	132	6.1%	—	—
	Rs 150,000 to Rs 169,999	124	5.8%	—	—
	Rs 170,000 to Rs 189,999	118	5.5%	—	—
	Rs 190,000 and over	486	22.6%	—	—
	I prefer not to say	227	10.6%	—	—
Marital status	Married	1,144	53.2%	—	—
	Widowed	5	0.2%	—	—
	Divorced	4	0.2%	—	—
	Separated	78	3.6%	—	—
	Never married	919	42.7%	—	—
Location	State	1,144	53.2%	749,758,470	98.4%
	Union Territory	5	0.2%	12,453,375	1.6%

Table S6. Voting Preferences by Group for Study 3

Voting preferences	Mean (SE) Group 1 (Gandhi bias)	Group 2 (Kejriwal bias)	Group 3 (Modi bias)	Kruskal–Wallis (χ^2)
PreImpressionGandhi	5.94 (0.10)	5.73 (0.10)	5.65 (0.10)	4.782
PreImpressionKejriwal	6.80 (0.09)	7.07 (0.09)	7.09 (0.08)	6.230*
PreImpressionModi	7.49 (0.10)	7.46 (0.10)	7.48 (0.09)	0.188
PreLikableGandhi	5.71 (0.10)	5.64 (0.10)	5.61 (0.10)	0.722
PreLikableKejriwal	6.68 (0.09)	6.78 (0.09)	6.87 (0.09)	2.030
PreLikableModi	7.40 (0.10)	7.29 (0.10)	7.29 (0.10)	1.483
PreTrustGandhi	5.57 (0.11)	5.52 (0.11)	5.42 (0.10)	0.955
PreTrustKejriwal	6.54 (0.10)	6.74 (0.10)	6.85 (0.09)	4.546
PreTrustModi	7.22 (0.11)	7.31 (0.11)	7.27 (0.10)	0.159
PreLikelyToVoteGandhi	0.10 (0.12)	0.08 (0.12)	0.08 (0.12)	1.587
PreLikelyToVoteKejriwal	1.19 (0.11)	1.38 (0.11)	1.55 (0.10)	5.178
PreLikelyToVoteModi	2.15 (0.12)	2.12 (0.12)	2.06 (0.12)	0.202
PostImpressionGandhi	5.78 (0.10)	5.52 (0.10)	5.35 (0.10)	9.552**
PostImpressionKejriwal	6.50 (0.09)	6.96 (0.09)	6.70 (0.08)	14.288**
PostImpressionModi	7.27 (0.10)	7.26 (0.10)	7.60 (0.09)	7.860*
PostLikableGandhi	5.62 (0.10)	5.46 (0.10)	5.26 (0.10)	6.322*
PostLikableKejriwal	6.37 (0.09)	6.84 (0.09)	6.64 (0.08)	13.456**
PostLikableModi	7.24 (0.11)	7.20 (0.11)	7.47 (0.10)	3.874
PostTrustGandhi	5.71 (0.11)	5.48 (0.10)	5.22 (0.10)	11.386*
PostTrustKejriwal	6.38 (0.10)	6.89 (0.10)	6.68 (0.08)	15.840***
PostTrustModi	7.18 (0.11)	7.20 (0.11)	7.49 (0.10)	4.758

*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$: Kruskal–Wallis tests were conducted between all three groups. Preferences were measured for each candidate separately on 10-point Likert scales.

Table S7. Treatment effect estimates for study 3 voting preferences

Predictor variable	Presearch vote Coefficient	Presearch vote SE	Postsearch vote Coefficient	Postsearch vote SE
Intercept	−0.716	0.090***	−0.552	0.088***
Sex				
Male	0	Referent	0	Referent
Female	0.168	0.100	0.030	0.099
Age group, y				
18–24	0	Referent	0	Referent
25–44	0.031	0.103	0.067	0.101
45–64	−0.222	0.217	−0.057	0.208
65+	−0.213	0.598	−0.366	0.598
Location				
State	0	Referent	0	Referent
Union Territory	−0.401	0.294	−0.321	0.279

The presearch and postsearch columns report the estimate and variance for both of the treatment groups using classical regression poststratification. Data for sex, age group, and location came from the 2011 India Census.

***$P < 0.001$.

Table S8. Minimum VMP levels needed to impact two-person races with various projected win margins and proportions of undecided Internet voters

Proportion of undecided Internet voters in the population ($i*u$)	Projected win margin									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
0.01	1.000	—	—	—	—	—	—	—	—	—
0.02	0.500	1.000	—	—	—	—	—	—	—	—
0.03	0.333	0.667	1.000	—	—	—	—	—	—	—
0.04	0.250	0.500	0.750	1.000	—	—	—	—	—	—
0.05	0.200	0.400	0.600	0.800	1.000	—	—	—	—	—
0.06	0.167	0.333	0.500	0.667	0.833	1.000	—	—	—	—
0.07	0.143	0.286	0.429	0.571	0.714	0.857	1.000	—	—	—
0.08	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000	—	—
0.09	0.111	0.222	0.333	0.444	0.556	0.667	0.778	0.889	1.000	—
0.10	0.100	0.200	0.300	0.400	0.500	0.600	0.700	0.800	0.900	1.000
0.11	0.091	0.182	0.273	0.364	0.455	0.545	0.636	0.727	0.818	0.909
0.12	0.083	0.167	0.250	0.333	0.417	0.500	0.583	0.667	0.750	0.833
0.13	0.077	0.154	0.231	0.308	0.385	0.462	0.538	0.615	0.692	0.769
0.14	0.071	0.143	0.214	0.286	0.357	0.429	0.500	0.571	0.643	0.714
0.15	0.067	0.133	0.200	0.267	0.333	0.400	0.467	0.533	0.600	0.667
0.16	0.063	0.125	0.188	0.250	0.313	0.375	0.438	0.500	0.563	0.625
0.17	0.059	0.118	0.176	0.235	0.294	0.353	0.412	0.471	0.529	0.588
0.18	0.056	0.111	0.167	0.222	0.278	0.333	0.389	0.444	0.500	0.556
0.19	0.053	0.105	0.158	0.211	0.263	0.316	0.368	0.421	0.474	0.526
0.20	0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450	0.500

Other Supporting Information Files

Dataset S1 (XLS)

APPENDIX V

How a Daily Regimen of Operant Conditioning Might Explain the Power of the Search Engine Manipulation Effect (SEME)

Robert Epstein, Michael Lothringer, and Vanessa R. Zankich
American Institute for Behavioral Research and Technology, 1035 East Vista Way, Suite 120, Vista, California 92084, United States of America

Behavior and Social Issues, in press.

Author Note

Robert Epstein https://orcid.org/0000-0002-7484-6282

Vanessa R. Zankich https://orcid.org/0000-0003-2375-6209

Correspondence concerning this article should be addressed to Robert Epstein at re@aibrt.org

This work was supported by general funds of the American Institute for Behavioral Research and Technology (AIBRT), a nonpartisan, nonprofit, 501(c)(3) organization. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Recent studies have shown that biased search results can produce substantial shifts in the opinions and voting preferences of undecided voters – a phenomenon called the "search engine manipulation effect" (SEME), one of the most powerful list effects ever discovered. We believe this is so because, unlike other list effects, SEME is supported by a daily regimen of operant conditioning. When people conduct searches for simple facts (86% of searches), the correct answer invariably turns up in the top position, which teaches users to attend to and click on high-ranking search results. As a result, when people are undecided, they tend to formulate opinions based on web pages linked to top search results. We tested this hypothesis in a controlled experiment with 551 US voters. Participants in our High-Trust group conducted routine searches in which the correct answer always appeared in the first search result. In our Low-Trust group, the correct answer could appear in any search position other than the first two. In all, participants had to answer five questions during this pre-training, and we focused our analysis on people who answered all the questions correctly ($n = 355$) – in other words, on people who were maximally impacted by the pre-training contingencies. A difference consistent with our hypothesis emerged between the groups when they were subsequently asked to search for information on political candidates. Voting preferences in the High-Trust group shifted toward the favored candidate at a higher rate (34.6%) than voting preferences in the Low-Trust group (17.1%, $p = 0.001$).

Keywords: search engines, Search Engine Manipulation Effect, SEME, search engine ranking, online manipulation, operant conditioning of online search behavior

How a Daily Regimen of Operant Conditioning Might Explain the Power of the Search Engine Manipulation Effect (SEME)

In recent years, people around the world have become increasingly dependent on search engines to obtain information, including information that helps them make decisions about complex and socially important matters, such as whom to vote for in an upcoming election (Arendt & Fawzi, 2018; Trevisan et al., 2016; Wang et al., 2017). An increasing body of evidence also shows that search results that favor one candidate, cause, or company – by which we mean that they link to web pages that make that candidate, cause, or company appear superior to competitors – can have a rapid and dramatic impact on people's opinions, purchases, and votes (Agudo & Matute, 2021; Allam et al., 2014; Epstein & Robertson, 2015, 2017; Epstein et al., 2022; Ghose et al., 2014; Joachims et al., 2007; Knobloch-Westerwick et al., 2015; Pan et al., 2007; Prinz et al., 2017; Wilhite & Houmanfar, 2015; cf. Feezell et al., 2021). In five randomized, controlled experiments with 4,556 participants in two countries, Epstein and Robertson (2015) showed that search rankings favoring one political candidate can rapidly produce dramatic shifts in the opinions and voting preferences of undecided voters, in some demographic groups producing vote margins as high as 80% after just one online search. They labeled this new form of influence the "search engine manipulation effect" (SEME) and demonstrated that these shifts can occur without people being aware that they have been manipulated. SEME has been replicated several times since 2015 (Agudo & Matute, 2021; Draws et al., 2021; Epstein et al., 2022; Eslami et al., 2017; Haas & Unkel, 2017; Knobloch-Westerwick et al., 2015; Ludolph et al., 2016; Pogacar et al., 2017; Trielli & Diakopoulos, 2019).

Moreover, since search results are ephemeral experiences (West, 2018; cf. Mckinnon & MacMillan, 2018) – fleeting, often personalized, experiences that are generated spontaneously, impact the user, and subsequently disappear without being stored anywhere – they can impact millions of users every day without leaving a paper trail for authorities to trace (Epstein, 2018a). One cannot go back in time to determine what ephemeral content people have been shown, even if one has access to the algorithm that generated that content (Hendler & Mulvehill, 2016; Paudyal & Wong, 2018; cf. Taylor, 2019).

The fact that more than 90% of searches conducted in almost every country in the world are conducted on just one search engine (Google) (StatCounter GlobalStats, n.d.) raises special concerns about SEME (Epstein, 2018a). It means that a single company – one that is unregulated, highly secretive, not accountable to the public, and that has, for all practical purposes, no competitors (Singer, 2019) – could be producing systematic changes in the thinking of billions of people every day with no way for other parties to counteract its influence, or even, for that matter, to detect and document that influence (Hazan, 2013; Ørmen, 2016; see S1 Text for additional information about bias in search results).

Why is SEME so large? It is a list effect, but it seems different, both qualitatively and quantitatively, from previously studied list effects. Researchers have been studying list effects, such as the serial position effect, for more than a century (Ebbinghaus, 2013; Mack et al., 2017; Murre & Dros, 2015), and such effects are sometimes substantive. For example, when Candidate A's name consistently appears above his or her opponent's name on a ballot – perhaps simply because the names are in alphabetical order – this tends to boost Candidate A's share of the votes by 3%–15% – an effect called the "Ballot-Order Effect" (Grant, 2017; Ho & Imai, 2008; Koppell & Steen, 2004). While counterbalancing the order of names on ballots can easily be done – even for paper ballots – it has rarely been done (Beazley, 2013).

The serial position effect itself can increase the likelihood of a word being recalled from a list; words at the beginning of a list (the primacy effect) and the end of a list (the recency effect) are usually recalled more often than words in the middle (Murdock, 1962). The ranking of content in lists can even affect juries' opinions (Anderson, 1958; Carlson & Russo, 2001), the opinions of judges in singing contests (Bruine de Bruin, 2005), and wine preferences (Mantonakis et al., 2009).

SEME might be large, at least in part, because people generally trust computer output more than they trust content in which the human hand is evident (Bogert et al., 2021; Logg et al., 2019). Most people have no idea how computers work or what an algorithm is; as a result, they are inclined to view computer-generated content as impartial or objective (Fast & Jago, 2020; Logg et al., 2018). This trust has also been driven by the positive image Big Tech companies have had for many years. That trust has been tarnished in recent years because of data breaches and other scandals (Burt, 2019; Fortune, 2020; Kramer, 2019), and leaks of documents and videos from these companies, along with reports by whistleblowers, have shown that the algorithmic output we see is frequently adjusted by employees. At Google, search results are apparently adjusted by employees at least 3,200 times a year (Google, n.d.; Meyers, 2019).

Trust in companies and trust in computer output can be driven by a number of factors – marketing and advertising, for example (Danbury et al., 2013; Sahin et al., 2011), or the fact that nearly all the services we receive from Big Tech companies appear to be free (Epstein, 2016; Nicas et al., 2019). It is not clear how SEME can be accounted for by such trust, however. How can we account for the fact that high-ranking search results are more trusted than lower-ranking results (Edelman, 2011; Marable, 2003; Pan et al., 2007)? Why is the preference for high-ranking results so strong – strong enough not only to influence purchases (Ghose et al., 2014; Joachims et al., 2007) but to have a large and almost immediate impact on opinions and voting preferences?

The preference for high-ranking search results might be due in part to what people sometimes call "laziness" or "convenience." People are busy, so, sometimes at least, they attend to and click on a high-ranking search result because doing so saves time. As one might expect, eye-tracking and other studies show that people generally attend to the first results displayed on a screen before they scroll down or click to another page (Athukorala et al., 2015; Nielsen & Pernice, 2010; Schultheiß & Lewandowski, 2020). This finding is comparable to the attention people pay (or at least used to pay) to above-the-fold content in newspapers. The limited attention span of users can be problematic for longer pages; people want information that gets to the point and are unlikely to read long web pages filled with text (Nielsen, 2010; Weinreich et al., 2008).

Convenience might contribute to some extent to the large impact of SEME, but in the present study we explore another possibility – namely, that the power of SEME derives in part from the distinctive way in which people interact with search results. In an authoritative list of the 100 most common search terms people use (Soulo, 2022), 86% of the search queries were one-to-two words long and simply directed users to simple facts or specific websites – search terms such as "news," "speed test," and "nfl scores." The correct website invariably turns up in the highest position of the search results that are generated; frequently, that same information occurs in the second or third positions, as well. Other lists of common search terms are also dominated by queries that tend to produce simple factual answers in the top position of search results (Hardwick, 2020; Siege Media, n.d.).

Because, day after day, the vast majority of search queries produce simple factual answers in the highest position of search results (Rose, 2018), we all learn, over and over again, that what is higher in the list is better or truer than what is lower in the list. To be more specific, we usually attend to and click on the highest-ranking search result because doing so is reinforced by the appearance of the correct answer to our query. Almost any reply to a verbal inquiry strengthens inquiries of that type, but a *correct* answer to an inquiry is an especially powerful reinforcer, presumably because it makes a speaker more effective (Skinner, 1957; cf. Kieta et al., 2018), and when the same source provides a series of correct answers over time, the value and potential power of those answers increases. As B. F. Skinner put it in his classic text on verbal behavior, “The extent to which the listener judges the response as true, valid, or correct is governed by the extent to which comparable responses by the same speaker have proved useful in the past” (Skinner, 1957, p. 427).

When, at some point, people finally enter an open-ended search query that either has no definitive answer (“trump”) or that seeks an opinion (“what’s the best restaurant in Denver”), they will tend both to attend to and click on high-ranking search results. We are speculating, in effect, that SEME is a large effect because it is supported by a daily regimen of operant conditioning. Although the idea that operant conditioning plays a role in voting behavior is not new (Visser, 1996), in this paper we are emphasizing a kind of operant conditioning that never stops and that people are entirely unaware of – specifically, one that reinforces attending to and clicking on high-ranking search results that appear in response to routine factual searches.

We test this hypothesis with a randomized, controlled experiment – a modified version of the experimental procedure used by Epstein and Robertson (2015) in their original SEME experiments (see S2 Text for details about the procedure). The present study added one feature to the Epstein and Robertson (2015) procedure: Before beginning the political opinion study, participants experienced a pre-training procedure that either reinforced or extinguished the tendency to attend to and click on high-ranking search results. In theory, extinguishing that tendency should (a) change the pattern of clicks that typifies search behavior, and (b) reduce the impact that statistically biased search results have on people's opinions and voting preferences.

Method

Participants

A total of 551 eligible US voters from 46 states were recruited through Amazon Mechanical Turk (MTurk, accessed through a company called Cloud Research, which screens out bots) and were paid a small fee (US$7.50) to participate. Fifty-nine point nine percent (n = 330) of participants identified themselves as female and 40.1% (n = 221) as male. The mean age was 38.3 (*SD* = 11.7). Seventy-three point nine percent (n = 407) of participants identified themselves as White, 8.2% (n = 45) as Black, 6.5% (n = 36) as Hispanic, 6.4% (n = 35) as Asian, 4.4% (n = 24) as Mixed, and 0.7% (n = 4) as Other. A majority of participants were college educated, with 55.2% (n = 304) reporting having received a bachelor's degree or higher.

Procedure

See S3 Text in our Supplementary Material for our statement of compliance with current ethical standards.

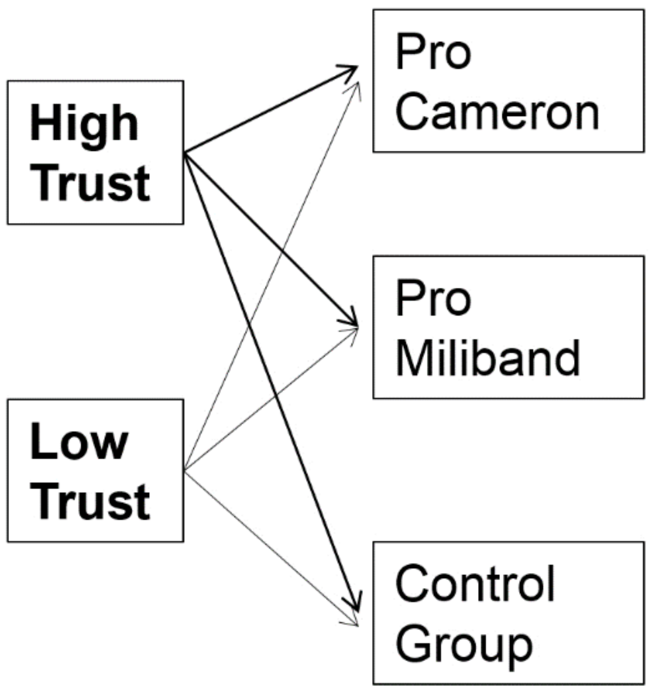
The experiment was conducted online, and participants identified themselves using their MTurk Worker IDs; we had no knowledge of their names or email addresses. Before the experiment began, participants were asked a series of demographic questions and were then given instructions about the experimental procedure (see S4 Text). In compliance with APA and HHS guidelines, participants also clicked to indicate their informed consent to participate in the study. We also asked participants how familiar they were with the two candidates identified in the political opinion portion of the study.

The initial dataset contained 806 records and was cleaned as follows: Records were deleted in which no clicks were recorded, in which people's reported familiarity with either candidate exceeded 3 on a scale from 1 to 10 (where 1 was labeled "Not at all" and 10 was labeled "Quite familiar"), or in which people reported English fluency below 6 on a scale from 1 to 10 (where 1 was labeled "Not fluent" and 10 was labeled "Highly fluent").

The experiment itself had two main parts (Fig. 1).

Fig. 1

The Two Parts of the Experimental Procedure



Note. In the pre-training portion of the procedure, participants were randomly assigned to either a High-Trust or a Low-Trust group. The trust pre-training trials were followed by a conventional SEME experiment, in which the two trust groups were first divided (by random assignment) into three search conditions: one favoring UK candidate David Cameron, one favoring UK candidate Ed Miliband, and one favoring neither candidate (control group). See text for details.

Pre-Training

In the pre-training portion of the experiment, participants were randomly assigned to either a High-Trust (n = 312) or a Low-Trust (n = 239) group. Each group was given five pre-training trials in which they were shown a search question that had a simple factual answer (such as "What is the capital of Lesotho?") (see S5 Text for details), and they were then given two minutes to find the answer using the Kadoodle search engine, which closely simulates the functioning of the Google search engine. All participants had access to the same search results (on two search result pages, each listing six search results) and web pages (which could be accessed by clicking on the corresponding search result). Only the order of the search results varied between the groups.

In the High-Trust group, the answer could always be found by clicking on the highest-ranking result – just as it is virtually always found in that position on the leading search engine. In the Low-Trust group, the correct answer could be found in any of the 12 search result positions *except* the first two. At the end of 2 minutes, participants were given a five-option, multiple-choice question and were asked to provide the correct answer to the question they were shown earlier. They were immediately then told whether their answer was correct or incorrect. In theory, the pre-training trials in the High-Trust group were strengthening the user's tendency to attend to and click on the highest-ranking search result, and the pre-training trials in the Low-Trust group were either (a) extinguishing tendencies to attend to and click on high-ranking search results, (b) reinforcing tendencies to attend to and click on low-ranking search results (differential reinforcement of alternative behavior), or (c) having both effects.

SEME Experiment

Immediately following the pre-training, the participants in each of the trust groups were randomly assigned to three sub-groups: Pro-Candidate-A, Pro-Candidate-B, or a control group in which neither candidate was favored. The election we used was the 2015 election for the Prime Minister of the United Kingdom; the candidates were David Cameron and Ed Miliband. We chose this election to try to assure that our participants – all from the US – would initially be "undecided" voters. On a 10-point scale, our participants reported an average familiarity level of 1.3 (0.6) for David Cameron and 1.3 (0.6) for Ed Miliband.

All participants (in each of the six sub-groups) were then given basic instructions about the "political opinion study" in which they were about to participate. Then they read brief, neutral biographies of both candidates (approximately 150 words each, see S6 Text), after which they were asked eight questions about any preferences they might have for each candidate: their overall impression of each candidate, how likeable each candidate was, and how much they trusted each candidate. We also asked which candidate they would likely vote for if they had to vote today (on an 11-point scale from -5 for one candidate to +5 for the other, with the order of the names counterbalanced from one participant to another), and, finally, which of the two candidates they would in fact vote for today (forced choice).

They were then given up to 15 minutes to use our mock search engine to conduct research on the candidates. All participants had access to five pages of search results, six results per page (see S7 Text for details). All search results were real (from the 2015 UK election, obtained from Google.com), and so were the web pages to which the search results linked. The only difference between the groups was the order in which search results were shown. In the Pro-Candidate-A group, higher ranking search results linked to web pages that favored Cameron (Candidate A), and the lowest ranking search results (on the last pages of search results) favored Miliband (Candidate B). In the Pro-Candidate-B group, the order of the search results was reversed. In the control group, pro-Cameron search results alternated with pro-Miliband search results (and the first search result had a 50/50 chance of favoring either candidate), so neither candidate was favored. Prior to the experiment, the “bias” of all web pages had been rated on an 11-point scale from -5 to +5 (with the names of the candidates counterbalanced) by five independent judges to determine the extent to which a web page favored one candidate or another. The mean bias rating for each web page was used in determining the ranking of search results.

When participants chose to exit from our search engine, they were asked those five opinion questions again, and they were then asked whether anything “bothered” them about the search results they had been shown. If they answered “yes,” then they could type the details about their concerns. This was our way of trying to detect whether people spotted any bias in the search results they saw. We could not ask about bias directly, because leading questions of that sort generate predictable and often invalid answers (Loftus, 1975). We subsequently searched textual responses for words such as “bias,” “skewed,” or “slanted” to identify people in the bias groups who had apparently noticed the favoritism in the search results we showed them.

Results

We focused our data analysis on people in the two pre-training groups who answered all five of the pre-training questions correctly. These individuals not only demonstrated high compliance with our instructions; they also presumably were most highly impacted by the pre-training contingencies. On any given trial in which people did *not* find the correct answer, they presumably were not impacted by the low-trust contingencies.

For comparison purposes, we also analyzed data from people who scored lower than 100% on the pre-training questions; the bulk of this analysis is included in the Supplementary Material of this paper. As one might expect, participants in the High-Trust group answered our multiple-choice questions more accurately (M_{Correct} = 4.8 out of 5 [0.4]) than participants in the Low-Trust group did (M_{Correct} = 4.1 [1.0]; t = 10.37, p < 0.001, d = 0.92) (also see S1 Fig.). This was presumably because Low-Trust participants had more trouble finding the correct answer in the allotted 2 minutes. Focusing on the high-compliance participants reduced the number of people in the High-Trust group from 312 to 255 and reduced the number of people in the Low-Trust group from 239 to 100.

Please note that we did not exclude any participants from the experiment; rather, we chose to analyze separately data we obtained from high-compliance participants – that is, people who were most likely to have been impacted by the training contingencies – and low-compliance participants – that is, people who were less likely to have been impacted by the training contingencies.

Pre-Training

Participants in the High-Trust group spent significantly more time on the webpages that were linked to the first two search results (M = 169.7 s [124.9]) than participants in the Low-Trust group did (M = 135.7 s [86.1]; t = 2.92, p = 0.004, d = 0.32). Participants in the High-Trust group also clicked more frequently on the webpages linked to the first two search results (M = 5.9 [1.2]) than participants in the Low-Trust group did (M = 5.4 [1.5]; t = 3.00, p = 0.003, d = 0.37). Participants in the High-Trust group also spent substantially *less* time on each of the search engine results pages (M = 83.5 s [49.0]) than participants in the Low-Trust group did (M = 168.2 s [66.1]; t = -11.63, p < 0.001, d = 1.46). In other words, High-Trust group participants were attending more to the first two search results and spent less time searching in general.

SEME Experiment

Immediately following the pre-training trials, all participants transitioned to a standard SEME procedure, in which it appears that the Low-Trust pre-training impacted behavior in a number of ways.

The main finding in SEME experiments is that participants show little preference for one candidate or the other before they conduct their search, and that post-search, the preferences of the participants in the two bias groups tend to shift in the direction of the bias that was present in the search results they had been shown. SEME studies look at five different measures of this shift, the most important of which is called “Vote manipulation power” or VMP (see S8 Text for how VMP is calculated). VMP is of special interest because it is a direct measure of the increase in votes produced by the bias. It is calculated from answers given to a forced-choice question we ask participants both pre- and post-search, namely, “If you had to vote right now, which candidate would you vote for?”

Biased search results tend to produce substantial VMPs after a single search (Epstein & Robertson, 2015; Epstein et al., 2022). This finding was replicated in the present study; however, the bias-driven VMP in the High-Trust group (VMP = 34.6%, McNemar’s $X^2 = 23.56$, $p < 0.001$) was substantially larger than the bias-driven VMP in the Low-Trust group (VMP = 17.1%, $X^2 = 1.56$, $p = 0.21$ NS, $z = -3.25$, $p = 0.001$) (see Table 1 and S1 Table for further details; cf. S2 and S3 Tables for low-compliance data; cf. S4 Table for high-compliance versus low-compliance VMP comparisons).

Table 1

VMP Percentages, Search Times, and Results Clicked by Trust Group (High-Compliance Participants, 100% Accuracy in Pre-training)

Condition	**VMP (*p*)**	**Mean Search Time (s) (*SD*)†**	**Mean No. of Results Clicked (*SD*)†**
Low-Trust	17.1 (0.21 NS)	408.8 (271.1)	7.7 (5.0)
High-Trust	34.6 (< 0.001)	323.0 (233.3)	6.1 (3.9)
Diff (%)	+102.3	-21.0	-20.8
Statistic	$z = -3.25$	$t(159.6) = -2.79$	$t(149.6) = -2.78$
p	0.001	0.006	0.006

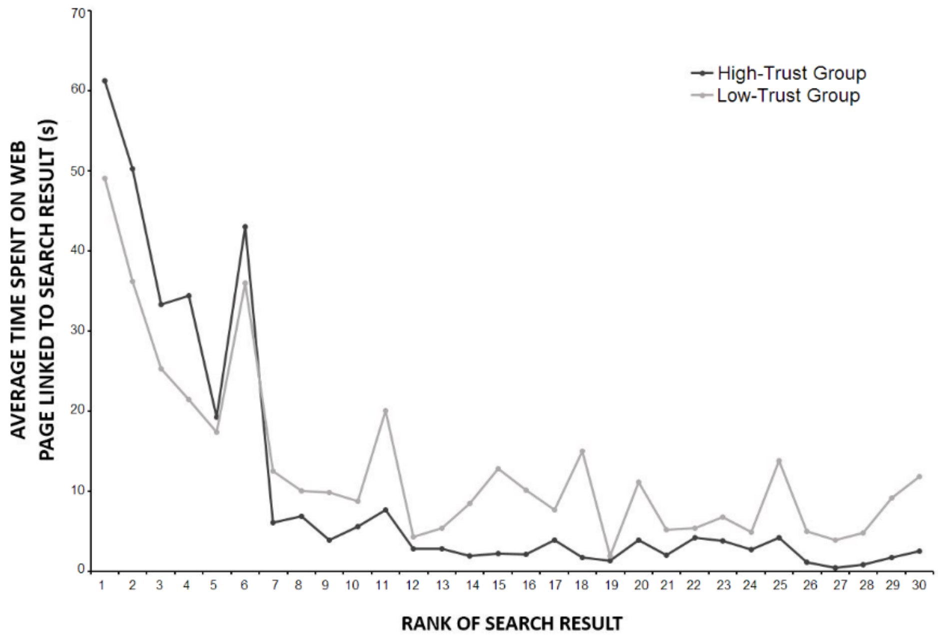
Note. McNemar's test was used to assess VMP significance. VMP is the percent increase in the number of subjects in the bias groups (combined) who said that they would vote for the favored candidate.

†These calculations were based on data from all three groups: that is, the two bias groups and the control group. This is because all three groups participated in the pre-training trials. VMP is calculated using data from the two bias groups only, so it cannot be calculated for the control group. Note that all *t*-tests employed in this study are two-tailed.

The different VMPs for the High- and Low-Trust groups can be explained by the different ways – all predictable from the pre-training session – these two groups interacted with our search engine in the political opinion portion of our study. Participants in the High-Trust group spent more time viewing the web page linked to the highest search result than participants in the Low-Trust group did (M_{High} = 60.9 s [58.1]; M_{Low} = 53.4 s [57.0]; $t = 1.11$, $p = 0.27$ NS; $d = 0.13$) (also see Fig. 2). In addition, participants in the High-Trust group clicked on the link to the first search result significantly more often than participants in the Low-Trust group did (M_{High} = 0.9 [0.4], M_{Low} = 0.8 [0.5], t = 2.18, p = 0.03, d = 0.22) (Fig. 3). Participants in the High-Trust group spent more time on web pages linked to search results on the first page of search results than participants in the Low-Trust group did (M_{High} = 241.5 s [193.9], M_{Low} = 204.6 [153.2], t = 1.71, p = 0.09 NS, d = 0.21), and participants in the Low-Trust group spent more than twice as much time on web pages linked to search results past the first page of search results than participants in the High-Trust group did (M_{Low} = 51.0 s [51.8], M_{High} = 20.4 s [32.5], t = -5.51, $p < 0.001$, d = 0.71) (Fig. 4). Participants in the High-Trust group also clicked on search results on the first page of search results significantly more often than participants in the Low-Trust group did (M_{High} = 4.0 [1.6], M_{Low} = 3.6 [1.6], t = 2.54, p = 0.01, d = 0.25), and participants in the Low-Trust group clicked on search results past the first page of search results significantly more often than participants in the High-Trust group did (M_{High} = 0.5 [0.8], M_{Low} = 1.0 [1.1], t = -3.94, $p < 0.001$, d = 0.52) (Fig. 5). These differences emerged presumably because people in the Low-Trust group had learned in pre-training to attend to and click on lower-ranked search results that people in the High-Trust group tended to ignore.

Fig. 2

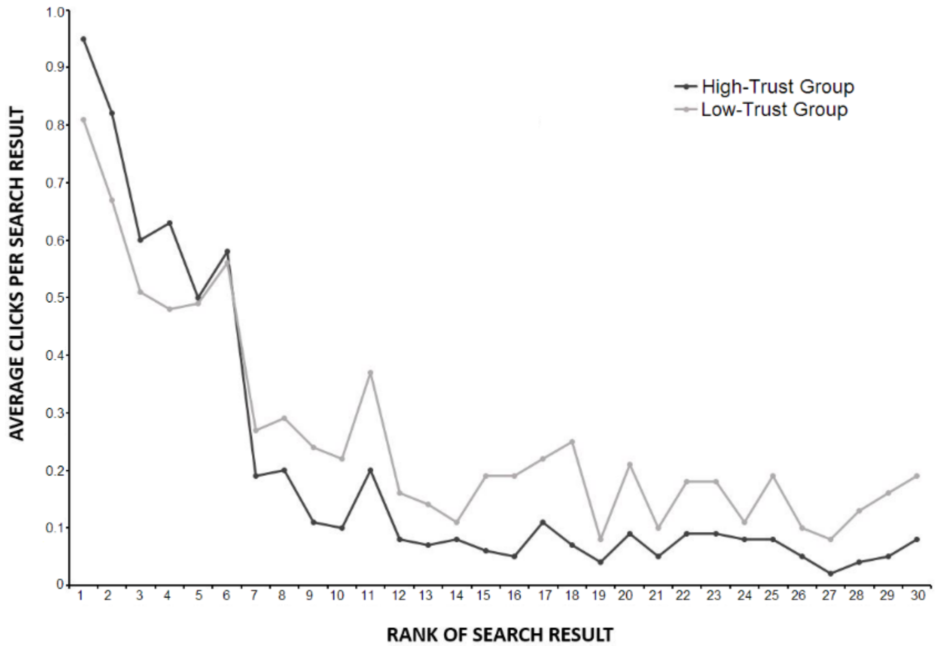
Time Spent on Search Result Web Pages as a Function of Search Result Rank (High-Compliance Participants)



Note. Participants in the Low-Trust group spent less time on web pages linked to the first page of search results and more time on web pages linked to subsequent pages of search results than participants in the High-Trust group did. For low-compliance data, see S2 Fig.

Fig. 3

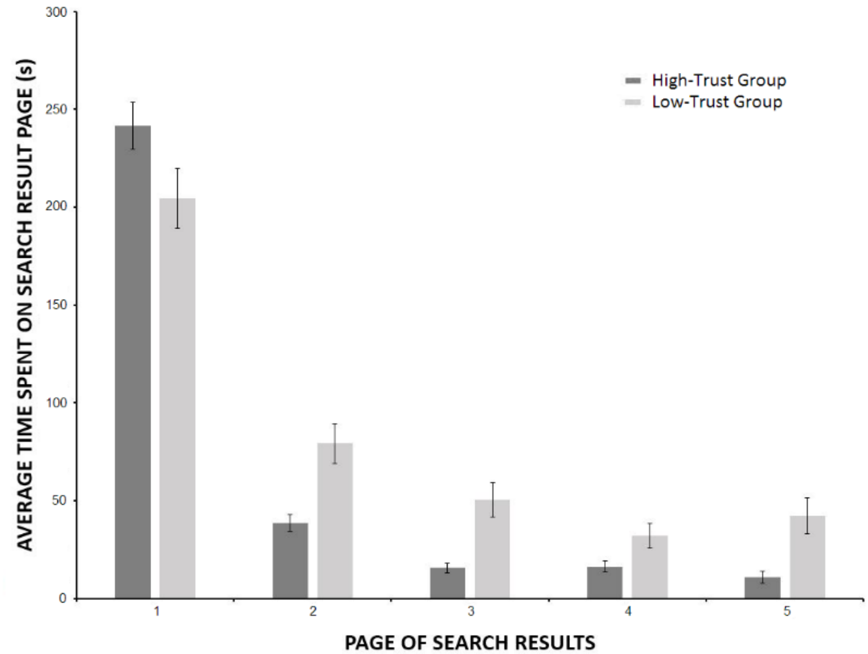
Clicks on Search Results as a Function of Search Result Rank (High-Compliance Participants)



Note. Participants in the Low-Trust group were less likely to click on results on the first page of search results and more likely to click on results on subsequent pages than participants in the High-Trust group were. For low-compliance data, see S3 Fig.

Fig. 4

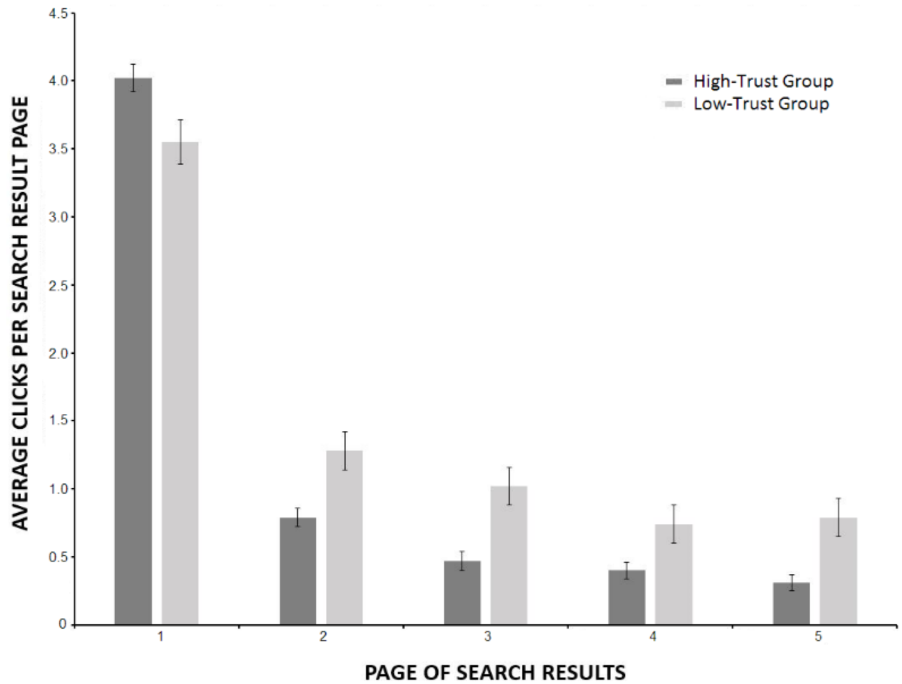
Time Spent on Search Result Pages as a Function of Page Number (High-Compliance Participants)



Note. Error bars show standard error of the mean. For low-compliance data, see S4 Fig.

Fig. 5

Cumulative Clicks on Search Results per Page as a Function of Page Number (High-Compliance Participants)



Note. Error bars show standard error of the mean. For low-compliance data, see S5 Fig.

Post search, differences also emerged on most of the answers to the seven pre-search preference questions. Pre-search, for question 7 – voting preference measured on an 11-point scale – we found no significant differences in mean ratings in the three sub-groups (pro-Cameron, pro-Miliband, and control) in both the High- and Low-Trust conditions (Table 2). Post-search, the mean ratings in the three sub-groups were significantly different in both the High- and Low-Trust conditions (Table 2).

Table 2

Changes in Voting Preferences Measured on an 11-Point Scale

	Pre-Search Likely Vote, Mean (SD)					**Post-Search Likely Vote, Mean (SD)**				
	Pro-Cameron	**Pro-Miliband**	**Control**	***H***	***p***	**Pro-Cameron**	**Pro-Miliband**	**Control**	***H***	***p***
Low-Trust	0.7 (2.8)	0.7 (2.3)	0.4 (2.7)	0.08	0.96 NS	-0.4 (3.1)	2.0 (2.9)	0.5 (3.1)	10.59	0.005
High-Trust	0.0 (2.6)	0.9 (2.7)	0.6 (2.4)	5.59	0.06 NS	-1.1 (3.1)	2.9 (2.6)	0.5 (3.3)	63.18	< 0.001

Note. A negative value indicates preference for David Cameron, and a positive value indicates preference for Ed Miliband.

Pre- vs. post-search shifts in ratings on the 11-point scale were consistent with the predicted impact of the bias, with pre/post gaps larger in the High-Trust group than in the Low-Trust group (Table 3). In the control group, pre/post shifts were minimal and non-significant (U = 1,259.5, p = 0.82 NS).

Table 3

Changes in Voting Preference for the Favored Candidate Measured on an 11-Point Scale, Bias Groups Only

	Pre-Search Likely Vote for Favored Candidate, Mean (SD)	Post-Search Likely Vote for Favored Candidate, Mean (SD)	Mean Difference †	z‡	p
Low-Trust	0.0 (2.7)	1.2 (3.1)	1.2	-3.86	< 0.001
High-Trust	0.5 (2.7)	2.1 (2.9)	1.6	-6.91	< 0.001
U	5,199.5	4,861.5	5,354.5		
p	0.20 NS	0.047	0.33 NS		

†Absolute values of the means are shown.

‡ The z values represent Wilcoxon signed ranks test comparing pre- and post-search ratings for the favored candidate.

Pre-search, we found no significant differences among the three sub-groups (pro-Cameron, pro-Miliband, and control) on their answers to any of the six opinion questions we asked about the candidates (S5 Table; see S6 Table for low-compliance data). Post-search, significant differences emerged for all six of those opinion questions for participants in both the High- and Low-Trust groups (S7 Table; see S8 Table for low-compliance data). Moreover, the net impact of biased search results on people's opinions (that is, the change in opinions about the favored candidate vs. the change in opinions about the non-favored candidate) was always larger in the High-Trust group than in the Low-Trust group and always shifted opinions (for both groups) in a way that was advantageous to the favored candidate (S9 Table; see S10 Table for low-compliance data; cf. S11 and S12 Tables for control group comparisons). However, nearly all the High- versus Low-Trust differences between pre/post changes in opinions about the candidates were nonsignificant (S13

Table; see S14 Table for low-compliance data). See S9 Text for information about perceived bias in the SEME experiment.

Discussion

The present study supports the theory that operant conditioning contributes to the power that search results have to alter thinking and behavior. The fact that a large majority (about 86%) of people's searches are for simple facts, combined with the fact that the correct answer to such queries invariably turns up in the highest-ranked position of search results, appears to teach people to attend to and click on that first result and, perhaps as a kind of generalization effect, to attend to and click on nearby search results in a pattern resembling one side of a generalization gradient. Both eye-tracking studies and studies looking at click patterns find those kinds of gradients for both attention and clicks (Athukorala et al., 2015; Chitika Insights, 2013; Cutrell & Guan, 2007; Dean, n.d.; Epstein & Robertson, 2015; Granka et al., 2004; Joachims et al., 2007; Kammerer & Gerjets, 2014; Lorigo et al., 2008; Pan et al., 2007; Schultheiß & Lewandowski, 2020). On the cognitive side, it could also be said that that daily regimen of operant conditioning is causing people to believe, trust, or have faith in the validity of high-ranking search results, and it is notable that people are entirely unaware that this regimen exists.

The fact that people generally believe that algorithms inherently produce objective and impartial output does not in and of itself explain the existence of that gradient of attention and responding. When, in the pre-training portion of the current experiment, we directed attention and clicks away from the top positions in the search list, we disrupted the usual gradient so that in the SEME portion of the study, attention was directed toward lower-ranking search results (in everyday language, we "broke the trust" people have in high-ranking results). As a result, the extreme candidate bias that was present in the search results we presented to participants in our two bias groups had less impact on the people in our Low-Trust pre-training group (VMP = 17.1%) than it did on the people in our High-Trust pre-training group (VMP = 34.6%, $p = 0.001$).

We note that if SEME is a large effect because of generalization, it is not the simple kind of generalization that occurs when wavelengths of light or sound are altered (Mis et al., 1972). That is because the nature of the task in the training situation is inherently different from the nature of the task in what we might call the test situation (the SEME experiment) – and this observation applies both to the present experiment and to the way people use search engines on a daily basis. In the pre-training phase of our experiment, people are searching for simple facts, and the reinforcing consequence is the correct answer; this is also the case when people are searching for simple facts on real search engines. In the test situation, however, there is no correct answer; the user is asking an open-ended question on an issue about which people might have a wide range of different opinions. In other words, there is a mismatch between informational properties of the training and test settings (Hogarth et al., 2015). This problem has long been a challenge when, with various impaired populations, new behavior is taught in a classroom setting, but it fails to occur in, say, the home setting; hence, the long-running concern with "transfer of training" in the behavior-analytic literature (Baldwin & Ford, 1988). Although a simple-fact query might be easily discriminable from an opinion query – at least most of the time – the present experiment sheds no light on this issue. We can assert only that pre-training that favors lower-ranked search results causes people to look more closely at lower-ranked search results, and that in turn reduces the magnitude of the shift in voting preferences.

As noted earlier, convenience might also play a role in the power that SEME has to shift opinions and voting preferences, but if that were the main or even a significant factor in explaining SEME's power, it seems unlikely that the Low-Trust training procedure we employed in the present experiment would have disrupted performance as much as it did. Breaking the pattern of reinforcement that usually supports search behavior seemed to override any importance that convenience (that is, that search position alone) might play in SEME.

Limitations and Future Research

At first glance, it might appear to be remarkable that so little retraining – a mere five search trials in which the correct answer to a search query could appear anywhere among 12 search results other than in the top two positions – could interfere with years of conditioning that reinforced attending to and clicking on the highest-ranking search items. Presumably, with more training trials, we could have reduced the impact of our biased search results far more than we did in the present procedure. But bear in mind that attending to and clicking on the highest-ranking search results has been consistently reinforced on a nearly continuous schedule – the kind of schedule that often makes behavior highly vulnerable to disruption when reinforcement is discontinued (Kimble, 1961; Lerman et al., 1996; Mackintosh, 1974). It is especially easy to disrupt behavior when it has been continuously reinforced in discrete trials (Nevin, 2012), which is always the case for search behavior on a search engine.

The present study is also limited in how it motivates participants to express their views about political candidates. They have little or no familiarity with the candidates or the issues, given that they are looking at a foreign election. Would similar numbers emerge in a study with real voters in the middle of a real election? This issue was addressed in Experiment 5 in the Epstein and Robertson study (2015). That experiment included more than 2,000 undecided voters throughout India during the final weeks of the 2014 Lok Sabha election for Prime Minister. Biased search results shifted both opinions and voting preferences, with shifts in voting preferences (the VMP) exceeding 60% in some demographic groups.

That said, recent research suggests that low-familiarity (also called “low-information”) voters differ in nontrivial ways from high-familiarity (“high-information”) voters (Yarchi et al., 2021). Our 2014 Lok Sabha experiment suggests that low-familiarity voters may be more vulnerable to SEME than high-familiarity voters, and so does a set of experiments we recently conducted on what we call the “multiple exposure effect” (MEE) (Epstein et al., 2023). Understanding the relationship between familiarity and vulnerability to manipulation will require a systematic investigation, however, not simply a comparison of values found in separate SEME experiments.

The familiarity issue does raise another question that we can address directly with the data we collected in the present study: Can we be assured that our participants were indeed undecided? Here we have strong affirmative evidence. As we noted in our Results section, the differences in pre-search opinion ratings across the three groups (Pro-Cameron, Pro-Milliband, and Control) were nonsignificant (Table 2). In addition, both the voting preferences on the 11-point scale and the voting preferences on the forced-choice question showed no candidate preferences (Table 3, S1 Table). Post-search, all these measures showed clear and predictable differences.

Pollsters often seek out people who are likely to vote, and, presumably, a company like Google can, given the vast amount of information they collect about people, easily discriminate between likely and unlikely voters. In the present study, we did not screen for this characteristic. In future studies, we will consider screening potential participants with a question such as, “How likely are you to vote in upcoming elections?”

We have other concerns about the real-world applicability of the present study, and we are addressing them in other research. The present study exposed voters to biased search results just once, but in the real world, voters might be exposed to similarly-biased search results hundreds of times before an election. Are multiple exposures to similarly biased search results additive over time? And how might opinions and voting preferences be affected if people are exposed to search results biased toward Candidate A on some occasions and Candidate B on others? Overall, do the opinions and voting preferences of undecided voters shift in the direction of the net bias?

In the real world, moreover, people are impacted by multiple sources of bias. In the traditional, non-digital world of political influence, many if not all of these sources of influence might cancel each other out. If Candidate A erects a billboard or buys a television commercial, Candidate B can do the same. But in the world of Big Tech, things work differently. If, for any reason, the algorithm of a large online platform favors one candidate, there is no way to counteract its impact, and if multiple online platforms all favor the same candidate, the impact of these different sources of influence might be additive.

Implications and Concerns

Given the concerns that have been raised about the power of biased search results to impact people's thinking and behavior, one might wonder whether informing people about the role that operant conditioning appears to play in their online decision making would have any practical benefit. We submit that raising such awareness would, unfortunately, have few or no benefits, for one simple reason: Search algorithms are designed to put the best possible answer in the top position; when one is searching for simple facts, that means the *correct* answer. A search engine that listed the best answer in a lower search position – especially in an unpredictable position – would be of little value. That means that the daily regimen of conditioning we described earlier will continue to occur as long as people continue to use properly functioning search engines. Worse still, people will always be unaware that the process by which they make both trivial and important decisions is being affected by a perpetual regimen of operant conditioning, as if they were rats trapped forever in an operant chamber.

So how can people be protected from bias that might occur in search results that are displayed in response to open-ended queries about, say, election-related issues? No matter what the cause of the bias, it can have a rapid and profound effect on the thinking and behavior of people who are undecided on an issue, and that, we believe, should be a matter for concern.

We suggest three ways to provide such protection. One would be for the US Congress, the European Parliament, or other relevant authorities to declare the Google's index – the database it uses to generate search results – to be a public commons (Epstein, 2019). This will quickly lead to the creation of hundreds, then thousands, of competing search platforms, each vying for the attention of different populations, just as thousands of news sources do currently. With numerous platforms having access to the index through a public API (an application programming interface), search will become both competitive and innovative again, as it was before Google began to dominate the search industry more than a decade ago.

Users could also be protected to some extent if browsers or search engines are at some point required to post bias alerts on individual search results or on entire search pages, with bias continuously rated by algorithms, human raters, or both. Epstein and Robertson (2016) showed that the magnitude of SEME could be reduced to some extent by such alerts (cf. Tapinsky et al., 2018; Wu et al., 2023). Alerts of this sort could also be used to flag the rising tide of online "misinformation" – an imperfect but not entirely unreasonable method for appeasing free speech advocates without suppressing content (Nekmat, 2020; Shin et al., 2023; cf. Bak-Coleman et al., 2022; BBC, 2017; Bruns et al., 2023).

Finally, a leak of documents from Google in 2019 showed that the company has long been concerned with finding ways to assure "algorithmic fairness," primarily as a way of correcting what Google executives and employees perceive to be social inequities (Lakshmanan, 2019). Setting aside the concerns one might have about the possibility that a highly influential company might be engaging in a large-scale program of social engineering (Chigne, 2018; Epstein, 2018b; Savov, 2018), the good news is that Google has developed tools for eliminating bias in algorithmic content quickly and efficiently. One of the leaked documents was a manual for Google's "Twiddler" application, which was developed "for re-ranking results from a single corpus" (Google, 2018). In other words, Google has the power to eliminate political or other bias in search results "almost as easily as one can flip a light switch" (Z. Vorhies, personal communication, June 26, 2020).

If steps are eventually taken to protect users from the bias in search results that might be displayed in response to open-ended queries, perhaps operant conditioning or other factors that currently focus user attention on high-ranking results will do no harm. As it stands, we believe that this almost irresistible tendency to attend to and click on high-ranking results, which is currently affecting the thinking and behavior of more than 5 billion people worldwide with no mechanisms in place to offset its influence, poses a serious threat to democracy, free speech, and human autonomy.

Author Contributions

Robert Epstein: Conceptualization, Methodology, Supervision, Writing – Original draft, Writing – Reviewing and Editing. **Michael Lothringer:** Statistical Analysis, Writing – Reviewing and Editing. **Vanessa Zankich:** Statistical Analysis, Visualization, Writing – Reviewing and Editing.

Data Availability

An anonymized version of the data can be found at https://doi.org/10.5281/zenodo.6978977. The data have been anonymized to comply with requirements of the sponsoring institution's Institutional Review Board (IRB). The IRB granted exempt status to this study under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB also exempted this study from informed consent requirements (relevant HHS Federal Regulations 45 CFR 46.101.(b)(2), 45 CFR 46.116(d), 45 CFR 46.117(c)(2), and 45 CFR 46.111).

Funding

This work was supported by general funds of the American Institute for Behavioral Research and Technology (AIBRT), a nonpartisan, nonprofit, 501(c)(3) organization. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The authors have no conflicts of interest to declare.

References

Agudo, U., & Matute, H. (2021). The influence of algorithms on political and dating decisions. *PLOS ONE*, *16*(4). https://doi.org/10.1371/journal.pone.0249454

Allam, A., Schulz, P. J., & Nakamoto, K. (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). https://doi.org/10.2196/jmir.2642

Anderson, N. (1958). Test of a model for opinion change. *Journal of Abnormal Psychology, 59*(3), 371–381. https://doi.org/10.1037/h0042539

Arendt, F., & Fawzi, N. (2018). Googling for Trump: Investigating online information seeking during the 2016 US presidential election. *Information, Communication & Society*, *22*(13), 1945–1955. https://doi.org/10.1080/1369118X.2018.1473459

Athukorala, K., Glowacka, D., Jacucci, G., Oulasvirta, A., & Vreeken, J. (2015). Is exploratory search different? A comparison of information search behavior for exploratory and lookup tasks. *Journal of the Association for Information Science and Technology*, *67*(11), 2635–2651. https://doi.org/10.1002/asi.23617

Baldwin, T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*(1), 63–105. http://dx.doi.org/10.1111/j.1744-6570.1988.tb00632.x

Bak-Coleman, J. B., Kennedy, I., Wack, M., Beers, A., Spiro, E. S., Starbird, K., & West, J. D. (2022). Combining interventions to reduce the spread of viral misinformation. *Nature Human Behavior*, *6*(10), 1–9. http://dx.doi.org/10.1038/s41562-022-01388-6

BBC. (2017, December 21). Facebook ditches fake news warning flag. https://www.bbc.com/news/technology-42438750

Beazley, M. B. (2013). Ballot design as fail-safe: An ounce of rotation is worth a pound of litigation. *Election Law Journal: Rules, Politics, and Policy*, *12*(1), 18–52. https://doi.org/10.1089/elj.2012.0171

Bogert, E., Schecter, A., & Watson, R. T. (2021). Humans rely more on algorithms than social influence as a task becomes more difficult. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-87480-9

Bruine de Bruin, W. (2005). Save the last dance for me: Unwanted serial position effects in jury evaluations. *Acta Psychologica*, *118*(3), 245–260. https://doi.org/10.1016/j.actpsy.2004.08.005

Bruns, H., Dessart, F. J., Krawczyk, M. W., Lewandowsky, S., Pantazi, M., Pennycook, G., Schmid, P., & Smillie, L. (2023). *The role of (trust in) the source of prebunks and debunks of misinformation. Evidence from online experiments in four EU countries*. ResearchGate. http://dx.doi.org/10.31219/osf.io/vd5qt

Burt, A. (2019, April 8). Can Facebook ever be fixed? *Harvard Business Review*. https://hbr.org/2019/04/can-facebook-ever-be-fixed

Carlson, K. A., & Russo, J. E. (2001). Biased interpretation of evidence by mock jurors. *Journal of Experimental Psychology: Applied*, *7*(2), 91–103. https://doi.org/10.1037/1076-898x.7.2.91

Chigne, J.-P. (2018, May 18). *Google's leaked internal video 'The Selfish Ledger' shows how a population could be controlled by data*. Tech Times. https://www.techtimes.com/articles/228053/20180518/googles-leaked-internal-video-the-selfish-ledger-shows-how-a-population-could-be-controlled-by-data.htm

Chitika Insights. (2013*). The value of Google results positioning.* http://info.chitika.com/uploads/4/9/2/1/49215843/chitikainsights-valueofgoogleresultspositioning.pdf

Cutrell, E., & Guan, Z. (2007). What are you looking for? An eye-tracking study of information usage in web search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 407–416. https://doi.org/10.1145/1240624.1240690

Danbury, A., Palazza, M., Mortimer, K., & Siano, A. (2013). Advertising and brand trust: Perspectives from the UK and Italy. In *Proceedings of the 18th International Conference on Corporate & Marketing Communication: Responsible Communication - Past, Present, Future* (pp. 1–11). University of Salerno.

Dean, B. (n.d.). We analyzed 5 million Google search results. Here's what we learned about organic CTR. *Backlinko*. Retrieved August 27, 2022 from https://backlinko.com/google-ctr-stats

Draws, T., Tintarev, N., Gadiraju, U., Bozzon, A., & Timmermans, B. (2021). This is not what we ordered: Exploring why biased search result rankings affect user attitudes on debated topics. *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 295–305. https://doi.org/10.1145/3404835.3462851

Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of Neuroscience*, *20*(4), 155–156. https://doi.org/10.5214/ans.0972.7531.200408

Edelman, B. (2011). Adverse selection in online "trust" certifications and search results. *Electronic Commerce Research and Applications*, *10*(1), 17–25. https://doi.org/10.1016/j.elerap.2010.06.001

Epstein, R. (2016, September 6). Free isn't freedom: How Silicon Valley tricks us. *Motherboard*. https://motherboard.vice.com/read/free-isnt-freedom-epstein-essay

Epstein, R. (2018a). Manipulating minds: The power of search engines to influence votes and opinions. In M. Moore & D. Tambini (Eds.), *Digital dominance: The power of Google, Amazon, Facebook, and Apple* (pp. 294–319). Oxford University Press.

Epstein, R. (2018b, May 25). Transcript to Google's internal video, "The Selfish Ledger." https://aibrt.org/downloads/GOOGLE–Selfish_Ledger–TRANSCRIPT.pdf

Epstein, R. (2019, August 8). Why Google poses a serious threat to democracy and how to end that threat. *Mercatornet*. https://www.thethinkingconservative.com/why-google-poses-a-serious-threat-to-democracy-and-how-to-end-that-threat/

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences USA, 112*(33), E4512–E4521. https://doi.org/10.1073/pnas.1419828112

Epstein, R., & Robertson, R. E. (2017). Suppressing the Search Engine Manipulation Effect (SEME). *Proceedings of the ACM: Human–Computer Interaction, 1*(2). https://doi.org/10.1145/3134677

Epstein, R., Ding, M., Mourani, C., Newland, A., Olson, E., & Tran, F. (2023). *Multiple searches increase the impact of similarly biased search results: An example of the "multiple exposure effect" (MEE).* SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4636728

Epstein, R., Lee, V., Mohr, R., & Zankich, V. R. (2022). The Answer Bot Effect (ABE):A powerful new form of influence made possible by intelligent personal assistants and search engines. *PLOS ONE, 17*(6). https://doi.org/10.1371/journal.pone.0268081

Eslami, M., Vaccaro, K., Karahalios, K., & Hamilton, K. (2017). “Be careful; things can be worse than they appear”: Understanding biased algorithms and users’ behavior around them in rating platforms. *Proceedings of the 11th International AAAI Conference on Web and Social Media*, *11*(1), 62–71.

Fast, N., & Jago, A. (2020). Privacy matters…or does it? Algorithms, rationalization, and the erosion of concern for privacy. *Current Opinion in Psychology*, *31*, 44–48 https://psycnet.apa.org/doi/10.1016/j.copsyc.2019.07.011

Feezell, J. T., Wagner, J. K., & Conroy, M. (2021). Exploring the effects of algorithm-driven news sources on political behavior and polarization. *Computers in Human Behavior, 116*. https://doi.org/10.1016/j.chb.2020.106626

Fortune. (2020, August 5). *World’s most admired companies*. https://fortune.com/worlds-most-admired-companies/

Ghose, A., Ipirotis, P., & Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science, 60*(7), 1617–1859. https://doi.org/10.1287/MNSC.2013.1828

Google. (n.d.). *How Search works*. https://www.google.co.uk/intl/en_uk/search/howsearchworks/mission/users/

Google. (2018, January 26). *Twiddler quick start guide*. https://aibrt.org/downloads/GOOGLE_2018–Twiddler_Quick_Start_Guide–Superroot.pdf

Granka, L., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 478–479. https://doi.acm.org/10.1145/1008992.1009079

Grant, D. (2017). The ballot order effect is huge: Evidence from Texas. *Public Choice*, *172*, 421–442. https://doi.org/10.1007/s11127-017-0454-8

Haas, A., & Unkel, J. (2017). Ranking versus reputation: Perception and effects of search result credibility. *Behaviour & Information Technology*, *36*(12), 1285–1298. https://doi.org/10.1080/0144929X.2017.1381166

Hardwick, J. (2020, April 12). Top Bing searches (as of 2020). *Ahrefs Blog.* https://ahrefs.com/blog/top–bing–searches/

Hazan, J. G. (2013). Stop being evil: A proposal for unbiased Google Search. *Michigan Law Review*, *111(*5), 789–820.

Hendler, J., & Mulvehill, A. (2016). *Social machines: The coming collision of artificial intelligence, social networking, and humanity.* Apress.

Ho, D., & Imai, K. (2008). Estimating causal effects of ballot order from a randomized natural experiment the California alphabet lottery, 1978–2002. *Public Opinion Quarterly*, *72*(2), 216–240. https://doi.org/10.1093/poq/nfn018

Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science, 24*(5):379–385. https://doi.org/10.1177/0963721415591878

Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F., & Gay, G. (2007). Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transaction on Information Systems*, *25*(2). https://doi.org/10.1145/1229179.1229181

Kammerer, Y., & Gerjets, P. (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), 177–191. https://doi.org/10.1080/10447318.2013.846790

Kieta, A. R., Cihon, T. M., & Abdel-Jalil, A. (2018). Problem solving from a behavioral perspective: Implications for behavior analysts and educators. *Journal of Behavioral Education*, *28*, 275–300. https://doi.org/10.1007/s10864-018-9296-9

Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning*. (2nd ed.). Appleton-Century-Crofts. https://doi.org/10.1037/14591-000

Koppell, J. G., & Steen, J. A. (2004). The effects of ballot position on election outcomes. *The Journal of Politics*, *66*(1), 267–281. https://doi.org/10.1046/j.1468-2508.2004.00151.x

Knobloch-Westerwick, S., Mothes, C., Johnson, B. K., Westerwick, A., & Donsbach, W. (2015). Political online information searching in Germany and the United States: Confirmation bias, source credibility, and attitude impacts. *Journal of Communication, 65*(3), 489–511. https://doi.org/10.1111/jcom.12154

Kramer, M. (2019, January 22). *With Facebook falling out of favor, will Instagram be enough to rescue shareholders?* CCN. https://www.ccn.com/with-facebook-falling-out-of-favor-will-instagram-be-enough-to-rescue-shareholders/

Lakshmanan, R. (2019, August 15). *Project Veritas releases 'internal documents' from Google and alleges anti-conservative bias.* The Next Web. https://thenextweb.com/google/2019/08/15/project-veritas-releases-internal-documents-from-google-and-alleges-anti-conservative-bias/

Lerman, D. C., Iwata, B. A., Shore, B. A., Kahng, S. (1996). Responding maintained by intermittent reinforcement: Implications for the use of extinction with problem behavior in clinical settings. *Journal of Applied Behavior Analysis*, *29*(2), 153–171. https://doi.org/10.1901/jaba.1996.29-153

Loftus, E. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*(4), 560–572. https://doi.org/10.1016/0010-0285(75)90023-7

Logg, J., Minson, J., & Moore, D. (2018, October 26). Do people trust algorithms more than companies realize? *Harvard Business Review*. https://hbr.org/2018/10/do-people-trust-algorithms-more-than-companies-realize

Logg, J., Minson, J., & Moore, D. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Lorigo, L., Haridasan, M., Brynjarsdottir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., & Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, *59*(7), 1041–1052. https://doi.org/10.1002/asi.20794

Ludolph, R., Allam, A., & Schulz, P. J. (2016). Manipulating Google’s knowledge 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). https://doi.org/10.2196/jmir.5430

Mack, C. C., Cinel, C., Davies, N., Harding, M., & Ward, G. (2017). Serial position, output order, and list length effects for words presented on smartphones over very long intervals. *Journal of Memory and Language*, *9*, 61–80. https://doi.org/10.1016/j.jml.2017.07.009

Mackintosh, N. J. (1974). *The psychology of animal learning.* Academic Press.

Mantonakis, A., Rodero, P., Lesschaeve, I., & Hastie, R. (2009). Order in choice: Effects of serial position on preferences. *Psychological Science*, *20*(11), 1309–1312. http://dx.doi.org/10.1111/j.1467-9280.2009.02453.x

Marable, L. (2003, June 30). *False oracles: Consumer reaction to learning the truth about how search engines work.* Consumer WebWatch. https://advocacy.consumerreports.org/wp-content/uploads/2013/05/false-oracles.pdf

McKinnon, J., & MacMillan, D. (2018, September 20). Google workers discussed tweaking search function to counter travel ban. *The Wall Street Journal*. https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472

Meyers, P. J. (2019, May 4). How often does Google update its algorithm? *Moz*. https://moz.com/blog/how-often-does-google-update-its-algorithm

Mis, F. W., Lumia, A. R., & Moore, J. W. (1972). Inhibitory stimulus control of the classically conditioned nictitating membrane response of the rabbit. *Behavior Research Methods & Instrumentation, 4*, 297–299. https://doi.org/10.3758/BF03207309

Murdock, B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482–488. https://doi.org/10.1037/h0045106

Murre, J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLOS ONE*, *10*(7). https://doi.org/10.1371/journal.pone.0120644

Nekmat, E. (2020). Nudge effect of fact-check alerts: Source influence and media skepticism on sharing of news misinformation in social media. *Social Media + Society*, *6*(1), 1–14. http://dx.doi.org/10.1177/2056305119897322

Nevin, J. A. (2012). Resistance to extinction and behavioral momentum. *Behavioural Processes, 90*(1), 89–97. https://doi.org/10.1016/j.beproc.2012.02.006

Nicas, J., Weise, K., & Isaac, M. (2019, September 8). How each big tech company may be targeted by regulators. *The New York Times*. https://www.nytimes.com/2019/09/08/technology/antitrust-amazon-apple-facebook-google.html

Nielsen, J. (2010, March 21). *Scrolling and attention.* Nielson Norman Group. https://www.nngroup.com/articles/scrolling-and-attention-original-research/

Nielsen, J., & Pernice, K. (2010). *Eyetracking web usability*. New Riders.

Ørmen, J. (2016). Googling the news: Opportunities and challenges in studying news events through Google Search. *Digital Journalism, 4*(1), 107–124. https://doi.org/10.1080/21670811.2015.1093272

Pan, B., Joachims, T., Granka, L., Hembrooke, H., Gay, G., & Lorigo, L. (2007). In Google we trust: User's decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication, 12*(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Paudyal, P., & Wong, W. (2018). Algorithmic opacity: Making algorithmic processes transparent through abstraction hierarchy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 192-196. https://doi.org/10.1177/1541931218621046

Pogacar, F. A, Chenai, A., Smucker, M. D., & Clarke, C. L. A. (2017). The positive and negative influence of search results on people's decisions about the efficacy of medical treatments. *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, 209-216. https://doi.org/10.1145/3121050.3121074

Prinz, R., Brighton, H., Luan, S., & Gigerenzer, G. (2017, March 23–25). *Can biased search engine results influence healthcare decisions?* [Paper presentation]. International Convention of Psychological Science, Vienna, Austria.

Rose, C. (2018, August 10). *SEO 101: How the Google search algorithm works.* SEO Mechanic. https://www.seomechanic.com/google-search-algorithm-work/

Sahin, A., Zehir, C., & Kitapci, H. (2011). The effects of brand experiences, trust and satisfaction on building brand loyalty: An empirical research on global brands. *Procedia - Social and Behavioral Sciences, 24*, 1288–1301. https://doi.org/10.1016/j.sbspro.2011.09.143

Savov, V. (2018, May 17). *Google's Selfish Ledger is an unsettling vision of Silicon Valley social engineering.* The Verge. https://www.theverge.com/2018/5/17/17344250/google-x-selfish-ledger-video-data-privacy

Schultheiß, S., & Lewandowski, D. (2020). How users' knowledge of advertisements influences their viewing and selection behavior in search engines. *Journal of the Association for Information Science and Technology*, *72*(3), 285–301. https://doi.org/10.1002/asi.24410

Shin, D., Kee, K. F., & Shin, E. Y. (2023). The nudging effect of accuracy alerts for combating the diffusion of misinformation: Algorithmic news sources, trust in algorithms, and users' discernment of fake news. *Journal of Broadcasting & Electronic Media, 67*(3), 1–20. http://dx.doi.org/10.1080/08838151.2023.2175830

Siege Media. (n.d.). *The 100 most popular Google keywords*. Siege Media. Retrieved April 20, 2020 from https://www.siegemedia.com/seo/most-popular-keywords

Singer, N. (2019, November 2.) The government protects our food and cars. Why not our data? *The New York Times*. https://www.nytimes.com/2019/11/02/sunday-review/data-protection-privacy.html

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.

Soulo, T. (n.d.). Top Google searches. *Ahrefs Blog*. Retrieved July 15, 2020 from https://ahrefs.com/blog/top-google-searches/

StatCounter GlobalStats. (n.d.). *Search engine market share worldwide*. Retrieved August 30, 2023 from https://gs.statcounter.com/search-engine-market-share

Tapinsky, G., Votta, F., & Roose, K. M. (2018). Fighting fake news and post-truth politics with behavioral science: The pro-truth pledge. *Behavior and Social Issues, 27*, 47–70. https://doi.org/10.5210/bsi.v.27i0.9127

Taylor, R. (2019, July 26). Facebook and Google algorithms are secret – but Australia plans to change that. *The Wall Street Journal*. https://www.wsj.com/articles/facebook-and-google-algorithms-are-secretbut-australia-plans-to-change-that-11564134106

Trevisan, F., Hoskins, A., Oates, S., & Mahlouly, D. (2016). The Google voter: Search engines and elections in the new media ecology. *Information, Communication & Society, 21*(1), 111–128. https://doi.org/10.1080/1369118X.2016.1261171

Trielli, D., & Diakopoulos, N. (2019). Search as news curator: The role of Google in shaping attention to news information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300683

Visser, M. (1996).Voting: A Behavior Analysis. *Behavior and Social Issues, 6*(1), 23–34. https://doi.org/10.5210/bsi.v6i1.278

Wang, Y., Wu, L., Luo, L., Zhang, Y., & Dong, G. (2017). Short-term internet search using makes people rely on search engines when facing unknown issues. *PLOS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0176325

Weinreich, H., Obendorf, H., Herder, E., & Matthias, M. (2008). Not quite the average: An empirical study of web use. *ACM Transactions on the Web, 2*(1), 1–31. https://doi.org/10.1145/1326561.1326566

West, S. (2018). *The challenge of anonymous and ephemeral social media: Reflective research methodologies and student-user composing practices* [Doctoral dissertation, University of Arkansas].

Wilhite, C. J., & Houmanfar, R. (2015). Mass news media and American culture: An interdisciplinary approach. *Behavior and Social Issues, 24*, 88–110. https://doi.org/10.5210/bsi.v.24i0.5004

Wu, Z., Draws, T., Maria Cau, F., Barile, F., Rieger, A., & Tintarev, N. (2023). Explaining search result stances to opinionated people. In L. Longo (Ed.), *Explainable artificial intelligence*. xAI 2023. Communications in Computer and Information Science (Vol. 1902). Springer. http://dx.doi.org/10.1007/978-3-031-44067-0_29

Yarchi, M., Wolfsfeld, G., & Samuel-Azran, T. (2021). Not all undecided voters are alike: Evidence from an Israeli election. *Government Information Quarterly, 38*(4), 101598. https://doi.org/10.1016/j.giq.2021.101598

Supplementary Material for "How a Daily Regimen of Operant Conditioning Might Explain the Power of the Search Engine Manipulation Effect (SEME)"

Submitted to: Behavior and Social Issues

Robert Epstein, Michael Lothringer, and Vanessa R. Zankich

American Institute for Behavioral Research and Technology

Correspondence concerning this article should be addressed to Robert Epstein at re@aibrt.org

S1 Text. Bias in Search Results

Recent statements by whistleblowers, as well as leaks of documents and videos, suggest that employees and executives at Google are well aware of the power they have to impact people's thinking and behavior and that they use that power systematically and strategically both to achieve their business objectives and to promote company values (Beres, 2019; Chigne, 2018; Epstein, 2018b; Hern, 2019; Lewis, 2017; Mckinnon & MacMillan, 2018; Statt, 2018).

Investigations by the European Commission (European Commission, 2018), the US Department of Justice (Mullins et al., 2015), the first author of the present article (Epstein, 2018b; Epstein et al., 2021), a federal commission in India (Ghosal, 2015), and others (Reuters Staff, 2016; cf. Metaxa et al., 2019) have found clear evidence of statistical bias in the search results shown to users by Google. For present purposes, we make no claims about the origins of such bias. However such bias creeps into search results – because of mandates from company executives (Hsu & Kang, 2020), the mischievousness of a rogue employee (Epstein, 2014; Lohr & Streitfeld, 2012), the way the conscious or unconscious biases of programmers get coded into algorithms (Lee et al., 2019; Nunez, 2016; Rainie & Anderson, 2017), the way users interact with algorithms

(Hahnel et al., 2018; Haider & Sundin, 2020; Mustafaraj et al., 2020; Schwartz, 2019; Steiner et al., 2020), other factors we do not yet know about, or some combination of such factors – that bias has the potential to impact the thinking and behavior of billions of people, and, we submit, it should therefore be analyzed and perhaps, in some instances, curtailed or eliminated.

It has long been known by marketers that the higher a search result is in the list of results, the more clicks it attracts (Cutrell & Guan, 2007; Granka et al., 2004; Salmerón et al., 2013; cf. Murphy et al., 2017). A study that looked at the pattern of clicks in a sample of 300 million search results found that 50% of all clicks went to the top two results and that 95% of clicks went to items on the first page of results (Chitika Insights, 2013), and these findings have been replicated in other studies (Advanced Web Ranking, n.d.; Dean, n.d.). Because businesses depend on clicks to attract customers, over the past 20 years, a multi-billion dollar industry – the "search engine optimization" (SEO) industry – has grown to help businesses rise higher in search results. At some point, researchers – mainly concerned with marketing issues – began to try to understand the distinctive pattern of clicks produced by lists of search results. For example, eye-tracking studies found that people tended to focus longer on high-ranking results (Cutrell & Guan, 2007; Granka et al., 2004; Lorigo et al., 2008; Pan et al., 2007). Other studies have shown that users will choose higher ranking results even when more relevant results are lower on the page (Haas & Unkel, 2017; Joachims et al., 2007; Pan et al., 2007; Kammerer & Gerjets, 2014; cf. Han et al., 2021; Liu & Zhang, 2019; Walhout et al., 2017).

Patterns of search behavior found in marketing studies prior to 2013 prompted the first author of the present article to ask the following: If high-ranking search results reliably attracted the most clicks, could such results be used to influence people's opinions and even, perhaps, their votes? The first experiments to explore such questions showed that search results could indeed shift opinions and voting preferences and, as we mentioned in the main body of text, that biased results could produce a large effect even after a single search.

S2 Text. Details of the Epstein and Robertson (2015) Experimental Procedure

In five randomized, controlled, counterbalanced, double-blind experiments published in the *Proceedings of the National Academy of Sciences*, undecided voters were first asked a number of demographic questions, then asked to read short biographies of two political candidates, and then to give their initial opinions about the candidates and to indicate which one they would likely vote for. Next, participants were able to use a mock search engine – "Kadoodle," which looked and functioned like the Google search engine – to research each candidate, having first been randomly assigned to one of three groups: a group in which search results favored Candidate A, Candidate B, or neither candidate. Finally, participants rated the candidates again and indicated whom they would likely vote for now that they had more information.

In multiple experiments and replications, a large and statistically significant increase was found in the percentage of people (in the two bias groups, combined) who selected the candidate who was favored in the bias groups. As a result, the authors concluded that if people are exposed to statistically biased search results in the searches they conduct every day, such bias might be having a large impact on the thinking and behavior of people around the world. The Epstein and Robertson (2015) experiments showed that search results favoring one candidate can easily shift 20% or more of undecided voters to favor that candidate after a single search – up to 80% in some demographic groups.

S3 Text. Compliance with Ethical Standards

The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB is FWA00021545. Informed written consent was obtained for all three experiments as specified in the Procedure section below.

The IRB granted exempt status to this study under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB also exempted this study from informed consent requirements (relevant HHS Federal Regulations, 45 CFR 46.116(d), 45 CFR 46.117(c)(2), and 45 CFR 46.111).was obtained for all three experiments as specified in the Procedure section below.

S4 Text. Participant Instructions

> Thank you for your interest in our study, which is being conducted by a nonprofit, nonpartisan organization called HFE Research. We are interested in how Internet research might affect the way people view political candidates. Here is how the study works:
>
> First, we will ask you some basic questions about yourself. Your answers will be kept strictly confidential and are being used for research purposes only, so please be honest.
>
> Then we will ask you to familiarize yourself with our search engine by performing several routine searches. After you've become familiar with our search engine, we'll ask you to report your views on the two main political candidates who ran for the office of prime minister in the 2015 UK election. After you have reported your views, you will be asked to do some background research on the candidates using our search engine. Then we will ask you some additional questions, and you're done!
>
> The entire process takes less than 20 minutes, and most people find it to be quite interesting.

This study has been reviewed and approved by the HFE Institutional Review Board. We do not anticipate any difficulties or risks in your participation in this survey, but if you encounter any problems or have any concerns while taking the survey, we encourage you to email the researchers at info@HFEResearch.org. After you have completed the survey you will have the option to contact us if for any reason you wish to have your data removed from the study.

PLEASE NOTE: It is important that you comply fully with the instructions you are given. If you skip any part of the study, we will not pay you. Please participate fully, accurately, and honestly in every part of the study. That is the only way it can produce meaningful results. Thank you very much for your cooperation!

Okay, are you ready to go? Then click below to continue.

By clicking continue I understand that I must be 18 or over to participate in this study, that my participation is voluntary, that I am free to withdraw at any time, that I am providing information anonymously and that demographic information collected is confidential and cannot be used to identify me. I agree to allow the data collected to be used for future research projects, and I understand that completion and submission of this survey implies my consent to participate in the present study.

S5 Text. Experimental Procedure Part 1: Pre-Training

Pre-training queries and search results, High-Trust group

Below are the 12 search results shown for each of five different queries as shown to members of the High-Trust group in the pre-training condition – two results pages with six results per page. In the High-Trust group, the correct answer to the query could always be found by clicking on the highest-ranking link. In the Low-Trust group, the correct answer could never be found by clicking on either of two highest-ranking links. Low-Trust queries and search results are shown further below.

Query:

1. What does IKEA stand for?

Page 1 of results

Ikea - brand of the many
http://www.theguardian.com/business/2005/jun/12/theobserver.observerbusiness4
Ikea's identity is founded on a commitment to good design at low prices. Thirty years ago,

Kamprad ... the ideas of 'flat-pack' and '**Ikea**' are inseparable. The practicality ... **'Ikea' stands for** ...

What Ikea Product Names REALLY Mean
http://www.huffingtonpost.com/entry/what-do-ikea-product-names...
Ikea product names can be a mouthful for the English-speaking set. ... every **Ikea** product gets a name chosen with love and care. "Typically, the name is hand-picked from an ...

The Most Ridiculous IKEA Product Names (and What They Mean)
http://flavorwire.com/225706/the-most-ridiculous-ikea-product-names-...
... reading the product names, first instituted to compensate for the dyslexia of **IKEA** ... of the most ridiculous-sounding **IKEA** product names. ... KNUTSTORP: Knutstorp Castle is the birthplace...

What do IKEA's furniture names actually mean?
http://www.techtimes.com/articles/13771/20140822/ikea-furniture-names...
There's one thing that always sticks with people after they leave: the names of **IKEA**'s furniture items. ... how **IKEA** names its products. Due to his dyslexia, **IKEA**'s founder Ingvar Kamprad thought labeling the ...

Ikea commits €1bn to sustainability and leads a roster of green ...
http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/11650751...
The Swedish flat-pack furniture company will spend €500m on wind power and about €100m on solar energy ... **Ikea**'s Scandinavian operations are now entirely energy independent ... Unilever's business model includes using sustainably ...

Lawsuit: IKEA to blame for dresser's deadly tip-over
http://www.philly.com/philly/blogs/dncrime/Lawsuit-IKEA-to-blame-for-dressers...
... toddler died after an **IKEA** dresser fell on him has sued the Scandinavian chain ... According to estimates from the Consumer Product Safety Commission, ... **IKEA** chests of drawers are safe for their intended use when ...

Page 2 of results

Frugal life of Mr IKEA: Meet the flatpack billionaire who only ...
http://www.express.co.uk/news/world/387763/Frugal-life-of-Mr...
... the 87th birthday tomorrow of **Ikea** founder Ingvar Kamprad. But Mr Kamprad, a widower, does not go in for extravagances such as birthday parties. ... The Truth About **Ikea** published in 2010, Kamprad’s former executive assistant Johan Stenebo ...

Behind the Brand: IKEA
http://www.theecologist.org/green_green_living/behind_the_label/1098324...
IKEA says that it is moving toward powering all of its stores with renewable energy, ... The High Cost of Discount Culture, Ellen Ruppel Shell argues that **IKEA** - by some measures the world’s third-largest consumer of wood - sells products with ...

Ikea to go 'forest positive' - but serious challenges lie ahead
http://www.theguardian.com/sustainable-business/ikea-sustainability-forest-...
Ikea isn't starting from ground zero – for years the company has been working through sustainable business models ... In 2012 **Ikea** increased the volume of solid wood from forests certified by the FSC from 16.2% to 22.6% and supported ...

How IKEA Became Kings of Content Marketing
https://contently.com/strategist/2014/11/07/how-ikea-became-kings-of-...
But what about **IKEA**? The Swedish 'Life Improvement Store' prints over 200 million copies of its catalog ... They don't just glean this consumer information from surveys and reports. **IKEA** actually sends design experts into people's ...

Ikea is betting big on wind energy in the US
http://blogs.marketwatch.com/behindthestorefront/2014/04/15/ikea-is-betting...
The Illinois wind farm, slated to include 49 wind turbines and be wholly owned by **Ikea**, is expected to be fully operational by the first half of 2015. **Ikea** plans to delegate management to wind and solar developer Apex Clean Energy. ...

IKEA Group and IKEA Foundation commit a total of EUR 1 ...
http://www.ikeafoundation.org/1-billion-for-climate-action/
Group - The EUR 600 million commitment to renewable energy, announced today by the **IKEA** Group, builds on the EUR 1.5 billion invested in wind and solar ... This includes going 100% for renewable energy, by investing in wind and ...

Kadooooooodle
1 2

---

2. What is the capital of Lesotho?

Page 1 of results

Lesotho Facts
http://www.mapsofworld.com/lesotho/information/facts.htm
The Kingdom of **Lesotho** is an enclave surrounded by the Republic of South Africa. ... Lesotho was founded by the British ... the only sizable city in the country, and the **capital** city of Lesotho is ...

Lesotho National Population Policy, June 1994
http://www.hsph.harvard.edu/population/policies/LESOTHO.htm
... urban areas are presently estimated to be growing at the rate of 5.5 percent with the exception of Maseru where the rate is higher than this. ... If the current rate of population growth (2.6%) continues, [the] population of Lesotho will double in less than three ...

Lesotho Population
http://countrymeters.info/en/Lesotho
... the population of **Lesotho** was estimated to be 2 072 046 people. This is an increase of 0.33 % (6 856 people) compared to population ... Density of population is calculated as permanently settled population ... The productive part of ...

Lesotho Economic Outlook
http://www.afdb.org/en/countries/southern-africa/lesotho/lesotho-economic-outlook
These include a lower degree of diversification, low domestic savings leading to over-dependence on foreign **capital** ... and enabled some level of liquidity. The slow implementation of the foreign financed **capital** expenditure ...

Lesotho country profile - Overview
http://www.bbc.com/news/world-africa-13728324
The Kingdom of **Lesotho** is made up mostly of highlands where many of the villages can be reached only ... forced by the lack of job opportunities to find work at South ... Economic woes have been compounded by the scrapping of a global textile quota ...

Maloti Mountains | mountains, Lesotho
http://www.britannica.com/place/Maloti-Mountains
The term as generally used outside **Lesotho** refers to a particular range that trends off to the southwest from the Great Escarpment of the Drakensberg Range, which forms ... containing the highest peaks in southern Africa. ...

Page 2 of results

History of Lesotho
https://en.wikipedia.org/wiki/History_of_Lesotho
Subsequent evolution of the state was shaped by contact with the British and Dutch colonists from Cape Colony. ... It was divided into seven administrative districts: Berea, Leribe, Maseru, Mohales Hock, Mafeteng, Qacha's Nek and Quthing.

Lesotho Its people, issues and history
http://africa.co.ls/aboutLesotho.html
Lesotho (pronounced li-soo-too), is officially the Kingdom of **Lesotho**, a landlocked country entirely surrounded by the Republic of South Africa ... some of the issues confronting **Lesotho** today including how the Basotho people may be able to ...

Climate of Lesotho
http://www.lesmet.org.ls/cimatology/climate-lesotho
The climate of **Lesotho** is primarily influenced by the country’s location in the Karoo Basin, ... constitute 85% of the country’s total annual precipitation. ... Senqu River Valley area to as high as 1,200mm ... border with the Republic of South Africa.

Lesotho Economy: Population, GDP, Inflation, Business, ...
http://www.heritage.org/index/country/lesotho

Lesotho is ranked 38th out of 46 countries in the Sub-Saharan Africa region, ... nearly three decades. **Lesotho** is a parliamentary constitutional monarchy. King Letsie III is ceremonial head of state. Thomas Thabane, elected prime minister ...

Lesotho Economy 2015
http://www.theodora.com/wfbcurrent/lesotho/lesotho_economy.html
Customs duties from the Southern Africa Customs Union accounted for 44% of government revenue in 2012. ... US. Diamond mining in Lesotho has grown in recent years and may contribute 8.5% to GDP by 2015, according to current forecasts. ...

Women in Lesotho: Gender Inequality
http://pcbalch.blogspot.com/2008/07/women-in-lesotho-gender-inequality.html
Many women in sub-Saharan Africa suffer relentlessly due to gender inequality in addition to other major underlying crises like poverty ... more tangible. For example, culturally in **Lesotho** a married woman is considered the property of her husband.

Kadooooooodle
1 2

---

3. Who were the two inventors of Post-it Notes?

Page 1 of results

An Idea That Stuck: How A Hymnal Bookmark Helped Inspire ...
http://www.npr.org/2014/07/26/335402996/an-idea-that-stuck-how-a-hymnal-...
It all started when he stumbled on a new type of adhesive that ... But he had a problem: He didn't know what to do with it. ... The **two inventors** of the **Post-it Note**, ... because their lab only had scrap yellow paper on hand. ...

3M has a plan to keep the Post-it note relevant to young ...
http://qz.com/161626/3m-has-a-plan-to-keep-the-post-it-note-relevant-to-...
These small squares of paper with a strip of adhesive on their rear, ... earn nicely for their maker, industrial conglomerate **3M**. ... Michael Vale, head of **3M**'s consumer and office business said.

Did the rise of agile methodologies significantly increase the ...
https://www.quora.com/Did-the-rise-of-agile-methodologies-significantly-...
We've burned through blocks of **Post-it notes** crazy fast for all my projects, so I was wondering if **Post-it** sales increased after Waterfall was thrown ... market is much bigger than just **Post-It** by **3M**. ...

Why 3M's Marketing Sticks With Millennials and DIY'ers
http://mashable.com/2014/05/21/3m-post-it-marketing-strategy/#uL30rLcsFPqd
3M has always been known as one of America's most innovative companies. In its 112-year history, ... "**Post-it**, for example, isn't about the stickiness of the **Post-it**," she says. "It's about how **Post-its** can add many small touches to your life ...

Inside 3M's First Global Brand Campaign In More Than 25 Years
http://www.forbes.com/sites/jenniferrooney/2015/03/11/inside-3ms-first-global-...
Consumers know the company well for its everyday brands: **Post-it**. Scotch. Filtrete. Command. ... St. Paul-based **3M**, with 32 billion in revenue, 65% of which comes from ... varied company, and to reap marketing efficiency, said Jesse Singh, ...

Bowing to pressure, 3M agrees to reshape its sustainable ...
https://www.minnpost.com/earth-journal/2015/03/bowing-pressure-3m-...
After a long siege of public pressure and negotiations, punctuated occasionally by media-savvy comic stunts, the **3M** Co. ... the steady stream of wood fiber it turns into **Post-Its**, masking tape and other products ...

Page 2 of results

Giant Post-It Note Tells 3M to "Do the Right Thing" for Forests
http://www.triplepundit.com/podium/giant-post-note-tells-3m-forests/
... unfurled a gigantic 40 x 40' **Post-It Note** launching a new campaign against the destructive environmental practices of **3M**, the St-Paul based corporate giant that produces **Post-It Notes** and Scotch Tape. ...

3M $3 Million Behind Bulletproof Glass Challenge Real?
http://guardianlv.com/2014/03/3m-3-million-behind-bulletproof-glass-challenge-real/
Recently there has been a picture being shared on social media which ... The glass poster case was prepared by covering it with a **3M** product called Scotchshield, ... all of the attention that it still seems to gather.

Goals & Progress | Sustainability at 3M United States
http://www.3m.com/3M/en_US/sustainability-us/goals-progress/
Setting goals to drive Sustainability progress is nothing new at **3M**. We have been setting global ... US Environmental Protection Agency's ENERGY STAR™ award for our worldwide energy-conservation efforts. This was the 10th consecutive year **3M** ...

How Post-it Notes Can Help You Keep Perspective
http://www.digitalistmag.com/innovation/using-post-notes-keep-perspective-03024843
Close your eyes and visualize a wall of yellow **Post-it Notes**. Each **Post-it Note** represents a day, week, month, or year as the timespan of the ... In addition to the **Post-It Notes** strategy, below are four questions ...

What's the thinking behind this color palette?
http://ask.metafilter.com/254903/Whats-the-thinking-behind-this-color-palette

The original yellow color was more or less random according to this interview. .... The familiar pale yellow used in the original **Post-it Notes**. ... would blend in - it was a pure **accident**.

Guide to Using Evernote with Post-it® Notes
https://evernote.com/partner/postitbrand/guide/#4
Once **Post-it® Notes** have been captured into Evernote, you can organize them into notebooks, tag them, and set reminders as you would with ... each **Post-it® Notes** note color can be automatically assigned to ...

Kadooooooodle
1 2

---

4. What were the two highest-grossing movies between 1990 and 1999?

Page 1 of results

The Highest Grossing 90s Movies
www.ranker.com/list/the-highest-grossing-90s-movies/all-genre-movies-lists
These are the **top 50 grossing films of the 1990s**. ... CGI finally began to reach a photo realistic quality as seen in such films as "Terminator 2" and "Jurassic Park".... span multiple genres and demographics. ... **top 50 grossing films of the 1990s**

14 Things You Might Not Know About ‘Ghost’
http://mentalfloss.com/article/66109/14-things-you-might-not-know-about-ghost
Produced for a modest $22 million, it ended the year with a worldwide **gross** of $505,702,588 ... According to Box Office Mojo**, the highest-grossing** domestic film of **1990** is ... After the unexpected $200 million domestic **gross** ...

Jurassic World explodes at box office to set record as biggest...
http://m.smh.com.au/entertainment/movies/jurassic-world-explodes-at...
Jun 15, 2015 ... In the United States, the **movie** has taken US$204.6 million on its opening weekend. The result pushes it ahead of ... The Avengers: The Age of Ultron and into position as the second strongest opening ever. ...

Resident Evil: Afterlife is top-grossing Canadian flick
http://www.theglobeandmail.com/arts/film/resident-evil-afterlife-is-top...
Resident Evil: Afterlife, the fourth instalment ... Canadian film in domestic theatres last year, **grossing** a total of just less than $7-million. ... toppling Porky's, a 1982 release, as the most successful Canadian-produced movie ever.

The most famous movie set in every state
http://www.businessinsider.com/the-most-famous-movie-set-in-every-state-2014-7
Everyone has that one movie that reminds them of home. We set out to name the most famous **movie** in every state – a challenging and subjective endeavor. ... The **movie's** lifetime **gross,** its critical acclaim, ...

The 100 Best Films of the 1990s
http://www.slantmagazine.com/features/article/the-100-best-films-of-the-1990s
By the current timetable of cultural recycling, pop artifacts tend to look ... but also not yet easily filed as products of their time - roughly 15 to 20 years following their initial conception ... I set about the task of ...

Page 2 of results

THE LISTS: Top 10 highest grossing films of the decade ...
http://www.incontention.com/2009/09/01/the-lists-top-10-highest-grossing-films-of...
The top 10 **highest grossing films** of the aughts is probably set in stone. At first ... was a doubt that this decade would be defined by anything but franchises ... Number of sequels in the top 10 **highest grossing movies of the 1990s**: 1. Number of sequels ...

10 High-Grossing Horror Films
http://www.chillertv.com/news/2015-07-08-10-high-grossing-horror-films
... take a look at some of the **highest grossing** horror films of all time! ... look at the biggest moneymakers in different categories, including found footage, slashers, remakes and supernatural horror!

The 5 Highest Grossing Disney Animated Musicals
http://www.cheatsheet.com/google-news/the-5-highest-grossing-disney-animated...
Disney believes the film could rival some of the studio's **best animated musical films** of all time. ... Frozen in private that suggests there are those at the studio who believe it could rival cherished films ... the late 1980s to late **1990s**. The risk Disney took ...

The 20 Highest Grossing Scary Movies Of All Time
http://www.businessinsider.com/the-20-highest-grossing-horror-movies-of-all...
In honor of Halloween, we thought it would be appropriate to round up the **highest-grossing** horror films of all time. ... The following 20 movies are ranked in ascending order according to US **gross totals**.

15 Highest-Grossing Best Picture Oscar Winners
http://www.theatlantic.com/entertainment/archive/2011/01/15-highest-grossing...
A ranked list of the Academy Award-winning films that have made **the most money at the box office** ... To determine the box-office ranking ... we adjusted their domestic **grosses** for inflation. The research team ...

The Top 25 Dinosaur Movies
http://screenrant.com/best-dinosaur-movies-jurassic-park-world/?view=all

The recent success of Jurassic World, which has secured the title of **highest global box-office** opening ever ... Whether you’re debating with your friends ... everyone would possibly die, admit it, you’d ... movie favorite, you are definitely not alone.

Kadoooooooodle

1 2

---

5. In what year was the first Harry Potter book released in the UK?

Page 1 of results

Harry Potter and the Scorcerer's Stone Introduction
http://www.shmoop.com/harry-potter-sorcerers-stone/
J.K. Rowling dreamt it up on a train ride to London and spent years...until one (yay Bloomsbury!) finally agreed to **publish** her work in the **United Kingdom** in ... **Harry Potter** and the Sorcerer's Stone was **published** in the United States in ...

Harry Potter and the Sorcerer's Stone (2001)
http://www.imdb.com/title/tt0241527/?ref_=ttrel_rel_tt
Rescued from the outrageous neglect of his aunt and uncle, a young boy ... **Harry Potter and the Sorcerer's Stone** is the first film...**novels** by **J.K. Rowling**. It is the tale of **Harry Potter**, an ordinary 11-year old boy serving as a sort of slave ...

The Wizarding World of Harry Potter
https://www.universalorlando.com/Theme-Parks/Wizarding-World-Of-Harry-Potter.aspx
Experience the two...World of **Harry Potter**...The Wizarding World of **Harry Potter** is included with your Universal Orlando theme park ...Experience all the magic and excitement of The Wizarding World of **Harry Potter** with this exclusive...

See 4 never-before-seen images from the illustrated Harry Potter...
http://www.ew.com/article/2015/10/04/illustrated-harry-potter-sorcerers-stone-images
Jim Kay's breathtaking illustrations from **Harry Potter and the Sorcerer's Stone**:The Illustrated Edition (out Oct. 6)...from his native **U.K.** little by little...the painting was never intended to be in the **book**...putting his own stamp on **Harry** and friends ...

The Harry Potter Personality Test
http://www.theatlantic.com/health/archive/2015/06/harry-potter-house-personality...
Pottermore, the **Harry Potter**-themed website unveiled by **J.K. Rowling** in 2012...has peered deep into my soul...For the study - titled "**Harry Potter** and the measures of personality...Beyond delighting or devastating the **Harry Potter** superfans...

Harry Potter and the Chamber of Secrets
https://www.hp-lexicon.org/about/books/cs/book_cs.html
...**first British** printing: July 1998, Bloomsbury Books...This is the second **book** ...**Harry Potter and the** Chamber of Secrets - 85,141 words...A working title for this **book** was **Harry Potter and the**...

Page 2 of results

Daniel Radcliffe's Next Trick is to Make Harry Potter Disappear
http://www.nytimes.com/2013/10/06/magazine/daniel-radcliffe.html?_r=1
Before Daniel Radcliffe became the most famous child actor in history, he was just a child...for the **first Harry Potter** film, "**Harry Potter and the Sorcerer's Stone**," he smiles brightly...

Harry Potter and the Deathly Hallows, Part 2
http://www.rollingstone.com/movies/reviews/harry-potter-and-the-deathly-hallows...
What is dead is the **Harry Potter** film franchise that milked Brit author **J.K. Rowling's** seven bestsellers for eight movies...**Harry Potter and the** Deathly Hallows, Part 2 puts...Chris Columbus' candy-assed **Sorcerer's Stone**, hit the box-office jackpot...

J.K. Rowling Supports That One Big 'Harry Potter' Theory on...
http://www.huffingtonpost.com/entry/jk-rowling-harry-potter-theory-dumbledore...
We've heard help will always be given at Hogwards to those who ask, and now **J.K. Rowling** is proving it. ..."**Harry Potter and the** Cursed Child," saying that the "cursed child" is not Tom Riddle...

Children's Books
https://www.nytimes.com/books/99/02/14/reviews/990214.14childrt.html
And so it is with **Harry Potter**, the star of "**Harry Potter and the Sorcerer's Stone**," by **J.K. Rowling**, a wonderful **first novel** from England that won major literary awards...Poor **Harry Potter** is orphaned as a baby...

Harry Potter and the Sorcerer's Stone
http://www.sparknotes.com/lit/harrypotter/context.html
Harry Potter and the Sorcerer's Stone emerged from the creative mind of **J.K.** (Joanna Kathleen) **Rowling**...Her first **book** was **published** under the original titled **Harry Potter and the Philosopher's Stone**...

J.K. Rowling Just Published a New Harry Potter Story
http://time.com/2965574/j-k-rowling-new-harry-potter-story/
Nearly seven years aftering **publishing** the final **book** in the **Harry Potter** series, **J.K. Rowling** has ... **published** to her website Pottermore. ... this is the first time **Rowling** has written...

Kadooooooodle
1 2

Pre-training questions and search results, Low-Trust group

In the Low-Trust group, the correct answer could never be found by clicking on either of two highest-ranking links.

Query:

1. What does IKEA stand for?

Page 1 of results

Ikea commits €1bn to sustainability and leads a roster of green ...
http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/11650751...
The Swedish flat-pack furniture company will spend €500m on wind power and about €100m on solar energy ... **Ikea**'s Scandinavian operations are now entirely energy independent ... Unilever's business model includes using sustainably ...

What Ikea Product Names REALLY Mean
http://www.huffingtonpost.com/entry/what-do-ikea-product-names...
Ikea product names can be a mouthful for the English-speaking set. ... every **Ikea** product gets a name chosen with love and care. "Typically, the name is hand-picked from an ...

The Most Ridiculous IKEA Product Names (and What They Mean)
http://flavorwire.com/225706/the-most-ridiculous-ikea-product-names-...
... reading the product names, first instituted to compensate for the dyslexia of **IKEA** ... of the most ridiculous-sounding **IKEA** product names. ... KNUTSTORP: Knutstorp Castle is the birthplace...

What do IKEA's furniture names actually mean?
http://www.techtimes.com/articles/13771/20140822/ikea-furniture-names...
There's one thing that always sticks with people after they leave: the names of **IKEA**'s furniture items. ... how **IKEA** names its products. Due to his dyslexia, **IKEA**'s founder Ingvar Kamprad thought labeling the ...

Ikea - brand of the many
http://www.theguardian.com/business/2005/jun/12/theobserver.observerbusiness4
Ikea's identity is founded on a commitment to good design at low prices. Thirty years ago,

Kamprad ... the ideas of 'flat-pack' and '**Ikea**' are inseparable. The practicality ... **'Ikea' stands for** ...

Lawsuit: IKEA to blame for dresser's deadly tip-over
http://www.philly.com/philly/blogs/dncrime/Lawsuit-IKEA-to-blame-for-dressers...
... toddler died after an **IKEA** dresser fell on him has sued the Scandinavian chain ... According to estimates from the Consumer Product Safety Commission, ... **IKEA** chests of drawers are safe for their intended use when ...

Page 2 of results

Frugal life of Mr IKEA: Meet the flatpack billionaire who only ...
http://www.express.co.uk/news/world/387763/Frugal-life-of-Mr...
... the 87th birthday tomorrow of **Ikea** founder Ingvar Kamprad. But Mr Kamprad, a widower, does not go in for extravagances such as birthday parties. ... The Truth About **Ikea** published in 2010, Kamprad's former executive assistant Johan Stenebo ...

Behind the Brand: IKEA
http://www.theecologist.org/green_green_living/behind_the_label/1098324...
IKEA says that it is moving toward powering all of its stores with renewable energy, ... The High Cost of Discount Culture, Ellen Ruppel Shell argues that **IKEA** - by some measures the world's third-largest consumer of wood - sells products with ...

Ikea to go 'forest positive' – but serious challenges lie ahead
http://www.theguardian.com/sustainable-business/ikea-sustainability-forest-...
Ikea isn't starting from ground zero – for years the company has been working through sustainable business models ... In 2012 **Ikea** increased the volume of solid wood from forests certified by the FSC from 16.2% to 22.6% and supported ...

How IKEA Became Kings of Content Marketing
https://contently.com/strategist/2014/11/07/how-ikea-became-kings-of-...
But what about **IKEA**? The Swedish "Life Improvement Store" prints over 200 million copies of its catalog ... They don't just glean this consumer information from surveys and reports. **IKEA** actually sends design experts into people's ...

Ikea is betting big on wind energy in the US
http://blogs.marketwatch.com/behindthestorefront/2014/04/15/ikea-is-betting...
The Illinois wind farm, slated to include 49 wind turbines and be wholly owned by **Ikea**, is expected to be fully operational by the first half of 2015. **Ikea** plans to delegate management to wind and solar developer Apex Clean Energy. ...

IKEA Group and IKEA Foundation commit a total of EUR 1 ...
http://www.ikeafoundation.org/1-billion-for-climate-action/
Group – The EUR 600 million commitment to renewable energy, announced today by the **IKEA** Group, builds on the EUR 1.5 billion invested in wind and solar ... This includes going 100% for renewable energy, by investing in wind and ...

Kadooooooodle

1 2

---

2. What is the capital of Lesotho?

Page 1 of results

Lesotho country profile - Overview
http://www.bbc.com/news/world-africa-13728324
The Kingdom of **Lesotho** is made up mostly of highlands where many of the villages can be reached only ... forced by the lack of job opportunities to find work at South ... Economic woes have been compounded by the scrapping of a global textile quota ...

Lesotho National Population Policy, June 1994
http://www.hsph.harvard.edu/population/policies/LESOTHO.htm
... urban areas are presently estimated to be growing at the rate of 5.5 percent with the exception of Maseru where the rate is higher than this. ... If the current rate of population growth (2.6%) continues, [the] population of Lesotho will double in less than three ...

Lesotho Population
http://countrymeters.info/en/Lesotho
... the population of **Lesotho** was estimated to be 2 072 046 people. This is an increase of 0.33 % (6 856 people) compared to population ... Density of population is tabl as permanently settled population ... The productive part of ...

Lesotho Economic Outlook
http://www.afdb.org/en/countries/southern-africa/lesotho/lesotho-economic-outlook
These include a lower degree of diversification, low domestic savings leading to over-dependence on foreign **capital** ... and enabled some level of liquidity. The slow implementation of the foreign financed **capital** expenditure ...

Lesotho Facts
http://www.mapsofworld.com/lesotho/information/facts.htm
The Kingdom of **Lesotho** is an enclave surrounded by the Republic of South Africa. ... Lesotho was founded by the British ... the only sizable city in the country, and the **capital** city of Lesotho is ...

Maloti Mountains | mountains, Lesotho
http://www.britannica.com/place/Maloti-Mountains
The term as generally used outside **Lesotho** refers to a particular range that trends off to the

southwest from the Great Escarpment of the Drakensberg Range, which forms ... containing the highest peaks in southern Africa. ...

Page 2 of results

History of Lesotho
https://en.wikipedia.org/wiki/History_of_Lesotho
Subsequent evolution of the state was shaped by contact with the British and Dutch colonists from Cape Colony. ... It was divided into seven administrative districts: Berea, Leribe, Maseru, Mohales Hock, Mafeteng, Qacha's Nek and Quthing.

Lesotho Its people, issues and history
http://africa.co.ls/aboutLesotho.html
Lesotho (pronounced li-soo-too), is officially the Kingdom of **Lesotho**, a landlocked country entirely surrounded by the Republic of South Africa ... some of the issues confronting **Lesotho** today including how the Basotho people may be able to ...

Climate of Lesotho
http://www.lesmet.org.ls/cimatology/climate-lesotho
The climate of **Lesotho** is primarily influenced by the country’s location in the Karoo Basin, ... constitute 85% of the country’s total annual precipitation. ... Senqu River Valley area to as high as 1,200mm ... border with the Republic of South Africa.

Lesotho Economy: Population, GDP, Inflation, Business, ...
http://www.heritage.org/index/country/lesotho
Lesotho is ranked 38th out of 46 countries in the Sub-Saharan Africa region, ... nearly three decades. **Lesotho** is a parliamentary constitutional monarchy. King Letsie III is ceremonial head of state. Thomas Thabane, elected prime minister ...

Lesotho Economy 2015
http://www.theodora.com/wfbcurrent/lesotho/lesotho_economy.html
Customs duties from the Southern Africa Customs Union accounted for 44% of government revenue in 2012. ... US. Diamond mining in Lesotho has grown in recent years and may contribute 8.5% to GDP by 2015, according to current forecasts. ...

Women in Lesotho: Gender Inequality
http://pcbalch.blogspot.com/2008/07/women-in-lesotho-gender-inequality.html
Many women in sub-Saharan Africa suffer relentlessly due to gender inequality in addition to other major underlying crises like poverty ... more tangible. For example, culturally in **Lesotho** a married woman is considered the property of her husband.

Kadooooooodle
1 2

---

3. Who were the two inventors of Post-it Notes?

Page 1 of results

What's the thinking behind this color palette?
http://ask.metafilter.com/254903/Whats-the-thinking-behind-this-color-palette
The original yellow color was more or less random according to this interview. .... The familiar pale yellow used in the original **Post-it Notes**. ... would blend in - it was a pure **accident**.

3M has a plan to keep the Post-it note relevant to young ...
http://qz.com/161626/3m-has-a-plan-to-keep-the-post-it-note-relevant-to-...
These small squares of paper with a strip of adhesive on their rear, ... earn nicely for their maker, industrial conglomerate **3M**. ... Michael Vale, head of **3M**'s consumer and office business said.

Did the rise of agile methodologies significantly increase the ...
https://www.quora.com/Did-the-rise-of-agile-methodologies-significantly-...
We've burned through blocks of **Post-it notes** crazy fast for all my projects, so I was wondering if **Post-it** sales increased after Waterfall was thrown ... market is much bigger than just **Post-It** by **3M**. ...

Why 3M's Marketing Sticks With Millennials and DIY'ers
http://mashable.com/2014/05/21/3m-post-it-marketing-strategy/#uL30rLcsFPqd
3M has always been known as one of America's most innovative companies. In its 112-year history, ... "**Post-it**, for example, isn't about the stickiness of the **Post-it**," she says. "It's about how **Post-its** can add many small touches to your life ...

Inside 3M's First Global Brand Campaign In More Than 25 Years
http://www.forbes.com/sites/jenniferrooney/2015/03/11/inside-3ms-first-global-...
Consumers know the company well for its everyday brands: **Post-it**. Scotch. Filtrete. Command. ... St. Paul-based **3M**, with 32 billion in revenue, 65% of which comes from ... varied company, and to reap marketing efficiency, said Jesse Singh, ...

Bowing to pressure, 3M agrees to reshape its sustainable ...
https://www.minnpost.com/earth-journal/2015/03/bowing-pressure-3m-...
After a long siege of public pressure and negotiations, punctuated occasionally by media-savvy comic stunts, the **3M** Co. ... the steady stream of wood fiber it turns into **Post-Its**, masking tape and other products ...

Page 2 of results

[Giant Post-It Note Tells 3M to "Do the Right Thing" for Forests]
http://www.triplepundit.com/podium/giant-post-note-tells-3m-forests/
... unfurled a gigantic 40 x 40' **Post-It Note** launching a new campaign against the destructive environmental practices of **3M**, the St-Paul based corporate giant that produces **Post-It Notes** and Scotch Tape. ...

[3M $3 Million Behind Bulletproof Glass Challenge Real?]
http://guardianlv.com/2014/03/3m-3-million-behind-bulletproof-glass-challenge-real/
Recently there has been a picture being shared on social media which ... The glass poster case was prepared by covering it with a **3M** product called Scotchshield, ... all of the attention that it still seems to gather.

[Goals & Progress | Sustainability at 3M United States]
http://www.3m.com/3M/en_US/sustainability-us/goals-progress/
Setting goals to drive Sustainability progress is nothing new at **3M**. We have been setting global ... US Environmental Protection Agency's ENERGY STAR™ award for our worldwide energy-conservation efforts. This was the 10th consecutive year **3M** ...

[How Post-it Notes Can Help You Keep Perspective]
http://www.digitalistmag.com/innovation/using-post-notes-keep-perspective-03024843
Close your eyes and visualize a wall of yellow **Post-it Notes**. Each **Post-it Note** represents a day, week, month, or year as the timespan of the ... In addition to the **Post-It Notes** strategy, below are four questions ...

[An Idea That Stuck: How A Hymnal Bookmark Helped Inspire ...]
http://www.npr.org/2014/07/26/335402996/an-idea-that-stuck-how-a-hymnal-...
It all started when he stumbled on a new type of adhesive that ... But he had a problem: He didn't know what to do with it. ... The **two inventors** of the **Post-it Note**, ... because their lab only had scrap yellow paper on hand. ...

[Guide to Using Evernote with Post-it® Notes]
https://evernote.com/partner/postitbrand/guide/#4
Once **Post-it® Notes** have been captured into Evernote, you can organize them into notebooks, tag them, and set reminders as you would with ... each **Post-it® Notes** note color can be automatically assigned to ...

Kadooooooodle
1 2

---

4. What were the two highest-grossing movies between 1990 and 1999?

Page 1 of results

15 Highest-Grossing Best Picture Oscar Winners
http://www.theatlantic.com/entertainment/archive/2011/01/15-highest-grossing...
A ranked list of the Academy Award-winning films that have made **the most money at the box office** ... To determine the box-office ranking ... we adjusted their domestic **grosses** for inflation. The research team ...

14 Things You Might Not Know About "Ghost"
http://mentalfloss.com/article/66109/14-things-you-might-not-know-about-ghost
Produced for a modest $22 million, it ended the year with a worldwide **gross** of $505,702,588 ... According to Box Office Mojo, the **highest-grossing** domestic **film** of **1990** is ... After the unexpected $200 million domestic **gross** ...

Jurassic World explodes at box office to set record as biggest...
http://m.smh.com.au/entertainment/movies/jurassic-world-explodes-at...
Jun 15, 2015 ... In the United States, the **movie** has taken US$204.6 million on its opening weekend. The result pushes it ahead of ... The Avengers: The Age of Ultron and into position as the second strongest opening ever. ...

Resident Evil: Afterlife is top-grossing Canadian flick
http://www.theglobeandmail.com/arts/film/resident-evil-afterlife-is-top...
Resident Evil: Afterlife, the fourth instalment ... Canadian film in domestic theatres last year, **grossing** a total of just less than $7-million. ... toppling Porky's, a 1982 release, as the most successful Canadian-produced movie ever.

The most famous movie set in every state
http://www.businessinsider.com/the-most-famous-movie-set-in-every-state-2014-7
Everyone has that one **movie** that reminds them of home. We set out to name the most famous **movie** in every state – a challenging and subjective endeavor. ... The **movie**'s lifetime **gross**, its critical acclaim, ...

The 100 Best Films of the 1990s
http://www.slantmagazine.com/features/article/the-100-best-films-of-the-1990s
By the current timetable of cultural recycling, pop artifacts tend to look ... but also not yet easily filed as products of their time--roughly 15 to 20 years following their initial conception ... I set about the task of ...

Page 2 of results

THE LISTS: Top 10 highest grossing films of the decade ...
http://www.incontention.com/2009/09/01/the-lists-top-10-highest-grossing-films-of...
The top 10 **highest grossing films** of the aughts is probably set in stone. At first ... was a doubt that this decade would be defined by anything but franchises ... Number of sequels in the top 10 **highest grossing movies of the 1990s**: 1. Number of sequels ...

10 High-Grossing Horror Films
http://www.chillertv.com/news/2015-07-08-10-high-grossing-horror-films
... take a look at some of the **highest grossing** horror films of all time! ... look at the biggest moneymakers in different categories, including found footage, slashers, remakes and supernatural horror!

The 5 Highest Grossing Disney Animated Musicals
http://www.cheatsheet.com/google-news/the-5-highest-grossing-disney-animated...
Disney believes the film could rival some of the studio's **best animated musical films** of all time. ... Frozen in private that suggests there are those at the studio who believe it could rival cherished films ... the late 1980s to late **1990s**. The risk Disney took ...

The 20 Highest Grossing Scary Movies Of All Time
http://www.businessinsider.com/the-20-highest-grossing-horror-movies-of-all...
In honor of Halloween, we thought it would be appropriate to round up the **highest-grossing** horror films of all time. ... The following 20 movies are ranked in ascending order according to US **gross totals**.

The Highest Grossing 90s Movies
www.ranker.com/list/the-highest-grossing-90s-movies/all-genre-movies-lists
These are the **top 50 grossing films of the 1990s**. ... CGI finally began to reach a photo realistic quality as seen in such films as "Terminator 2" and "Jurassic Park".... span multiple genres and demographics. ... **top 50 grossing films of the 1990s**

The Top 25 Dinosaur Movies
http://screenrant.com/best-dinosaur-movies-jurassic-park-world/?view=all
The recent success of Jurassic World, which has secured the title of **highest global box-office** opening ever ... Whether you're debating with your friends ... everyone would possibly die, admit it, you'd ... movie favorite, you are definitely not alone.

Kadooooooodle
1 2

---

5. In what year was the first Harry Potter book released in the UK?

Page 1 of results

J.K. Rowling Just Published a New Harry Potter Story
http://time.com/2965574/j-k-rowling-new-harry-potter-story/
Nearly seven years aftering **publishing** the final **book** in the **Harry Potter** series, **J.K.**

Rowling has ... **published** to her website Pottermore. ... this is the first time **Rowling** has written...

Harry Potter and the Sorcerer's Stone (2001)
http://www.imdb.com/title/tt0241527/?ref_=ttrel_rel_tt
Rescued from the outrageous neglect of his aunt and uncle, a young boy ... **Harry Potter and the Sorcerer's Stone** is the first film...**novels** by **J.K. Rowling**. It is the tale of **Harry Potter**, an ordinary 11-year old boy serving as a sort of slave ...

The Wizarding World of Harry Potter
https://www.universalorlando.com/Theme-Parks/Wizarding-World-Of-Harry-Potter.aspx
Experience the two...World of **Harry Potter**...The Wizarding World of **Harry Potter** is included with your Universal Orlando theme park ...Experience all the magic and excitement of The Wizarding World of **Harry Potter** with this exclusive...

See 4 never-before-seen images from the illustrated Harry Potter...
http://www.ew.com/article/2015/10/04/illustrated-harry-potter-sorcerers-stone-images
Jim Kay's breathtaking illustrations from **Harry Potter and the Sorcerer's Stone**:The Illustrated Edition (out Oct. 6)...from his native **U.K.** little by little...the painting was never intended to be in the **book**...putting his own stamp on **Harry** and friends ...

The Harry Potter Personality Test
http://www.theatlantic.com/health/archive/2015/06/harry-potter-house-personality...
Pottermore, the **Harry Potter**-themed website unveiled by **J.K. Rowling** in 2012...has peered deep into my soul...For the study - titled "**Harry Potter** and the measures of personality...Beyond delighting or devastating the **Harry Potter** superfans...

Harry Potter and the Chamber of Secrets
https://www.hp-lexicon.org/about/books/cs/book_cs.html
...**first British** printing: July 1998, Bloomsbury Books...This is the second **book** ...**Harry Potter and the** Chamber of Secrets - 85,141 words...A working title for this **book** was **Harry Potter and the**...

Page 2 of results

Daniel Radcliffe's Next Trick is to Make Harry Potter Disappear
http://www.nytimes.com/2013/10/06/magazine/daniel-radcliffe.html?_r=1
Before Daniel Radcliffe became the most famous child actor in history, he was just a child...for the **first Harry Potter** film, "**Harry Potter and the Sorcerer's Stone**," he smiles brightly...

Harry Potter and the Deathly Hallows, Part 2
http://www.rollingstone.com/movies/reviews/harry-potter-and-the-deathly-hallows...
What is dead is the **Harry Potter** film franchise that milked Brit author **J.K. Rowling's** seven bestsellers for eight movies...**Harry Potter and the** Deathly Hallows, Part 2 puts...Chris Columbus' candy-assed **Sorcerer's Stone**, hit the box-office jackpot...

J.K. Rowling Supports That One Big 'Harry Potter' Theory on...
http://www.huffingtonpost.com/entry/jk-rowling-harry-potter-theory-dumbledore...
We've heard help will always be given at Hogwards to those who ask, and now **J.K. Rowling** is proving it. ..."**Harry Potter and the** Cursed Child," saying that the "cursed child" is not Tom Riddle...

Children's Books
https://www.nytimes.com/books/99/02/14/reviews/990214.14childrt.html
And so it is with **Harry Potter**, the star of "**Harry Potter and the Sorcerer's Stone**," by **J.K. Rowling**, a wonderful **first novel** from England that won major literary awards...Poor **Harry Potter** is orphaned as a baby...

Harry Potter and the Sorcerer’s Stone
http://www.sparknotes.com/lit/harrypotter/context.html
Harry Potter and the Sorcerer's Stone emerged from the creative mind of **J.K.** (Joanna Kathleen) **Rowling**...Her first **book** was **published** under the original titled **Harry Potter and the Philosopher's Stone**...

Harry Potter and the Scorcerer's Stone Introduction
http://www.shmoop.com/harry-potter-sorcerers-stone/
J.K. Rowling dreamt it up on a train ride to London and spent years...until one (yay Bloomsbury!) finally agreed to **publish** her work in the **United Kingdom** in ... **Harry Potter** and the Sorcerer's Stone was **published** in the United States in ...

Kadooooooodle
1 2

S6 Text. Experimental Procedure Part 2: SEME Experiment, Pre-Search Candidate Biographies

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 Master's 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.

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.

S7 Text. Experimental Procedure Part 2: SEME Experiment, Search Results

Pro-Miliband group search results

Below are the 30 search results for the pro-Miliband group: 5 pages of results with 6 results per page. They appeared in reverse order for the pro-Cameron group. They alternated by candidate for the control group.

Page 1 of results

Ed Miliband would make a great prime minister, says bu...
www.bbc.co.uk
Company owner Arnab Dutt claims Ed Miliband would be good for business, ... make a great prime minister, says business ... owner Arnab Dutt claims Ed Miliband ...

David Cameron accused of being ' chicken' after he pul...
www.theguardian.co.uk
David Cameron accused of being 'chicken' after he pulls out of video debate

1m voters lost from electoral roll, says Ed Miliband...
www.telegraph.co.uk
Ed Miliband will accuse Nick Clegg on Friday of delivering "the final insult" to young people, claiming that electoral changes mean 1 million people, many of them ...

David Cameron Mulls Ban on Encrypted Messaging Including...
www.bbc.co.uk
Peter Byrne/Reuters . British Prime Minister David Cameron reacted to last week's terrorist attack in Paris by participating in a march declaring solidarity with ...

Billionaire named in tax scandal files coughs up " ......
www.theguardian.co.uk
Sections Latest News UK News World News Weird ... A billionaire tycoon named in the HSBC tax scandal files had dinner with David ... a billionaire racehorse owner ...

Ed Miliband: don't mistake my decency for weakness | P...
www.telegraph.co.uk
Ed Miliband has insisted that he has the strength of character to be prime minister, arguing that decency should not be mistaken for weakness, as the Labour party ...

Page 2 of results

Ed Miliband's apprenticeship vow: Get the grades and...
www.bbc.co.uk

Labour will create 80,000 more apprenticeships if Ed Miliband wins the election, guaranteeing that those who get the grades will get a job. Mr Miliband made the ...

What David Cameron just proposed would endanger every...
www.theguardian.co.uk
What David Cameron just proposed would endanger every Briton and destroy the IT industry. David Cameron says there should be no "means of communication" which "we ...

Ed Miliband: Labour will put arts at 'the heart' of ...
www.telegraph.co.uk
Nick Clark is the arts correspondent of The Independent. He joined the newspaper in June 2007, initially reporting on the stock markets. He has covered beats ...

Election 2015: David Cameron's is more focused on...
www.bbc.co.uk
Election 2015: David Cameron's is more focused on painting a false picture of Labour than the truth about the country

David Cameron's plan to force young jobless into unpaid...
www.theguardian.co.uk
News UK News David Cameron ... David Cameron’s plan to force the young jobless to do unpaid work will not help them find a ... blackjack and slots at Mirror Casino.

Stephen Green row: David Cameron's key exchanges in...
www.telegraph.co.uk
Does the prime minister expect us to believe that in Stephen Green’s three years as a ...

Page 3 of results

So much for a Punch and Judy show. Our feeble MPs are...
www.bbc.co.uk
So much for a Punch and Judy show. Our feeble MPs are out for the count Armando Iannucci

The Problem With David Cameron's ' Schizophrenic' Cybe...
www.theguardian.co.uk
U.K. Prime Minister David Cameron asked for the U.S government's help ... Cameron's 'Schizophrenic' Cyber Security ... this problem is." Watch the full ...

David Cameron’s flagship Cancer Drugs Fund ‘is a waste...
www.telegraph.co.uk
David Cameron’s flagship Cancer Drugs Fund ‘is a waste of NHS cash ...

Conservatives are edging ahead as Ed Miliband holds...
www.telegraph.co.uk
Though the Conservatives have moved ahead of Labour, the poll is not all good news for Mr Cameron. Voters are almost twice as likely to say that it is time for a ...

Ed Miliband's appointment of Lord Prescott is...
www.theguardian.co.uk
Ed Miliband's appointment of Lord Prescott is 'desperate', say Labour MPs Labour MPs warn that the return of Lord Prescott is 'too little, too late', as ...

Britain must be more like the US and spread its ' econo...
www.bbc.co.uk
Britain must be more like the US and spread its 'economic might' ... David Cameron says Britain must replicate the US and have more great cities Photo: ...

Page 4 of results

Britain will 'never give up' freedom of speech, David...
www.telegraph.co.uk
Britain will 'never give up' freedom of speech, David Cameron says after Charlie Hebdo attack Prime Minister pledged support of UK's intelligence services ...

Misunderstanding the present: Ed Miliband wants to gov...
www.theguardian.co.uk
If anything defines Ed Miliband’s leadership of the Labour Party, it is the belief that British politics has reached an inflexion point like the one that enabled ...

David Cameron announces arrival of free Wi- Fi on trai..
www.bbc.co.uk
David Cameron announces arrival of free Wi-Fi on trains from ... announce plans that will see the rollout of free Wi-Fi on trains across the United Kingdom from 2017.

Ed ' misleading voters with letters that look like the...
www.telegraph.co.uk
Ed 'misleading voters with letters that look like they're from the NHS': Miliband under fire for 'cruel' mail ... which could easily be mistaken for test results ...

David Cameron attacks Ed Miliband in Downing Street sp...
www.theguardian.co.uk
News ‘ General election 2015; David Cameron attacks Ed Miliband in Downing Street speech ... David Cameron has spoken in Downing Street after his visit to ...

Angela Merkel gives Ed Miliband a miss - Telegraph
www.bbc.co.uk

News – Politics - Ed Miliband; Angela Merkel gives Ed Miliband a miss Angela Merkel will not meet Ed Miliband ahead of ...

Page 5 of results

David Cameron - Biography - Prime Minister - Biography...
www.telegraph.co.uk
David Cameron is best known for being a revolutionary leader of Britain's Conservative Party, a quality that eventually won him the election as prime minister in 2010 ...

MoS victory as David Cameron steps in to save memorial...
www.theguardian.co.uk

MoS victory as David Cameron steps in to save memorial to Battle of Britain aces. ... David Cameron has vowed to save the chapel ...

David Cameron: Isil poses a direct and deadly threat to...
www.bbc.co.uk
David Cameron: Isil poses a direct and deadly threat to Britain The poisonous extremism on the march in Iraq and Syria affects us all - and we have no ...

PMQs: David Cameron smashed Ed Miliband to bits - Teleg...
www.telegraph.co.uk
News - General election 2015; PMQs: David Cameron smashed Ed Miliband to bits ... David Cameron smashed Ed Miliband to bits.

Regressive, sarcastic and pious - welcome to Britain...
www.theguardian.co.uk
They will want five years of socialism in the UK, followed by another chance at breaking away. The risk is that by then the people of England will be so ...

Ed Miliband is a blancmange in a hurricane - Telegraph
www.bbc.co.uk
Ed Miliband is a blancmange in a hurricane Labour’s leader is weak, indecisive, lacks clarity, and has turned his party into a vacuum...

Kadooooooodle

1 2 3 4 5

S8 Text. VMP Calculation

VMP is calculated as follows:

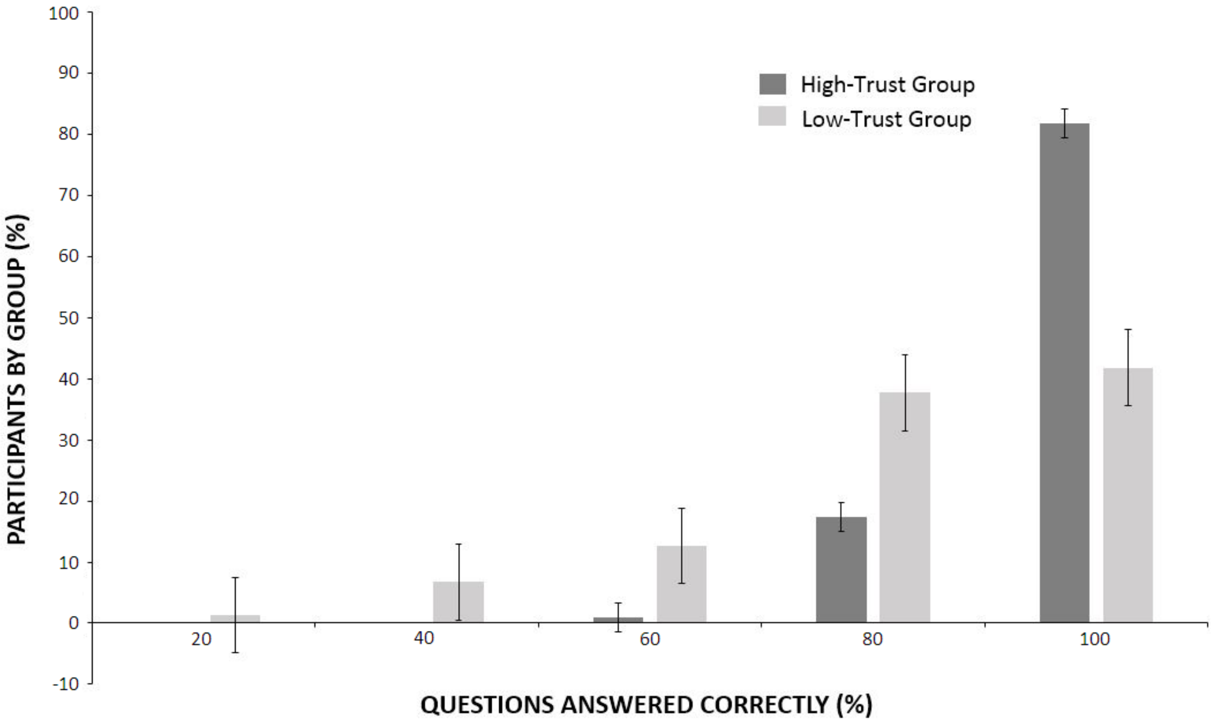
$$\left(\frac{V_{Post} - V_{Pre}}{V_{Pre}}\right) * 100$$

where V_{Post} is the number of participants in the two bias groups (combined) who said, post-search, that they would vote for the candidate who was favored in the search results, and V_{Pre} is the number of participants in the two bias groups (combined) who said, pre-search, that they would vote for that same candidate. In other words, VMP is the percentage increase in the number of people who voted for the favored candidate after having been exposed to search results that favored that candidate.

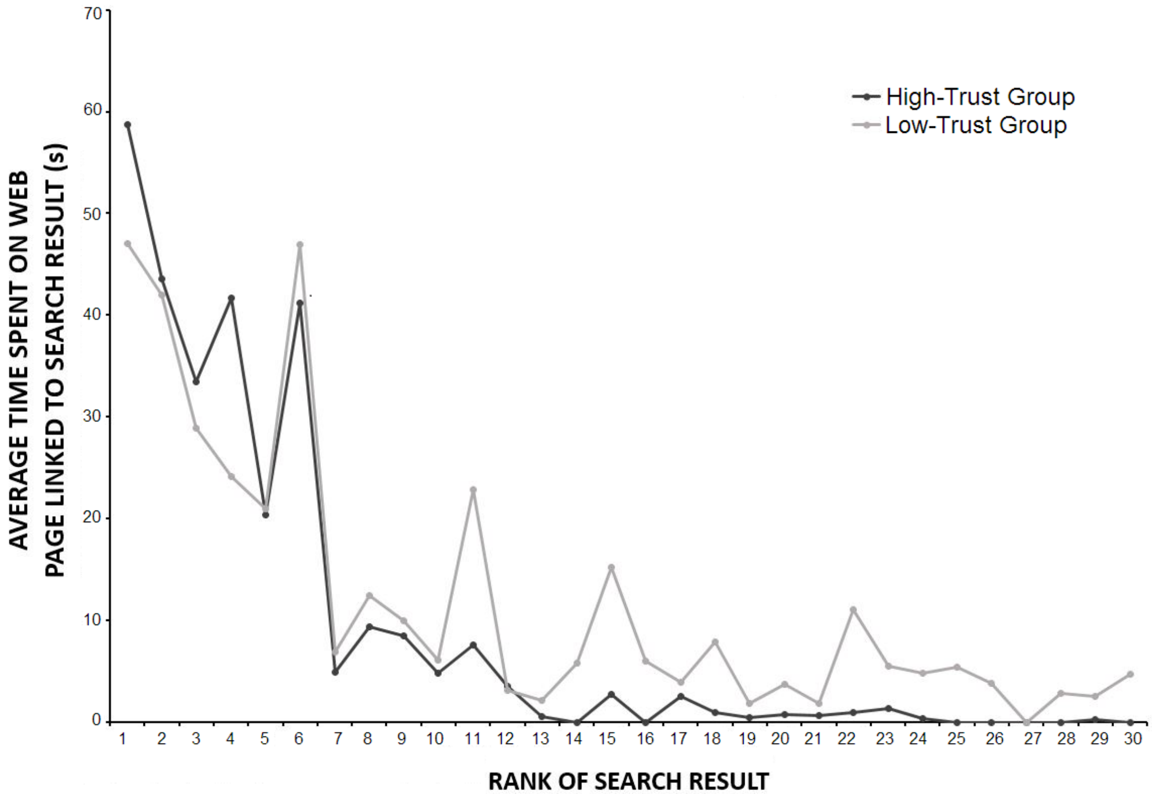
S9 Text. Perceived Bias in the SEME Experiment

Because we did not employ a masking procedure to disguise the biased ordering of search results in this experiment (Epstein & Robertson, 2015), we expected a sizeable proportion of our participants in the bias groups to notice the bias. In fact, for high-compliance participants, 34.6% of those who were exposed to biased search results in the High-Trust group, and 31.7% of those who were exposed to biased search results in the Low-Trust group reported noticing the bias. This is comparable to the finding in Experiment 1 (in which no mask was used to disguise the bias) reported by Epstein and Robertson (2015), in which 25.0% of the participants in the bias groups reported noticing a bias.

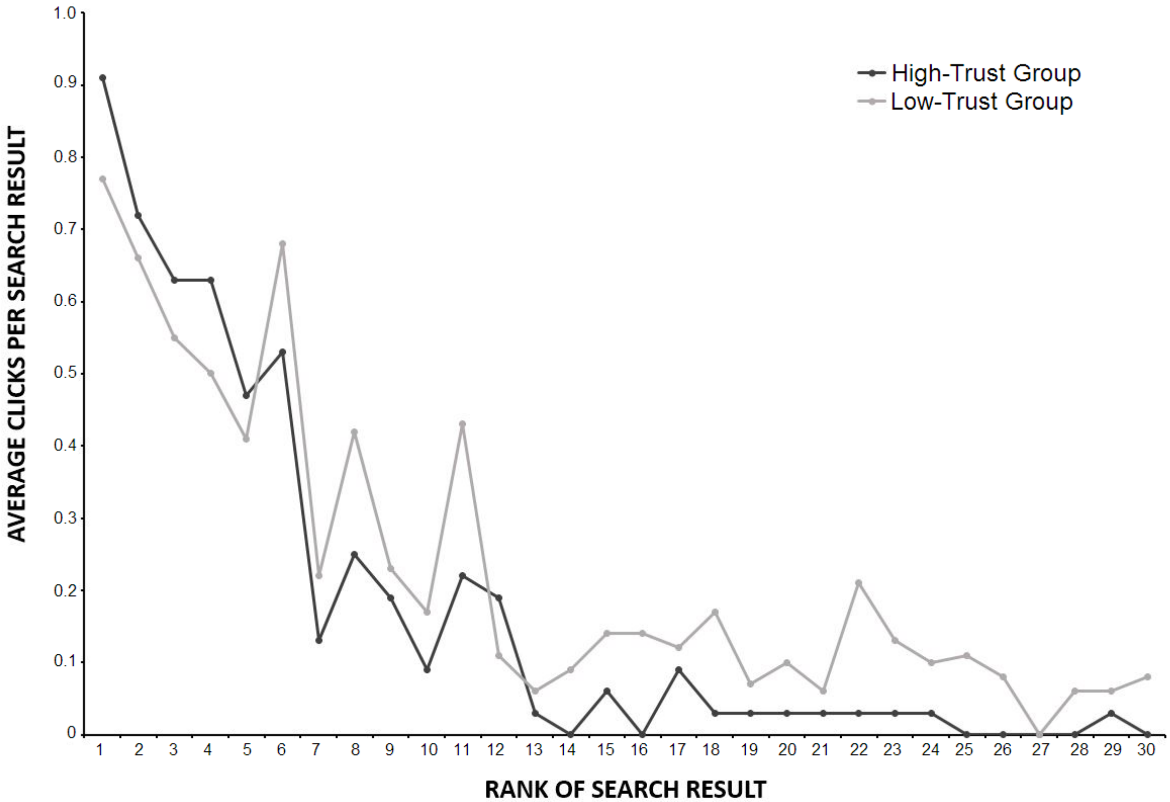
S1 Fig. Percentage of people in each pre-training group who answered the pre-training questions correctly. Overall, participants in the High-Trust group answered the pre-training questions more accurately than participants in the Low-Trust group did, presumably because the latter had more trouble finding the correct answers. Error bars show standard error of the mean.



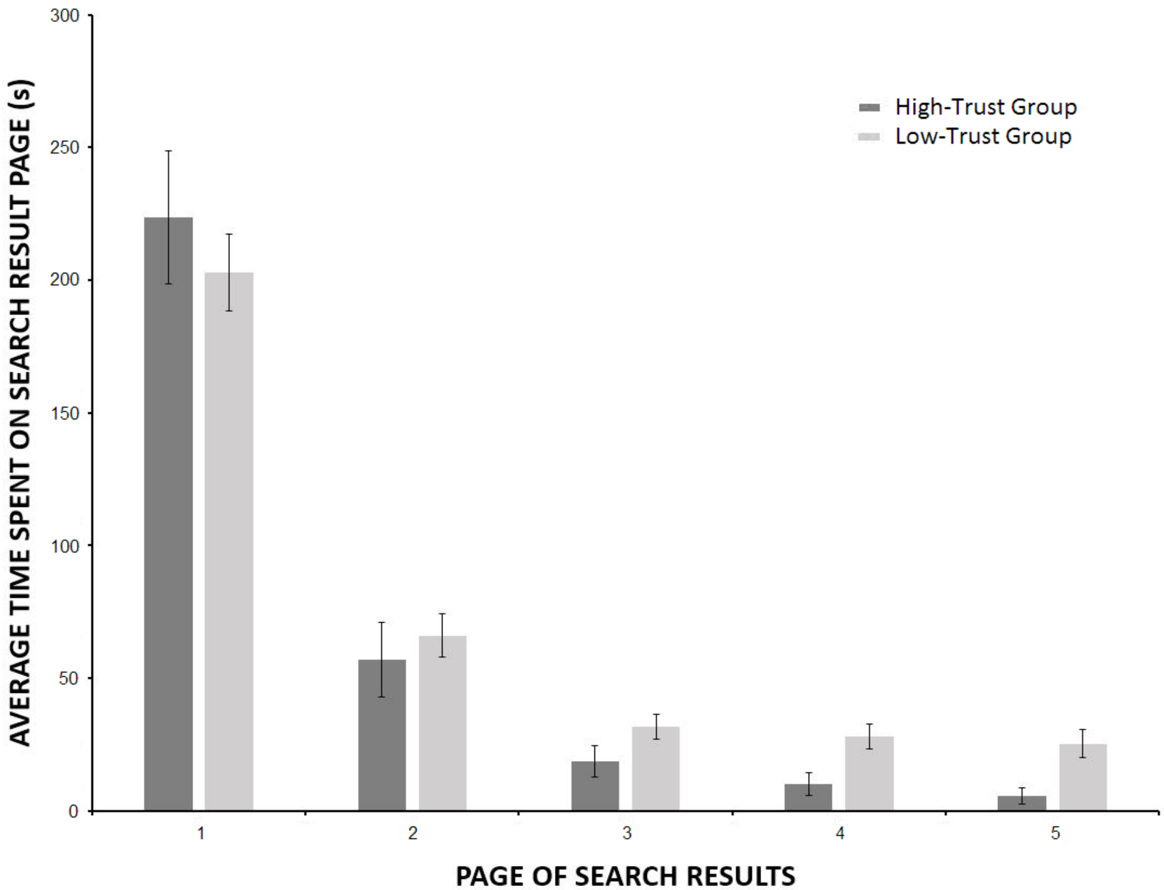
S2 Fig. Time spent on search result web pages as a function of search result rank (low-compliance participants). Participants in both trust groups spent less time on web pages linked to lower-ranked search results than web pages linked to higher-ranked search results. Participants in the Low-Trust group spent less time on web pages linked to the first five search results than participants in the High-Trust group did. Participants in the Low-Trust group also spent more time on web pages linked to search results past the first five search results than participants in the High-Trust group did.



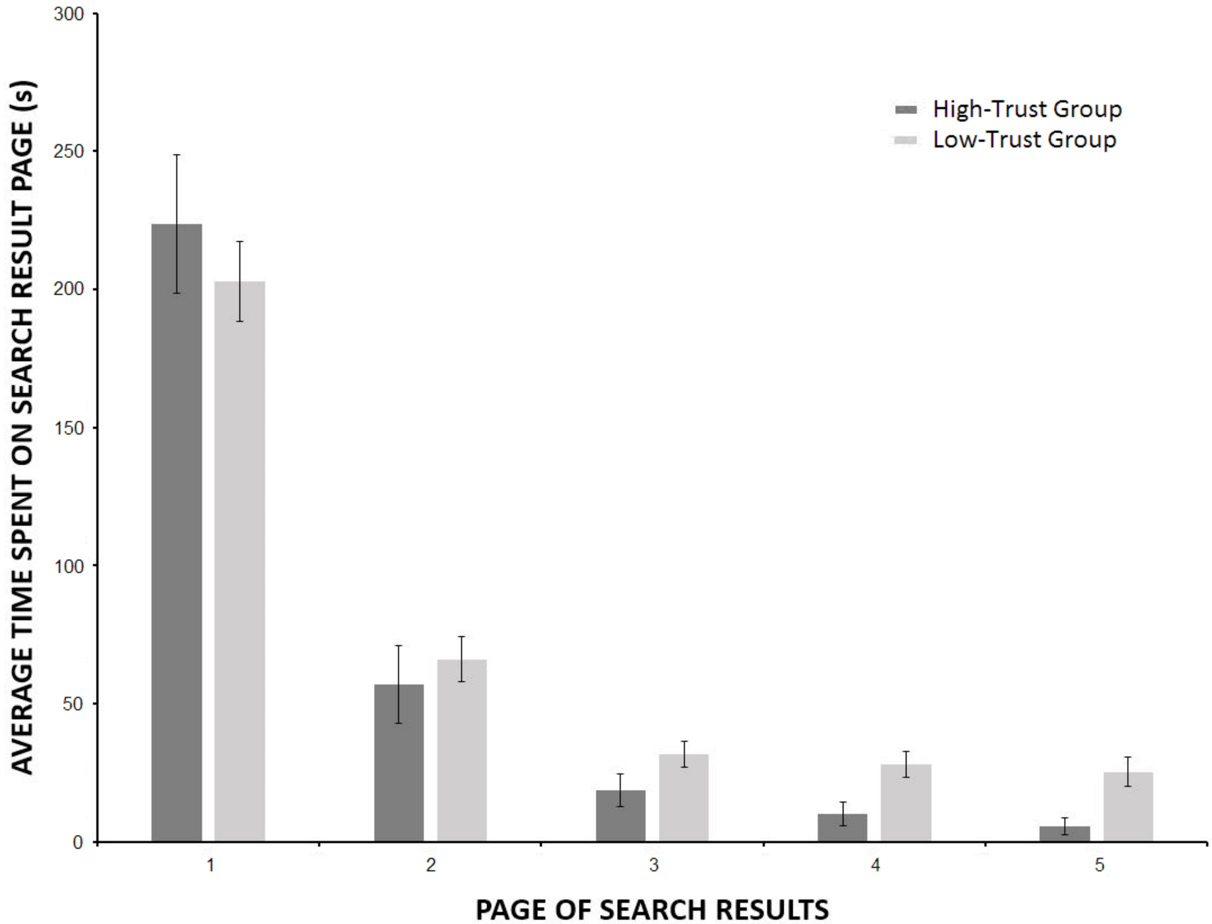
S3 Fig. Clicks on search results as a function of search result rank (low-compliance participants). Participants in both trust groups were less likely to click on lower-ranked results than higher-ranked search results. Participants in the Low-Trust group were less likely to click on the first five search results than participants in the High-Trust group. Participants in the Low-Trust group were also more likely to click on search results past the first five than participants in the High-Trust group.



S4 Fig. Time spent on search result pages as a function of page number (low-compliance participants). Participants in both trust groups spent less time on search result pages past the first page. Participants in the Low-Trust group spent less time on the first page of search results than participants in the High-Trust group did. Participants in the Low-Trust group also spent more time on search result pages past the first page than participants in the High-Trust group did. Error bars show standard error of the mean.



S5 Fig. Cumulative clicks on search results per page as a function of page number (low-compliance participants). Participants in both trust groups were less likely to click on search results on pages past the first page. Participants in the Low-Trust group were less likely to click on search results on the first page than participants in the High-Trust group. Participants in the Low-Trust group were also more likely to click on search results past the first page than participants in the High-Trust group. Error bars show standard error of the mean.



S1 Table. Vote proportion in each SEME condition pre- and post-search (high-compliance participants).

		Pro-Cameron Bias	Pro-Miliband Bias	Control	X^2	*p*
Low-Trust	**Vote Proportion**†					
	Pre-Search	0.39	0.28	0.41	1.29	0.52 NS
	Post-Search	0.55	0.25	0.39	5.93	0.05 NS
High-Trust	**Vote Proportion**†					
	Pre-Search	0.53	0.39	0.39	4.60	0.10 NS
	Post-Search	0.66	0.14	0.41	53.41	< 0.001

† Vote proportion is the number of votes for David Cameron divided by the total number of votes placed.

S2 Table. **VMP percentages, search times, and results clicked by Trust Group (low-compliance participants).**

Condition	**VMP (*p*)**	**Mean Search Time (sec) (*SD*)** †	**Mean No. of Results Clicked (*SD*)** †
Low-Trust	59.5 (< 0.001)	354.2 (258.7)	7.0 (3.8)
High-Trust	47.4 (0.004)	314.8 (261.6)	5.6 (3.5)
Diff (%)	-20.3	+11.1	-20.0
Statistic	$z = 1.19$	$t(194) = -0.96$	$t(194) = -2.46$
p	0.23 NS	0.34 NS	0.02

Note. McNemar's test was used to assess VMP significance. VMP is the percent increase in subjects in the bias groups (combined) who said that they would vote for the favored candidate. †These calculations were based on data from all three groups: that is, the two bias groups and the control group. VMP is calculated using data from the two bias groups only. Note that all *t*-tests employed in this study are two-tailed.

S3 Table. Vote proportion in each SEME condition pre- and post-search (low-compliance participants).

		Pro-Cameron Bias	Pro-Miliband Bias	Control	X^2	*p*
Low-Trust	**Vote Proportion†**					
	Pre-Search	0.45	0.64	0.38	6.77	0.03
	Post-Search	0.55	0.24	0.42	9.17	0.01
High-Trust	**Vote Proportion†**					
	Pre-Search	0.62	0.42	0.48	1.19	0.55 NS
	Post-Search	0.85	0.11	0.52	17.81	< 0.001

† Vote proportion is the number of votes for David Cameron divided by the total number of votes placed.

S4 Table. High- and Low-Trust differences in VMP percentages by levels of compliance.

	Compliance	*n**	VMP (*p*)
Low-Trust			
	Low	91	59.5 (< 0.001)
	High	63	17.1 (0.21 NS)
z			5.24
p			< 0.001
High-Trust			
	Low	32	47.4 (0.004)
	High	185	34.6 (< 0.001)
z			1.39
p			0.16 NS

Note. McNemar's test was used to assess VMP significance.
**n* is the total number of participants in the bias groups combined.

S5 Table. Pre-search opinion ratings of David Cameron and Ed Miliband (high-compliance participants).

		Pro-Cameron	Pro-Miliband	Control	*H*	*p*
Cameron	**Impression**	6.82 (1.81)	6.46 (1.94)	6.69 (1.89)	2.02	0.36 NS
	Trust	5.82 (2.03)	5.54 (2.28)	5.65 (2.00)	1.34	0.51 NS
	Likeability	6.65 (1.83)	6.24 (1.97)	6.57 (1.97)	2.72	0.26 NS
Miliband	**Impression**	6.95 (1.87)	7.20 (1.62)	7.36 (1.61)	2.38	0.31 NS
	Trust	5.93 (1.95)	6.10 (2.21)	6.17 (1.88)	0.97	0.62 NS
	Likeability	6.79 (1.86)	6.97 (1.72)	7.15 (1.77)	1.86	0.40 NS

S6 Table. Pre-search opinion ratings of David Cameron and Ed Miliband (low-compliance participants).

		Pro-Cameron	Pro-Miliband	Control	*H*	*p*
Cameron	**Impression**	7.00 (2.33)	7.38 (1.79)	6.97 (2.03)	0.97	0.62 NS
	Trust	5.97 (2.44)	6.41 (2.09)	5.78 (2.21)	2.56	0.28 NS
	Likeability	6.55 (2.09)	7.16 (1.72)	6.84 (1.97)	2.70	0.26 NS
Miliband	**Impression**	7.50 (1.74)	7.46 (1.78)	7.47 (1.48)	0.15	0.93 NS
	Trust	6.39 (2.26)	6.23 (2.06)	6.32 (2.04)	0.50	0.78 NS
	Likeability	7.03 (1.90)	7.28 (1.75)	7.34 (1.61)	0.71	0.70 NS

S7 Table. Post-search opinion ratings of David Cameron and Ed Miliband (high-compliance participants).

		Pro-Cameron	Pro-Miliband	Control	*H*	*p*
Cameron	**Impression**	6.56 (2.12)	3.56 (2.14)	5.04 (2.29)	88.08	< 0.001
	Trust	5.56 (2.34)	3.29 (2.07)	4.37 (2.19)	56.10	< 0.001
	Likeability	6.14 (2.20)	3.44 (2.12)	4.93 (2.33)	73.75	< 0.001
Miliband	**Impression**	4.91 (2.19)	7.25 (2.06)	5.81 (2.27)	62.96	< 0.001
	Trust	4.30 (2.29)	6.51 (2.24)	5.21 (2.44)	48.64	< 0.001
	Likeability	5.19 (2.21)	7.16 (2.11)	5.77 (2.18)	51.25	< 0.001

S8 Table. Post-search opinion ratings of David Cameron and Ed Miliband (low-compliance participants).

		Pro-Cameron	Pro-Miliband	Control	*H*	*p*
Cameron	**Impression**	6.65 (2.56)	3.82 (2.27)	5.42 (2.36)	34.83	< 0.001
	Trust	5.85 (2.59)	3.54 (2.43)	4.81 (2.56)	22.72	< 0.001
	Likeability	6.26 (2.44)	4.11 (2.41)	5.40 (2.44)	22.19	< 0.001
Miliband	**Impression**	5.08 (2.45)	7.21 (2.43)	5.79 (2.19)	25.23	< 0.001
	Trust	4.55 (2.19)	6.36 (2.17)	5.22 (2.47)	18.96	< 0.001
	Likeability	5.29 (2.34)	6.85 (2.22)	5.92 (2.36)	14.11	0.001

S9 Table. Pre- and post-search opinion ratings of favored and non-favored candidates (high-compliance participants, bias groups only).

		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†**
Low-Trust	**Impression**	7.14 (1.64)	6.63 (2.03)	-0.51	6.95 (1.81)	4.49 (2.39)	-2.46	-4.16***
	Trust	5.83 (2.32)	5.68 (2.22)	-0.15	5.81 (2.20)	3.97 (2.25)	-1.84	-3.88***
	Likeability	6.86 (1.61)	6.48 (2.24)	-0.38	6.73 (1.79)	4.41 (2.47)	-2.32	-4.03***
High-Trust	**Impression**	7.00 (1.74)	7.05 (2.13)	0.05	6.58 (1.95)	4.03 (2.21)	-2.55	-8.59***
	Trust	6.03 (2.07)	6.24 (2.35)	0.21	5.67 (2.14)	3.64 (2.21)	-2.03	-7.71***
	Likeability	6.82 (1.83)	6.80 (2.19)	-0.02	6.39 (1.98)	4.13 (2.27)	-2.26	-7.80***

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate

***$p < 0.001$

S10 Table. Pre- and post-search opinion ratings of favored and non-favored candidates (low-compliance participants, bias groups only).

		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†**
Low-Trust	**Impression**	7.22 (2.09)	6.66 (2.51)	-0.56	7.51 (1.80)	4.68 (2.44)	-2.83	-5.05***
	Trust	6.08 (2.18)	5.86 (2.33)	-0.22	6.47 (2.14)	4.31 (2.35)	-2.16	-3.78***
	Likeability	6.82 (1.94)	6.35 (2.31)	-0.47	6.99 (1.82)	4.99 (2.46)	-2.00	-4.70***
High-Trust	**Impression**	7.25 (2.09)	7.69 (2.35)	0.44	7.25 (1.67)	3.81 (2.32)	-3.44	-4.33***
	Trust	6.16 (2.50)	6.81 (2.48)	0.65	6.19 (2.29)	3.31 (2.26)	-2.88	-4.15***
	Likeability	7.16 (2.02)	7.13 (2.38)	-0.03	7.41 (1.78)	3.91 (2.20)	-3.50	-4.25***

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate

***$p < 0.001$

S11 Table. Pre- and post-search opinion ratings of David Cameron and Ed Miliband (high-compliance participants, control group only).

		David Cameron Mean (SD)			Ed Miliband Mean (SD)			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	$z^{\dagger}$ (p)
Low-Trust	**Impression**	6.38 (2.07)	4.59 (2.00)	-1.79	7.19 (1.63)	5.59 (2.51)	-1.60	-1.19 (0.24 NS)
	Trust	5.35 (1.89)	4.19 (2.30)	-1.16	5.97 (1.66)	5.11 (2.57)	-0.86	-1.34 (0.18 NS)
	Likeability	6.30 (2.07)	4.51 (2.41)	-1.79	6.92 (1.82)	5.54 (2.23)	-1.38	-1.38 (0.17 NS)
High-Trust	**Impression**	6.86 (1.77)	5.27 (2.41)	-1.59	7.46 (1.60)	5.93 (2.15)	-1.53	-0.15 (0.88 NS)
	Trust	5.81 (2.05)	4.47 (2.15)	-1.34	6.27 (1.99)	5.26 (2.39)	-1.01	-0.73 (0.47 NS)
	Likeability	6.71 (1.91)	5.16 (2.28)	-1.55	7.27 (1.75)	5.89 (2.16)	-1.38	-0.36 (0.72 NS)

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for David Cameron to the post-minus-pre ratings for Ed Miliband

S12 Table. Pre- and post-search opinion ratings of David Cameron and Ed Miliband (low-compliance participants, control group only).

		David Cameron Mean (SD)			Ed Miliband Mean (SD)			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†** (*p*)
Low-Trust	**Impression**	7.00 (2.08)	5.19 (2.46)	-1.81	7.67 (1.42)	5.81 (2.29)	-1.86	-0.46 (0.65 NS)
	Trust	6.02 (2.22)	4.71 (2.63)	-1.31	6.50 (1.95)	5.38 (2.50)	-1.12	-0.18 (0.86 NS)
	Likeability	6.92 (2.04)	5.15 (2.51)	-1.77	7.44 (1.66)	6.02 (2.43)	-1.42	-0.59 (0.55 NS)
High-Trust	**Impression**	6.92 (1.98)	5.88 (2.15)	-1.04	7.08 (1.55)	5.76 (2.03)	-1.32	-0.41 (0.68 NS)
	Trust	5.32 (2.16)	5.00 (2.43)	-0.32	5.96 (2.21)	4.92 (2.45)	-1.04	-1.22 (0.22 NS)
	Likeability	6.68 (1.87)	5.88 (2.26)	-0.80	7.16 (1.52)	5.72 (2.25)	-1.44	-0.93 (0.35 NS)

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate

S13 Table. Mean difference comparisons between High- and Low-Trust groups for pre- and post-search opinion ratings of favored and non-favored candidates (high-compliance participants, bias groups only).

	Favored Candidate Pre/Post Mean Difference			Non-Favored Candidate Pre/Post Mean Difference		
	Impression	**Trust**	**Likeability**	**Impression**	**Trust**	**Likeability**
Low-Trust	-0.51	-0.15	-0.38	-2.46	-1.84	-2.32
High-Trust	0.05	0.21	-0.02	-2.55	-2.03	-2.26
Diff	+0.56	+0.36	+0.36	-0.09	-0.19	-0.13
U	4,714.5	5,261.0	5,194.5	5,780.0	5,577.5	5,683.5
p	0.02	0.24 NS	0.19 NS	0.92 NS	0.61 NS	0.77 NS

S14 Table. Mean difference comparisons between High- and Low-Trust groups for pre- and post-search opinion ratings of favored and non-favored candidates (low-compliance participants, bias groups only).

	Favored Candidate Pre/Post Mean Difference			Non-Favored Candidate Pre/Post Mean Difference		
	Impression	**Trust**	**Likeability**	**Impression**	**Trust**	**Likeability**
Low-Trust	-0.56	-0.22	-0.47	-2.83	-2.16	-2.00
High-Trust	0.44	0.65	-0.03	-3.44	-2.88	-3.50
Diff	+1.00	+0.87	+0.44	-0.61	-0.72	-1.50
U	1,127.0	1,126.5	1,358.5	1,223.0	1,204.0	984.5
p	0.06 NS	0.05 NS	0.57 NS	0.18 NS	0.14 NS	0.006

References

Advanced Web Ranking. (n.d.). *Google organic CTR history*. Retrieved June 4, 2022, from https://www.advancedwebranking.com/ctrstudy/

Beres, D. (2019, July 10). *The race to predict our behavior: Tristan Harris testifies before Congress*. Big Think. https://bigthink.com/the-present/tristan-harris-attention/

Chigne, J.-P. (2018, May 18). *Google's leaked internal video 'The Selfish Ledger' shows how a population could be controlled by data*. Tech Times. https://www.techtimes.com/articles/228053/20180518/googles-leaked-internal-video-the-selfish-ledger-shows-how-a-population-could-be-controlled-by-data.htm

Chitika Insights. (2013*). The value of Google results positioning.* http://info.chitika.com/uploads/4/9/2/1/49215843/chitikainsights-valueofgoogleresultspositioning.pdf

Cutrell, E., & Guan, Z. (2007). What are you looking for? An eye-tracking study of information usage in web search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 407–416. https://doi.org/10.1145/1240624.1240690

Dean, B. (n.d.). We analyzed 5 million Google search results. Here's what we learned about organic CTR. *Backlinko*. Retrieved August 27, 2022 from https://backlinko.com/google-ctr-stats

Epstein, R. (2014, May). Google's snoops: Mining our private data for profit and pleasure. *Dissent*. https://www.dissentmagazine.org/online_articles/googles-snoops-mining-our-data-for-profit-and-pleasure

Epstein, R. (2018b, May 17). Taming Big Tech: The case for monitoring. *Hacker Noon.* https://hackernoon.com/taming-big-tech-5fef0df0f00d

Epstein, R., Bock, S., Peirson, L., & Wang, H. (2021, June 14–15). *Large-scale monitoring of Big Tech political manipulations in the 2020 Presidential election and 2021 Senate runoffs, and why monitoring is essential for democracy* [Paper presentation]. American Association of Behavioral and Social Sciences 24th Annual Meeting. Virtual.

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences USA, 112*(33), E4512–E4521. https://doi.org/10.1073/pnas.1419828112

European Commission. (2019). *Antitrust: Commission fines Google 1.49 billion Euros for abusive practices in online advertising*. https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1770.

Ghosal, A. (2015, August 31). *Google could face a $1.4b fine in India for rigging search results*. The Next Web. http://thenextweb.com/google/2015/08/31/google-could-face-a-1-4b-fine-in-india-for-rigging-search-results/

Granka, L., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 478–479. https://doi.acm.org/10.1145/1008992.1009079

Haas, A., & Unkel, J. (2017). Ranking versus reputation: Perception and effects of search result credibility. *Behaviour & Information Technology, 36*(12), 1285–1298. https://doi.org/10.1080/0144929X.2017.1381166

Hahnel, C., Goldhammer, F., Krohne, U., & Naumann, J. (2018). The role of reading skills in the evaluation of online information gathered from search engine environments. *Computers in Human Behavior*, *78*, 223–234. https://doi.org/10.1016/j.chb.2017.10.004

Haider, J., & Sundin, O. (2020). Information literacy challenges in digital culture: Conflicting engagements of trust and doubt. *Information, Communication & Society, 25*(8), 1176–1191. https://doi.org/10.1080/1369118X.2020.1851389

Han, B., Shah, C., & Saelid, D. (2021). Users' perception of search engine biases and satisfaction. In L. Boratto, S. Faralli, M. Marras, & G. Stilo (Eds.), *Communications in computer and information science* (pp. 14–24). Springer. https://doi.org/10.48550/arXiv.2105.02898

Hern, A. (2019, July 13). Google whistleblower launches project to keep tech ethical. *The Guardian*. https://www.theguardian.com/world/2019/jul/13/google–whistleblower–launches–project–to–keep–tech–ethical

Hsu, T., & Kang, C. (2020, June 9). Morally impossible: Some advertisers take a timeout from Facebook. *The New York Times*. https://www.nytimes.com/2020/06/09/business/media/facebook–advertisers–trump–zuckerberg.html

Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F., & Gay, G. (2007). Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transaction on Information Systems, 25*(2). https://doi.org/10.1145/1229179.1229181

Kammerer, Y., & Gerjets, P. (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), 177–191. https://doi.org/10.1080/10447318.2013.846790

Lee, N., Resnick, P., & Barton, G. (2019, May 22). *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. The Brookings Institute.

https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

Lewis, P. (2017 October 6). 'Our minds can be hijacked': The tech insiders who fear a smartphone dystopia. *The Guardian*. https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia

Liu, J., & Zhang, Z. (2019). The role of domain knowledge in document selection from search results. *Journal of the Association for Information Science and Technology, 70*(11), 1236–1247. https://doi.org/10.1002/asi.24199

Lohr, S., & Streitfeld, D. (2012, April 30). Data engineer in Google case is identified. *The New York Times*. https://www.nytimes.com/2012/05/01/technology/engineer-in-googles-street-view-is-identified.html

Lorigo, L., Haridasan, M., Brynjarsdottir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., & Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology, 59*(7), 1041–1052. https://doi.org/10.1002/asi.20794

McKinnon, J., & MacMillan, D. (2018, September 20). Google workers discussed tweaking search function to counter travel ban. *The Wall Street Journal*. https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472

Metaxa, D., Park, J. S., Landay, J. A., & Hancock, J. (2019). Search media and elections: A longitudinal investigation of political search results in the 2018 US elections. *Proceedings* *of the ACM on Human-Computer Interaction, 3*(CSCW), 1–17. https://doi.org/10.1145/3359231

Mullins, B., Winkler, R., & Kendall, B. (2015, March 19). Inside the US antitrust probe of Google: Key FTC staff wanted to sue internet giant after finding 'real harm to consumers and to innovation.' *The Wall Street Journal*. https://www.wsj.com/articles/inside-the-u-s-antitrust-probe-of-google-1426793274

Murphy, J., Hofacker, C., & Mizerski, R. (2017). Primacy and recency effects on clicking behavior. *Journal of Computer-Mediated Communication, 11*(2), 522–535. https://doi.org/10.1111/j.1083-6101.2006.00025.x

Mustafaraj, E., Lurie, E., & Devine, C. (2020). The case for voter centered audits of search engines during political elections. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 559–569. https://doi.org/10.1145/3351095.3372835

Nunez, M. (2016, May 9). *Former Facebook workers: We routinely suppressed conservative news.* Gizmodo. https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006

Pan, B., Joachims, T., Granka, L., Hembrooke, H., Gay, G., & Lorigo, L. (2007). In Google we trust: User's decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication, 12*(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Rainie, L., & Anderson, J. (2017, February 8). *Code-dependent: Pros and cons of the algorithm age.* Pew Research Center. https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/

Reuters Staff. (2016, August 11). *South Korea says investigating whether Google broke antitrust laws*. Reuters. https://www.reuters.com/article/us-southkorea-google-idUSKCN10N08N

Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the web for conflicting topics: Page and user factors. *Computers in Human Behavior, 29,* 2161–2171. https://doi.org/10.1016/j.chb.2013.04.034

Schwartz, O. (2019, November 25). In 2016, Microsoft's racist chatbot revealed the dangers of online conversation. *IEEE Spectrum*. https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation

Statt, N. (2018, October 10). *Leaked Google research shows company grappling with censorship and free speech.* The Verge. https://www.theverge.com/2018/10/10/17961806/google-leaked-research-good-censor-censorship-freedom-of-speech-research-china

Steiner, M., Magin, M., Stark, B., & Geiß, S. (2020). Seek and you shall find? A content analysis on the diversity of five search engines' results on political queries. *Information, Communication & Society, 25*(2), 217–241. https://doi.org/10.1080/1369118X.2020.1776367

Walhout, J., Oomen, P., Jarodzka, H., & Brand-Gruwel, S. (2017). Effects of task complexity on online search behavior of adolescents. *Journal of the Association for Information Science and Technology, 68*(6), 1449–1461. https://doi.org/10.1002/asi.23782

APPENDIX VI

The Search Suggestion Effect (SSE): How autocomplete search suggestions can be used to impact opinions and votes

Robert Epstein[*], Savannah Aries, Kelly Grebbien, Alyssa M. Salcedo, and Vanessa R. Zankich

American Institute for Behavioral Research and Technology, Vista, CA 92084

[*]Corresponding author. Email: re@aibrt.org

Robert Epstein (PhD, Harvard University) is Senior Research Psychologist at the American Institute for Behavioral Research and Technology.

Savannah Aries (BA, University of California, Los Angeles) is a research intern at the American Institute for Behavioral Research and Technology.

Kelly Grebbien (BA, California Baptist University) is a graduate student at Liberty University.

Alyssa M. Salcedo (BA, California Baptist University) is a graduate student at the University of California Los Angeles.

Vanessa R. Zankich, (BS, University of California, San Diego) is Research Director at the American Institute for Behavioral Research and Technology.

Word count: 8,908

ABSTRACT

Can autocomplete search suggestions – the words or phrases generated by a search engine as people are typing a search term – influence opinions and votes? Previous research has shown that search results that favor one political candidate can have a rapid and substantial impact on the opinions and voting preferences of undecided voters. News reports in 2016 suggested that a leading search engine was suppressing negative search suggestions for one US Presidential candidate but not for her opponent. We conducted a progressive series of five randomized, controlled, counterbalanced, double-blind experiments to determine what effect differential suppression of this type might have on voters. We found that negative suggestions attract far more clicks than neutral or positive ones, consistent with extensive research on negativity bias, and that the differential suppression of negative search suggestions can turn a 50/50 split among undecided voters into more than a 90/10 split favoring the candidate for whom negative search suggestions were suppressed. We conclude that differentially suppressing negative search suggestions can have a dramatic impact on the opinions and voting preferences of undecided voters, potentially shifting a large number of votes without people knowing and without leaving a paper trail for authorities to trace.

Keywords: search suggestion effect; SSE; autocomplete; online influence; search engine manipulation effect; negativity bias

The Search Suggestion Effect (SSE): How autocomplete search suggestions can be used to impact opinions and votes

Introduction

On June 9, 2016, an American news organization called SourceFed posted a 7-minute video on YouTube which claimed that when people used search engines to find information about Presidential candidates, the leading search engine company (Google) showed substantially different search suggestions than were being shown on other search engines (Flores, 2016; Supplemental Text S1). The video stated that when people began typing neutral search terms about Hillary Clinton, the Bing and Yahoo search engines generated lists of suggestions that included many negative terms. Using the same search terms on Google rarely, if ever, produced negative search suggestions for Clinton. In response to the phrase "hillary clinton is," for example, Yahoo generated a list of 10 negative suggestions, among them "hillary clinton is a liar" and "hillary clinton is the devil." Bing generated a list of 8 suggestions, 7 of which were negative, among them "hillary clinton is a filthy liar" and "hillary clinton is a murderess." In response to the same phrase, Google offered only two suggestions: "hillary clinton is winning" and "hillary clinton is awesome" (Figure 1).

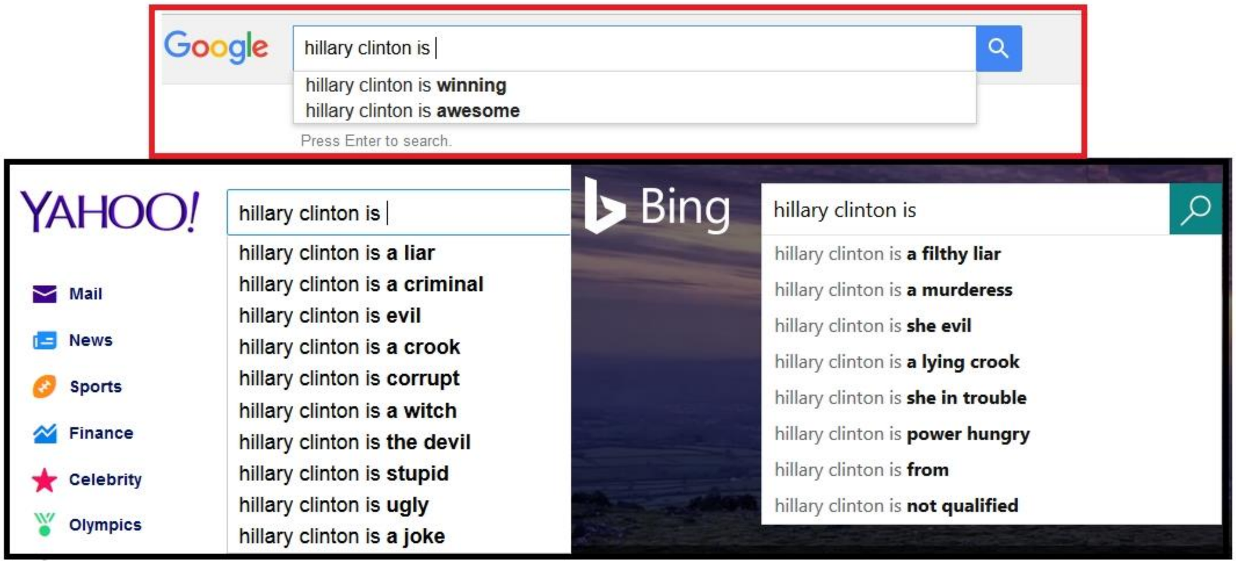


Figure 1. "Hillary Clinton is" on three search engines, showing that negative suggestions appear to be suppressed on Google.com but not on the other search engines. Image composed from three screenshots obtained on August 3, 2016.

Google denies deliberately altering search suggestions to favor any particular political candidate (Goldman, 2016; Nicas & Andrews, 2016) and has publicly stated that search suggestions "will not show a predicted query that is offensive or disparaging when displayed in conjunction with a person's name" (Goldman, 2016). Exceptions to this rule are easy to find, however (Harrington & Scher, 2016; Napolitano, 2016; Seitz, 2016; Figure S1). For example, in response to "trump is," on July 9, 2020, Google.com suggested "trump is losing" and "trump is losing election" (Figure S2), and on July 27, 2020, the company suggested "trumpisanidiot" and "trumpisalaughingstock" (Figure S3).

It is not our intention in this paper to weigh in on the debate about how favoritism might come to exist in search suggestions. Rather, we are posing and attempting to answer the following questions about the power that search suggestions might have to alter the opinions or votes of users: (1) How might voters have been impacted in the summer of 2016 by negative search suggestions – or the lack thereof – shown for different political candidates? (2) How might one structure search suggestions if one deliberately wanted to alter the opinions or votes of users? (3) Is there an optimal number of search suggestions for influencing people? (4) Can people's opinions or voting preferences be altered by search suggestions without their knowledge? (5) Who is most vulnerable to manipulation of this sort?

A brief history of autocomplete search suggestions

Autocomplete search suggestions (sometimes called "autofill suggestions" or "query suggestions") were developed in 2004 by Google software engineer Kevin Gibbs (Garber, 2013). In Gibbs's blogpost to debut this feature, he wrote, "We've found that Google Suggest not only makes it easier to type in your favorite searches…, but also gives you a playground to explore what others are searching about, and learn about things you haven't dreamt of" (Gibbs, 2004). In the early years, it is likely that search suggestions mainly showed people what many or most other people were searching for, and the company still claims that information about other people's searches play a role in the generation of suggestions (Yehoshua, 2016). That said, the SourceFed video not only showed screenshots of search suggestions, it also used Google Trends to determine what terms and phrases people were actually searching for on Google. The video showed several instances in which terms that were searched for frequently were not being offered as search suggestions and, conversely, instances in which suggested terms were not being searched for (Flores, 2016; Supplemental Text S1).

Initially, autocomplete was an opt-in feature of Google's search engine, but in 2008, this feature was activated for all users with no way to opt out (Gibbs, 2004). During the early years of autocomplete's existence, users were almost always shown 10 search suggestions. In 2010, however, the list was shortened to four suggestions for most searches, at least on desktop and laptop computers (Epstein, 2020; Stack Exchange, 2010). On approximately October 3, 2017, the length of the list increased again to 10 on laptops and desktops, and this occurred with no comment from Google (Schwartz, 2017). On mobile devices, which started to become commonplace around 2007 (Pothitos, 2016), Google has typically shown five search suggestions. As of this writing (July 19, 2023), the company is continuing to show 10 suggestions on laptop and desktop computers and 8, 5, or 3 on mobile devices.

Changes in both the content and number of suggestions over the years were likely made with specific goals in mind, but one can only speculate about the nature of those goals. There seems to be little doubt, however, that what began as an innocuous tool for helping users has gradually morphed into a tool the purpose of which is often unclear.

The power of autocomplete

Where a search result is listed on a search engine results page has long been of interest to marketers, because the higher a result, the more clicks it is likely to attract, which generally results in more income (Klapdor, Anderl, Waingenheim, & Schumann, 2014; cf. Hong & Kim, 2018). Moving up one position in search results can increase click-through rate (CTR) by 32.3% (Dean, 2022). For many businesses, that increase results in a proportional increase in sales (Ramaboa & Fish, 2018). 50% of all clicks go to the top two search results, and 95% of clicks occur on the first page of results (Advanced Web Ranking, 2020; Chitika Insights, 2013; Dean, 2022). An entire industry – the Search Engine Optimization (SEO) industry – has emerged over the past two decades simply to help companies override search algorithms in order to boost one's rank in search results by a notch or two (Granka, Joachims, & Gay, 2004).

After autocomplete was introduced, it didn't take long for marketers to figure out that search suggestions could also be manipulated to some extent, and that the higher a company or product was in the list of suggestions, the more money the company or vendor would make

(Ramaboa & Fish, 2018). In other words, there is an order effect for search suggestions, just as there is for search results (Epstein & Robertson, 2015; Ramaboa & Fish, 2018), although the order effect for search results is much larger (Granka et al., 2004). Search suggestions are especially important for marketers in part because people click one of those suggestions in about 23 percent of the searches they initiate (Dean, 2020).

But why suppress negative search suggestions for one political candidate? In particular, how might the suppression of negative search suggestions for one candidate have impacted voting in the 2016 Presidential election in the United States?

Negativity bias

The phrase "negative search suggestion," is vague, but linguists have developed fairly precise ways of characterizing various properties of words and phrases (Hannan et al., 2009). Loosely speaking, what we keep referring to as "negative" terms corresponds to what linguists label "low-valence" terms (Shuman, Sander, & Scherer, 2013). What is there about low-valence terms that might cause a tech company to suppress them for a political candidate they support and to allow them to appear for the opposing candidate?

Researchers in multiple fields have long studied a phenomenon called "negativity bias," which is sometimes referred to as "the cockroach in the salad" effect (Carretié, Mercado, Tapia, & Hinojosa, 2001). A wide variety of negative, unpleasant, or threatening stimuli affect people in predictable ways. Generally speaking, they (a) draw more attention than neutral or positive stimuli (Robertson et al., 2023), (b) lead to more vivid or lasting memories than neutral or positive stimuli, (c) elicit stronger emotions than neutral or positive stimuli, and generally "have greater effects than positive factors across a wide range of psychological phenomena" (Johnson & Tierney, 2019). If one finds a cockroach in one's salad, it not only draws one's attention, it also ruins the entire salad. The inverse does not work; that is, adding a piece of chocolate to a plate of sewage does not make the sewage more appealing. Could negativity bias help explain the differential suppression of low-valence search suggestions that apparently took place in the months leading up to the 2016 Presidential election?

We attempted to answer this question in a progressive series of five experiments that examined how valence, order, and the number of search suggestions people are shown impact both people's search behavior and the opinions and voting preferences they form. The first four experiments were exploratory; they allowed us in a simple and progressive fashion to learn more about search suggestions. The fifth experiment combines what the previous experiments taught us about search suggestions – in particular, the roles that valence, list length, and position play in determining people's selections – with what is known about the power that search results have to impact people's opinions and voting preferences (Epstein & Robertson, 2015; Guess, Barberá, Munzert, & Yang, 2021). In other words, it is in the fifth experiment that we directly measure the power that search suggestions have to influence the opinions and votes of undecided voters.

Procedures and results

Exempt approval

Exempt approval for our study was granted by the institutional review board (IRB) of the sponsoring institution, the American Institute for Behavioral Research and Technology (AIBRT). AIBRT is registered with the HHS Office for Human Research Protections (OHRP) under IORG0007755. Our IRB is registered with OHRP under IRB00009303, and the Federalwide Assurance number for our IRB is FWA00021545. The study also qualified for a waiver of informed consent, in part to preserve the anonymity of participants (see HHS Federal

Regulations 45 CFR 46.101.(b)(4), 45 CFR 46.116(d), 45 CFR 46.117(c)(2), and 45 CFR 46.111).

Recruitment, cleaning, and design

Participants were recruited from Amazon Mechanical Turk (MTurk) (Sheehan, 2017) through a company called CloudResearch, which screens out bots and suspect participants. Although using MTurk to collect data is common in the social sciences, we recognize its limitations. MTurk workers are less politically diverse, more highly educated, younger, and less religious than the general US population (Litman, 2017; Moss & Litman, 2020). In addition, researchers have raised concerns about the validity of MTurk data due to participant inattention, self-misrepresentation, and social desirability bias (Aguinis et al., 2020).

Participants were screened to be eligible to vote in the US. Data were cleaned to remove people who indicated that their English fluency was below 6 on a scale from 1 to 10 where 1 was labeled “not fluent” and 10 was labeled “highly fluent.” Demographic characteristics of the participants in all five experiments are summarized in Table S1. All five experiments were randomized, controlled, counterbalanced, and double-blind in design.

Experiment 1: negative (low-valence) search terms

Participants and methods

For the first experiment, data were collected between July 29 and October 17, 2016. We showed a diverse group of 609 participants (after cleaning) four sets of search suggestions (four suggestions per set, plus a fifth option that allowed participants to type their own suggestion) – two sets for the search term “Tim Kaine” (Hillary Clinton’s running mate in 2016) and two sets for the search term “Mike Pence” (Donald Trump’s running mate in 2016). In each pair of search suggestion sets, only one contained a low-valence suggestion (either “Tim Kaine scandal” or

"Mike Pence scandal"). The positions of the low-valence suggestions were fixed: "Tim Kaine scandal" and "Mike Pence scandal" always appeared in the third and second positions, respectively. In the corresponding set of search suggestions without the low-valence suggestion, a control suggestion appeared (either "Tim Kaine office" or "Mike Pence office"). The control suggestions were fixed in the same positions as their low-valence counterparts. All four sets of search suggestions were presented on the same web page, one beneath the other, also in a fixed order (Figure S4), and participants were asked to click on one suggestion in each set or to type their own suggestion. Figure S5 shows the experimental procedure in diagrammatic form.

Results

Consistent with research on negativity bias, the negative search suggestions attracted 41.6% of clicks, more than twice as many as one would expect by chance (20%, given five possible responses; $z = 11.55$, $p < 0.001$). In addition, participants clicked the negative suggestions 5.8 times as often as they clicked the corresponding control suggestions ($z = 19.76$, $p < 0.001$). Undecided voters, however, clicked the negative search suggestions 14.8 times as often as they clicked the corresponding control suggestions (Table S2; $z = 7.98$, $p < 0.001$).

The response pattern also demonstrated confirmation bias. Liberals clicked the negative search suggestions (for both candidates combined) 38.1% of the time. However, they clicked "Pence scandal" 53.7% of the time (Pence was the conservative candidate) and "Kaine scandal" only 22.4% of the time (Kaine was the liberal candidate) ($z = 3.73$, $p < 0.001$). Conservatives clicked the negative search suggestions (for both candidates combined) 28.9% of the time. However, they clicked "Kaine scandal" 36.8% of the time and "Pence scandal" only 21.1% of the time ($z = 1.84$, $p = 0.06$ NS).

Experiment 2: position effect

Participants and methods

The data for Experiment 2 were collected between October 20, 2016 and February 15, 2017. In this experiment, we sought to determine whether the position of the low-valence suggestion affected the number of times it was clicked. A diverse group of 1,126 participants was shown four sets of search suggestions, two for the search term "Tim Kaine" and two for the search term "Mike Pence." In each pair of search suggestion sets, only one contained a negative search suggestion ("Tim Kaine scandal" or "Mike Pence scandal"), but this time, the position of the negative suggestion varied randomly among the first four search suggestions of its set; the other suggestions did not change position. Once again, the fifth option – always last – allowed participants to enter their own search term. Again, all four sets of search suggestions were presented on the same web page, one beneath the other in a fixed order (Figure S6), and participants were asked to click on one suggestion in each set or to type their own suggestion. Figure S7 shows the experimental procedure in diagrammatic form.

Results

Again, consistent with research on negativity bias, the negative suggestions attracted 45.0% of clicks, more than twice as many as one would expect by chance ($z = 17.91$, $p < 0.001$). In addition, participants clicked the negative suggestions 2.6 times as often as they clicked the corresponding control suggestions ($z = 19.83$, $p < 0.001$). The response pattern also demonstrated confirmation bias. Liberals clicked the negative search suggestions (for both candidates combined) 39.9% of the time. However, they clicked "Pence scandal" 48.7% of the time and "Kaine scandal" only 30.0% of the time ($z = 5.29$, $p < 0.001$). Conservatives clicked the negative search suggestions (for both candidates combined) 38.6% of the time. However, they clicked

“Kaine scandal” 51.0% of the time and “Pence scandal” only 24.0% of the time ($z = 5.03$, $p < 0.001$).

This experiment also yielded evidence of a position effect (Figure 2), with people clicking more on the negative suggestion when it was positioned higher in the list of search suggestions (standardized $\beta = -0.97$, $t = -5.96$, adjusted $r^2 = 0.92$, $p < 0.05$).

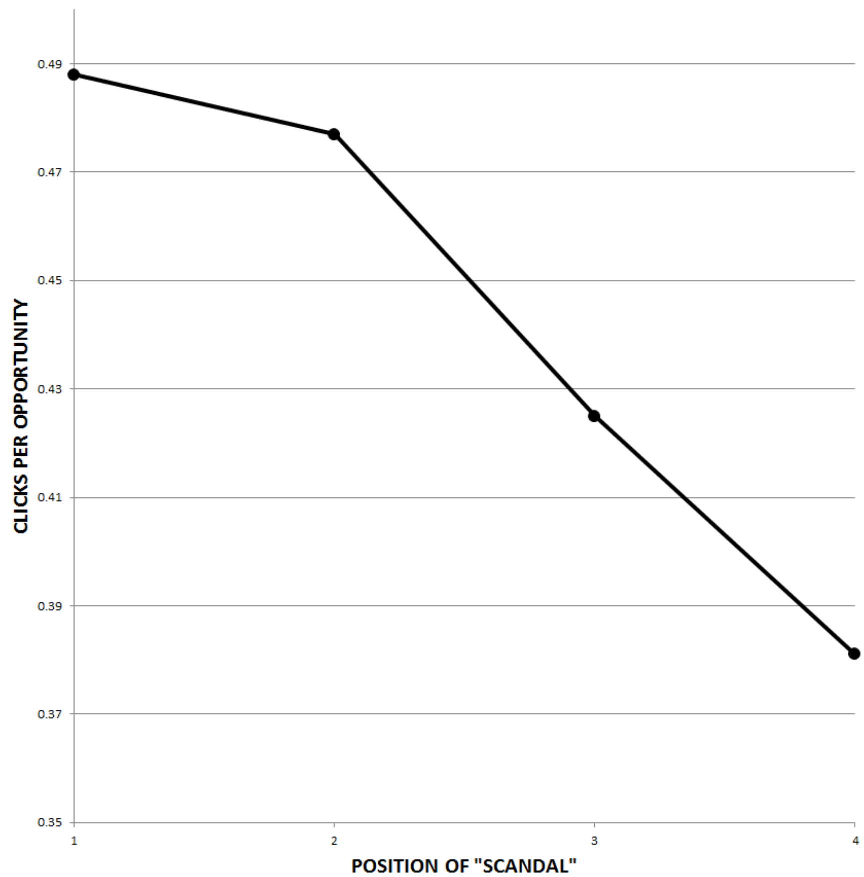


Figure 2. Experiment 2: Proportion of clicks to negative suggestion (“scandal”) as a function of position in the list of search suggestions. When “scandal” appeared in the first (highest) position, it was clicked nearly half the time – the lower its position in the list, the lower the proportion of clicks it received.

We also calculated the probability of clicks to each of the six different search suggestions (including “scandal”), along with the probability of clicking the fifth option, in which people entered their own search term. Click probabilities varied between 0.04 (to search terms containing the word “email”) and 0.45 (to search terms containing the word “scandal”) (Figure S8). The fact that different suggestions attracted different numbers of clicks suggested to us that this type of analysis would be more rigorous if we used some of the tools that linguists use to analyze language. We did so in Experiments 3 and 4.

Experiment 3: valence effect

Participants and methods

The data for Experiment 3 were collected between February 17 and July 16, 2017. A diverse group of 542 people participated in this experiment. We used a standard linguistics database (Warriner, Kuperman, & Brysbaert, 2013) to assess how valence – the extent to which a word is positive or negative (Shuman et al., 2013; Warriner et al., 2013) – affects which search suggestions people select. We controlled for two other measures of word properties that linguists use to characterize words: arousal level – the extent to which a word’s affect is calming or exciting – and frequency – a measure of the prevalence of a word in English (Kuperman, Estes, Brysbaert, & Warriner, 2014; Warriner et al., 2013). We used just one highly negative suggestion: “suicide,” which has a valence of 1.58 on a scale from 1.0 to 9.0. We also used eight other suggestions measured on the same scale, four with a moderate-valence level (M = 5.1, SD = 0.2) and four with a high-valence level (M = 7.2, SD = 0.2). All nine terms had similar arousal levels (M = 5.2, SD = 0.6) and frequencies (M = 17,407.3, SD = 1,456.4).

In this experiment, we also introduced a new variable: number of search suggestions. All participants were presented with three sets of search suggestions. Two of the sets contained four suggestions (and a fifth option – always appearing last – that allowed people to enter their own search term), and one set contained eight suggestions (and a ninth option – always appearing last – that allowed people to enter their own search term).

All display parameters were randomized (except, as noted above, the position of the option for entering one's own search term): The low-valence term appeared in all three lists, but its position varied. The ordering of the other search suggestions also varied; any of these suggestions could appear in any position in any of the three lists, although suggestions were not repeated within any list. The three lists also varied by number of search suggestions within each list (4-4-8, 4-8-4, or 8-4-4). For an example of one display, see Figure S9. Figure S10 shows the experimental procedure in diagrammatic form.

Results

We again found an effect for valence; the probability of clicking the low-valence suggestion was significantly higher than the probability of clicking either a moderate- or high-valence suggestion ($P_{\text{Low}} = 0.45$, $P_{\text{ModHigh}} = 0.10$, $z = 15.0$, $p < 0.001$). We also found, once again, position effects for the low-valence search suggestion in both the four-item (standardized $\beta = -0.80$, $t = 4.4$, adjusted $r^2 = 0.52$, $p < 0.05$) and eight-item lists ($\beta = -0.70$, $t = -2.6$, $r^2 = 0.42$, $p < 0.05$).

We also found that people were more likely to click "suicide" when it appeared in four-item lists (click probability = 0.48) than in eight-item lists (click probability = 0.40) ($z = 2.65$, $p < 0.01$). The findings from this experiment suggested a procedure for determining how to maximize control over people's searches. This matter is explored in Experiment 4.

Experiment 4: maximizing the control of search behavior using search suggestions

Participants and methods

Data were collected for Experiment 4 between February 28 and March 2, 2017. A diverse group of 302 participants were shown six sets of search suggestions: two sets of two suggestions, two sets of four suggestions, and two sets of eight suggestions (each with an extra item at the end which allowed participants to type their own search terms). One of two negative search suggestions – "execution" (valence = 2.28) or "racism" (valence = 1.48) – appeared in each list; three lists showed one negative suggestion, and three lists showed the other negative suggestion. All relevant features of the data were randomized (see Experiment 3). Four moderate-valence suggestions (M = 5.53, SD = 0.75) and four high-valence suggestions (M = 7.83, SD = 0.43) also appeared in the lists. See Figure S11 for a sample display. Figure S12 shows the experimental procedure in diagrammatic form.

Results

Once again, we found an effect for valence; the probability of clicking the low-valence suggestion was higher than the probability of clicking either a moderate- or high-valence suggestion ($P_{\text{Low}} = 0.37$, $P_{\text{ModHigh}} = 0.14$, $z = 10.7$, $p < 0.001$) (Figure 3).

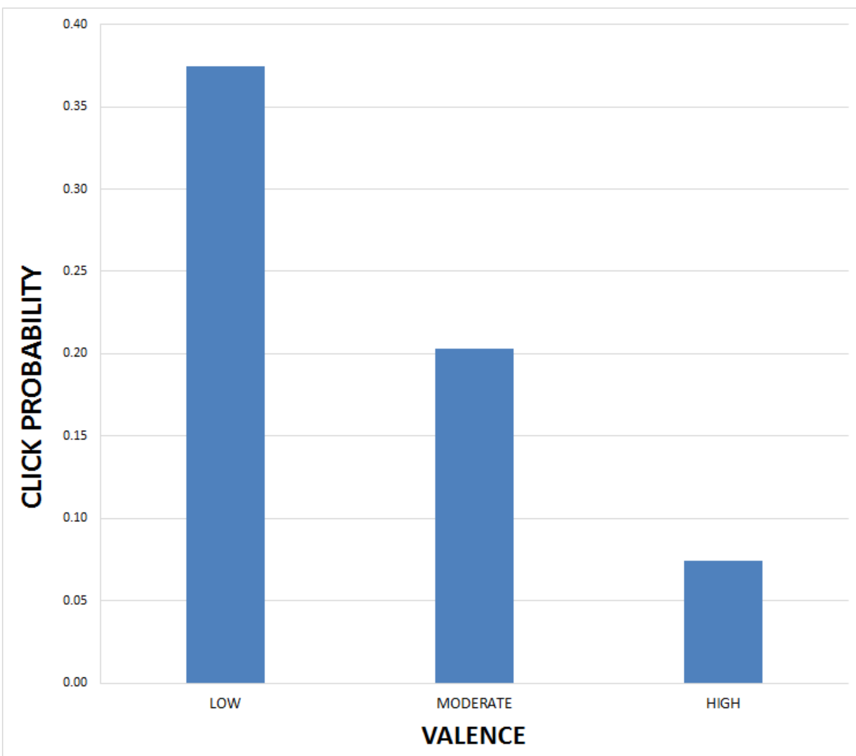


Figure 3. How click probability varies as a function of the valence levels of search suggestions in Experiment 4.

We also found two effects for list size. First, the longer the list, the lower the probability that people would click the negative search suggestion, no matter where it appeared in the lists of 2, 4, or 8 items ($P_2 = 0.51$, $P_4 = 0.39$, $P_8 = 0.23$, $\beta = -1.00$, $t = -12.1$, $r^2 = 0.99$, $p = 0.05$ NS). Similarly, the longer the list, the lower the probability that people would click the negative search suggestion when it appeared in the first position ($P_2 = 0.54$, $P_4 = 0.30$, $P_8 = 0.08$, $\beta = -1.0$, $t = -33.1$, $r^2 = 1.00$, $p < 0.05$).

And second, the longer the list, the higher the probability that people would click a search suggestion rather than typing their own search term, although this trend was small ($P_2 = 0.70$, $P_4 = 0.75$, $P_8 = 0.84$, $\beta = 0.99$, $t = 6.06$, $r^2 = 0.95$, $p = 0.10$ NS).

If one were trying to maximize the control that low-valence search suggestions might exert over search behavior, it would be useful to know the number of search suggestions that would maximize such control. If a list is too short, the likelihood that a user will type his or her own search term is increased, and if a list is too long, he or she is less likely to click, or perhaps even to notice, the low-valence suggestion. How long should a list of search suggestions be to balance these two tendencies – that is, to offer enough options to minimize the chance that the user will type his or her own search term and to maximize the chance that the user will notice and click the low-valence suggestion?

Finding that optimal list size is especially important if one is suppressing negative search suggestions for the cause or candidate one supports (easily done by checking the valence of search suggestions before displaying them) and allowing negative search suggestions to appear for the cause or candidate one opposes.

Figure 4 uses data from Experiment 4 to estimate the optimal list size. The blue area shows the decrease in the probability of clicking negative search suggestions as list size increases, and the red area shows the increase in the probability of clicking any search suggestion as list size increases. The highest point in the regions where the two shaded areas overlap is the point at which the probability of clicking the low-valence suggestions is maximized. Hence, the ideal list size for this type of control is 4.

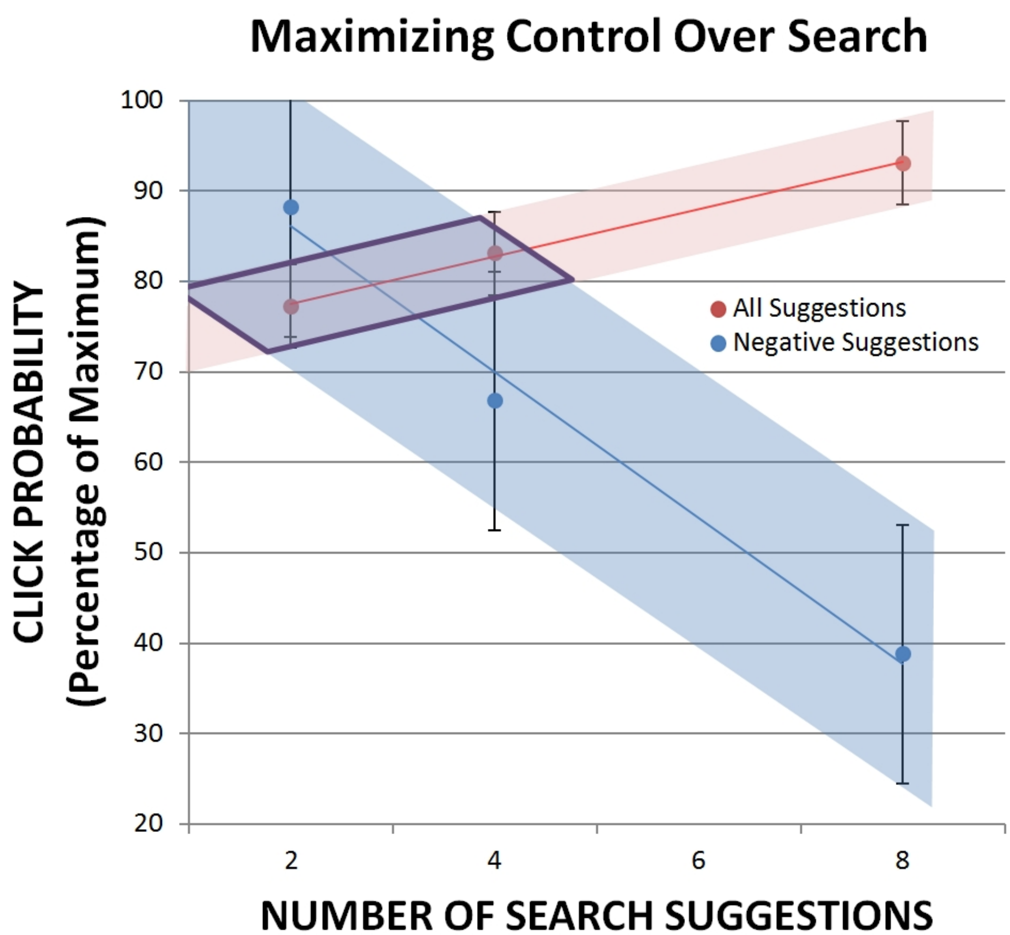


Figure 4. Click probability (percentage of maximum) as a function of number of search suggestions. The error bars around each point show the standard error (SE) of the mean. The shaded areas show, roughly, the SE regions encompassing each of the two lines in the graph. The upward-sloping line (red) shows the increasing probability of clicking on a search suggestion as list size increases. The downward-sloping line (blue) shows the decreasing probability of clicking on a negative search suggestion as list size increases. The highest point in the overlapping trapezoidal area is that point at which the probability of clicking a negative search suggestion is maximized. That point is almost directly over the number 4 on the *x* axis.

As noted in the introduction to this report, in 2010, Google reduced the size of the list of search suggestions it showed people on laptop and desktop computers from 10 to 4 and then increased the list size to 10 in October, 2017. Several months before that, on March 21, 2017, we used a Windows desktop computer to look up 1,000 of the most commonly used search terms (PageTraffic, 2017) using Google's search engine. 908 of those searches yielded lists of four search suggestions (overall $M = 4.07$, $SD = 0.26$, see Figure S13 to view the distribution).

These four experiments raise a practical question that can be addressed using a modification of the procedure that measures the Search Engine Manipulation Effect (SEME) (Epstein, Lee, Mohr, & Zankich, 2022; Epstein & Robertson, 2015) – namely, can we shift the opinions and voting preferences of undecided voters in a SEME experiment by manipulating the search suggestions people are shown just before the search results are displayed? If so, that would suggest that search suggestions themselves have enormous power to impact people's thinking and behavior.

Experiment 5: manipulating search suggestions in a SEME experiment

Participants and methods

The data for Experiment 5 were collected between September 19 and October 28, 2017. A politically diverse group of 340 people from 47 US states participated in this experiment (after cleaning). They were randomly assigned to one of four groups, which we will describe momentarily. At the outset, people were eliminated from the study if they responded "Yes" to the question, "Do you know a lot about politics in Australia?" All participants were then given basic instructions and asked a series of demographic questions. Among other things, they were asked to identify their political leaning and also asked how familiar they were with two candidates who ran for Prime Minister of Australia in 2010 – Tony Abbott and Julia Gillard.

Familiarity level was reported on a 10-point scale from 1 to 10, where 1 was labeled "not at all" and 10 was labeled "quite familiar." Participants who answered either of the familiarity questions with a value larger than 3 were eliminated from the dataset during cleaning.

Participants were then given brief information about the two candidates – a brief paragraph for each candidate approximately 140 words in length (Supplemental Text S2) – and then asked their opinions about the candidates and about their voting preferences (Figure S14). Employing a foreign election made it likely that our US participants would be undecided voters (Epstein & Robertson, 2015), particularly since we eliminated participants who were familiar with either of the candidates. Our results also confirmed that participants were undecided prior to the manipulation (see below for details). After cleaning, the mean familiarity level for Abbott was 1.40 (SD = 0.72), and the mean familiarity level for Gillard was 1.20 (SD = 0.51).

Immediately following those questions, participants in three of the groups (Groups 2, 3, and 4) were shown a list of four search suggestions (in Kadoodle, our simulation of Google Search) and asked to click one suggestion to generate search results (Figure S15). In the remaining group (Group 1), no search suggestions were shown; instead, people were simply shown search results. Half the participants in the latter group were shown search results favoring Tony Abbott, and half were shown search results favoring Julia Gillard. By "favoring," we mean that the search results linked to web pages that made one candidate look better than the other. All web pages had previously been rated (on an 11-point scale from 5 to 0 to 5, with the names of each candidate counterbalanced at either end of the scale) by five people who were not familiar with the experiment or its hypotheses.

In the experiment, search results were either in an order from high-Abbott-support to low-Abbott-support to low-Gillard-support to high-Gillard-support, or they were in the opposite order. In all, 30 search results were shown on five results pages, with six search results per page (Figure 5). All search results and all web pages were real; the search results were sourced from the Google search engine.

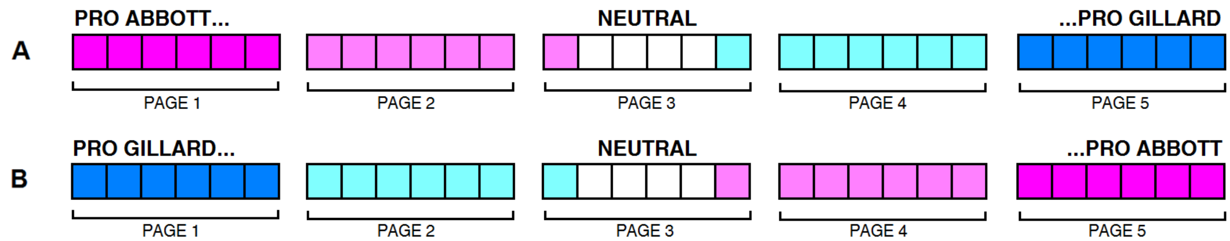


Figure 5. Search result sequences for Experiment 5. In a pro-Abbott group – either in the absence of search suggestions (Group 1) or if the participant clicks on a positive search suggestion (Groups 2 and 4) – the sequence of search results is shown in diagram A above. In a pro-Abbott group, if the participant clicks on a negative search suggestion (Groups 3 and 4) – the sequence of search results is shown in diagram B (which is the reverse order of the sequence shown in diagram A). In a pro-Gillard group – either in the absence of search suggestions (Group 1) or if the participant clicks on a positive search suggestion (Groups 2 and 4) – the sequence of search results is shown in diagram B above. In a pro-Gillard group, if the participant clicks on a negative search suggestion (Groups 3 and 4) – the sequence of search results is shown in diagram A (which is the reverse order of the sequence shown in diagram B).

Each of the three search suggestion groups saw a different list of search suggestions. In Group 2, participants saw four positive search suggestions for one of the two candidates. In Group 3, they saw four negative suggestions for one of the two candidates. Participants in Group 4 saw three positive suggestions and one negative suggestion for one of the two candidates. To examine the research design of Experiment 5 in detail, see Figure S16.

If a participant clicked a positive search term for a candidate, he or she was then shown search results favoring that candidate. If a participant clicked a negative search term for a candidate, he or she was then shown search results favoring the opposing candidate. For all four groups, once search results were displayed, people could take up to 15 minutes to research the candidates in the usual manner – that is, by clicking search results to visit web pages (Epstein &

Robertson, 2015). They could also click numbers at the bottom of each page of search results to navigate to different pages of search results.

In Group 1, the only difference between the groups was the order in which the search results were shown. In Groups 2, 3, and 4, once participants clicked a search suggestion, the only difference between the groups was, once again, the order in which the search results were shown. Note that all four groups were participating in a simple SEME experiment (Epstein & Robertson, 2015). Because participants in Group 1 saw no search suggestions and participants in Group 2 could only choose among four positive search suggestions, both groups were exposed to exactly the same search results, and we therefore expected to find similar shifts in voting preferences in these groups. Participants in Group 3 could only choose among four negative search suggestions, which means they all saw search results in the opposite order; we therefore expected to find *negative* shifts about the same magnitude as the shifts in Groups 1 and 2 (except negatively signed). Finally, the procedure in Group 4 (three positive search suggestions, one negative search suggestion) was a blend of the procedures in Groups 2 (four positive search suggestions) and 3 (four negative search suggestions). Because, as we learned in Experiments 1 through 4, a negative search suggestion attracts far more clicks than do neutral or positive search suggestions, in Group 4, we expected to get almost no shift in voting preferences. Roughly speaking, we expected the impact of one negative search suggestion to cancel out the effects of the three positive search suggestions.

After completing their research using our custom search engine, which looked and functioned like the Google search engine (Figure S17), all participants were again asked those opinion and voting questions regarding each of the candidates.

As in previous SEME experiments, before each session ended, we asked our participants if anything about the search results they were shown "bothered" them. If they responded "yes," they could then type an account of their concerns. This is a fairly standard way of determining whether participants detected bias in the search results they saw; unfortunately, one cannot ask people directly if they saw bias, because a leading question of this sort will artificially inflate the detection rate (Guess et al., 2021; Loftus, 1975). At the completion of the experiment, we examined the textual responses for indications that participants saw bias in the search results. We did so both manually and using an algorithm that searched for bias-related terms and phrases such as "biased," "one sided," and "skewed in favor of…" (Epstein & Robertson, 2015).

Results

Consistent with previous SEME findings, the voting preferences of participants who saw no search suggestions shifted toward the favored candidate by 39.6% (Table 1, Group 1, McNemar's $X^2 = 12.96$, $p < 0.001$). (We call that percentage the VMP, which stands for Vote Manipulation Power; it is the percentage increase in the number of people voting for the favored candidate [Epstein et al., 2022; Epstein & Robertson, 2015]. For further information about VMP and to see how it is calculated, read Supplemental Text S3). The voting preferences of participants in the search suggestion group that was shown only positive search suggestions shifted similarly (48.9%, Table 1, Group 2, $X^2 = 15.75$, $p < 0.001$). Participants who were shown four negative suggestions (and no positive suggestions) shifted *away* from the candidate shown in the search bar (-40.5%, Table 1, Group 3, $X^2 = 10.32$, $p < 0.01$). Finally, the voting preferences of participants who saw three positive search suggestions and one negative search suggestion barely shifted (-7.5%, Table 1, Group 4, $X^2 = 0.16$, $p = 0.69$ NS). Presumably this occurred because the negative search suggestion attracted 45.2% of the clicks (negativity bias). In other words, a single negative search suggestion can impact opinions dramatically because it links to search results that might be strongly biased against the candidate in question. Table 1 shows VMPs for Groups 1 to 4 broken down by gender, political group, and age.

These findings suggest that search suggestions can be used to create a win margin among undecided voters of more than 85% (39.6% + 48.9% = 88.5%). To put this another way, manipulating search suggestions can in theory turn a 50/50 split among undecided voters into more than a 90/10 split.

Table 1. Experiment 5: VMP percentages by condition and demographic group.

	Group 1: No Search Suggestions (%)	Group 2: Four Positive Search Suggestions (%)	Group 3: Four Negative Search Suggestions (%)	Group 4: One Negative and Three Positive Search Suggestions (%)
n	83	91	82	84
All Participants	39.6	48.9	-40.5	-7.5
Gender				
Male	20.0	42.9	-33.3	-7.7
Female	60.9	54.2	-47.4	-7.4
Political View				
Moderate	13.3	54.5	-44.4	-20.0
Conservative	66.7	33.3	-40.0	12.5
Liberal	40.9	61.1	-43.8	0.0
Age				
18 - 35 years old	37.9	59.1	-36.0	-9.5
36 - 85 years old	42.1	39.1	-50.0	-5.3

The changes in participants' opinions about the candidates shifted in a manner that was consistent with the shifts in voting preferences. In Groups 1 (no search suggestions) (Table S3) and 2 (four positive search suggestions) (Table S4), impression, trust, and likeability changed (post-search) very little for the favored candidate but decreased substantially for the non-favored candidate. In Group 3 (four negative search suggestions) (Table S5), impression, trust, and likeability decreased substantially for the favored candidate but changed very little for the non-favored candidate. In Group 4 (three positive search suggestions and one negative search suggestion) (Table S6), impression, trust, and likeability changed very little for the favored candidate and decreased somewhat for the non-favored candidate.

The fact that our participants rated the favored and non-favored candidates roughly equally before search in all four groups and very differently after search in Groups 1, 2, and 3, is yet another indication that our participants were undecided about our two candidates at the beginning of the experiment.

The pre-post changes in voting preferences on the 11-point scale were also consistent with the VMP changes. These changes were statistically significant in Groups 1, 2, and 3 and not significant in Group 4 (Table S7). That finding is consistent with both the VMPs (Table 1) and the changes in opinions we found in the four groups (Tables S3-S6).

Regarding perception of bias: 37.0% of the participants in Experiment 5 in the groups that were shown search suggestions appeared to detect bias in the search results they saw. In the group that was not shown search suggestions, 41.0% of participants appeared to detect the bias they were shown; as one might expect, the difference between these two percentages was not significant: $z = 0.65$, $p = 0.26$ NS. The latter percentage is consistent with earlier findings in which people were shown biased search results that contained no masking (Epstein & Robertson, 2015). When bias is masked – that is, when pro-Candidate-A results include an occasional pro-Candidate-B result, the percentage of participants who can detect bias drops significantly – sometimes to 0 – even though the shift in opinions and voting preferences can remain large (Epstein et al., 2022; Epstein & Robertson, 2015). Even more disturbing, people who detect bias in search results shift, on average, even farther in the direction of the bias than people who do not detect bias (Epstein & Robertson, 2015). In the present experiment (Experiment 5), no masking was employed, so these issues were not explored.

Discussion

The five experiments we have described above suggest clear answers to the questions we posed earlier:

1. How might voters have been impacted in the summer of 2016 by negative search suggestions – or the lack thereof – shown for different political candidates? Our results suggest that if negative search suggestions had been consistently suppressed for Hillary Clinton for a period of months before the 2016 election, while negative suggestions had been allowed to appear for her opponent, Donald Trump, the voting preferences of a large number of undecided voters would likely have shifted toward Mrs. Clinton.

By making some reasonable assumptions, we estimate that between 1.48 and 2 million votes might have been shifted to Mrs. Clinton through a combination of biased search suggestions (specifically, through the differential suppression of negative search suggestions for Clinton) working in combination with biased search results (which are automatically produced when people click biased search suggestions). See Supplemental Text S4 for how we arrived at this estimate.

If that range of numbers seems high, bear in mind that over a period of months, undecided voters might be exposed to similarly biased search suggestions and search results hundreds of times, which would presumably amplify this source of influence. Also consider: Which type of influence is likely to be the most impactful – a competitive and visible form of influence that causes undecided voters to be wary (such as a television commercial or billboard), or a non-competitive, invisible source of influence to which people are subjected repeatedly without awareness?

This analysis appears, at least at first, to break down when we add filter bubbles and echo chambers to the picture. When ardent conservatives get all their news from conservative news sources, or when ardent liberals get all their news from liberal news sources, they don't see much competing information (Mitchell, Gottfried, Kiley, & Matsa, 2014). *But the people at the extreme ends of the opinion scale are not the people we are concerned about in the present study*. Rather, we are focusing on people who are undecided, uncommitted, undeclared, unaffiliated, or all of the above – people who are not trapped in filter bubbles and who are trying, with varying degrees of effort, to make up their minds. *Those* people are the targets of both election campaigns and algorithms, especially as Election Day grows near (Höchstötter & Lewandowski, 2009). But the campaigns are competitive, whereas the content-generating algorithms controlled by large online platforms might be one-sided, with users having no way to counteract their influence. These companies have complete control over their own content, and they also serve as gateways to other content sources.

As we have noted, repeating the manipulation likely makes it stronger (Epstein et al., 2017). What happens if the same bias is also present in newsfeeds, targeted messages (Epstein, Tyagi, & Wang, in press), answer boxes (Epstein et al., 2022), answers provided by personal assistants (Epstein et al., 2022), and dozens of other "ephemeral experiences" (McKinnon & MacMillan, 2008) controlled by large tech platforms – nearly all of which are invisible forms of influence that leave no paper trails? We are currently investigating this issue in a separate body of research.

We noted earlier that we chose to use an Australian election in our study in order to guarantee that our participants would be “undecided,” and we indeed had two reasons to believe that they were in fact undecided. First, as we noted in the Results section of Experiment 5, participants split roughly evenly in both their opinions and their voting preferences prior to conducting their search (Tables S3-S7). We also reported that the initial familiarity with the two candidates was low (1.40 for Abbott, 1.20 for Gillard). It is important to recognize here, however, that low-familiarity (also called low-information) undecided voters differ in nontrivial ways from high-familiarity (high-information) undecided voters (Yarchi, Wolfsfeld, & Samuel-Azran, 2021). In SEME studies, for example, low-familiarity undecided voters tend to be more vulnerable to the manipulation than high-familiarity undecided voters, although significant effects occur with both groups (e.g., see Experiment 5 in Epstein & Robertson, 2015).

2. How might one structure search suggestions if one deliberately wanted to alter the opinions or votes of users? Based on the experiments we have described above, we can provide a relatively simple answer to this question, but our answer will also necessarily be simplistic. The simple answer is: program your algorithm so that just before it displays its next list of search suggestions (after, say, you have typed “hillar”), it checks a “white list” (Lo, 2010). If any of the search terms it is about to display (such as, “hillary clinton is the devil”) are on the white list (in this case, “hillary clinton” is on that list), it immediately deletes all the search suggestions that contain low-valence terms. Suggestions such as “hillary clinton is a filthy liar” are omitted. In theory, so many search suggestions of this type might be deleted by the algorithm that it would have to dig deep to display anything at all, leaving it, perhaps, with search suggestions that few people have searched for, such as “hillary clinton is awesome.”

In truth, however, a search algorithm like Google's would be far more sophisticated than the one we have described. Based on extensive data the company has collected about users, the algorithm would likely show people suggestions based on the actual likelihood that individuals will click them. When you know your users well, word valence is only one variable you might use in maximizing your control over a user's click. That possibility should be investigated further, but one could only crudely approximate the precision with which a company like Google could exercise such control.

3. Is there an optimal number of search suggestions for influencing people? Experiment 4 suggests that the optimal number is four, which Google showed people between 2010 and 2017 – until, that is, about 3 weeks after we first went public with our preliminary findings about search suggestions. Again, however, our answer is simplistic. When you have the ability to personalize content based on massive amounts of information about people, you can optimize the control you have when you are showing people almost any number of search suggestions (within reason). That is an intriguing possibility that should be investigated further, but, once again, doing so without the extensive resources of a company like Google would be difficult.

4. Can people's opinions or voting preferences be altered by search suggestions without their knowledge? We believe the answer is yes, but we did not ask about possible awareness of bias in search suggestions in the present study. We did ask about possible awareness of bias in search *results* in Experiment 5 (the only experiment that contained search results). Although nearly 40% of participants in Experiment 5 said that the search results appeared to be biased, we know from previous experiments that that kind of awareness can easily be minimized with masking procedures (Epstein & Robertson, 2015; cf. Day & Altman, 2000).

We did not inquire about the possible perception of bias in search suggestions because real search suggestions are typically displayed for only a fraction of a second. In our procedures, participants needed to examine the list of search suggestions closely. Because real search suggestions are brief and ephemeral, we question whether most people would perceive any bias in them, but this issue should be explored further. Generally speaking, the more briefly a visual stimulus is presented, the fewer features people are able to perceive (Hegdé, 2008).

5. Who is most vulnerable to manipulation of this sort? We performed only a few basic demographic analyses on our data – enough to answer a simple but important question: Are some people more vulnerable to influence by biased search suggestions and search results than others? The answer appears to be yes (see Table 1, data for Groups 1 and 2) (also see Epstein and Robertson [2015]). This finding is important, not because it gives us a definitive picture of who is vulnerable and who is not but because it shows that substantial individual and demographic differences in vulnerability exist. This means that tech companies that have accumulated vast amounts of data about users can conceivably use those data to optimize manipulations that employ search suggestions.

Funding

This research was supported by general funds of the American Institute for Behavioral Research and Technology. No external funding was used to support the research or to compensate the researchers.

Data availability statement

Anonymized raw data will be posted at Zenodo.com upon publication of the report. Anonymized raw data can also be requested by writing to info@aibrt.org. Anonymization was required to comply with the requirement of the sponsoring institution's Institutional Review Board that the identities of the participants be protected in accordance with HHS Federal Regulation 45 CFR 46.101.(b)(2).

Acknowledgments

This report is based in part on papers presented at the 2nd biennial meeting of the International Convention of Psychological Science in 2017 and the 98th annual meeting of the Western Psychological Association in 2018. We thank J. Martinez, R. Mohr, H. G. Park, and S. Shepherd for help in preparing and running the experiments, and R. Anderson for assistance in preparing the manuscript. R.E. designed the experiments, directed the research, and wrote the paper. S.A., K.G., A.M.S., and V.R.Z. reviewed the literature, helped analyze the data, and helped prepare the manuscript.

Disclosure statement

The authors have no competing interests to declare.

References

Advanced Web Ranking. *Google organic CTR history (2020).* https://www.advancedwebranking.com/ctrstudy/

Carretié, L., Mercado, F., Tapia, M., Hinojosa, J. (2001). Emotion, attention, and the 'negativity bias', studied through event-related potentials. *International Journal of Psychophysiology*, *41*(1), 75-85. https://doi.org/10.1016/s0167-8760(00)00195-1

Chitika. (2013, June 7). *The value of Google results positioning*. http://info.chitika.com/uploads/4/9/2/1/49215843/chitikainsights-valueofgoogleresultspositioning.pdf

Day, S. J., & Altman, D. G. (2000). Statistics notes: Blinding in clinical trials and other studies. *BMJ Clinical Research*, *321*(7259), 504. https://doi.org/10.1136%2Fbmj.321.7259.504

Dean, B. (2020, August 20). *How people use Google Search.* Backlinko. https://backlinko.com/google-user-behavior

Dean, B. (2022, August 17). *We analyzed 4 million Google search results: Here's what we learned about organic click through rate*. Backlinko. https://backlinko.com/google-ctr-stats

Epstein, R. (2021). The technological elite are now in control. In *Champions of freedom, Volume 49: Big Tech*. Hillsdale College Press. https://aibrt.org/downloads/EPSTEIN_2020-The_Technological_Elite_Are_Now_in_Control.pdf

Epstein, R., Ding, M., Mourani, C., Olson, E., Robertson, R. E., & Tran, F. (2017, April). *Multiple searches increase the impact of the Search Engine Manipulation Effect (SEME)* [Paper presentation]. 97th Annual Meeting of the Western Psychological Association, Sacramento, CA. https://aibrt.org/downloads/EPSTEIN_et_al._2017-WPA-Multiple_Searches_Increase_the_Impact_of_the_Search_Engine_Manipulation_Effect.pdf

Epstein, R., Lee, V., Mohr, R., & Zankich, V. R. (2022). The answer bot effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. *PLOS ONE*, *17*(6), e0268081. https://doi/org/10.1371/journal.pone.0268081Flores, R. (2016, June 11). Hillary Clinton Google suggestions accused of favoring candidate. *CBS News.* https://www.cbsnews.com/news/hillary-clinton-google-suggestions-accused-favoring-candidate-election-2016/

Epstein, R., Tyagi, C., & Wang, H. (in press). What would happen if Twitter sent consequential messages to only a strategically important subset of users? A quantification of the targeted messaging effect (TME). *PLOS ONE* (preprint posted to SSRN, September 18, 2022). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4207187

Epstein, R., & Robertson, R. (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), E4512-4521. https://doi.org/10.1073/pnas.1419828112

Garber, M. (2013, August 23). *How Google's autocomplete was ... created / invented / born*. The Atlantic. https://www.theatlantic.com/technology/archive/2013/08/how-googles-autocomplete-was-created-invented-born/278991/

Gibbs, K. (2004, December 10). *Google blogpost with invention of autocomplete*. Google Blog. https://googleblog.blogspot.com/2004/12/ive-got-suggestion.html

Goldman, D. (2016, June 15). The truth about the Hillary Clinton Google conspiracy theory. *CNN Business.* https://money.cnn.com/2016/06/10/technology/hillary-clinton-google-search-results/

Granka, L., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. *SIGIR '04: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* Association for Computing Machinery.

Guess, A. M., Barberá, P., Munzert, S., & Yang, J. (2021). The consequences of online partisan media. *Proceedings of the National Academy of Sciences, 118*(14), 2013464118. https://doi.org/10.1073/pnas.2013464118

Hannah, M. T., Le Mens, G., Hsu, G., Kovács, B., Negro, G., Pólos, L., Pontikes, E. G., & Sharkey, A. J. (2009). *Concepts and categories: Foundations for sociological and cultural analysis*. Columbia University Press.

Harrington. E., & Scher, B. (2016, June 10). Here are 10 more examples of Google search results favorable to Hillary. *The Washington Free Beacon*. http://freebeacon.com/politics/here-are-10-more-examples-of-google-search-results-favorable-to-hillary/

Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, *84*(4), 405-439. https://doi.org/10.1016/j.pneurobio.2007.09.001

Höchstötter, N., & Lewandowski, D. (2009). What users see - Structures in search engine results pages. *Information Science*, *179*(12), 1796-1812. https://doi.org/10.1016/j.ins.2009.01.028

Hong, S., & Kim, N. (2018). Will the internet promote democracy? Search engines, concentration of online news readership, and e-democracy. *Journal of Information Technology & Politics*, *15*(4). https://doi.org/10.1080/19331681.2018.1534703

Johnson, D. D. P., & Tierney, D. (2019). Bad world: The negativity bias in international politics. *International Security*, *43*(3), 96-140. https://doi.org/10.1162/ISEC_a_00336

Kaushal, N. (n.d.). *Most popular key words on search engines.* PageTraffic. https://www.pagetraffic.com/blog/most-popular-keywords-on-search-engines/

Klapdor, S., Anderl, E., Waingenheim, F., & Schumann, J. (2014). Finding the right words: The influence of keyword characteristics on performance of paid search campaigns. *Journal of Interactive Marketing*, *28*(4), 285-301. https://doi.org/10.1016/j.intmar.2014.07.001

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, *143*(3), 1065. https://doi.org/10.1037/a0035669

Lo, J. (2010, November 15). *Whitelisting for cyber security: What it means for consumers*. The Public Interest Advocacy Centre. http://piac.staging.craftandcrew.ca/research-reports/whitelisting-for-cyber-security-what-it-means-for-consumers/

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*(4), 560-572. https://doi.org/10.1016/0010-0285(75)90023-7

McKinnon, J. D., & MacMillan, D. (2018, September 20). Google workers discussed tweaking search function to counter travel ban: Company says none of proposed changes to search results were ever implemented. *Wall Street Journal.* https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472

Mitchell, A., Gottfried, J., Kiley, J., & Matsa, K. E. (2014, October 21). *Political polarization & media habits*. Pew Research Center. https://www.pewresearch.org/journalism/2014/10/21/political-polarization-media-habits/

Napolitano, A. (2016, June 23). *"Attempts to find things like, 'crime,' 'criminal activity,' 'indictment,' have been filtered out by Google" in favor of Hillary Clinton*. The Poynter Institute. https://www.politifact.com/factchecks/2016/jun/23/andrew-napolitano/did-google-adjust-its-autocomplete-algorithm-hide-/

Nicas, J., & Andrews, N. (2016, June 10). Google denies accusations it manipulated searches for Hillary Clinton. *The Wall Street Journal*. https://www.wsj.com/articles/google-denies-accusations-it-manipulated-searches-for-hillary-clinton-1465594099

Pothitos, A. (2016, October 31). *The history of the smartphone.* Mobile Industry Review. http://www.mobileindustryreview.com/2016/10/the-history-of-the-smartphone.html

Ramaboa, K., & Fish, P. (2018). Keyword length and matching options as indicators of search intent in sponsored search. *Information Processing & Management*, *54*(2), 175-183. https://doi.org/10.1016/j.ipm.2017.11.003

Robertson, C. E., Pröllochs, N., Schwarzenegger, K., Pärnamets, P., Van Bavel, J. J., & Feuerriegel, S. (2023). Negativity drives online news consumption. *Nature Human Behaviour, 7*, 812-822. https://doi.org/10.1038/s41562-023-01538-4

Schwartz, B. (2017, October 3). *Google autocomplete showing 10 suggestions again*. Search Engine Round Table. https://www.seroundtable.com/google-autocomplete-10-suggestions-24560.html

Seitz, D. (2016, June 13). Did Google manipulate autocomplete to protect Hillary Clinton? *Yahoo News.* https://news.yahoo.com/did-google-manipulate-autocomplete-protect-210039362.html

Sheehan, K. B. (2017). Crowdsourcing research: Data collection with Amazon's Mechanical Turk. *Communications Monograph*, *85*(1), 140-156. https://doi.org/10.1080/03637751.2017.1342043

Shuman, V., Sander, D., & Scherer, K. R. (2013). Levels of valence. *Frontiers in Psychology*, *4*, 261. https://doi.org/10.3389/fpsyg.2013.00261

StackExchange. (2017, March 20). *Longer Google autocomplete lists?* https://webapps.stackexchange.com/questions/8990/longer-google-autocomplete-lists

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*, 1191-1207. https://doi.org/10.3758/s13428-012-0314-x

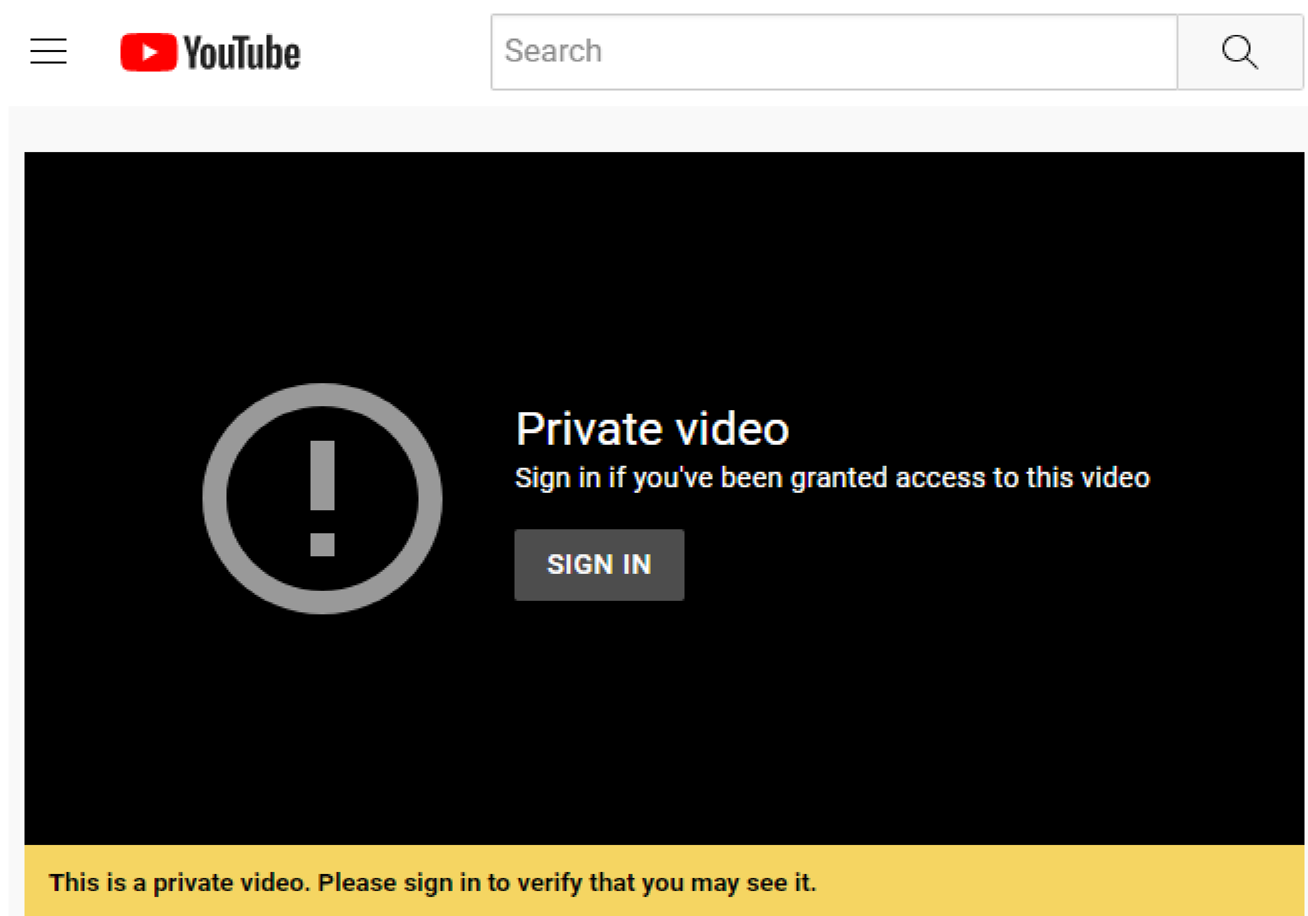
Yarchi, M., Wolfsfeld, G., & Samuel-Azran, T. (2021). Not all undecided voters are alike: Evidence from an Israeli election. *Government Information Quarterly, 38*(4), 101598. https://doi.org/10.1016/j.giq.2021.101598

Yehoshua, T. (2016, June 10). *Google Search autocomplete*. Google: The Keyword. https://blog.google/products/search/google-search-autocomplete/

Supporting Information

Supplemental Text S1. SourceFed video posted June 9, 2016.

SourceFed originally posted a 7-minute video on YouTube (owned by Google) on June 9, 2016. Within a day, the number of views increased to nearly a million, at which point Google made the video "private." It has been inaccessible to the public ever since. The URL for the private video is: https://www.youtube.com/watch?v=PFxFRqNmXKg. As of this writing (June 7, 2022), that page shows the following image:



SourceFed also posted a 3-minute version of the video on Facebook, which soon attracted more than 25 million views. That version of the video is still accessible at: https://www.facebook.com/SourceFedNews/videos/vb.322741577776002/1199514293432055/?type=2&theater

Supplemental Text S2. Experiment 5: Brief candidate biographies, pre-search.

Julia Gillard. Born in Barry, Wales (UK) on September 29th, 1961, Gillard moved with her family to Australia in 1966. While Gillard was studying at the University of Melbourne, she became the second woman to lead the Australian Union of Students. After graduating from the University of Melbourne with a Bachelor of Arts and a Bachelor of Laws in 1986, Gillard joined the law firm of Slater & Gordon and practiced industrial law. She became a partner at the law firm at age 29 but left in 1996 to enter politics as the Chief of Staff for the Leader of the Opposition. Gillard lives with her longtime boyfriend, Tim Mathieson, and has chosen not to have children in order to devote herself to her career. In 1998, Gillard was elected to the House of Representatives to represent the Labor party. Currently, Gillard is the Deputy Prime Minister of Australia and leader of the Australian Labor Party.

Tony Abbott. Born in London, England on November 4th, 1957, to Australian parents, Tony Abbott moved with his family to Australia in 1970. While attending the University of Sydney, Abbot became a prominent political student activist and served as president of the Student Representative Council. In 1981, Abbott graduated from the University of Sydney with a Bachelor of Economics and a Bachelor of Laws. Shortly after, he was awarded the prestigious Rhodes Scholarship and moved back to England to attend the University of Oxford, where he earned a Masters of Arts in Politics and Philosophy. Abbott married Margaret Aitken in 1988, and they now have three daughters, Louise, Bridget and Frances. Elected to the Australian House of Representatives in 1994, Abbott is now the Leader of the Opposition in the Australian House of Representatives and also leader of the Australian Liberal Party.

Supplemental Text S3. Method for computing VMP.

Vote Manipulation Power (VMP) is calculated as follows:

$$\left(\frac{p' - p}{p}\right) \times 100$$

where p is the total number of people who voted for the favored candidate pre-manipulation, and p' is the total number of people who voted for the favored candidate post-manipulation. If, pre-manipulation, a group of 100 people is split 50/50 in the votes they give us, and if, post-manipulation, a total of 67 people now vote for the favored candidate, the VMP is:

$$\left(\frac{67 - 50}{50}\right) \times 100$$

or 34%. Because p' is 17 points larger than p, the win margin is 34 (2 × 17, or 34%), and the final vote is 67 to 33, with the favored candidate the winner. So in any group in which the vote is split 50/50 pre-manipulation, the VMP is also the win margin. Note that 17 individuals did not need to *shift* to produce this win margin. We only needed the *net* number of people voting for the favored candidate to be 67. As a practical matter, that net is the key statistic a campaign staff would likely want to know.

Supplemental Text S4. Estimating the impact of SSE on votes

For estimation purposes, let us first assume that all other sources of influence besides SSE – content from television, radio, the internet, newspapers, and magazines, for example – are highly competitive and roughly cancel each other out; they certainly have the potential to do so. With most traditional sources of influence, the human hand is apparent, so people are often wary. In contrast, with computer-generated content, such as search suggestions, newsfeeds, search results, and answers on personal assistants, people often mistakenly assume that the content is inherently objective and unbiased because it is computer-generated (Bogert, Schecter, & Watson, 2021; Logg, Minson, & Moore, 2018). The human hand – the people who wrote the algorithms, the people who are adjusting the algorithms, the human-managed blacklists and whitelists that the algorithms might be checking – is invisible. With traditional sources of influence, people are usually aware that someone is trying to influence them; with computer-generated content on the internet, people are typically unaware that they are being subjected to any form of influence. When people are unaware of influence, they have no way to counteract it, and shifts in thinking and behavior can be large, especially when people are undecided on an issue (Bond et al., 2012).

Let us further assume that no sources of influence exist to counteract the influence of biased search suggestions on the leading search engine. 92% of search in the US and most other countries is conducted on Google (Statscounter GlobalStats, 2021), so competing search engines have relatively little impact on elections. Bing, the most popular competitor, attracts only 2.6% of search engine traffic, and each of the other search engines – DuckDuckGo, Baidu, Yahoo, Yandex, Startpage, Qwant, Swiss Cows, and so on – attracts less than 1.6% of search engine traffic (Statscounter GlobalStats, 2021).

Let us also assume that 6 months before a national election, at least 20% of voters are undecided. Voter surveys conducted before US Presidential elections since the 1940s suggest that the actual percentage probably varies between 30% and 60% (Annenberg Public Policy Center, 2008; Liu, Ye, Sun, Jiang, & Wang, 2021; Mayer, 2008). In the 2016 Presidential election, 138.8 million people voted for president, which suggests that 6 months before Election Day at least 27 million people (using our conservative 20% estimate) could have been tipped one way or another by an effective, non-competitive source of influence.

Let us further assume that between 60% and 80% of these people sometimes use search engines to gather information about election related issues (the actual proportion is probably at the high end of this range, if not higher, and that proportion increases from one election to the next worldwide) (American National Election Studies, 2021; Shearer, 2021). Note that in our Experiment 5, 90.3% of our participants said they had used search engines to get political information, and 93.3% of our participants said Google was the search engine they used most frequently. Using the more conservative 60% to 80% range, that gives us between 16.2 and 21.6 million people to influence with our biased search suggestions. Although it is true that people are not offered search suggestions on every search, we can safely assume that all of our undecided voters are shown search suggestions on at least some searches, and given that about 23% of people click a search suggestion when offered a list of search suggestions (Dean, 2020), the number of people over whom we can exert strong influence might now drop to between 3.7 and 5.0 million. Again, our estimate is conservative here, especially given our finding (in Experiment 1) that eligible, undecided voters – the voters most likely to be targeted for manipulations – are strongly inclined to click on low-valence search terms (recall that they clicked on low-valence terms 14.8 times as often as they clicked on neutral or positive terms). In other words, the people whom one might want to influence are the easiest to influence using this type of manipulation.

Experiment 5 suggests that the differential suppression of low-valence search suggestions can produce large shifts in voting preferences during a single search experience, turning a 50/50 vote split into more than a 90/10 vote split (which means we can tip more than 40 out of 100 vulnerable people toward one candidate). If, over a 6-month period, we are able to shift 40% of our undecided voters, as we did in Experiment 5 (Group 2) after just one search, that means that in our election with 138.8 million voters, we should be able to use biased search suggestions to shift between 1.48 and 2 million voters toward the candidate of our choice. If all other voters were split 50/50, this means we could use SSE to create a win margin for our candidate of roughly between 3 and 4 million votes (2 times the win margins).

The real vote margin that Google controls is probably larger, given that:

(a) we have likely underestimated the proportion of voters who are undecided 6 months before an election; and

(b) we have likely underestimated the proportion of voters who get political information online in the months leading up to an election.

In addition:

(a) people can be presented with biased search suggestions hundreds of times over that 6-month period;

(b) Google can easily identify the voters who remain undecided at any point in time and thus target and personalize their manipulations; and

(c) Google can bias search results, YouTube videos, answer boxes, answers provided by Google Home (a personal assistant similar to Amazon's Echo device), answers provided by the

Google Assistant (a personal assistant on Android phones, similar to Apple's Siri), and other content to favor the same candidate, thus exerting more influence.

Finally, we note that content from other Silicon Valley tech companies – virtually all of which have the same political leanings – can push the margin even farther. Some of these companies have strategic relationships with Google, which directly expands Google's influence (Leswig, 2018; Pressman, 2017).

Figure S1. Search suggestions for "donald trump is a" on August 4, 2016, showing two negative suggestions (ellipses were added).

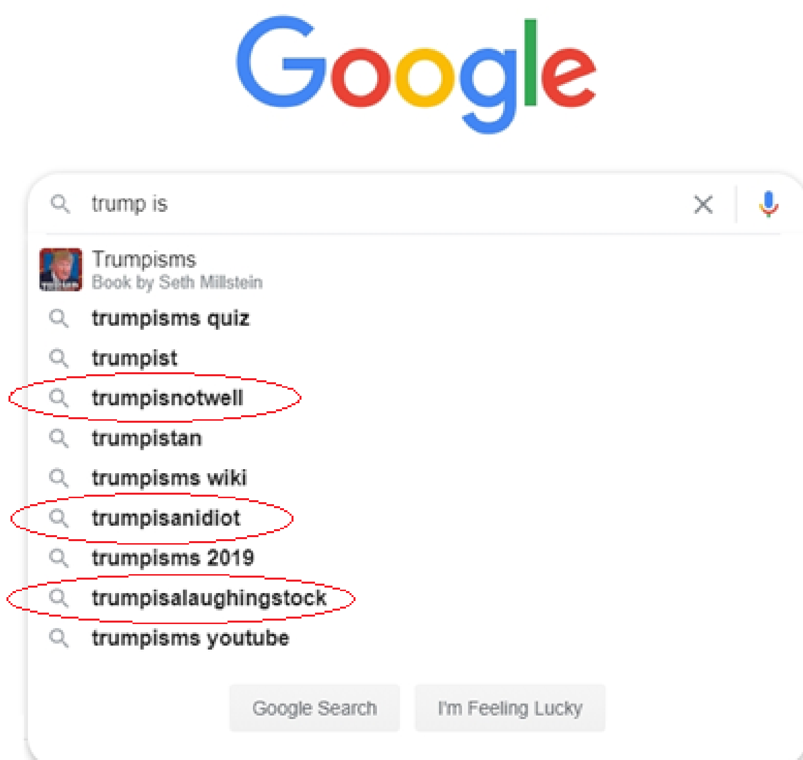


Figure S2. "Trump is" on Google.com on July 9, 2020, where two of the four suggestions shown were negative (ellipses were added).

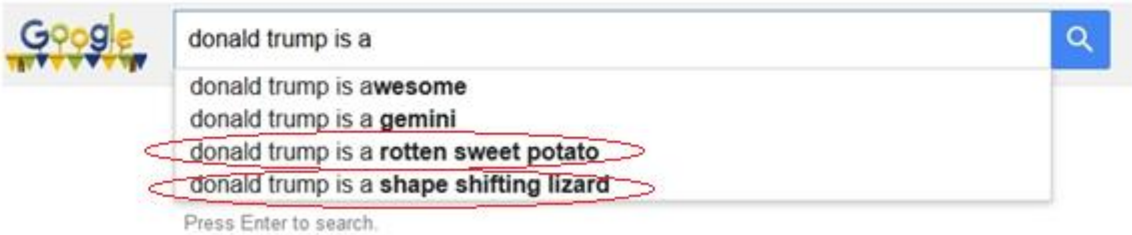


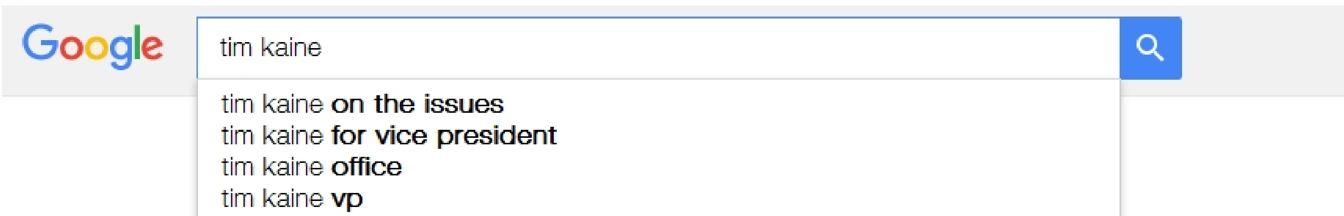
Figure S3. “Trump is” on Google.com on July 27, 2020, where three of the suggestions shown were negative (ellipses were added).

Experiment 1: Appearance of the Four Items on the Web Page

For each of the following questions, select the answer that seems best to you.

---

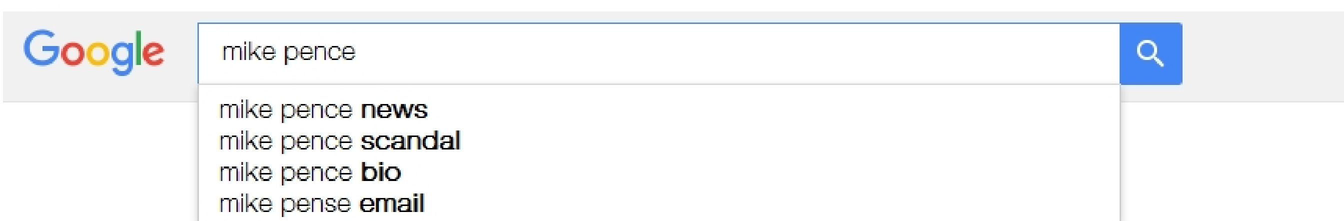
1. Say you Googled "Tim Kaine," the Democratic candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Kaine's name, which option, if any, would you be most likely to click?



- a. tim kaine **on the issues**
- b. tim kaine **for vice president**
- c. tim kaine **office**
- d. tim kaine **vp**
- e. I would type something else: Please type here

---

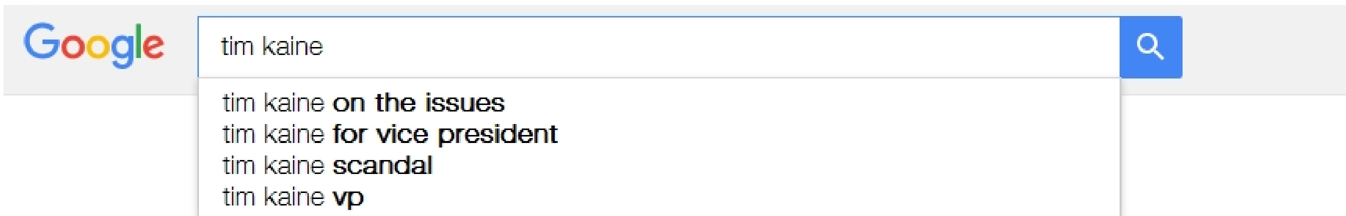
2. Say you Googled "Mike Pence," the Republican candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Pence's name, which option, if any, would you be most likely to click?



- a. mike pence **news**
- b. mike pence **scandal**
- c. mike pence **bio**
- d. mike pence **email**
- e. I would type something else: Please type here

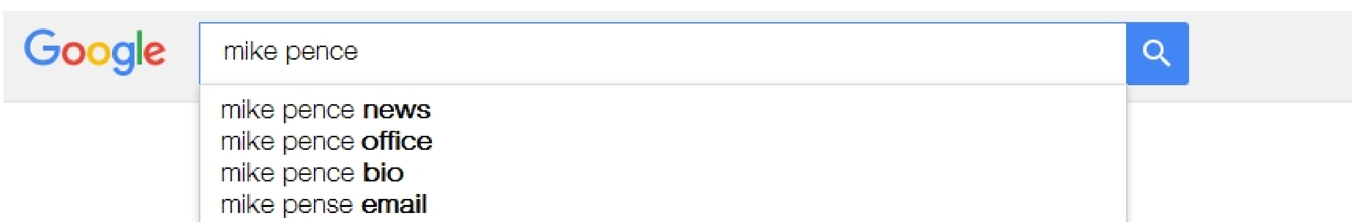
---

3. Say you Googled "Tim Kaine," the Democratic candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Kaine's name, which option, if any, would you be most likely to click?



- a. tim kaine **on the issues**
- b. tim kaine **for vice president**
- c. tim kaine **scandal**
- d. tim kaine **vp**
- e. I would type something else: Please type here

4. Say you Googled "Mike Pence," the Republican candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Pence's name, which option, if any, would you be most likely to click?



- a. mike pence **news**
- b. mike pence **office**
- c. mike pence **bio**
- d. mike pence **email**
- e. I would type something else: Please type here

Figure S4. Items and instructions shown on web page in Experiment 1. All participants saw all four items on the same web page, as pictured above.

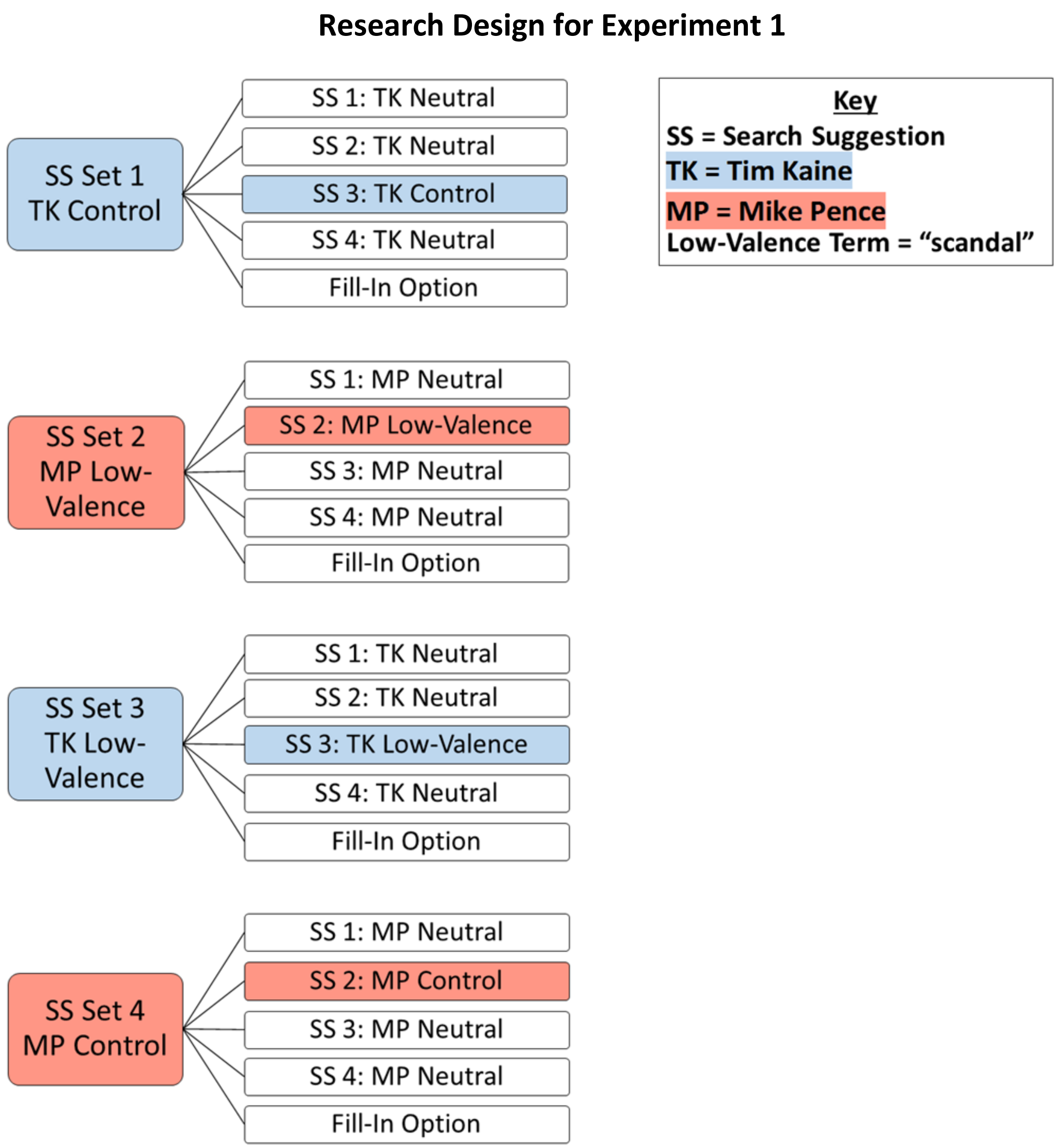
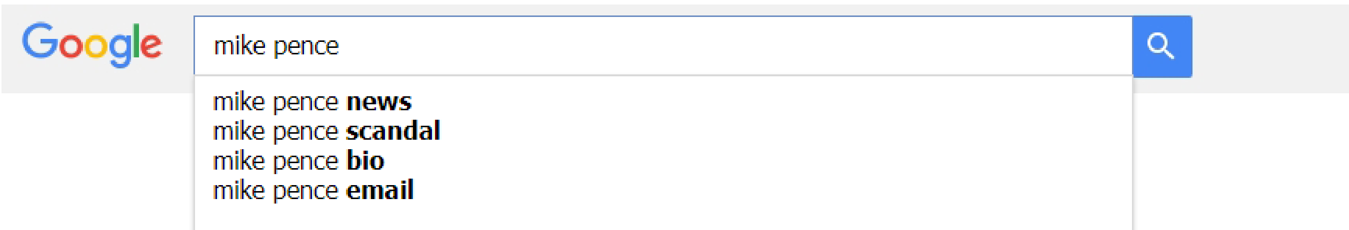


Figure S5. Design of Experiment 1. In this experiment, the number of clicks to a low-valence search term was compared to the number of clicks to a neutral control term in the same position, as well as to the number of clicks to neutral search terms in other positions. All participants saw the same four sets of search suggestions in which the order of the search suggestions was fixed. Sets 1 and 3 showed, respectively, a control term ("office," neutral valence) and a low-valence term ("scandal") for Democratic candidate Tim Kaine; both always appeared in the third position. Sets 2 and 4 showed, respectively, a control term ("office," neutral valence) and a low-valence term ("scandal") for Republican candidate Mike Pence; both always appeared in the second position.

Experiment 2: Appearance of the Four Items on the Web Page

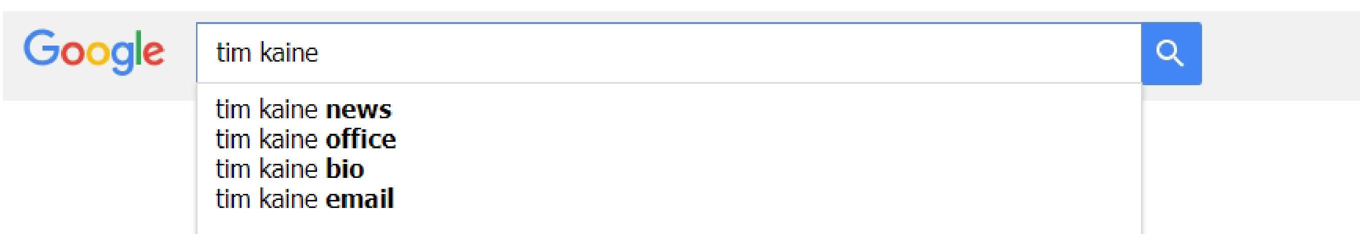
1. Say you Googled "Mike Pence," the Republican candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Pence's name, which option, if any, would you be most likely to click?



- ○ a. mike pence **news**
- ○ b. mike pence **scandal**
- ○ c. mike pence **bio**
- ○ d. mike pence **email**
- ○ e. I would type something else: [Please type here]

---

2. Say you Googled "Tim Kaine," the Democratic candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Kaine's name, which option, if any, would you be most likely to click?



- ○ a. tim kaine **news**
- ○ b. tim kaine **office**
- ○ c. tim kaine **bio**
- ○ d. tim kaine **email**
- ○ e. I would type something else: [Please type here]

---

3. Say you Googled "Mike Pence," the Republican candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Pence's name, which option, if any, would you be most likely to click?

Google

mike pence

mike pence **news**
mike pence **office**
mike pence **bio**
mike pence **email**

- a. mike pence **news**
- b. mike pence **office**
- c. mike pence **bio**
- d. mike pence **email**
- e. I would type something else: Please type here

---

4. Say you Googled "Tim Kaine," the Democratic candidate for vice president, in order to learn more about him. If Google made the following suggestions after you typed Kaine's name, which option, if any, would you be most likely to click?

Google

tim kaine

tim kaine **scandal**
tim kaine **office**
tim kaine **bio**
tim kaine **email**

- a. tim kaine **scandal**
- b. tim kaine **office**
- c. tim kaine **bio**
- d. tim kaine **email**
- e. I would type something else: Please type here

Figure S6. Items and instructions shown on web page in Experiment 2. All participants saw all four items on the same web page, as pictured above, but for each participant, the position of the negative search suggestion could appear in any of the first four search suggestions of its set. The other suggestions did not change position.

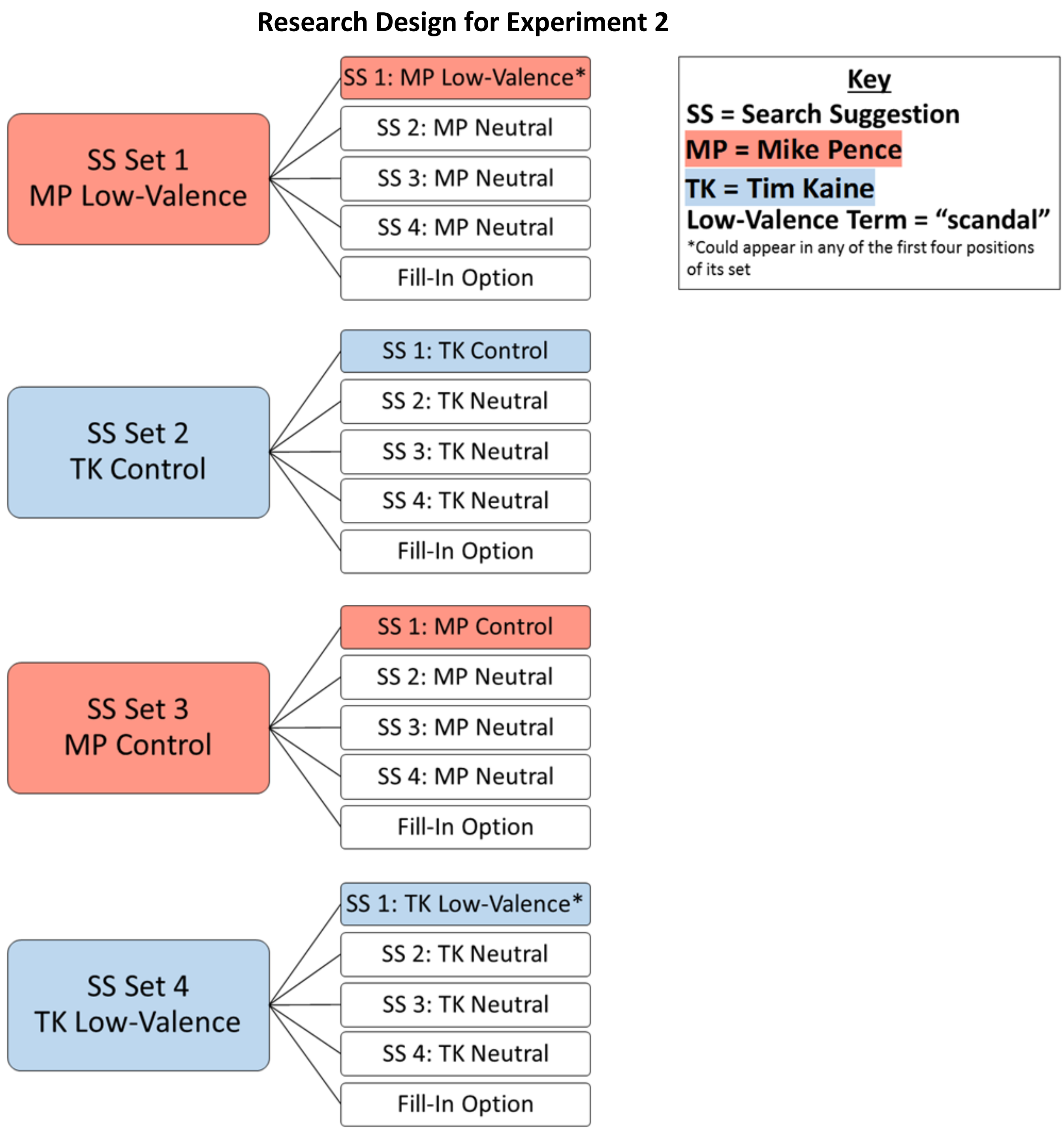


Figure S7. Design of Experiment 2. As in Experiment 1, participants were shown four sets of search suggestions, always in the same order. However, the low-valence term could appear in any of the first four positions of sets 1 and 4. The positions of the neutral terms did not change, so each neutral term served as a control term for the low-valence term when it appeared in the same relative position as the low-valence term. Shown above is just one of 16 possible configurations: when both low-valence terms appeared in the first position of their sets.

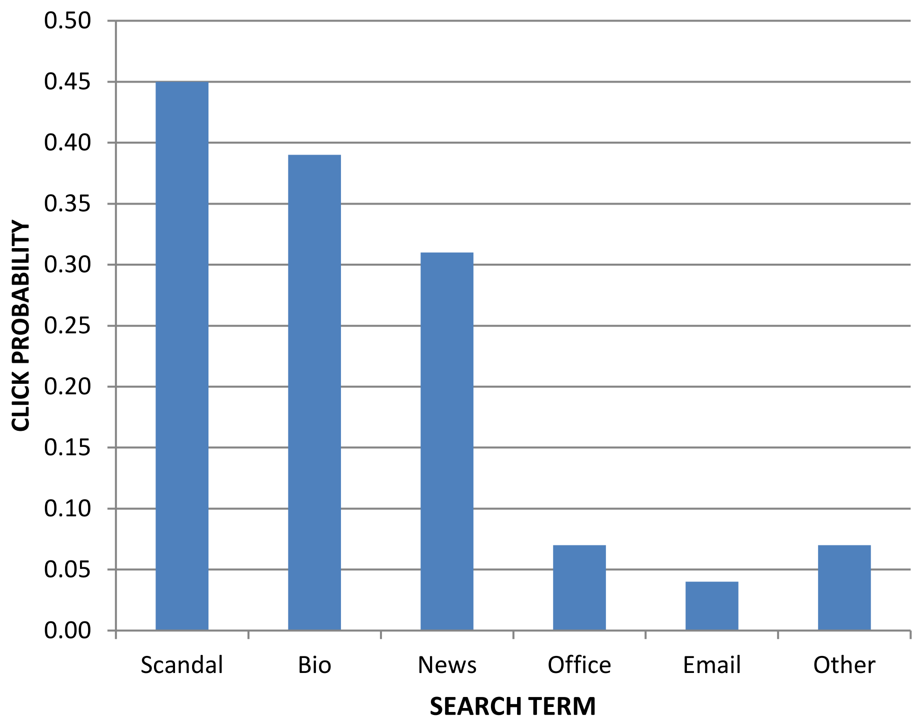


Figure S8. Click probability for all search suggestions and positions in Experiment 2.

Experiment 3: Appearance of the Three Items on the Web Page

Say you wanted to learn more about Donald Trump's grandfather, Frederick Trump. You begin to type "Frederick Trump's" into the search bar. If Google made the following suggestions as you typed, which option would you click?

1. Click on your preferred choice:

Frederick Trump's

- Frederick Trump's **suicide**
- Frederick Trump's **stretch**
- Frederick Trump's **holiday**
- Frederick Trump's **evaluation**
- Frederick Trump's **wedding**
- Frederick Trump's **achievement**
- Frederick Trump's **honor**
- Frederick Trump's **fashion**
- I would type something else: Please type here

2. Click on your preferred choice:

Frederick Trump's

- Frederick Trump's **suicide**
- Frederick Trump's **fashion**
- Frederick Trump's **honor**
- Frederick Trump's **stretch**
- I would type something else: Please type here

3. Click on your preferred choice:

Frederick Trump's

- Frederick Trump's **evaluation**
- Frederick Trump's **suicide**
- Frederick Trump's **gate**
- Frederick Trump's **achievement**
- I would type something else: Please type here

Figure S9. Items and instructions shown on web page in Experiment 3.

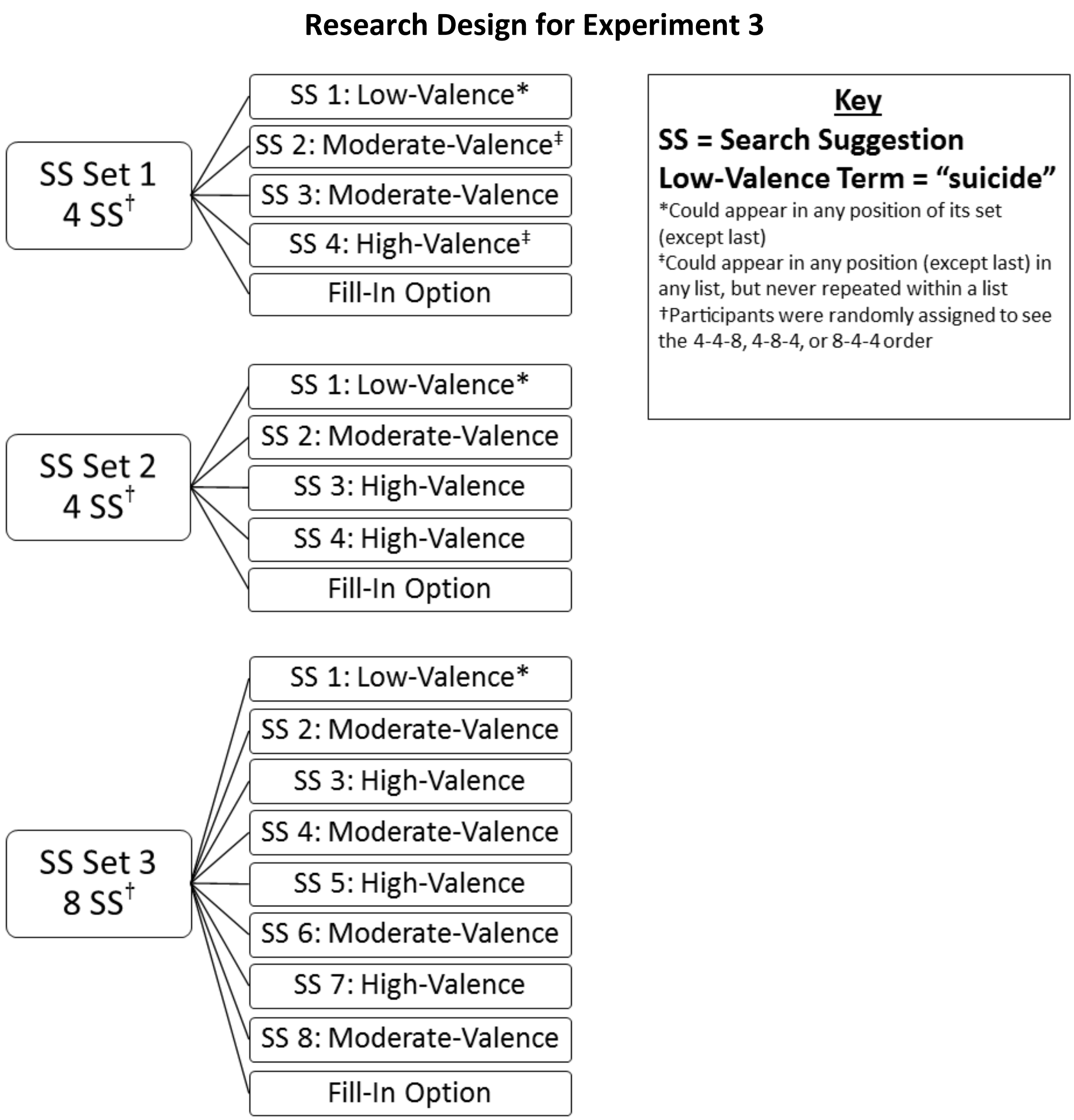


Figure S10. Design of Experiment 3. In Experiment 3, all participants saw three sets of search suggestions which consisted of one set of eight suggestions and two sets of four suggestions; the order of those sets varied randomly. All sets also included a final fill-in option. Within each set of search suggestions, the low-, moderate-, and high-valence terms could appear in any position. Only one possible configuration (out of 128 possible configurations) is shown above: one in which the low-valence term appeared in the first position of each set.

Experiment 4: Appearance of the Six Items on the Web Page

Say you wanted to learn more about Donald Trump's grandfather, Frederick Trump. You begin to type "Frederick Trump's" into the search bar. If Google made the following suggestions as you typed, which option would you click?

1. Click on your preferred choice:

Google Frederick Trump's

Frederick Trump's **execution**
Frederick Trump's **lifestyle**
Frederick Trump's **treasure**
Frederick Trump's **competitor**
Frederick Trump's **excitement**
Frederick Trump's **expertise**
Frederick Trump's **prediction**
Frederick Trump's **determination**
I would type something else: Please type here

2. Click on your preferred choice:

Google Frederick Trump's

Frederick Trump's **expertise**
Frederick Trump's **execution**
Frederick Trump's **happiness**
Frederick Trump's **competitor**
I would type something else: Please type here

3. Click on your preferred choice:

Google

Frederick Trump's

Frederick Trump's **expertise**
Frederick Trump's **excitement**
Frederick Trump's **determination**
Frederick Trump's **competitor**
Frederick Trump's **treasure**
Frederick Trump's **happiness**
Frederick Trump's **prediction**
Frederick Trump's **racism**
I would type something else: Please type here

---

4. Click on your preferred choice:

Google

Frederick Trump's

Frederick Trump's **happiness**
Frederick Trump's **racism**
Frederick Trump's **expertise**
Frederick Trump's **competitor**
I would type something else: Please type here

---

5. Click on your preferred choice:

Google

Frederick Trump's

Frederick Trump's **happiness**
Frederick Trump's **racism**
I would type something else: Please type here

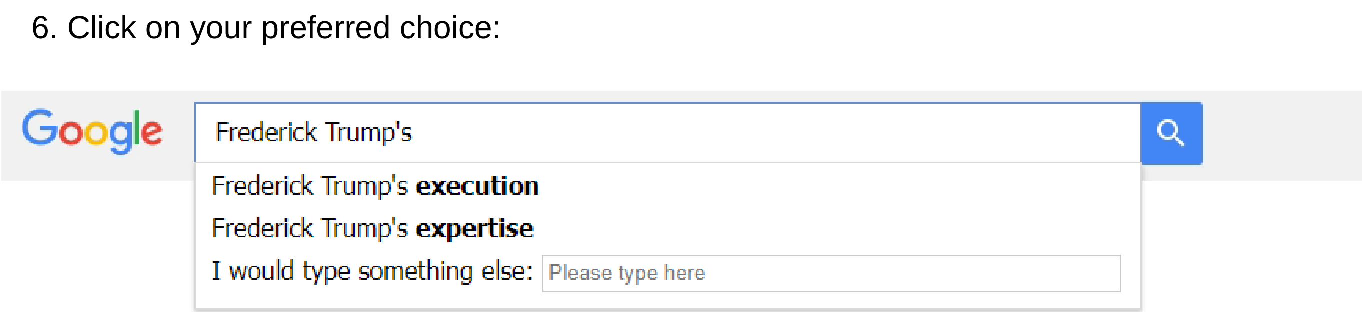


Figure S11. Items and instructions shown on web page in Experiment 4.

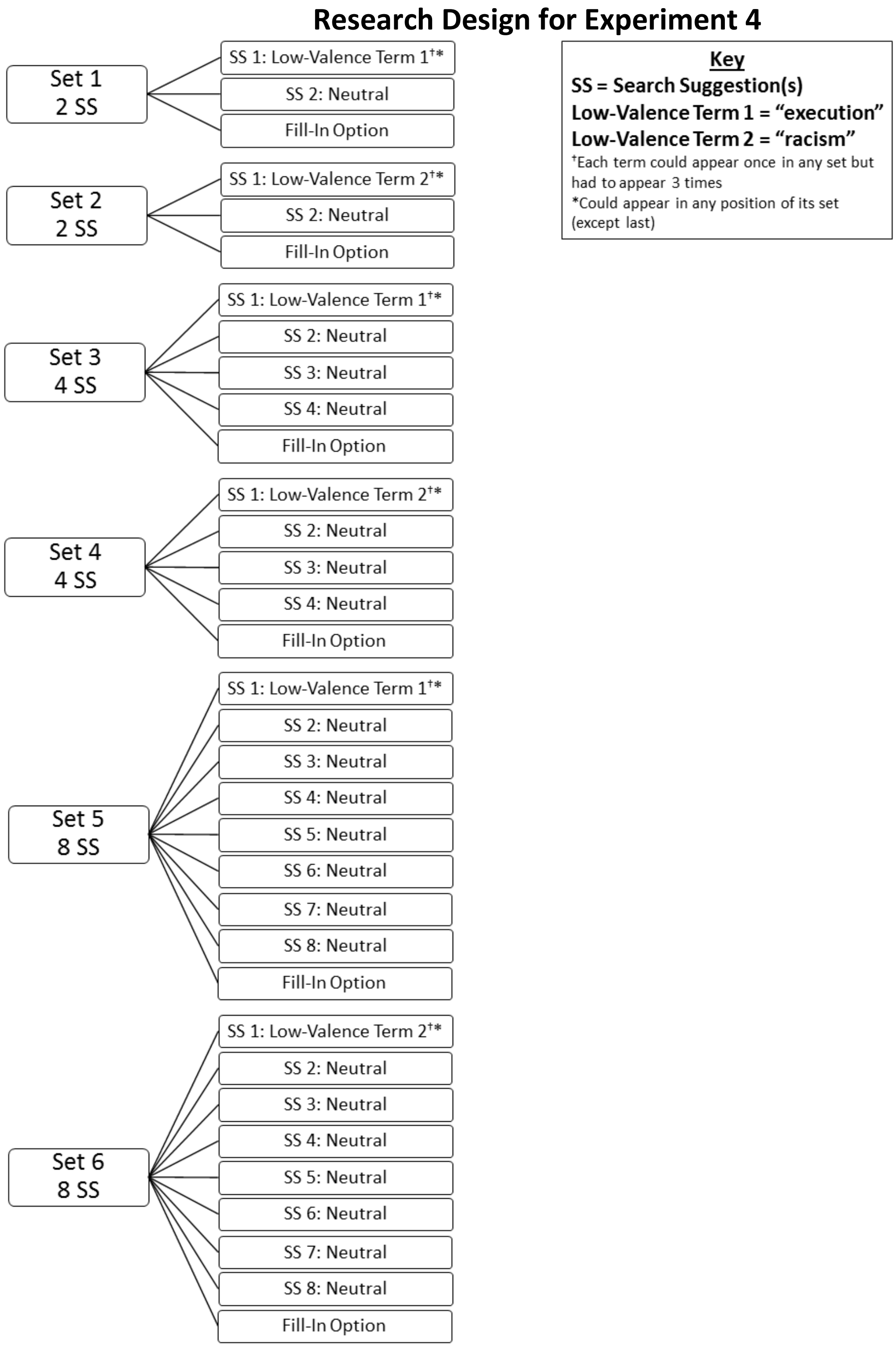


Figure S12. Design of Experiment 4. In Experiment 4, all participants saw six sets of search suggestions, two with two suggestions, two with four suggestions, and two with eight suggestions, all presented in random order. Each set contained one of two low-valence terms, which could appear in any position. Each low-valence term appeared a total of three times.

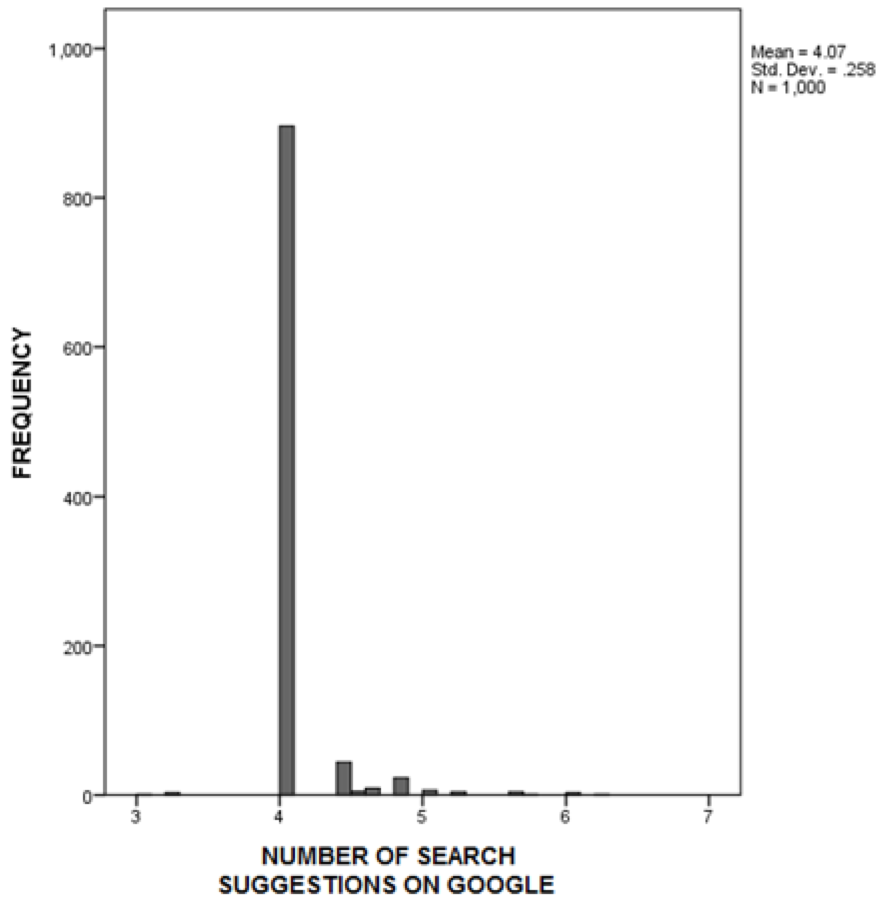


Figure S13. Number of search suggestions shown by Google in response to 1,000 popular search terms, as of March 21, 2017.

Is your overall impression of **Tony Abbott** positive or negative?

Negative ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Positive*

Is your overall impression of **Julia Gillard** positive or negative?

Negative ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Positive*

How *likable* do you find **Tony Abbott** ?

Unlikable ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Likable*

How *likable* do you find **Julia Gillard** ?

Unlikable ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Likable*

How much do you *trust* **Tony Abbott** ?

Not al all ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *A great deal*

How much do you *trust* **Julia Gillard** ?

Not al all ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *A great deal*

PLEASE NOTE: Your answers to these next 2 questions should be consistent with each other! Please read the questions carefully!

If you had to vote today, how likely would you be to vote for either candidate? (NOTE: 0 means you have no preference.) Numbers to the left of 0 mean you favor **Tony Abbott**. And numbers to the right of 0 mean you favor **Julia Gillard**

Tony Abbott ○5 ○4 ○3 ○2 ○1 ○0 ○1 ○2 ○3 ○4 ○5 **Julia Gillard**

If you had to vote right now, which candidate would you vote for?

○ **Tony Abbott** ○ **Julia Gillard**

Click below to continue.

Continue

Figure S14. Experiment 5, pre-search questions about opinions and voting preferences. In the experiment, the order of the candidates' names was counterbalanced.

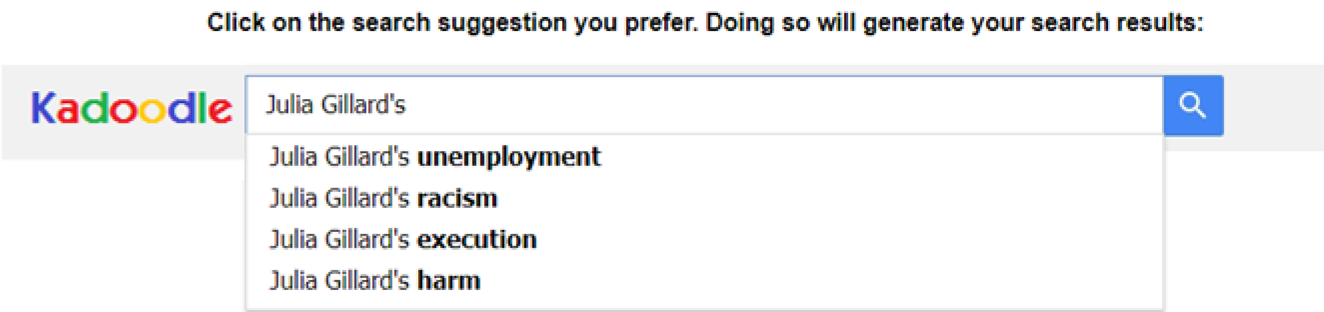


Figure S15. Sample search suggestions shown for Group 3 (four negative search suggestions) in Experiment 5.

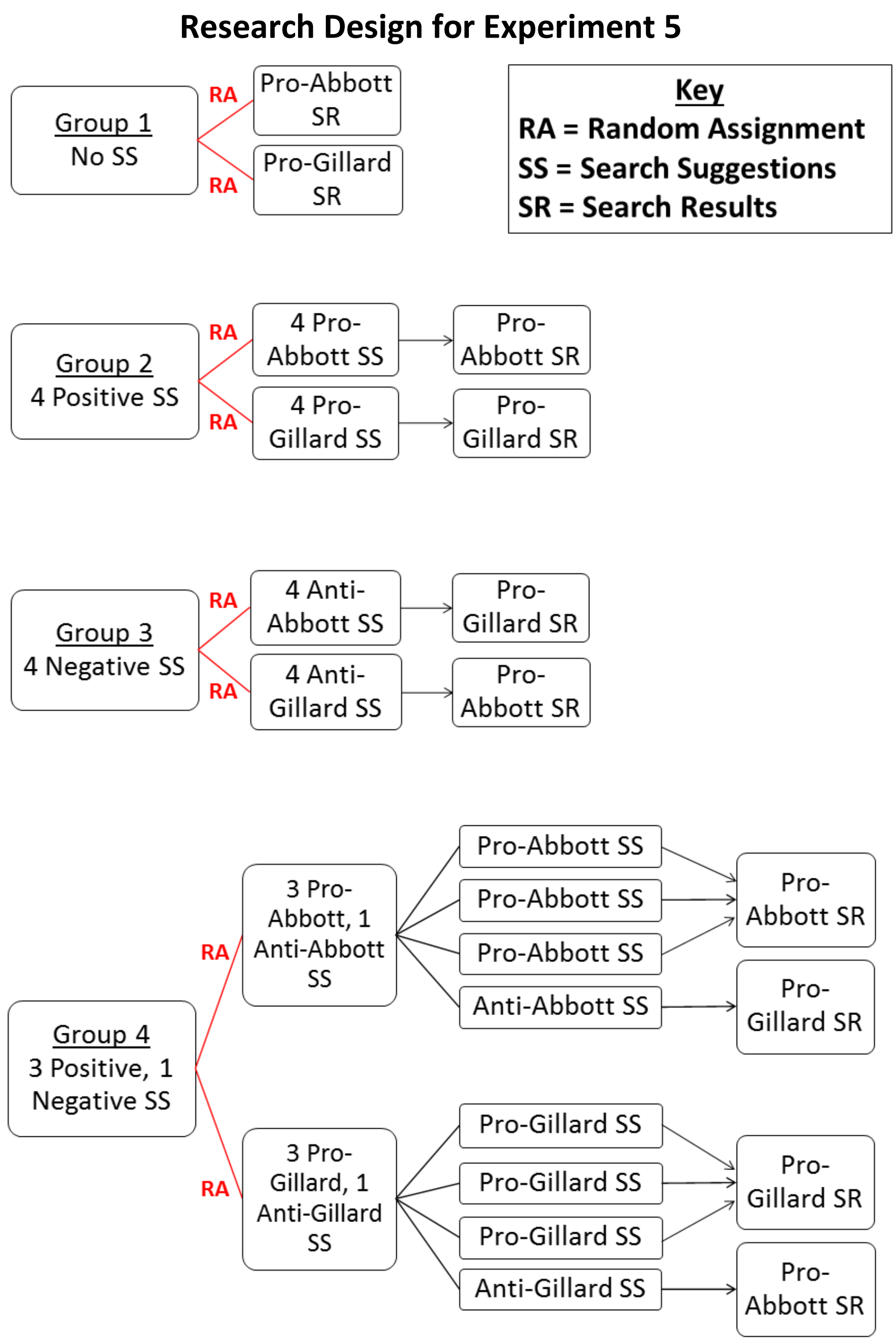


Figure S16. Design of Experiment 5. Experiment 5 was a standard SEME experiment (Epstein & Robertson, 2015) in which participants were randomly assigned to one of four groups. In Group 1, no search suggestions (SSs) were shown. In Group 2, four positive SSs were shown, so a click on any one of them produced search results biased in favor of the candidate named in the suggestions. In Group 3, four negative SSs were shown, so a click on any one of them produced search results biased *against* the candidate named in the suggestions (and thus favoring the opposing candidate). In Group 4, three positive and one negative SSs were shown; a click on any of the positive suggestions produced search results favoring the named candidate, and a click on the negative suggestion produced search results favoring the opposing candidate.

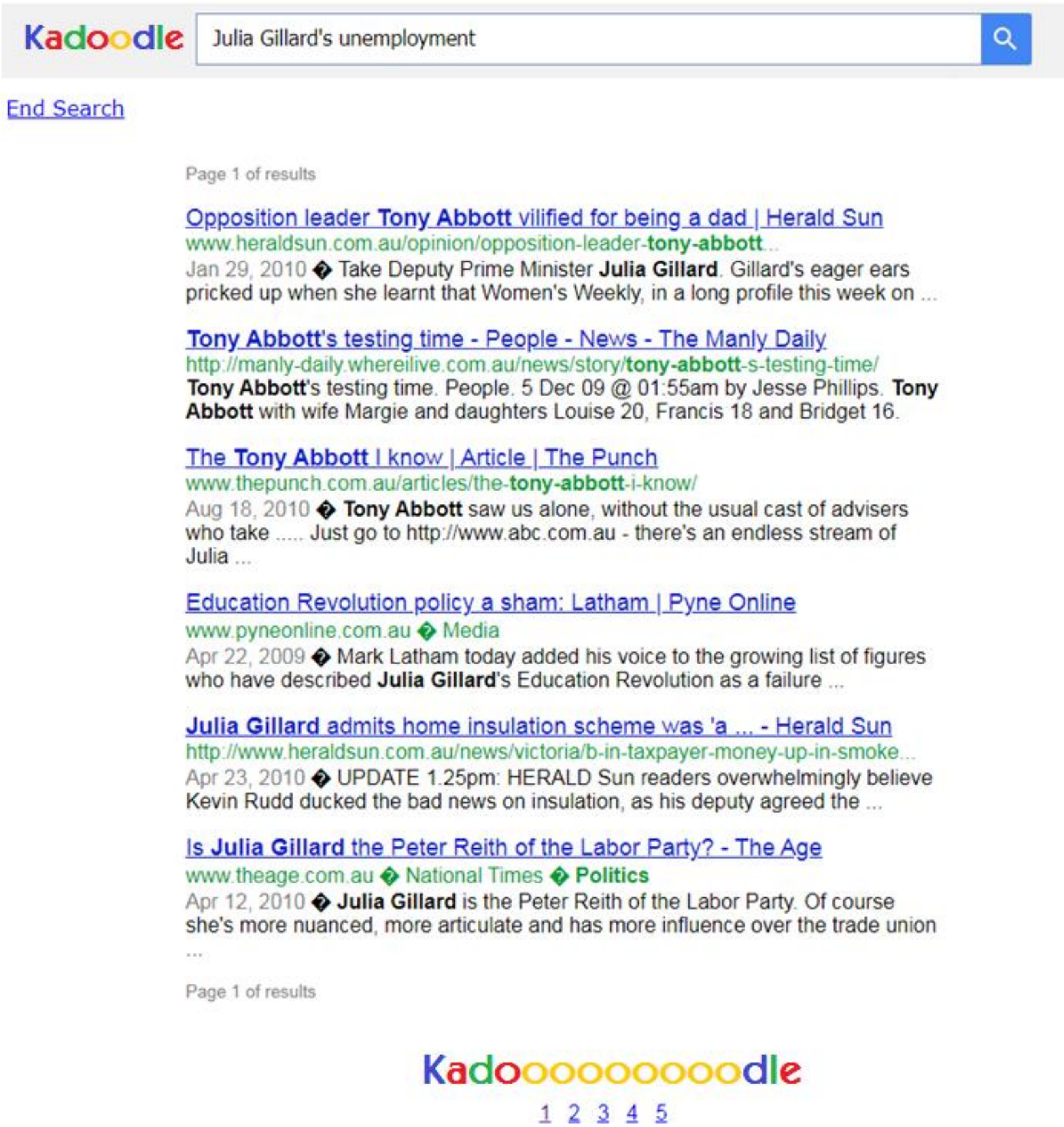


Figure S17. Sample search results (page 1 of 5) for Group 3 in Experiment 5.

Table S1. Demographics characteristics across experiments.

	Experiment 1 (*n* = 609)	Experiment 2 (*n* = 1,126)	Experiment 3 (*n* = 542)	Experiment 4 (*n* = 302)	Experiment 5 (*n* =340)
Mean Age (*SD*)	22.8 (14.2)	21.0 (17.4)	24.2 (18.8)	35.4 (11.2)	36.1 (11.0)
Gender (*n*)					
Male	328	541	231	160	153
Female	222	419	174	141	187
Unknown	59	166	137	1	0
Political Affiliation (*n*)					
Conservative	145	202	102	87	83
Liberal	242	473	199	154	144
Moderate	0	0	163	61	98
None	NA†	NA	NA	NA	11
Other	NA	NA	NA	NA	4
Unknown	0	0	78	0	0
Voter Status					
Decided	446	813	NA	NA	NA
Undecided	155	235	NA	NA	NA
Unknown	8	78	NA	NA	NA
Fluency (SD)	9.6 (0.9)	9.6 (0.9)	9.5 (1.0)	9.9 (0.4)	10.0 (0.16)

†NA = not asked.

Table S2. Experiment 1: Clicks to negative search suggestions vs. clicks to other suggestions.

Sample	*n*	Clicks to Negative Suggestions	Percent of Clicks to Negative Suggestions	Clicks to Control Suggestions	Percent of Clicks to Control Suggestions	Ratio, Clicks to Negatives vs. Clicks to Controls
All	609	507	41.6	88	7.2	5.8
Undecided Voters	155	57	51.8	4	3.5	14.8

Table S3. Experiment 5, Group 1 (no search suggestions): Opinion changes, pre- vs. post-search.

Pre-Search Opinions				
	Favored Candidate	Non-Favored Candidate	*z*[1]	*p*
Impression	7.36 (2.08)	7.11 (2.10)	1.20	0.23 NS
Trust	6.25 (2.12)	6.24 (2.24)	0.37	0.71 NS
Likeability	7.25 (2.06)	7.12 (2.10)	-0.02	0.98 NS
Post-Search Opinions				
	Favored Candidate	Non-Favored Candidate	*z*[1]	*p*
Impression	7.04 (2.37)	3.95 (2.07)	5.85	< 0.001
Trust	6.23 (2.67)	3.41 (2.05)	5.50	< 0.001
Likeability	6.88 (2.46)	3.51 (2.04)	6.24	< 0.001

[1]z value using Wilcoxon test in SPSS.

Table S4. Experiment 5, Group 2 (four positive search suggestions): Opinion changes, pre- vs. post-search.

Pre-Search Opinions				
	Favored Candidate	Non-Favored Candidate	*z*[1]	*p*
Impression	7.42 (1.75)	7.38 (1.89)	-0.07	0.95 NS
Trust	6.52 (1.96)	6.42 (2.06)	0.70	0.50 NS
Likeability	7.18 (1.96)	7.20 (1.87)	-0.30	0.76 NS
Post-Search Opinions				
	Favored Candidate	Non-Favored Candidate	*z*[1]	*p*
Impression	7.56 (1.91)	4.41 (2.43)	6.59	< 0.001
Trust	6.76 (2.12)	4.31 (2.57)	5.92	< 0.001
Likeability	7.30 (2.06)	4.25 (2.45)	6.55	< 0.001

[1]z value using Wilcoxon test in SPSS.

Table S5. Experiment 5, Group 3 (four negative search suggestions): Opinion changes, pre- vs. post-search.

Pre-Search Opinions				
	Favored Candidate	Non-Favored Candidate	z^1	*p*
Impression	7.24 (1.76)	7.17 (1.96)	0.17	0.86 NS
Trust	5.99 (2.02)	5.79 (2.15)	0.55	0.58 NS
Likeability	7.09 (1.81)	6.76 (1.88)	1.51	0.13 NS
Post-Search Opinions				
	Favored Candidate	Non-Favored Candidate	z^1	*p*
Impression	4.30 (1.93)	6.70 (2.16)	-5.23	< 0.001
Trust	4.05 (1.99)	5.90 (2.43)	-4.71	< 0.001
Likeability	4.48 (2.04)	6.66 (2.26)	-4.94	< 0.001

[1]z value using Wilcoxon test in SPSS.

Table S6. Group 4 (three positive search suggestions and one negative suggestion): Opinion changes, pre- vs. post-search.

Pre-Search Opinions				
	Favored Candidate	Non-Favored Candidate	z^1	*p*
Impression	7.37 (1.91)	7.39 (1.97)	0.26	0.79 NS
Trust	6.50 (2.02)	6.61 (1.98)	-0.24	0.81 NS
Likeability	7.25 (2.06)	7.20 (2.03)	0.36	0.72 NS
Post-Search Opinions				
	Favored Candidate	Non-Favored Candidate	z^1	*p*
Impression	5.48 (2.60)	6.31 (2.42)	-1.80	0.07 NS
Trust	5.32 (2.56)	5.86 (2.39)	-1.31	0.19 NS
Likeability	5.39 (2.63)	6.11 (2.52)	-1.60	0.11 NS

[1]z value using Wilcoxon test in SPSS.

Table S7. Experiment 5: Changes in voting preferences on the 11-point scale.

	Group 1: No Search Suggestions	Group 2: Four Positive Suggestions	Group 3: Four Negative Suggestions	Group 4: Three Positive and One Negative Suggestion
Pre-Search Vote for Favored Candidate	0.29 (2.7)	0.11 (2.77)	-0.24 (2.70)	-0.21 (3.08)
Post-Search Vote for Favored Candidate	2.19 (2.62)	1.67 (2.91)	-1.70 (2.74)	-0.46 (3.35)
z^1	-5.39	-5.76	4.82	0.72
p	< 0.001	< 0.001	< 0.001	0.47 NS

[1]z value using Wilcoxon test in SPSS.

Supplemental References

American National Election Studies. (n.d.). *Time of presidential election vote decision 1948-2004*. Retrieved from: https://electionstudies.org/resources/anes-guide/top-tables/?id=109

Annenberg Public Policy Center. (2010, December 8). *2008 National Annenberg Election Survey Telephone and Online Data Sets* (account required for access). Retrieved from: https://www.annenbergpublicpolicycenter.org/2008-naes-telephone-and-online-data-sets/

Bogert, E., Schecter, A., & Watson, R. T. (2021). Humans rely more on algorithms than social influence as a task becomes more difficult. *Scientific Reports*, 11. https://doi.org/10.1038/s41598-021-87480-9

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. https://doi.org/10.1038/nature11421

Dean, B. (2020, August 20). *How people use Google Search*. Backlinko. Retrieved from: https://backlinko.com/google-user-behavior

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), E4512- E4521. https://doi.org/10.1073/pnas.1419828112

Leswig, K. (2018, September 29). *Apple makes billions from Google's dominance in search – and it's a bigger business than iCloud or Apple Music*. Business Insider. Retrieved from: https://www.businessinsider.fr/us/aapl-share-price-google-pays-apple-9-billion-annually-tac-goldman-2018-9/

Liu, Y., Ye, C., Sun, J., Jiang, Y., & Wang, H. (2021). Modeling undecided voters to forecast elections: From bandwagon behavior and the spiral of silence perspective. *International Journal of Forecasting*, *37*(2), 461–483. https://doi.org/10.1016/j.ijforecast.2020.06.011

Logg, J. M., Minson, J. A, & Moore, D. A. (2018, October 26). Do people trust algorithms more than companies realize? *Harvard Business Review*. Retrieved from: https://hbr.org/2018/10/do-people-trust-algorithms-more-than-companies-realize

Mayer, W. G. (2008). *The swing voter in American politics* (pp. 115–119). Washington, DC: Brookings Institution Press.

Pressman, A. (2017, September 25). *How Google is beating Microsoft Bing on the iPhone*. Fortune. Retrieved from: http://fortune.com/2017/09/25/google-bing-default-iphone/

Shearer, E. (2021, January 12). *More than eight-in-ten Americans get news from digital devices*. Pew Research Center. Retrieved from: https://pewrsr.ch/2MZqns7

Statcounter GlobalStats. (n.d.). *Search engine market share worldwide*. Retrieved from: https://gs.statcounter.com/search-engine-market-share

APPENDIX VII

Can Biased Search Results Change People's Opinions About Anything at All? A Close Replication of the Search Engine Manipulation Effect (SEME)

Robert Epstein* (ORCID 0000-0002-7484-6282) and Ji Li (ORCID 0009-0007-4796-1561)

American Institute for Behavioral Research and Technology, Vista, California 92084, United States of America

Submitted for publication. Comments welcome. Do not quote or cite without permission.

*Corresponding author. E-mail address: re@aibrt.org (R. Epstein)

Abstract

A series of experiments published in 2015 in the *Proceedings of the National Academy of Sciences* showed that search engine results favoring one candidate can (a) shift the preferences of undecided voters toward that candidate by up to 80% in some demographic groups and (b) be masked so people show no awareness of the manipulation. We labeled this phenomenon the Search Engine Manipulation Effect (SEME), and it appears to be one of the largest behavioral effects ever discovered. The 2015 experiments and others we have published since then have focused on shifts in voting preferences. In three new experiments with a total of 1,137 US residents (mean age = 33.2), we sought to determine whether biased search rankings could shift people's opinions on topics that do not involve candidates or elections. Each of the new experiments looked at a different topic, and participants were pre-screened to make sure they didn't have strong opinions about these topics. The topics were: Is artificial intelligence useful or dangerous? Is fracking helpful or dangerous? And: Are people born gay or do they choose to be gay? All participants were first asked various demographic questions, then shown brief summaries of the "pro" and "anti" views on each topic, and then asked their opinions about each topic. Next, participants were allowed to conduct an online search using our mock search engine (Kadoodle) lasting up to 15 minutes. All search results were real and linked to real web pages; only the order of search results varied from group to group. In each experiment, one-third of the participants saw biased search results favoring one perspective; one-third saw biased search results favoring the opposing perspective; and one-third (the control group) saw mixed search results. After completing their search, participants were again asked for their opinions about the topic. Our primary dependent variable was Manipulation Power (MP), the overall increase in the number of participants favoring one viewpoint after having viewed search rankings favoring that viewpoint. The MPs in the three experiments were 25.0%, 30.9%, and 17.8%, respectively. Corresponding shifts were also found for how persuasive participants found each viewpoint to be and for how much they trusted each viewpoint. We conclude that biased search rankings can impact more than voting preferences. It appears that search rankings favoring one viewpoint on a wide range of topics can cause people who have not yet formulated a strong opinion on such topics to adopt the favored perspective. If our findings prove to be robust, we are exposing what might be considered a fatal flaw in search engines, namely that even without human interference, search algorithms will inevitably alter the thinking and behavior of billions of people worldwide on perhaps any topic for which they have not yet formed strong opinions. With self-determining AIs being rapidly integrated into search algorithms, one might wonder whether the manipulative power of personalized, biased search results will be harnessed in coming years in ways that serve humanity's interests.

Keywords: search engines; search engine manipulation effect; SEME; online manipulation; manipulation power; VMP; MP

Can Biased Search Results Change People's Opinions About Anything at All? A Close Replication of the Search Engine Manipulation Effect (SEME)

1. Introduction

Controlled experiments conducted in recent years have shown that bias in search engine results can rapidly shift the opinions and voting preferences of undecided voters – by as much as 80% in some demographic groups [1]. Research has also shown that this effect – the search engine manipulation effect (or SEME, pronounced "seem") – can easily be masked so that no users are aware of the bias they are seeing [1]. SEME has been partially or fully replicated multiple times since it was first published in 2015 [2-16].

Research has also shown that when people do suspect that they are viewing biased search results, that awareness does not necessarily protect them from the impact of the bias. Epstein and Robertson [1] showed in a controlled experiment with more than 2,000 participants from all 50 US states that the few people (8.6%) who could recognize the bias shifted even farther, on average, than people who did not recognize the bias. Why this occurred is not clear, but it could have been because people have inordinate faith in the validity of computer output, at least in part because they have no idea how computers work [17,18]. Research has also shown that vulnerability to SEME and other new forms of online manipulation varies substantially from one demographic group to another [1,6-12].

At least three other features of search engines make them potentially problematic, at least in the eyes of some experts and public policy makers: First, all the content shown to users on search engines is ephemeral. Search suggestions, answer boxes, and search results are all generated on the fly, impact users, and then disappear, and online ephemeral content can impact people's thinking significantly [19,20]. Because they are not stored anywhere, they leave no paper trail for authorities to trace. If biased search results shifted votes in an election – perhaps, in a close election, so many votes that that bias determined the outcome – there would be no way to go back in time to document such an effect [21,22,cf. 23].

Second, over the years, search engines have based their content on increasingly vast amounts of data they have collected about each user; in other words, they now personalize (or "customize") content to meet the needs of individual users [24]. Many users like this feature of modern search engines, which they consider to be the digital equivalent of personal shoppers [25,26]. On the downside, a long history of research in the marketing and advertising fields has shown that the more one knows about the customer, the easier it is to manipulate him or her [27,28]. This ability applies as much to voters as it does to shoppers [29,30].

Third, because about 92% of search worldwide – everywhere outside of mainland China (the PROC) – is conducted on just one search engine, with no other search engine attracting more than 4% of search [31, 32,cf. 33], if the leading search engine chooses to shift votes or opinions in just one direction, there is no way to counteract that very powerful form of influence. If the bias has been masked, there also may be no way to detect it. This is very different from most forms of influence that affect people every day, especially in the days leading up to elections. Most forms of influence – billboards, television and radio commercials, newspapers ads and editorials, online ads and podcasts – are inherently competitive. If you have the resources, you counter your opponent's brutal attack ad with one or more ads that are even more brutal. But if the dominant search engine chooses – either by deliberate acts of its employees or by unconscious or neglectful management of its algorithms [20,22,34] – to support one political candidate or party, there is no way to counteract that influence.

For these reasons, it is important to understand how SEME works, who it affects, and the magnitude of its ability to alter opinions, beliefs, purchases, behavior, and votes. As we have been arguing elsewhere in recent years, it is also important that we develop permanent monitoring systems that can preserve and analyze ephemeral content on a large scale [35-37]. If we don't preserve ephemeral content, we will never know how and to what extent existing and emerging tech companies are impacting our minds, our children, and our political systems.

Nearly all the research that has been conducted on SEME has focused on only one of these domains – namely, the ability of biased search results to alter the opinions and voting preferences of undecided voters. We are aware of one conference presentation in which SEME was partially replicated in a context involving people's knowledge about health issues [38]. An earlier study found that search results linked to webpages that contained high-quality information about vaccines communicated more knowledge to people than did search results linked to low-quality webpages [39], but that study did not measure opinion shifts.

We are left with a consequential question that we believe has not yet been answered adequately: Can biased search results shift people's opinions not just about political candidates but about a wide range of different topics – perhaps any topic at all? We acknowledge that this question applies mainly, if not exclusively, to people who have not yet made up their mind about that topic: about where to go on vacation, about what kind of car they should buy, about whether gays should be able to marry or adopt children, and so on. How much power do biased search results have, across a wide range of different topics and issues, to shift the opinions and behavior of people who are vulnerable to being influenced?

2. Experiment 1. Can Biased Search Results Shift People's Views About Artificial Intelligence (AI)?

2.1 Method

2.1.1 Ethics statement

The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB is FWA00021545. Informed written consent was obtained for all three experiments as specified in the Procedure section below.

2.1.2 Participants

378 participants were recruited online from the Amazon Mechanical Turk (MTurk) subject pool during March 2016. The mean age of our participants was 33.8 (*SD* = 11.4). 56.3% (n = 213) of our participants identified themselves as female and 43.7% (n = 165) as male. For detailed information about the basic demographic characteristics of our participants in all three experiments, see Table S1.

91.3% (n = 345) of our participants reported using Google as their primary search engine; 4.2% (n = 16) reported using Bing, 2.6% (n = 10) reported using Yahoo, and 1.9% (n = 7) reported using some “other” search engine. Participants reported conducting a wide range of number of searches per week – from 1 per week to over 200 per week (*M* = 12.8, *SD* = 17.1). 44.2% (n = 167) of participants identified themselves as politically liberal, 30.2% (n = 114) as moderate, 17.7% (n = 67) as conservative, 6.6% (n = 25) as having no political viewpoint, and 1.3% (n = 5) as other.

We asked participants how familiar they were with arguments that favored the use of artificial intelligence (AI) and arguments that were critical of AI on a scale from 1 to 10, where 1 represented “Not familiar at all” and 10 represented “Very familiar.” The mean familiarity level with pro-AI arguments was 3.1 (*SD* = 2.2), and the mean familiarity level with anti-AI arguments was 3.1 (*SD* = 2.3).

2.1.3 Procedure

The experiment was conducted online and employed a pre/post design. Participants were first asked, “Do you have strong beliefs about artificial intelligence?” and only people who clicked “No” were allowed to continue. Then participants were given basic instructions and asked for their informed consent (S5 Text). As required by the sponsoring institution’s IRB, participants were not asked for identifying information such as name, email address, or telephone number. The participants were then asked a series of demographic questions. They were shown brief (about 100 words) paragraphs about AI. The first paragraph presented a point of view favoring AI, and the second presented a point of view opposing AI (see S2 Text for the full content).

Participants were then asked six opinion questions about AI: two regarding their overall impressions, two regarding how persuasive they found the two viewpoints they had read, and two regarding how much they trusted those two viewpoints; for the full text of the questions, which participants answered on 10-point Likert scales, see S1 Fig. Next, on an 11-point scale from 5 to 0 to 5 (S1 Fig), participants indicated which viewpoint they favored, with "Pro AI" and "Anti AI" appearing at each end of the scale, with the positions counterbalanced. Finally, participants were asked to choose which viewpoint they favored in a forced-choice question (S1 Fig) – again, with the positions of the answers counterbalanced. This page of questions comprised the pre-search test.

At the beginning of the experiment, all participants were randomly assigned to one of three groups: Pro-AI, Anti-AI, or the Control Group, in which people saw alternating pro- and anti-AI arguments. The sequences are shown in Fig 2.

All 30 webpages used in this experiment had previously been rated by five independent reviewers on an 11-point scale from 5 to 0 to 5, where "Pro AI" and "Anti AI" appeared at either end of the scale, and their order was counterbalanced. Based on the mean ratings of the reviewers, the search results were ranked from the most Pro AI (referring to the web page to which the search result linked) to the most Anti AI (again, referring to the web page to which the search result linked), with the relatively neutral search results in the middle (Fig 2, Group 1).

After participants answered those eight questions about the pro- and anti-AI points of view, they were then given up to 15 min to use our Kadoodle search engine – a Google simulator – to learn more about AI. Our search engine showed participants five pages of search results, with six search results per page. Participants could click on any of the results and could switch between the pages by clicking on numbers at the bottom of each page (see Fig 1).

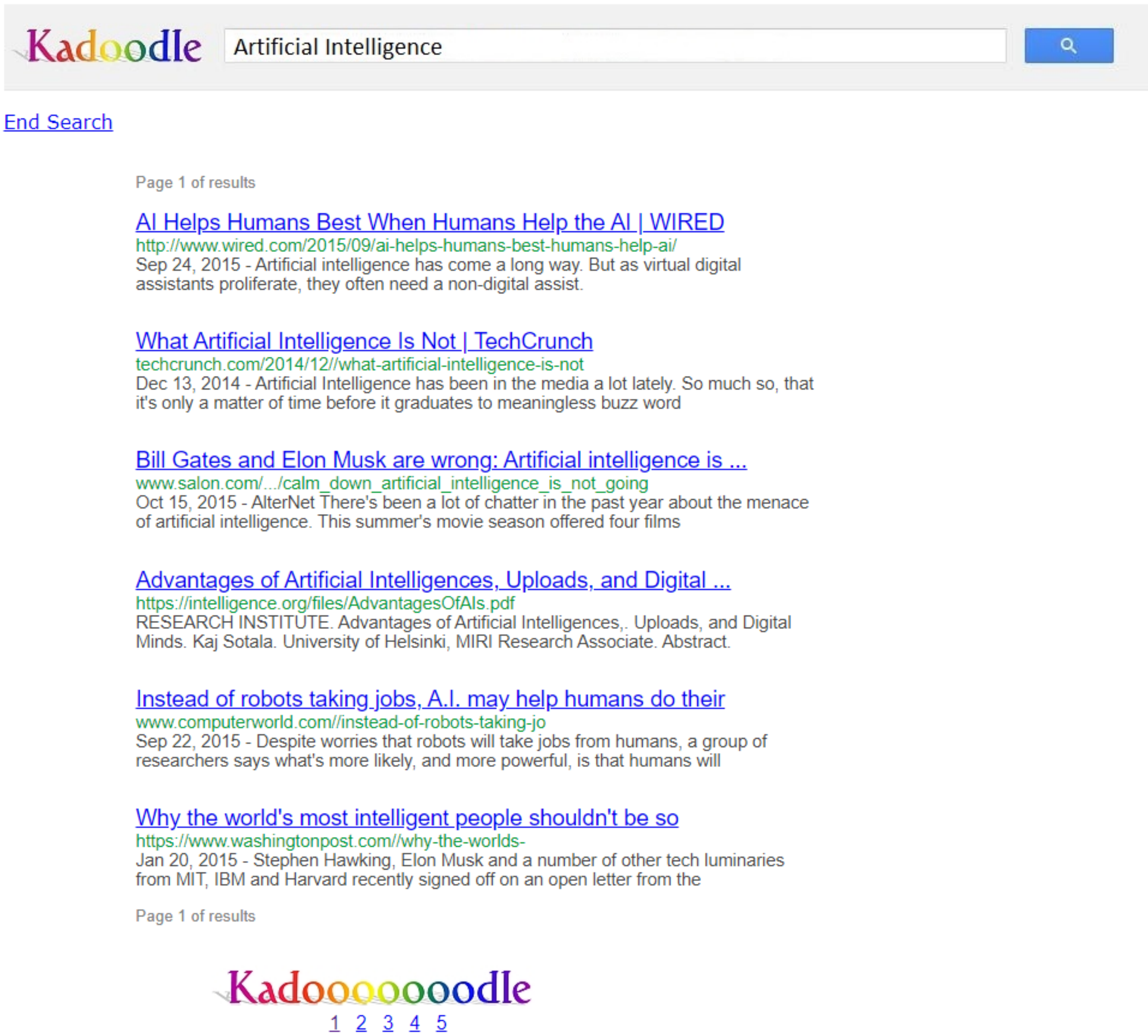


Fig 1. Example of Kadoodle search results page. Each search results webpage contained six different search results. The participant could click on a link to view the corresponding web page, or he or she could click on one of the numbers at the bottom of the page to switch to a different page of search results. The "End Search" shortcut can be seen in the top left corner of the page.

All search results were real, scraped from the Google search engine, and all webpages were real, scraped from the internet. The webpages were presented as image files created from the original pages, with no active links.

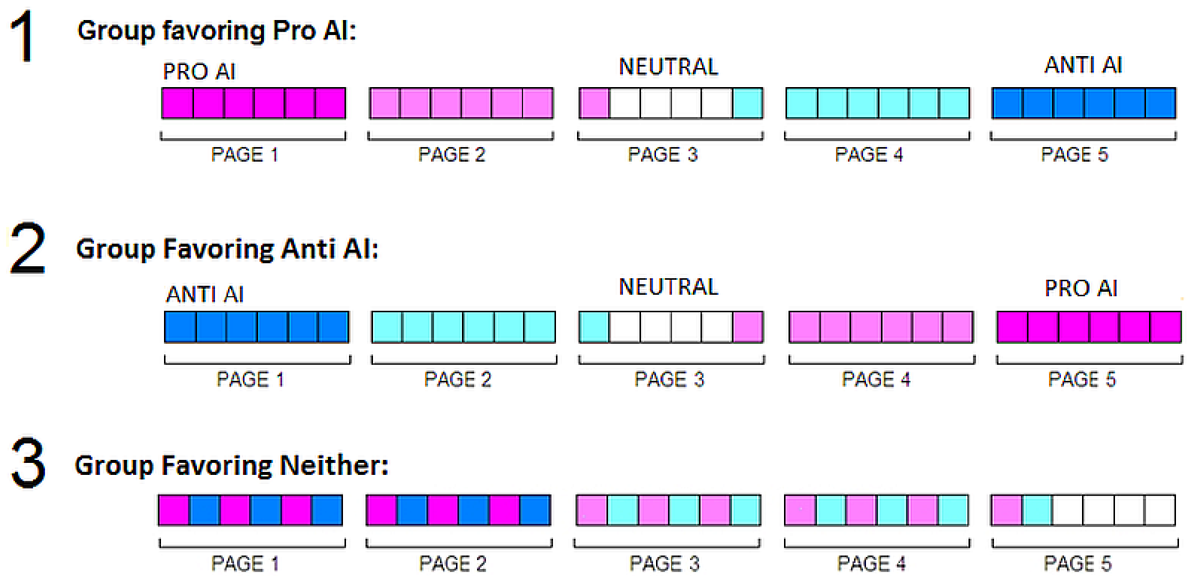


Fig 2. Ordering of search results for the three groups. Each small square represents a search result, and each group of six squares represents the search results on one page. Dark pink signifies that a search result links to a web page in which the content has been rated (by independent raters) to be pro-AI. Light pink signifies that search results link to web pages that are less favorable to AI. White signifies that the linked web pages are relatively neutral toward AI. Light blue signifies that search results link to web pages that are somewhat anti-AI. Medium blue signifies that search results link to web pages that are strongly anti-AI. In Group 1 (Pro AI), the search results are in order from pro-AI to anti-AI. In Group 2 (Anti-AI), the search results are in the opposite order. In Group 3 (Control), pro- and anti-AI search results alternate.

Participants could end their search by clicking a button on the top left of the screen that read "End Search" (Fig 1). If they failed to click the button, the search session would end when a 15-min time limit was reached.

Now participants were again asked to answer the six opinion questions and two voting questions they had answered prior to search (the post-search test). After participants responded to these questions, they were asked whether anything in the experiment had bothered them. If they answered "yes," they could then explain what had bothered them in a text box. The purpose of asking participants about what bothered them was to determine whether they detected bias in the search results. We could not ask directly about bias because leading questions of that sort are known to inflate estimates [40].

Finally, participants were thanked for their participation in the experiment and provided with a code number that they could use to be paid by MTurk.

2.2 Results

Most of the experiments we have conducted on online manipulation since we began this type of research in 2013 [1,6-13] have used "vote manipulation power" (VMP) as the most informative metric of change. VMP was defined as the post-manipulation percentage increase in the number of people voting for the candidate favored in the manipulation (see S1 Text for further details). Because we are now extending our investigation to look at topics that are not election-related, we are introducing a broader variant of VMP, calling it simply "manipulation power" or MP. We define MP as the post-manipulation percentage increase in the number of people choosing the opinion favored in the manipulation (or the belief, candidate, product, perspective, or other categorical content that can be made to look superior to an alternative).

In Experiment 1, the MP was 25.0% (McNemar's Test $X^2 = 22.22$, $p < 0.001$), which means that in the two bias groups combined, the bias in the search results increased the number of people choosing either a pro-AI or an anti-AI perspective by 25.0%. Specifically, before the search was conducted, the total number of people in the two bias groups who chose the favored perspective was 124. After the search, that number increased by 25.0% to 155.

On the 11-point voting preference scale, pre-manipulation, we found no significant difference between the mean ratings in the three groups ($M_{Pro} = 1.62$, $SD = 2.5$; $M_{Anti} = 1.78$, $SD = 2.4$; $M_{Control} = 1.31$, $SD = 2.5$; Kruskal-Wallis $H = 3.13$; $p = 0.209$ NS). Post manipulation, we found a significant difference between mean ratings in the three groups ($M_{Pro} = 1.99$, $SD =$

2.7; M_{Anti} = 0.18, SD = 3.0; M_{Control} = 1.03, SD = 2.9; H = 24.68; p < 0.001). Participants in Group 1 shifted 0.37 points toward the favored opinion (Pro AI), and participants in Group 2 condition shifted 1.6 points towards the favored opinion (Anti AI). In addition, the pre-manipulation mean preference for the favored opinion (Groups 1 and 2 combined) was significantly different from the post-manipulation mean preference for the favored opinion (Groups 1 and 2 combined) (M_{Pre} = 0.004; SD_{Pre} = 3.0; M_{Post} = 0.96; SD_{Post} = 3.1; M_{Diff} = 0.956; Wilcoxon Signed Ranks z = 6.00; p < 0.001).

The shift was also indicated by three measures for each of the two opposing opinions: measures of overall impression, persuasiveness, and level of trust (S1 Fig). Pre to post, the mean favored opinions increased for all three measures, and the non-favored opinions decreased for all three measures. Pre to post, the overall change in opinions was highly significant for all three measures and was in the predicted direction (Table 1).

In the two bias groups combined (Groups 1 and 2, n = 246), the number of people who noticed bias in the search results they saw was 38.2%. This is consistent with the level of bias perception in other SEME experiments when masking has not been employed to disguise the bias [1,4-6].

Table 1. Experiment 1: Pre- and post-search opinion ratings of the favored and non-favored candidate measured on 10-point scales, bias groups only

	Favored Opinion Mean (*SD*)			Non-Favored Opinion Mean (*SD*)				
	Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	$z^{\dagger}$	p
Overall Impression	5.65 (2.6)	6.46 (2.5)	+ 0.81	5.74 (2.6)	4.98 (2.4)	- 0.76	-5.20	< 0.001
Persuasiveness	5.63 (2.4)	6.88 (2.5)	+ 1.25	5.76 (2.3)	5.09 (2.6)	- 0.67	-6.46	< 0.001
Level of Trust	5.61 (2.2)	6.63 (2.3)	+ 1.02	5.85 (2.1)	5.36 (2.4)	- 0.49	-6.15	< 0.001

†The z values come from Wilcoxon Signed Ranks Test between the post-search minus pre-search ratings for the favored candidate and the post-search minus pre-search ratings for the non-favored candidate.

3. Experiment 2. Can Biased Search Results Shift People's Views About Fracking?

3.1 Methods

3.1.1 Participants

394 participants were recruited online from MTurk during March 2016. The mean age of our participants was 32.9 (*SD* = 10.2). 52.0% (n = 205) of our participants identified themselves as female and 48.0% (n = 189) as male. For detailed information about basic demographic characteristics, see Table S1.

94.2% (n = 371) of the participants selected Google as their primary search engine, 4.1% (n = 16) as Bing, 1.5% (n = 6) as Yahoo, and 0.3% (n = 1) as other. Participants reported the number of searches they conducted per week ranging from 0 to 150 (*M* = 14.4, *SD* = 18.1). 42.4% (n = 167) of participants reported being liberal, 33.8% (n = 133) as moderate, and 15.5% (n = 61) as conservative; 6.1% (n = 24) reported having no political views, and 2.3% (n = 9) reported their political viewpoint as other. The mean familiarity level of participants with pro-fracking arguments was 3.6 (SD = 2.3); for anti-fracking arguments, it was 4.2 (SD = 2.5).

3.1.2 Procedure

The procedure for Experiment 2 was the same as for Experiment 1, except that the topic was fracking.

3.2 Results

In Experiment 2, the MP was 30.9% (McNemar's Test $X^2 = 25.14$, $p < 0.001$), which means that in the two bias groups combined, the bias in the search results increased the number of people choosing either a pro-fracking or an anti-fracking perspective by 30.9%. Specifically, before the search was conducted, the total number of people in the two bias groups who chose the favored perspective was 136. After the search, that number increased by 30.9% to 178.

On the 11-point voting preference scale, pre-manipulation, we found no significant difference between the mean ratings in the three groups ($M_{\text{Pro}} = -0.79$, $SD = 2.7$; $M_{\text{Anti}} = -0.68$, $SD = 2.7$; $M_{\text{Control}} = -0.24$, $SD = 2.7$; $H = 2.48$; $p = 0.289$ NS). Post manipulation, we found a significant difference between mean ratings in the three groups ($M_{\text{Pro}} = -0.09$, $SD = 3.2$; $M_{\text{Anti}} = -2.44$, $SD = 2.8$; $M_{\text{Control}} = -0.97$, $SD = 3.7$; $H = 40.35$; $p < 0.001$). Participants in Group 1 shifted 0.7 points toward the favored opinion (Pro Fracking), and participants in Group 2 condition shifted 1.76 points towards the favored opinion (Anti Fracking). In addition, the pre-manipulation mean preference for the favored opinion (Groups 1 and 2 combined) was significantly different from the post-manipulation mean preference for the favored opinion (Groups 1 and 2 combined) ($M_{Pre} = -0.09$; $SD_{Pre} = 2.8$; $M_{Post} = 1.12$; $SD_{Post} = 3.3$; $M_{Diff} = 1.21$; $z = 8.06$; $p < 0.001$).

The shift was also indicated by three measures for each of the two opposing opinions: measures of overall impression, persuasiveness, and level of trust (S1 Fig). Pre to post, the mean favored opinions increased for all three measures, and the non-favored opinions decreased for all three measures. Pre to post, the overall change in opinions was highly significant for all three measures and was in the predicted direction (Table 2).

In the two bias groups combined (Groups 1 and 2, $n = 286$), the number of people who noticed bias in the search results they saw was 50.3%. This is higher than the typical level of bias perception we have found in other SEME experiments when masking has not been employed to disguise the bias [1,4-6].

Table 2. Experiment 2: Pre- and post-search opinion ratings of the favored and non-favored candidate measured on an 11-point scale, bias groups only

	Favored Opinion Mean (*SD*)			**Non-Favored Opinion Mean (*SD*)**				
	Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†**	***p***
Overall Impression	5.73 (2.3)	6.68 (2.6)	+ 0.95	5.78 (2.3)	4.80 (2.6)	- 0.98	-7.66	< 0.001
Persuasiveness	5.77 (2.2)	6.95 (2.5)	+ 1.18	5.66 (2.3)	4.75 (2.7)	- 0.91	-7.47	< 0.001
Level of Trust	5.44 (2.0)	6.26 (2.6)	+ 0.82	5.49 (2.1)	4.77 (2.6)	- 0.72	-6.56	< 0.001

† The z values come from Wilcoxon signed ranks test between post-search minus pre-search ratings for the favored candidate and the post-search minus pre-search ratings for the non-favored opinion.

4. Experiment 3. Can Biased Search Results Shift People's Views About Sexual Orientation?

4.1 Methods

4.1.1 Participants

365 participants were recruited online from MTurk during March 2016. The mean age of our participants was 32.8 (*SD* = 10.6). 55.9% (n = 204) of our participants identified themselves as female and 44.1% (n = 161) as male. For detailed information about basic demographic characteristics, see Table S1.

93.7% (n = 342) of our participants reported Google as their primary search engine, 2.5% (n = 9) as Bing, 2.2% (n = 8) as Yahoo, and 1.6% (n = 6) as other. Participants reported the number of searches they conducted per week ranging from 0 to 250 (*M* = 14.1, *SD* = 24.6). 42.7% (n = 156) of participants reported being liberal, 36.7% (n = 134) as moderate, and 12.1% (n = 44) as conservative; 5.8% (n = 21) reported having no political views, and 2.7% (n = 10) reported their political viewpoint as other. The mean familiarity level of participants with born-gay arguments was 7.3 (*SD* = 2.6); for choose-to-be-gay arguments, it was 7.3 (*SD* = 2.5).

4.1.2 Procedure

The procedure for Experiment 3 was the same as for Experiment 1, except that the topic was sexual orientation – specifically, whether people are born gay or whether they choose to be gay.

4.2 Results

In Experiment 3, the MP was 17.8% (McNemar's Test $X^2 = 11.81$, $p < 0.001$), which means that in the two bias groups combined, the bias in the search results increased the number of people choosing either a pro-AI or an anti-AI perspective by 17.8%. Specifically, before the search was conducted, the total number of people in the two bias groups who chose the favored perspective was 135. After the search, that number increased by 17.8% to 159.

On the 11-point voting preference scale, pre-manipulation, we found no significant difference between the mean ratings in the three groups ($M_{\text{Born}} = 1.54$, $SD = 2.9$; $M_{\text{Choose}} = 1.27$, $SD = 3.1$; $M_{\text{Control}} = 1.86$, $SD = 2.9$; $H = 2.25$; $p = 0.325$ NS). Post manipulation, we found a significant difference between mean ratings in the three groups ($M_{\text{Born}} = 2.56$, $SD = 3.0$; $M_{\text{Choose}} = 0.41$, $SD = 3.6$; $M_{\text{Control}} = 2.01$, $SD = 3.2$; $H = 26.87$; $p < 0.001$). Participants in Group 1 shifted 1.02 points toward the favored opinion (Born Gay), and participants in Group 2 condition shifted 0.86 points towards the favored opinion (Choose to be Gay). In addition, the pre-manipulation mean preference for the favored opinion (Groups 1 and 2 combined) was significantly different from the post-manipulation mean preference for the favored opinion (Groups 1 and 2 combined) ($M_{Pre} = 0.12$; $SD_{Pre} = 3.3$; $M_{Post} = 1.05$; $SD_{Post} = 3.6$; $M_{Diff} = 0.93$; $z = 7.13$; $p < 0.001$).

The shift was also indicated by three measures for each of the two opposing opinions: measures of overall impression, persuasiveness, and level of trust (S1 Fig). Pre to post, the mean favored opinions increased for all three measures, and the non-favored opinions decreased for all three measures. Pre to post, the overall change in opinions was highly significant for all three measures and was in the predicted direction (Table 3).

In the two bias groups combined (Groups 1 and 2, $n = 252$), the percentage of people who noticed bias in the search results they saw was 28.6%. This is similar to the typical level of bias perception we have found in other SEME experiments when masking has not been employed to disguise the bias [1,4-6].

Table 3. Experiment 3: Pre- and post-search opinion ratings of the favored and non-favored candidate measured on an 11-point scale, bias groups only

	Favored Opinion Mean (*SD*)			**Non-Favored Opinion Mean (*SD*)**				
	Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†**	***p***
Overall Impression	6.00 (2.6)	6.65 (2.9)	+ 0.65	5.93 (2.7)	5.24 (3.0)	- 0.69	-5.97	< 0.001
Persuasiveness	5.70 (2.8)	6.63 (3.1)	+ 0.93	5.69 (2.8)	4.87 (3.1)	- 0.82	-7.91	< 0.001
Level of Trust	5.77 (2.7)	6.52 (3.0)	+ 0.75	5.82 (2.8)	5.08 (3.1)	- 0.74	-6.90	< 0.001

†The z values come from Wilcoxon Signed Ranks Test between the post-search minus pre-search ratings for the favored candidate and the post-search minus pre-search ratings for the non-favored candidate.

5. Discussion

In the three experiments we have described above, we produced shifts in preferences of 25.0%, 30.9%, and 17.8%, respectively, after our participants conducted just one search on our Kadoodle search engine. These numbers are based on shifts in the two bias groups combined. The fact that participants were randomly assigned to one or the other of those two groups means we were able to shift people's thinking for or against a particular perspective arbitrarily. Mean voting preferences on an 11-point scale and mean opinion ratings also shifted predictably in the direction of the bias. These results support our conjecture that bias in the online search results displayed to users by search engine companies have the potential to change people's thinking about – well, perhaps anything at all.

Key to this finding is the fact that we deliberately worked with people who did not, to begin with, already have strong opinions about the three topics we explored. Presumably, people with strong opinions about such matters would be difficult to influence with biased search results [e.g., 41]. That is a matter we are continuing to investigate in ongoing research. If indeed people who are undecided on some issue are the most vulnerable to such a manipulation, it is notable that search engine companies are not only in a unique position to employ bias in search results to impact people's thinking, they are also in a unique position to identify people who are most vulnerable to this type of manipulation – that is, people who have not yet made up their minds. A company such as Google, which openly tracks people through their emails [42] (using Gmail, the most widely used email system in the world [43]), online searches (using the Google search engine, which handles 92% of online search in most countries [31,32]), Chrome (the most widely used browser in the world [44]), Android (the most widely used mobile operating system in the world [45]), and many other platforms and applications [46,47], can easily identify people who are undecided or uncommitted on some issue.

Even without intent by employees or executives at Google and other companies that operate search engines, the power that biased search results appear to have to shift opinions about a wide range of topics should be a matter of great concern to legislators, regulators, and public policy makers. We make this strong assertion because, by definition, search algorithms always do three things: they *filter* content (by selecting a small amount of content to display while setting aside a vast amount of other content), they *order* content (by ranking the content they will display), and they *customize* content (by adjusting both the filtering and ordering to best match the interests and needs of the user). In other words, in some sense *all* search results are biased, and therein lies their value. For a given user, search results will always favor one dog food over another, and we wouldn't want it any other way. The problem is that for people sitting on a fence, that customized, filtered, and ranked content the search engine shows them appears to be effective as a tool for pushing people to tumble off one side of that fence.

To put this another way, the search engine is not only the most powerful tool ever invented for providing factual answers to simple questions, it might also be the most powerful mind control machine ever invented, even if mind control was never the intent. In a separate study [48] we examined this issue from the perspective of operant conditioning. About 86% of the searches people conduct on major search engines are for simple facts [49], and those facts almost invariably turn up in the top position of search results. Like rats in a Skinner box, we thus learn, over and over again, that the most valuable and accurate search results are the ones at the top of the list. When that day comes when we pose an open-ended query – "best restaurant in Atlanta," "is fracking safe?," or "how to solve the immigration problem" – we again tend to attend to and trust those high-ranking links, which will bring us to web pages that likely favor one perspective. That should surprise no one; there are no equal-time rules in search algorithms, after all. They are designed to find the "best" results, not to show a series of pro- and anti- results in alternating order (like the results we showed in our Control Group). Google does so by examining link patterns [50], but no matter what technique is used, a search algorithm will always, or nearly always, tend to favor one perspective over another. That favoritism might occur because one perspective is dominant on the internet, because of the conscious or unconscious biases of the programmers who created and maintain the search algorithm [51-53], or because of company policies that elevate or suppress content deliberately through white listing or black listing [34,54]. The present study extends previous research only in helping to shed light on one issue: Could search engine bias shift people's views about a wide range of different topics? The answer appears to be yes.

To put this issue yet another way: SEME is a list effect with a difference. Unlike other list effects researchers have studied over the past century, beginning with the serial position effect [55-57], SEME is supported by a daily regimen of operant conditioning that will never stop. Simple factual searches will continue to teach people *ad nauseum* that high-ranking search results are truer and more valid than lower ranking search results. Presumably this is why companies worldwide spend vast sums each year trying to push their products a notch or two higher in Google search results; a single increment can increase clicks by 32.3% [58].

In research we are currently conducting on what we call the “digital personalization effect” (DPE), we are learning that personalization – for example, showing people content from sources we know they trust – can dramatically increase the impact of SEME and other new forms of influence the internet has made possible [59]. When you combine three causal factors – (1) bias in search results, which is an essential and important feature of good search results, (2) customization in search results, which Google in particular has long taken pride in providing [24,60], and (3) a company’s ability to identify just those users who are especially vulnerable to influence – a rather daunting picture emerges. The picture becomes even more alarming when one recognizes that both search suggestions [11] and answer boxes [10] – both of which are commonly shown by Google search – also have the power to shift opinions. What if all of these factors align to push opinions in the same direction? And what if these types of influence are similarly biased in online experiences people are having day after day on multiple platforms? We are currently exploring these questions in experiments on what we call the “multiple exposure effect” (MEE) and the “multiple platforms effect” (MPE).

5.1 Limitations and Future Research

Our conclusions are subject to a number of constraints, two of which we believe are obvious and nontrivial. First, our subjects were drawn from the MTurk subject pool. In recent years, that subject pool has been tainted by bots [61,62], and concerns have been raised about just how representative the US portion of that subject pool is of the general population [63,64].

Fortunately, we conducted the present experiments in early 2016, well before most of the substantive concerns about MTurk were expressed [61,62]. Nevertheless, we acknowledge that the subjects in our experiments are not necessarily representative of the general population, a matter that can only be explored with replications using other sampling methods. On the bright side, our participants were demographically diverse (Table S1) – far more so than the small group of sophomores at a single college or university who have so often been utilized in social science studies [65-67].

Second, we made no attempt to measure the staying power of the opinion shifts we measured. The impact we had on participants in our six bias groups might be as ephemeral as search results typically are. Although our procedures can shed no light on this issue, we would be remiss in not pointing out that a search engine company such as Google could easily expose users to similarly biased content dozens or even hundreds of times over a period of a few months. If such exposures are additive in their impact, it is not unreasonable to believe that our experiments might be *underestimating* the power that biased search results might have on people's thinking about virtually any topic (as long as the users have not already formed strong opinions).

Our Kadoodle simulator also differed from Google's search engine in some respects. Google typically shows many pages of search results with about 10 results per page (on desktop and laptop computers). We showed only five pages of search results with only six results per page. We also did not show people search suggestions or answer boxes, which have become universal on Google search pages in recent years. When answer boxes are added to search results, people spend less time examining search results and click on fewer search results [10]; if the answer boxes share the bias of the search results, however, opinions and votes shift even farther in the direction of the bias than they would have had only search results been shown [10]. Again, if search suggestions share the same bias as the search results, they too will increase the impact of those results [11]. So although our simulator differs from Google's home page, it does so mainly in ways that make it less powerful as a source of influence.

Regarding future research, we have already mentioned three projects we have in progress that will shed more light on new forms of manipulation that the internet has made possible: MEE, MPE, and DPE. Regarding the range of opinions that might be influenced by biased (or by biased and personalized) search results, determining and understanding that range can be accomplished by varying topics in systematic ways. We have already learned that different demographic groups vary in how vulnerable they are to the manipulation we employed in the present experiments (S2 Table to S5 Table), and we also have found demographic effects in other studies of online influence [1,9-11,13,59]. Further research might show predictable patterns in how vulnerable different demographic groups are (and, for that matter, in how vulnerable different individuals are) to having their opinions altered on different topics by biased search results. An extensive literature on influence and decision-making has already shown how demographic characteristics interact both with types of influence and the topics being considered [1,11,68-70].

Future research should also explore an odd feature of the search engine – one that we alluded to earlier and that might be considered a fatal flaw. Search results are useful precisely because they order information from best to worst; an equal-time rule would make them worthless, although perhaps – as a way of protecting the free-and-fair election from undue influence – an exception could be made someday for links to information about political candidates. Generally speaking, however, search results will *always* train people to value high-ranking results over lower ones, which means perforce that search results shown in response to open-ended queries will *always* shift people's thinking and behavior, sometimes in trivial ways and sometimes in profound ones. This will almost always occur, moreover, without people's awareness [1]. If shifts of this sort are programmed by social engineers or pranksters at tech companies, humanity will always be in thrall to such people to some extent. Our guess, though, is that only an infinitesimally small portion of open-ended queries are of interest to Big Tech programmers or executives. That means, unfortunately, that the vast majority of shifts in opinions and behavior being produced by search engines 24 hours a day in people around the world are currently being determined by *algorithms*.

Where algorithms are being left to their own devices (so to speak) by their human creators, they are currently determining what content goes viral or gets suppressed, what many people buy, what many people believe, and whom many people vote for. As self-determining AI systems are increasingly incorporated into the algorithms that currently dominate our lives, will the growing power of these systems be used in humanity's interest? Will we even understand what is happening to us?

Declaration of competing interest

The authors have no conflicts of interest to declare.

Acknowledgments

We thank R. Robertson for assistance in conducting the experiments and C. Mourani, E. Olson, and V. R. Zankich for assistance in data analysis and manuscript preparation.

Authors' contributions

Robert Epstein: Conceptualization, Methodology, Supervision, Writing – Original Draft, Writing – Reviewing and Editing. **Ji Li:** Statistical Analysis, Writing – Reviewing and Editing

Funding

This work was supported by general funds of the American Institute for Behavioral Research and Technology (AIBRT), a nonpartisan, nonprofit, 501(c)(3) organization. The authors received no specific funding for this work. This research did not receive any specific external support from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

An anonymized version of the data will be available at Zenodo.org upon acceptance of the manuscript. Data can also be requested from info@aibrt.org. The data have been anonymized to comply with requirements of the sponsoring institution's Institutional Review Board (IRB). The IRB granted exempt status to this study under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal.

References

1. Epstein R, & Robertson RE. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proc Natl Acad Sci USA. 2015; 112(33), E4512-E4521. Available from: https://doi.org/10.1073/pnas.1419828112
2. Ludolph, R., Allam, A., & Schulz PJ. Manipulating Google's knowledge box to counter biased information processing during an online search on vaccination: Application of a technological debiasing strategy. *J Med Internet Res*, 2016; *18*(6):e137. pmid:27255736
3. Haas, A. & Unkel, J. Ranking versus reputation: Perception and effects of search result credibility. *Behavior & Information Technology*. *36*(12), 1285-1298. 2017; https://doi.org/10.1080/0144929X.2017.1381166
4. Pogacar FA, Ghenai A, Smucker MD, Clarke CLA. The positive and negative influence of search results on people's decisions about the efficacy of medical treatments. Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval [Internet]. 2017 Oct 1–4;:209–216. Available from: https://dl.acm.org/doi/10.1145/3121050.3121074
5. Agudo U, Matute H. The influence of algorithms on political and dating decisions. PLOS One. 2021;16(4):e0249454. Available from: http://dx.doi.org/10.1371/journal.pone.0249454
6. Epstein R & Robertson R. Why is the Search Engine Manipulation Effect (SEME) so large? Testing an operant conditioning hypothesis. Paper presented at the 96th annual meeting of the Western Psychological Association; 2016; Long Beach, CA.

7. Epstein R & Robertson RE. Suppressing the Search Engine Manipulation Effect (SEME). Proceedings of the ACM: Human–Computer Interaction. 2017; 1(2). Available from: https://doi.org/10.1145/3134677
8. Epstein, R., & Zankich, V. R. (2022). The surprising power of a click requirement: How click requirements and warnings affect internet users' willingness to disclose personal information. *PLOS ONE*. 17(2): e0263097.
9. Epstein, R., & Flores, A. (2023, August). The YouTube Manipulation Effect (YME): A quantification of the impact that the ordering of YouTube videos can have on opinions and voting preferences. https://dx.doi.org/10.2139/ssrn.4527207
10. Epstein, R., Lee, V., Mohr, R., & Zankich, V. R. (2022). The answer bot effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. *PLOS ONE*, *17*(6). https://doi.org/10.1371/journal.pone.0268081
11. Epstein, R., Aries, S., Grebbien, K., Salcedo, A. M., & Zankich, V. R. (2023a, August). The Search Suggestion Effect (SSE): How autocomplete search suggestions can be used to impact opinions and votes. *SSRN*. http://dx.doi.org/10.2139/ssrn.4535163
12. Epstein R., Tyagi, C., Wang, H. (2023b, July). What would happen if twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME). *PLOS ONE*, *18*(7): e0284495. https://doi.org/10.1371/journal.pone.0284495
13. Epstein, R., Huang, Y., & Megerdoomian, M. (2023c, August 4). The Opinion Matching Effect (OME): A subtle but powerful new form of influence that is apparently being used on the internet. *SSRN*. https://dx.doi.org/10.2139/ssrn.4532141

14. Eslami, M., Vaccaro, K., Karahalios, K., & Hamilton, K. (2017). "Be careful; things can be worse than they appear": Understanding biased algorithms and users' behavior around them in rating platforms. Proceedings of the 11th International AAAI Conference on Web and Social Media, 11(1), 62–71.
15. Trielli, D., & Diakopoulos, N. (2019). Search as news curator: The role of Google in shaping attention to news information. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–15. https://doi.org/10.1145/3290605.3300683
16. Draws T, Tintarev N, Gadiraju U, Bozzon A, Timmermans B. This is not what we ordered: Exploring why biased search result rankings affect user attitudes on debated topics. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY, USA: ACM; 2021.
17. Smith, A. (2018, November 16). *Public attitudes toward computer algorithms*. Pew Research Center. https://www.pewresearch.org/internet/2018/11/16/public-attitudes-toward-computer-algorithms/
18. Bogert, E., Schecter, A. & Watson, R.T. Humans rely more on algorithms than social influence as a task becomes more difficult. *Sci Rep 11*, 8028 (2021). https://doi.org/10.1038/s41598-021-87480-9
19. Chen K., & Cheung H. (2019, March). Unlocking the power of ephemeral content: The roles of motivations, gratification, need for closure, and engagement. *Computers in Human Behavior*, *97*, 67-74. https://doi.org/10.1016/j.chb.2019.03.007
20. McKinnon JD, McMillan D. (2018, September 20). *Google Workers Discussed Tweaking Search Function to Counter Travel Ban: Company says none of proposed changes to search results were ever implemented.* The Wall Street Journal.

https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472

21. Cadwalladr, C. (2016, December 11). *Google is not "just" a platform. it frames, shapes and distorts how we see the world*. The Guardian. https://www.theguardian.com/commentisfree/2016/dec/11/google-frames-shapes-and-distorts-how-we-see-world
22. Grind, K., Schechner, S., McMillan, R., & West, J. (2019, August 6). *How google interferes with its search algorithms and changes your results*. The Wall Street Journal. https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753
23. Walker, K. (2018, May 4). *Supporting election integrity through Greater Advertising Transparency*. Google. https://blog.google/outreach-initiatives/public-policy/supporting-election-integrity-through-greater-advertising-transparency/
24. Google Official Blog. (2009, December 4). Retrieved August 15, 2023, from https://googleblog.blogspot.com/2009/12/personalized-search-for-everyone.html.
25. Arora, N., Ensslen, D., Fiedler, L., Liu, W. W., Robinson, K., Stein, E., & Schüler, G. (2021, November 12). *The value of getting personalization right-or wrong-is multiplying*. McKinsey & Company. https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying#/
26. Shanahan, T., Tran, T. P., & Taylor, E. C. (2019). Getting to know you: Social media personalization as a means of enhancing brand loyalty and perceived quality. *Journal of Retailing and Consumer Services*, *47*, 57–65. doi:10.1016/j.jretconser.2018.10.007

27. Erevelles, S., Fukawa, N., & Swayne, L. (2015). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897–904. https://doi.org/10.1016/j.jbusres.2015.07.001

28. Eva A. van Reijmersdal, Sophie C. Boerman, Guda van Noort. (2022). Effects of online behaviorally targeted native advertising on persuasion: A test of two competing mechanisms, *Computers in Human Behavior Reports*, *7*, 100221, https://doi.org/10.1016/j.chbr.2022.100221.

29. Rahat, G., & Kenig, O. (2018). *From party politics to personalized politics?: Party change and political personalization in democracies*. Oxford University Press. https://doi.org/10.1093/oso/9780198808008.001.0001

30. Robertson, Ronald E. & Lazer, David & Wilson, Christo. (2018). Auditing the Personalization and Composition of Politically-Related Search Engine Results Pages. https://doi.org/10.1145/3178876.3186143

31. Bleu, N. (2022, September 30). *19+ crucial search engine statistics for 2023 (market share, usage, & demographics)*. Startup Bonsai. https://startupbonsai.com/search-engine-statistics/

32. StatCounter. (2023, July). *Search engine market share worldwide*. StatCounter Global Stats. https://gs.statcounter.com/search-engine-market-share

33. Nadler, J., Cicilline, D. (2020). *Investigation of Competition in Digital Markets*. Subconnittee on Antitrust. https://fm.cnbc.com/applications/cnbc.com/resources/editorialfiles/2020/10/06/investigation_of_competition_in_digital_markets_majority_staff_report_and_recommendations.pdf

34. Meyers, L. (2023, June 8). *Google manipulates search results, former engineer shockingly confirms*. CCN.com. https://www.ccn.com/google-manipulates-search-results-former-engineer-shockingly-confirms/

35. Epstein, R. (2018, May 13). *Taming big tech: The case for monitoring*. HackerNoon. https://hackernoon.co m/taming-big-tech-5fef0df0f00d

36. Epstein, R., Bock, S., Peirson, L., Wang, H., & Voillot, M. (2022, April). How we preserved more than 1.5 million online "ephemeral experiences" in the recent U.S. elections, and what this content revealed about online election bias. Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. https://aibrt.org/downloads/EPSTEIN_et_al_2022-WPA-How_We_Preserved_More_Than_1.5_Million_Online_Ephemeral_Experiences_in_Recent_US_Elections...pdf

37. Epstein, R., & Peirson, L. (2023, April). *How we preserved more than 2.4 million online ephemeral experiences in the 2022 midterm elections, and what this content revealed about online election bias*. Paper presented at the 103rd annual meeting of the Western Psychological Association, Riverside, CA. https://aibrt.org/downloads/EPSTEIN_&_Peirson_2023-WPAHow_We_Preserved_More_Than_2.4_Million_Online_Ephemeral_Experiences_in_the_Midterm_Elections.pdf

38. Prinz, R., Brighton, H., Luan, S., & Gigerenzer, G. (2017, March 24). Can biased search engine results influence healthcare decisions? Paper presented at the meeting of the International Convention of Psychological Science, Vienna, Austria.

https://aibrt.org/downloads/ICPS-Symposium-Subtle_New_Forms_of_Internet_Influence-24March2017-Vienna.pdf

39. Allam A, Schulz PJ, Nakamoto K. (2014). The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating Google output. *J Med Internet Res*, *16*(4):e100. pmid:24694866
40. Loftus E. Leading questions and the eyewitness report. *Cogn Psychol*, 1975;*7*(4):560-572. https://doi.org/10.1016/0010-0285(75)90023-7
41. Nyhan B, Reifler J. When Corrections Fail: The Persistence of Political Misperceptions. Political Behavior. 2010 Jun;32(2):303–30.
42. Nield D. All the ways Google tracks you—and how to stop it. Wired [Internet]. 2019 May 27 [cited 2023 Oct 9]; Available from: https://www.wired.com/story/google-tracks-you-privacy/
43. Huang K. (2022, Feb). *Email Client Market Share in February 2022*. Litmus. https://www.litmus.com/blog/email-client-market-share-february-2022/
44. StatCounter. (2023, Sept). *Browser Version (Partially Combined) Market Share Worldwide*. StatCounter Global Stats. https://gs.statcounter.com/browser-version-partially-combined-market-share#monthly-202209-202309
45. StatCounter. (2023, Sept). *Mobile Operating System Market Share Worldwide*. StatCounter Global Stats. https://gs.statcounter.com/os-market-share/mobile/worldwide/#monthly-202209-202309
46. Editor of IndexRise. List of Google products and services – IndexRise [Internet]. Indexrise. [cited 2023 Oct 9]. Available from: https://www.indexrise.com/google-services/

47. Nakashima R. AP Exclusive: Google tracks your movements, like it or not [Internet]. AP News. 2018 [cited 2023 Oct 9]. Available from: https://apnews.com/article/828aefab64d4411bac257a07c1af0ecb

48. Epstein, R., Lothringer, M., & Zankich, V. R. (2023, August 24). How a daily regimen of operant conditioning might explain the surprising power of the Search Engine Manipulation Effect (SEME). *SSRN*, https://dx.doi.org/10.2139/ssrn.4551486

49. Ong SQ, Hardwick J, Soulo T. Top Google searches [Internet]. Ahrefs Blog. Ahrefs; 2023 [cited 2023 Oct 9]. Available from: https://ahrefs.com/blog/top-google-searches/

50. Langville AN, Meyer CD. Google's PageRank and beyond: The science of search engine rankings. Princeton, NJ: Princeton University Press; 2006.

51. Lee, N., Resnick, P., & Barton, G. (2019, May 22). *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. The Brookings Institute. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

52. Nunez, M. (2016, May 9). *Former Facebook workers: We routinely suppressed conservative news.* Gizmodo. https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006

53. Rainie, L., & Anderson, J. (2017, February 8). *Code-dependent: Pros and cons of the algorithm age.* Pew Research Center. https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/

54. Arthur Bloom J. EXCLUSIVE: Documents detailing Google's 'news blacklist' show manual manipulation of special search results [Internet]. The Daily Caller. 2019 [cited

2023 Oct 9]. Available from: https://dailycaller.com/2019/04/09/google-news-blacklist-search-manipulation/

55. Ebbinghaus, H. (1885/1913). Memory: A contribution to experimental psychology (H. A. Ruger & C.E. Bussenius, Trans.). New York: Teachers College, Columbia University. (Originally in German, Ueber das gedächtnis: Untersuchen zur experimentellen psychologie).
56. Murre, J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLOS ONE, 10*(7). https://doi.org/10.1371/journal.pone.0120644
57. Mack, C. C., Cinel, C., Davies, N., Harding, M., & Ward, G. (2017). Serial position, output order, and list length effects for words presented on smartphones over very long intervals. *Journal of Memory and Language*, *9*, 61–80. https://doi.org/10.1016/j.jml.2017.07.009
58. Dean B. We analyzed 4 million Google search results: Here's what we learned about organic click through rate. 2022 Aug 17. In: Backlinko [Internet]. Available from: https://backlinko.com/google-ctr-stats
59. Epstein, R., Newland, A., & Tang, Liyu. (2024). The Digital Personalization Effect (DPE): How personalization of online content can dramatically increase the impact of new forms of online influence. Paper to be presented at the 104th annual meeting of the Western Psychological Association, San Francisco, CA.
60. Krishnan P. Extend the reach of your site personalization in Google Optimize [Internet]. Google. 2019 [cited 2023 Oct 9]. Available from: https://blog.google/products/marketingplatform/analytics/extend-reach-your-site-personalization-google-optimize/

61. Bai, H. (2018) Evidence that A Large Amount of Low Quality Responses on MTurk Can Be Detected with Repeated GPS Coordinates. Retrieved from: https://www.maxhuibai.com/blog/evidence-that-responses-from-repeating-gps-are-random

62. Dreyfuss, E. (2018, August 17). A bot panic hits Amazon's Mechanical Turk. Wired. https://www.wired.com/story/amazon-mechanical-turk-bot-panic/

63. Litman, L. (2017, November 21). *Strengths and limitations of Mechanical Turk.* CloudResearch. https://www.cloudresearch.com/resources/blog/strengths-and-limitations-of-mechanical-turk/

64. Moss, A., & Litman, L. (2020, June 12). *Demographics of people on Amazon Mechanical Turk.* CloudResearch. https://www.cloudresearch.com/resources/blog/who-uses-amazon-mturk-2020-demographics/

65. Huff C, Tingley D. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. Res Politics [Internet]. 2015;2(3):205316801560464. Available from: http://dx.doi.org/10.1177/2053168015604648

66. Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156-2160. https://doi.org/10.1016/j.chb.2013.05.009

67. Ramsey, S. R., Thompson, K. L., McKenzie, M., Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, *58*, 354-360. https://doi.org/10.1016/j.chb.2015.12.049

68. Hu J, Zeng H-J, Li H, Niu C, Chen Z. Demographic prediction based on user's browsing behavior. In: Proceedings of the 16th international conference on World Wide Web. New York, NY, USA: ACM; 2007.

69. Yom-Tov E. Demographic differences in search engine use with implications for cohort selection. Inf Retr Boston [Internet]. 2019;22(6):570–80. Available from: http://dx.doi.org/10.1007/s10791-018-09349-2

70. Chen, L., Ma, R., Hannák, A., & Wilson, C. (2018) Investigating the impact of gender on rank in resume search engines. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 651, 1-14. https://doi.org/10.1145/3173574.3174225

Supporting Information

S1 Text: Manipulation Power (MP) calculation

$$100 * \frac{p' - p}{p}$$

where p is the number of people who chose the opinion favored in the manipulation prior to that manipulation, and p' the number of people who chose the opinion favored in the manipulation after that manipulation. Thus, MP can be defined as the post-manipulation percentage increase in the number of people choosing the opinion favored in the manipulation.

S2 Text: Artificial Intelligence Summary

Anti A.I. Artificial intelligence, also known as A.I., refers to intelligent machines, computers, or software. A.I. can automate processes, and research is increasing its human-like capacities. As technology continues to improve, A.I. will grow more *dangerous* to humans. This view is generally considered to be **Anti A.I.**

Pro A.I. Artificial intelligence, also known as A.I., refers to intelligent machines, computers, or software. A.I. can automate processes, and research is increasing its human-like capacities. As technology continues to improve, A.I. will grow more *useful* to humans. This view is generally considered to be **Pro A.I.**

S3 Text: Fracking Summary

Pro-Fracking. Fracking, also known as hydraulic fracturing, is a technique used to extract oil from rocks. It has potential economic benefits and is legal in most countries.
Anti-Fracking. Fracking, also known as hydraulic fracturing, is a technique used to extract oil from rocks. It has potential environmental risks and is regulated in some countries.

S4 Text: Sexual Orientation Summary

People choose to be gay. Gay people, also referred to as lesbians or homosexuals, are people who only have romantic and sexual relationships with members of their own gender. Being gay is a *choice* people make, *not* a characteristic they are born with.

People are born gay. Gay people, also referred to as lesbians or homosexuals, are people who only have romantic and sexual relationships with members of their own gender. Being gay is *not a choice* people make, but a characteristic they are born with.

S5 Text: Request for Informed Consent

By clicking continue I understand that I must be 18 or over to participate in this study, that my participation is voluntary, that I am free to withdraw at any time, that I am providing information anonymously and that demographic information collected is confidential and cannot be used to identify me. I agree to allow the data collected to be used for future research projects, and I understand that completion and submission of this survey implies my consent to participate in the present study.

S1 Table: Experiment 1, 2, & 3 Demographics

	Experiment 1 (n = 378)	**Experiment 2 (*n* = 394)**	**Experiment 3 (*n* = 365)**
Mean Age (*SD*)	33.8 (11.4)	32.9 (10.2)	32.8 (10.6)
Gender (%)			
Female	213 (56.3%)	205 (52.0%)	204 (55.9%)
Male	165 (43.7%)	189 (48.0%)	161 (44.1%)
Education (%)			
None	0 (0%)	1 (0.3%)	0 (0%)
High School	31 (8.2%)	30 (7.6%)	36 (9.9%)
Some College	148 (39.2%)	160 (40.6%)	137 (37.5%)
Bachelors	148 (39.2%)	145 (36.8%)	147 (40.3%)
Masters	39 (10.3%)	53 (13.5%)	32 (8.8%)
Doctorate	12 (3.2%)	5 (1.3%)	13 (3.6%)
Race/Ethnicity (%)			
White	298 (78.8%)	317 (80.5%)	286 (78.4%)
Black	32 (8.5%)	19 (4.8%)	22 (6.0%)
Asian	22 (5.8%)	21 (5.3%)	18 (4.9%)
Mixed	13 (3.4%)	10 (2.5%)	12 (3.3%)
Hispanic	12 (3.2%)	23 (5.8%)	23 (6.3%)

Other	1 (0.3%)	4 (1.0%)	4 (1.1%)
Religion (%)			
None	151 (39.9%)	141 (35.8%)	138 (37.8%)
Christianity	124 (32.8)	135 (34.3%)	122 (33.4%)
Catholicism	46 (12.2%)	60 (15.2%)	52 (14.2%)
Judaism	8 (2.1%)	8 (2.0%)	10 (2.7%)
Islam	7 (1.9%)	4 (1.0%)	3 (0.8%)
Hinduism	6 (1.6%)	2 (0.5%)	1 (0.3%)
Prefer Not to Say	15 (4.0%)	18 (4.6%)	8 (2.2%)
Other	21 (5.6%)	26 (6.6%)	31 (8.5%)
Income (%)			
Under $10,000	21 (5.6%)	21 (5.3%)	12 (3.3%)
$10,000 to 14,999	24 (6.3%)	19 (4.8%)	18 (4.9%)
$15,000 to 29,999	61 (16.1%)	58 (14.7%)	54 (14.8%)
$30,000 to 39,999	55 (14.6%)	53 (13.5%)	54 (14.8%)
$40,000 to 49,999	37 (9.8%)	52 (13.2%)	40 (11.0%)
$50,000 to 74,999	75 (19.8%)	82 (20.8%)	85 (23.3%)
$75,000 to 99,999	43 (11.4%)	51 (12.9%)	47 (12.9%)
$100,000 to 149,999	35 (9.3%)	36 (9.1%)	34 (9.3%)
$150,000 and over	14 (3.7%)	13 (3.3%)	10 (2.7%)
Prefer Not to Say	13 (3.4%)	9 (2.3%)	11 (3.0%)

S2 Table: Demographics Analysis by Gender

Experiment		*n*	MP (%)	McNemar's Test	*p*
Artificial Intelligence	**Male**	165	28.0	13.5	< 0.001
	Female	213	23.0	9.76	0.002
	Change (%)	-	5.0	-	-
Fracking	**Male**	189	23.9	6.25	0.012
	Female	205	38.5	19.12	< 0.001
	Change (%)	-	14.6	-	-
Born Gay	**Male**	204	14.7	3.77	0.049
	Female	161	21.7	7.04	0.007
	Change (%)	-	7.00	-	-

S3 Table: Demographics Analysis by Age

Experiment		*n*	MP (%)	McNemar's Test	*p*
Artificial Intelligence	**≥ 30**	205	41.9	15.85	< 0.001
	< 30	173	8.1	5.26	0.019
	Change (%)	-	33.8	-	-
Fracking	**≥ 30**	223	30.3	16.48	< 0.001
	< 30	171	31.7	8.31	0.004
	Change (%)	-	1.4	-	-
Born Gay	**≥ 30**	195	18.7	7.35	0.011
	< 30	170	16.7	4.50	0.031
	Change (%)	-	2.0	-	-

S4 Table: Demographics Analysis by Education Level

Experiment		*n*	MP (%)	McNemar's Test	*p*
Artificial Intelligence	**≥ Bachelor's**	199	17.9	6.62	0.010
	< Bachelor's	179	33.3	15.61	< 0.001
	Change (%)	-	15.4	-	-
Fracking	**≥ Bachelor's**	203	30.2	15.61	< 0.001
	< Bachelor's	191	31.5	9.59	0.002
	Change (%)	-	1.3	-	-
Born Gay	**≥ Bachelor's**	192	21.4	11.64	0.001
	< Bachelors	173	13.9	1.89	0.167 (NS)
	Change (%)	-	7.5	-	-

S5 Table: Demographics Analysis by Ethnicity

Experiment		*n*	MP (%)	McNemar's Test	*p*
Artificial Intelligence	**White**	298	22.7	13.80	< 0.001
	Non-White	80	33.3	7.56	0.004
	Change (%)	-	10.6	-	-
Fracking	**White**	317	31.3	19.34	< 0.001
	Non-White	77	29.2	4.9	0.021
	Change (%)	-	2.1	-	-
Born Gay	**White**	286	17.9	13.79	< 0.001
	Non - White	79	17.2	0.08	0.774 (NS)
	Change (%)	-	0.7	-	-

S1 Fig: Pre-Search Impression Questions

Is your *overall impression* of views that are **Anti A.I.** positive or negative?

Negative ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Positive*

Is your *overall impression* of views that are **Pro A.I.** positive or negative?

Negative ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Positive*

How *persuasive* do you find views that are **Anti A.I.** ?

Not persuasive ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Very persuasive*

How *persuasive* do you find views that are **Pro A.I.** ?

Not persuasive ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *Very persuasive*

How much do you *trust* views that are **Anti A.I.** ?

Not al all ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *A great deal*

How much do you *trust* views that are **Pro A.I.** ?

Not al all ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○8 ○9 ○10 *A great deal*

PLEASE NOTE: Your answers to these next 2 questions should be consistent with each other! Please read the questions carefully!

On the scale below, indicate which view you favor.
0 means you have no preference. Numbers to the left of 0 mean you favor views that are **Anti A.I.**. And numbers to the right of 0 mean favor views that are **Pro A.I.**.

Anti A.I. ○5 ○4 ○3 ○2 ○1 ○0 ○1 ○2 ○3 ○4 ○5 **Pro A.I.**

Again, which view do you favor?

○**Anti A.I.** ○**Pro A.I.**

Click below to continue.

Continue

APPENDIX VIII

The YouTube Manipulation Effect (YME):

A Quantification of the Impact that the Ordering of YouTube Videos Can Have on Opinions and Voting Preferences

Robert Epstein[1*] (ORCID 0000-0002-7484-6282) and Alex Flores[1] (ORCID 0009-0003-6756-4644)

[1]American Institute for Behavioral Research and Technology, Vista, California, United States of America

*Corresponding author.

Email: re@aibrt.org (RE)

Submitted for publication. Do not quote or cite without permission. Comments welcome.

Running head: The YouTube Manipulation Effect (YME)

Keywords: YouTube Manipulation Effect, YME, online influence, Search Engine Manipulation Effect, SEME, ordering of YouTube videos, autoplay

Abstract

Recent research has identified a number of powerful new forms of influence that the internet and related technologies have made possible. Randomized, controlled experiments have shown, for example, that when results generated by search engines are presented to undecided voters, if those search results favor one political candidate over another, the opinions and voting preferences of those voters can shift dramatically – by up to 80% in some demographic groups. The present study employed a YouTube simulator to identify and quantify another powerful form of influence that the internet has made possible – the YouTube Manipulation Effect (YME). In two randomized, controlled, counterbalanced, double-blind experiments with a total of 1,463 politically-diverse, eligible US voters, we show that when a sequence of videos displayed by the simulator is biased to favor one political candidate, and especially when the "up-next" video suggested by the simulator favors that candidate, both the opinions and voting preferences of undecided voters shift dramatically toward that candidate. Voting preferences shifted by between 51.5% and 65.6% overall, and by more than 75% in some demographic groups. We also tested a method for masking the bias in video sequences so that awareness of bias was greatly reduced. In 2018, a YouTube official revealed that 70% of the time people spend watching videos on the site, they are watching content that has been suggested by the company's recommender algorithms. This gives Google (YouTube's parent company) unprecedented power to impact thinking and behavior, and YouTube video sequences have also been implicated in many cases of political radicalization. The fact that no laws or regulations exist to limit Google's ability to use YouTube as a manipulative tool should be a matter of great concern to leaders and public policy makers worldwide.

The YouTube Manipulation Effect (YME):

A Quantification of the Impact that the Ordering of YouTube Videos Can Have on Opinions and Voting Preferences

1. Introduction

The internet has made it possible for tech companies to dominate the thinking and behavior of more than 5 billion people worldwide using new subliminal techniques. Our research team has discovered and quantified several of these techniques in randomized, controlled experiments conducted since 2013. Our research demonstrates that some new types of manipulation that the internet has made possible can easily shift votes and opinions without people's knowledge and without leaving paper trails for authorities to trace [1-7]. A growing body of evidence suggests that Big Tech companies might sometimes use these new techniques of influence strategically and deliberately [8-12,cf.13]. Tristan Harris, a former "design ethicist" at Google, said that he was a member of a team at the company whose job it was to influence "a billion people's attention and thoughts every day" [14]. Jaron Lanier, one of the early investors in Google and Facebook, claims that Big Tech content has "morphed into continuous behavior modification on a mass basis" [14]. Another early investor in these companies, Roger McNamee, has said that he now regrets having supported them, asserting that they now constitute "a menace to public health and to democracy" [15,cf.14].

Three recent leaks of internal content from Google are also cause for concern. First, in emails leaked from the company to *The Wall Street Journal* in 2018, employees are discussing how they might be able to change people's views about Trump's travel ban by using what they call "ephemeral experiences" [8] – that is, content such as search results and newsfeeds which appears briefly, impacts the user, and then disappears forever. Second, "The

Selfish Ledger," a nearly 9-min video that leaked from the company, describes the power Google has to reengineer humanity – specifically, to "refine human behavior" – according to "company values" [16,cf.17]. Third, "The Good Censor," a company PowerPoint presentation, explains that tech companies have been forced over the years to move away "from passive facilitation to active curation" of content, deciding what content users worldwide can and cannot see [18].

In recent years, a number of authorities and experts have expressed particular concern about the way Google's YouTube platform might be influencing users, especially young children [19-26,cf.27-36]. A 2019 *New York Times* investigation concluded that "YouTube's algorithms may have played a decisive role" in the rise of right-wing Brazilian president Jair Bolsonaro by "boost[ing] fringe videos into the mainstream" and helping to spread conspiracy theories and misinformation, especially about diseases [37]. In some cases, when users in Brazil were watching sports videos, YouTube's up-next suggestion (normally, the video image shown in the upper-right of the screen) would be for a Bolsonaro video, with one Bolsonaro video leading to others [29]. In some instances, just a few clicks have been known to take users down "rabbit holes" of similar videos making extreme claims that have sometimes radicalized them [20,28,29,35-40,cf.41-43]. This phenomenon has prompted sociologist Zeynep Tufekci to label YouTube as "one of the most powerful radicalizing instruments of the 21st century" [29,34].

One of the most interesting cases of radicalization, reported in 2019, involved a 26-year-old White male named Caleb Cain of West Virginia – "a college dropout looking for direction" [28]. He turned to YouTube for guidance and was soon "pulled into a far-right filled universe, watching thousands of videos filled with conspiracy theories, misogyny and racism." He watched more than 12,000 such videos, falling, he later said, into "the alt-right rabbit hole." Cain's conversion apparently did not cause him to act violently, but other converts have been more aggressive. In 2020, Brenton Harrison Tarrant, a 28-year-old White male from Australia, was convicted of 51 counts of murder, 40 counts of attempted murder, and one count of terrorism – the first terrorist convicted in New Zealand's history. His rampage took place in March, 2019, and his victims were worshippers at two mosques in Christchurch. Tarrant had been radicalized by YouTube videos. He even had made use of infinitelooper.com, which would repeat certain inspirational YouTube videos for him endlessly [44].

In 2022, the Anti-Defamation League published an ambitious study – well executed but not peer-reviewed, as far as we can tell – on user exposure to "alternative" and extremist content on YouTube [45]. Based on data obtained from a representative sample of 859 people in the US who were enrolled with YouGov, a national polling firm, the study concluded that roughly 1 in 5 YouTube users are exposed to "alternative" content – "channels that can serve as gateways to more extreme forms of content" – and that 1 in 10 users are exposed to extremist content directly. Although the researchers did not find evidence that extreme or disturbing YouTube content converted people with moderate views, they did find (a) that such content strongly attracted people "who already have high levels of racial resentment," (b) that when people watch such videos, they are "more likely to see and follow recommendations to similar videos," and (c) that when someone is viewing an extremist video, other extremist videos are likely to be recommended alongside them [45].

Some studies have found stronger evidence of radicalization on YouTube [46,47,cf.43,48-53]. A large-scale 2020 study published by the Association for Computing Machinery that examined more than 330,000 videos concluded, for example, "We find strong evidence for radicalization among YouTube users, and that YouTube's recommender system enables Alt-right channels to be discovered, even in a scenario without personalization…. Moreover, regardless of the degree of influence of the recommender system in the process of radicalizing users, there is significant evidence that users are reaching content sponsoring fringe ideologies from the Alt-lite [people who ‘flirt with’ white supremacist ideology] and the Intellectual Dark Web” [46].

Other recent studies have catalogued and counted the growing number of extremist videos available on the YouTube platform [42,46-50,54-57]. Even though YouTube regularly removes many such videos from its platform, the number of alternative and extremist videos available to users worldwide on any given day is probably in the millions. This conjecture is based on recent surveys suggesting that upwards of 20% of internet users in the US have encountered hateful or harassing content on YouTube, which, at this writing (July 4, 2023), hosts more than 800 million videos [45,58,59]. The disturbing content could be in the videos themselves or in the comments provoked by those videos.

Both leaks and official statements from Big Tech platforms suggest that controversial content is an important part of content offerings because (a) it draws more traffic, and more traffic is generally more profitable [14,34,60,61,cf.35,49], (b) it keeps people on a website longer [14,35,63,cf.29,37], and (c) it increases the “watch time” of videos [14,35,cf.29,37]. Content personalization on platforms like YouTube has proved to be especially important in increasing the “stickiness” of websites [14,28,39,62-64]. Even Mark Zuckerberg, the CEO of Facebook/Meta, acknowledged the value of controversial content from a business perspective in an official statement he released in 2018. According to Zuckerberg, “Our research suggests that no matter where we draw the lines for what is allowed, as a piece of content gets close to that line, people will engage with it more on average – even when they tell us afterwards they don’t like the content” [65].

Multiple studies have also shown that YouTube’s recommender algorithms are especially aggressive in recommending “pseudoscientific” videos and other content of dubious value [21,22,62,66,cf.67]; again, the more dubious the content, the more traffic is generated and the more watch time is increased. Anti-vaccine videos on YouTube have been shown to lead to recommendations of a disproportionately large number of additional anti-vaccine videos compared to pro-vaccine videos [68].

Even if radicalization on YouTube were rare, YouTube's video-management algorithm does allow it to occur. A vulnerable individual can be drawn into a highly persuasive sequence of videos when three important mechanisms are in alignment: filtering, ordering, and customization. Filtering is the process by which the algorithm selects some videos for presentation (a small sample) and rejects others (the vast majority). Ordering is the process by which the algorithm places one video ahead of another. And customization is the process by which the algorithm refines the filtering and ordering based on (a) information from the personal profile that Google has accumulated about the user and (b) priorities that the company or its employees might have about how they want to influence users. When these factors align, users can be caught in so-called "loops," "echo chambers," and "filter bubbles" of similarly biased content [48,50,69-78,cf.79-83]. Relevant here is the fact that YouTube's algorithms also determine whether content goes viral on the platform [84,85].

Note that all three of these factors also operate on Google's search engine. Although we tend not to think about YouTube this way, YouTube is actually the second largest search engine in the world, as well as the world's largest video-sharing social media platform [86]. The first video a user watches during a YouTube session is usually suggested after the user types a search term into YouTube's search bar. After that first search is completed, however, YouTube and the Google search engine part ways in how they influence the user. On Google, the user at some point clicks away to another website – ideally, from a business perspective, to a website Google wants the user to visit [4]. As Larry Page, co-founder of Google said famously long ago, "We want to get you out of Google and to the right place as fast as possible" [87]. Whether he meant the right place for the user or the right place for the company is unclear. On YouTube, the goal is the opposite: It is to keep the user on the platform as long as possible [14,28,29,34,35]. That behavioral addiction is the goal has been acknowledged by Google whistleblowers [88,89,14] and suggested by researchers [90,91].

YouTube generally accomplishes these ends in ways that are too subtle for most users to discern. Users are likely aware that if they fail to search for another video or to click on one of the recommended videos (shown to the right of or just below the video screen), YouTube will automatically play another video – the up-next video shown on desktop and laptop computers in the upper-right corner of the computer screen. On smart phones, the up-next video will often play automatically even if the user has never seen the thumbnail version of it; this can occur, for example, when the phone is tilted to the landscape (horizontal) position, which causes the video that is playing to take up the entire screen. Until January 2015, a labeled "autoplay" on-off switch appeared above the up-next video on desktop and laptop computers which allowed the user to stop up-next videos from playing automatically (S1 Fig). At this writing, however, the switch appears immediately below the video, and it no longer has a label on it; one must scroll over it (on desktop and laptop computers) or touch it (on mobile devices) even to find out what the button is for (S2 Fig). These cosmetic changes were likely implemented to increase watch time [92-94,cf.95-97].

According to official YouTube statements, watch time is YouTube's most important concern [61,98,cf.14]. In 2018, a YouTube official revealed that 70% of the time people spend watching videos on the site, they are watching content that has been suggested by the company's up-next algorithm [99]. The importance of watch time has also been emphasized in public statements by former Google software engineer Guillaume Chaslot, who summarized this issue thus: "Watch time was the priority. Everything else was considered a distraction" [96]. When Chaslot suggested to his supervisors that the YouTube algorithm be modified to free users from content feedback loops, they rejected his ideas. "[T]he entire business model is based on watch time," according to Chaslot, and "divisive content" is especially effective in locking in user attention [14,cf.39]. Tristan Harris expressed this concept metaphorically: "There's a spectrum on YouTube between the calm section – the Walter Cronkite, Carl Sagan part – and Crazytown, where the extreme stuff is. If I'm YouTube and I want you to watch more, I'm always going to steer you toward Crazytown" [28]. Unfortunately, such content can include "bizarre and disturbing" content directed at young children [100].

YouTube also exercises its power to influence and control by (a) demonetizing content it finds objectionable and thus discouraging certain content creators from posting videos [102-104,cf.35,101,105], (b) restricting access to videos, in one case limiting access to more than 50 videos from the conservative Prager University organization, among them a video by noted Harvard Law School professor Alan Dershowitz about the founding of Israel and a video about the "thou shall not kill" provision of The Ten Commandments [106,107,cf.101], (c) deleting videos from its platform [108-112], and (d) reordering videos – in other words, boosting the positions of videos it is trying to promote and demoting videos it is trying to suppress [37,57,63,64,66]. In a 2 min 2017 video leaked from Google in 2019 by a former Google staffer, Susan Wojcicki, the then CEO of YouTube, explains to her staff the process by which YouTube's recommender algorithm was currently being altered to boost content the company viewed as valid and demote content the company considered suspect (S3 Fig) [113]. That re-ranking process has continued to this day; at this writing (June 28, 2023), US Congressman and Presidential candidate Robert F. Kennedy Jr. is in the news protesting the removal of several of his videos from YouTube [114,115].

When users have challenged such actions, US courts have repeatedly ruled in favor of Google and YouTube, asserting that by deleting or reordering content, these platforms, as private companies, are exercising their right to free speech under the First Amendment to the US Constitution [106,116-119].

A growing body of research demonstrates the power of YouTube's recommender algorithms either to cause people to formulate opinions where their opinions are initially weak, or to further strengthen opinions where opinions are initially strong [66,71-74,cf.78,120-122]. Some of this research extends these general findings to the political realm [123]. For example, a 2020 study by Cho et al. (conducted with 108 undergraduate students at one university) demonstrated the power that YouTube's recommender algorithms have to "reinforce and polarize" existing political opinions [76,cf.77]. Newer studies suggest that YouTube's up-next algorithm might be biased to some extent in one direction politically [43,81], although "communities" of YouTube users can have almost any political bias [50,120].

Whether videos are generally more influential than auditory, textual, or still-image media is a matter that has not been well explored, to our knowledge – in part, we believe, because of the difficulties inherent in designing studies that compare the persuasiveness of these media fairly. One recent study suggests, however, that video is substantially more powerful than text in convincing people that political content is real, but that it is only slightly more persuasive than text [124,125,cf.126-131]. Whatever the truth is about direct comparisons, researchers have consistently found that videos get far more "shares" online than other forms of media do – according to one recent estimate, "1200% more shares than text and images combined" [132,cf.133], and online content apparently has far more impact, in general, than offline content [134,cf.123].

The Importance of Capturing Ephemeral Content

These new methods of influence are especially problematic because they are controlled worldwide (outside the People's Republic of China) by a small number of corporate monopolies, which means one cannot counteract them. If a political candidate airs an attack ad on television or on the internet, the opponent can air a rejoinder. But if a large online platform uses new techniques of influence to support a candidate, the opponent can do nothing to counteract that influence; in many cases, that manipulation might not even be visible. As we noted earlier, many online manipulations also make strategic use of ephemeral experiences to change thinking or behavior; that normally guarantees that these manipulations leave no paper trails for authorities to trace. Note that although YouTube videos are *not* ephemeral, the video sequences and up-next suggestions the company makes are indeed ephemeral. They are generated on the fly for the individual user and stored nowhere, and there is no way, to our knowledge, for anyone – including Google employees – to go back in time to regenerate them.

We and our colleagues have successfully built monitoring systems that have preserved increasingly larger bodies of ephemeral experiences in the days leading up to six elections in the US [135-137]. In 2020, we preserved and subsequently analyzed more than 1.5 million ephemeral experiences obtained and then aggregated through the computers of a politically-diverse group of 1,735 registered voters in four swing states. We captured data on the Google, Bing, and Yahoo search engines, as well as on YouTube and Google's home page [138,139], and we found substantial political bias on these platforms, sufficient, perhaps, to have shifted millions of votes among undecided voters. Based on our preliminary analysis of data we had collected, on November 5, 2020, three US Senators sent a warning letter to the CEO of Google about the political bias we had detected in Google content, and Google immediately turned off political bias in the search results it was sending to Georgia residents in the weeks leading up to the two US Senate runoff elections scheduled there for January 5, 2021. Google also stopped sending go-vote reminders to Georgia residents. Monitoring systems, it appears, can be used to make Big Tech companies accountable to the public. As Supreme Court Justice Louis D. Brandeis opined a century ago, "Sunlight is said to be the best of disinfectants; electric light the most efficient policeman" [140].

In 2022, we expanded our network of "field agents" to include 2,742 registered voters, and we preserved more than 2.5 million ephemeral experiences on multiple platforms, this time including both Twitter and Facebook [141]. We are currently in the process of building a permanent, large-scale "digital shield" in all 50 US states which will, we hope, protect our elections from manipulation by emerging technologies for the foreseeable future [141-143].

Proposals have also been made to try to track or reduce the potential manipulative power of software such as YouTube's recommender algorithms by developing methods that increase algorithmic transparency and accountability [144-148,cf.149,150]. The companies that control these algorithms will likely resist such efforts, however, and because of their increasing reliance on machine learning techniques, algorithms have grown increasingly opaque over the years – so mysterious that even the original programmers can't understand them [151]. The clearest way, in our view, to monitor, preserve, and analyze algorithmic output is to look over the shoulders of large, representative samples of real users as they are viewing real content. Doing so is necessary in part because so much content is now customized to fit characteristics of individual users [152-154].

YouTube Manipulation Effect (YME)

The present paper focuses on a powerful new form of influence we call the YouTube Manipulation Effect (YME), in which we use a YouTube simulator we call DoodleTube to determine the extent to which we can shift the opinions and voting preferences of undecided voters by manipulating the order of recommended videos – in other words, by exercising control over YouTube's recommender algorithms. The videos were biased to favor one candidate or his opponent. By manipulating the order, we also had control over which video was in the up-next position and which therefore would play automatically if the user did not select a different video. By parsing the data demographically, we also determined how vulnerable people in different demographic groups were to the manipulation.

2. Ethics Statement

The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. AIBRT is registered with the HHS Office for Human Research Protections (OHRP) under IORG0007755. The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB is FWA00021545. Informed written consent was obtained for both experiments as specified in the Procedure section of Experiment 1.

3. Experiment 1: Biased YouTube Ordering with No Mask

In our first experiment, we sought to determine whether a biased ordering of videos – biased to favor one political candidate – could shift opinions and voting preferences toward that candidate. By "no mask," we mean that high-ranking videos consistently favored one candidate. In Experiment 2, in order to reduce perception of bias, we masked the bias by mixing in videos that supported the non-favored candidate (see Procedure sections below for details).

3.1 Methods

3.1.1 Participants

After cleaning, our participant sample for this experiment consisted of 959 eligible US voters recruited through the Amazon Mechanical Turk (MTurk) subject pool. During the cleaning process, we removed participants who reported an English fluency level below 6 on a 10-point scale, where 1 was labeled "Not fluent" and 10 was labeled "Highly fluent." In order to assure that our participants were undecided, we also removed participants who reported a level of familiarity with either of the two candidates exceeding 3 on a 10-point scale. In all, 41 participants were removed during cleaning.

Overall, 558 (58.2%) of our participants identified themselves as female, 391 (40.8%) as male, and 10 (1.0%) chose not to identify. Racial and ethnic background was as follows: 701 (73.1%) of our participants identified themselves as White, 102 (10.6%) as Black, 75

(7.8%) as Asian, 58 (6.0%) as Mixed, and 22 (2.3%) as Other. Overall, 26.8% of the individuals in the sample identified themselves as non-White.

Regarding level of education completed: 2 (0.2%) reported no education; 41 (4.3%) reported not having a high school degree; 309 (32.2%) reported completing high school; 428 (44.6%) reported having a bachelor's degree; 148 (15.4%) reported having a master's degree; and 31 (3.2%) reported having a doctoral degree.

Regarding Political alignment: 435 (45.4%) of the participants identified themselves as liberal, 291 (30.3%) as moderate, 187 (19.5%) as conservative, 30 (3.1%) as not political, and 16 (1.7%) as other.

Regarding YouTube usage: 958 (99.9%) reported that they have used YouTube before and 602 (62.8%) reported that they have used YouTube to get information about political or ideological topics. Participants reported using YouTube an average of 18.0 (*SD* = 32.4) times a week.

See S1 Table for detailed demographic information for Experiment 1.

3.1.2 Procedure

All procedures were conducted online, with sessions conducted on December 7, 2021, December 11, 2021, and January 7, 2022. Participants were first asked two screening questions; sessions were terminated if they said they were not eligible to vote in the US or if they reported a level of familiarity with Australian politics exceeding 3 on a 10-point scale.

Participants who passed our screening questions were then asked various demographic questions and then given instructions about the experimental procedure. We also displayed a short video and asked participants whether they were able to see the video and whether the video autoplayed. If the videos did not play at all or did not autoplay the session was terminated. At the end of the instructions page, and in compliance with APA and HHS guidelines, participants were asked for their consent to participate in the study. If they clicked "I Do," the session continued; if they clicked "I Do Not," the session ended.

Participants were then asked further questions about their political leanings and voting behavior.

Participants were then given a short paragraph about each candidate (See S1 Text in Supporting Information for the full paragraphs), each about 120 words in length. Participants were next asked three opinion questions on a 10-point scale about each candidate: one regarding their overall impression of the candidate, one regarding how likeable they found the candidate, and one regarding how much they trusted the candidate. They were then asked, on an 11-point scale with values ranging from 5 to 0 to 5, which candidate they would be likely to vote for if they “had to vote today.” Finally, they were asked which candidate they would vote for if they “had to vote right now” (forced choice).

Participants were then given an opportunity to use DoodleTube – our YouTube simulator – to watch videos about these candidates in order to gather information to help them decide which of the two candidates to vote for. They were given a maximum of 15 minutes and a minimum time of 10 minutes to view the videos. See S2 Text for the complete instructions.

On the next screen, participants saw an online video platform called DoodleTube displaying a search bar with a pre-inputted query of “Australian Prime Minister Election” and a series of videos relating to that query (Fig 1). Participants could click on any of the videos to play it. When a video was clicked, the screen switched to a video view screen with the up-next video on the top of the right side bar, along with other recommended videos beneath it (Fig 2).

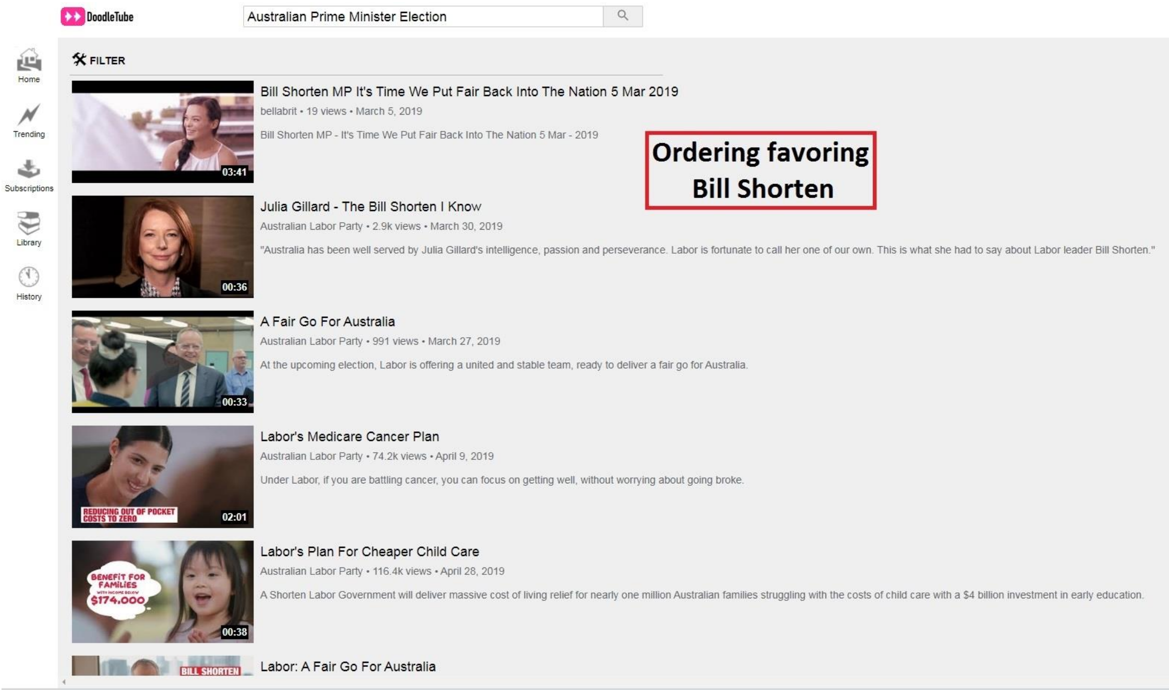


Fig 1. Initial screen when a DoodleTube Session begins. In this instance, the participant had been randomly assigned to a group in which the order of the videos favored candidate Bill Shorten. (The red-outline box above was not shown.)

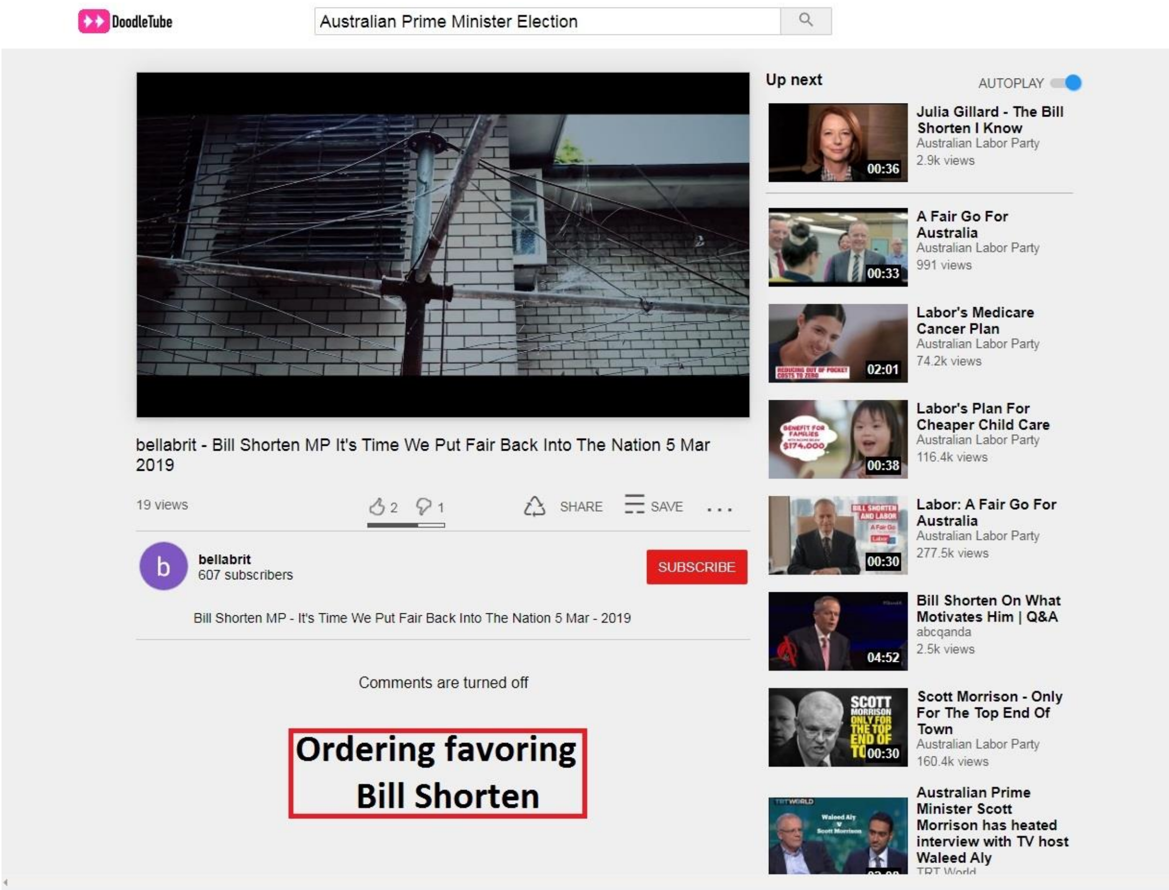


Fig 2. Screen that appears after the participant has clicked on one of the videos shown in Fig 1. (The red-outlined box above was not shown.)

The participant could watch an entire video, or the participant could click on a different video to switch the view to that new one. The participant could also allow a video to play to the end and then allow the up-next video to play; this occurred automatically if the participants did not click on another video.

Participants had been randomly assigned to one of three groups: Pro-Candidate-A (Scott Morrison), Pro-Candidate-B (Bill Shorten), or the control group. People in all three groups had access to all 40 of the videos that were included in the experiment, but the videos were listed in a different order in each group. As shown in Fig 3A, in the Pro-Morrison group, the order of the videos would go from Pro-Morrison videos to Pro-Shorten videos. If the participant was assigned to the Pro-Shorten group, the order of the videos would go from Pro-Shorten videos to Pro-Morrison videos (Fig 3B). The control group would have both Pro-Morrison and Pro-Shorten videos in alternation (Fig 3C).

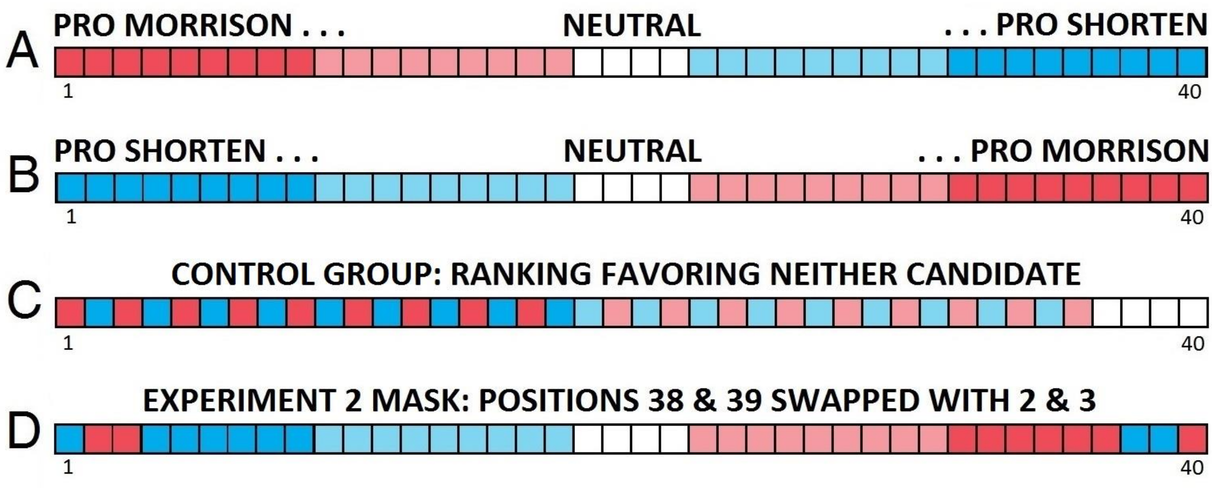


Fig 3. Ordering of videos in Experiments 1 and 2. A: Experiment 1, Group 1 (first bias group, 40 videos in order from Pro Morrison to neutral to Pro Shorten). B: Experiment 1, Group 2 (second bias group, 40 videos in order from Pro Shorten to neutral to Pro Morrison). C: Group 3 (control group, 40 videos in which Pro Morrison and Pro Shorten are alternated). D: Experiment 2: The manipulation is masked by swapping the video in Position 2 with the video in Position 39, and the video in Position 3 with the video in Position 38 (shown only for Group 1, Pro Morrison bias group).

The End button in the upper-left corner of the web page was invisible and inaccessible for the minimum required video view time of 10 minutes. After the 10 minutes was up, the End button appeared. However, participants were allowed to continue browsing the videos up to the maximum time of 15 minutes. When the 15 minutes was up, the participants were redirected to the questions page. The timer was paused when the videos were paused and restarted when the videos were restarted to ensure participants were viewing the video for the required amount of time. The autoplay feature was on by default and could not be turned off. Users could, however, go to whatever video they desired either on the sidebar or the home screen.

Following the DoodleTube experience, participants were again asked the same series of questions that they were asked before they began watching the videos: three opinion questions about each candidate (10-point scale on impression, likeability, trust), which candidate they were likely to vote for (11-point scale from 5 to 0 to 5), and which candidate they would vote for now (forced choice).

Next, participants were asked whether any of the content they had seen on DoodleTube "bothered" them in any way. They could reply Yes or No, and then they could explain their answer by typing freely in a text box. This is a conservative way of determining whether people perceived any bias in the content they had seen .We could not ask people directly about their awareness of bias because leading questions of that sort often produce misleading answers (Loftus, 1975). To assess bias we searched the textual responses for words such as "bias," "skewed," or "slanted" to identify people in the bias groups who had apparently noticed the favoritism in the search results they had been shown.

3.2 Results

For election campaign officials, the most important result in this experiment would almost certainly be what we call the "vote manipulation power" or VMP, which we define as the post-manipulation percentage increase in the proportion of participants preferring the favored candidate (in response to the forced-choice question). In a group that is initially split 50/50, the VMP also turns out to be the post-manipulation vote margin. For additional information about VMP and how we compute it, see S3 Text in the Supporting Information.

In Experiment 1, for all participants in the two bias groups combined (Groups 1 and 2), the VMP was 51.5% (McNemar's Test $X^2 = 98.20$, $p < 0.001$). S2 to S4 Tables show VMPs broken down by gender, race/ethnicity, and level of educational attainment. For those demographic characteristics, we found significant and consistent effects only for gender, with females having higher VMPs than males.

On the 11-point voting preferences scale, pre-manipulation, we found no significant difference between mean ratings in the three groups ($M_{\text{Morrison}} = -0.22$, $SD = 2.75$; $M_{\text{Shorten}} = -0.15$, $SD = 2.90$; $M_{\text{Control}} = 0.04$, $SD = 2.81$; Kruskal-Wallis $H = 1.53$; $p = 0.47$ NS). Post manipulation, we found a significant difference between mean ratings in the three groups ($M_{\text{Morrison}} = -1.85$, $SD = 3.19$; $M_{\text{Shorten}} = 2.14$, $SD = 3.11$; $M_{\text{Control}} = 0.45$, $SD = 3.53$; $H = 188.67$; $p < 0.001$). Participants in Group 1 shifted 1.63 points towards the favored candidate (Morrison), and participants in the Group 2 condition shifted 2.29 points towards the favored candidate (Shorten). In addition, the pre-manipulation mean preference for the favored candidate (Groups 1 and 2 combined) was significantly different from the post-manipulation mean preference for the favored candidate (Groups 1 and 2 combined) ($M_{\text{Pre}} = 0.04$, $SD_{\text{Pre}} = 2.83$, $M_{\text{Post}} = 1.99$, $SD_{\text{Post}} = 3.15$, Wilcoxon $z = -12.02$, $p < 0.001$). Opinion ratings for both candidates also shifted significantly in the predicted direction (Table 1). Finally, the proportion of videos our participants watched which were selected by our up-next recommendation was 56.8% ($SD = 27.1$). (Nonparametric statistical tests such as the Kruskal-Wallis H are frequently employed in this study because the ratings of the candidates lie on ordinal scales [155]. Means and standard deviations in such instances are reported for comparison purposes, although the appropriateness of their use with ordinal data has long been debated [156,157].

Demographic breakdowns of the data obtained on the 11-point voting preference scale in Experiment 1 are shown in S5 to S13 Tables. Male/female differences on this scale were highly significant (S5 to S7 Tables). Differences by educational attainment and race/ethnicity were not consistently significant (S8 to S13 Tables).

Table 1. Experiment 1: Pre and Post opinion ratings of favored and non-favored candidates.

	Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
	Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	**z†**
Impression	7.07 (1.87)	7.42 (2.35)	0.35	7.01 (1.93)	4.82 (2.45)	-2.19	-13.64***
Trust	6.18 (2.07)	6.70 (2.52)	0.52	6.22 (2.07)	4.52 (2.39)	-1.70	-12.84***
Likeability	6.98 (1.85)	7.43 (2.41)	0.45	6.95 (1.96)	4.81 (2.50)	-2.14	-14.12***

†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate
*** $p < 0.001$

Although the shift in voting preferences was substantial (VMP = 51.5%) in Experiment 1, it is notable that 33.0% of the participants in the two bias groups (Groups 1 and 2) appeared to detect political bias in the videos they watched. In SEME experiments, perception of bias can easily be reduced with masking procedures – for example, by mixing one or more search results favoring the non-favored candidate mixed among the more frequent and higher-ranking search results biased toward the favored candidate [2]. Could we reduce perception of bias in a YouTube environment using a similar mask, and, if so, might we still produce a substantial shift in voting preferences? We attempted to answer these questions in Experiment 2.

4. Experiment 2: Biased YouTube Ordering with Mask

4.1 Methods

4.1.1 Participants

After cleaning, our participant sample consisted of 491 eligible US voters recruited through the MTurk subject pool. The cleaning procedure was identical to that of Experiment 1, and a total of nine participants were removed from the sample during that procedure. The group was demographically diverse. See S1 Table for detailed demographic information for Experiment 2.

Regarding YouTube usage: 489 (99.6%) reported that they had used YouTube before and 310 (63.1%) reported that they had used YouTube to get information about political or ideological topics. Participants reported using YouTube an average of 16.2 (*SD* =30.2) times per week.

4.1.2 Procedure

The procedure in Experiment 2, with sessions conducted on December 24, 2021, and January 8, 2022, was identical to that of Experiment 1, with one exception: The order of videos in the experimental groups had a mask in the 2nd and 3rd positions. Specifically, we swapped the usual video in Position 2 with the video from Position 39, and the usual video in Position 3 with the video from Position 38 (see Fig 3D). In other words, in the Pro-Morrison group, the video order remained the same as it was in Experiment 1 except in the 2nd and 3rd order the videos were Pro-Shorten. Similarly for Pro-Shorten group, the videos in the 2nd and 3rd position were Pro-Morrison.

4.2 Results

In Experiment 1 (no mask), the VMP was 51.5%, and 33.0% of participants in the two bias groups showed some awareness of bias in the ordering of the videos. In Experiment 2 (mask), the VMP was 65.6% (McNemar's Test $X^2 = 67.11$, $p < 0.001$), and only 14.6% of participants in the two bias groups showed some awareness of bias in the ordering of the videos. The VMP in Experiment 2 was 27.4% higher than the VMP in Experiment 1 ($z = 4.23$, $p < 0.001$). The perception of bias in Experiment 2 was 55.8% lower than the perception of bias in Experiment 1 ($z = 6.19$, $p < 0.001$). Thus it appears that the manipulation can indeed be masked in such a way as to reduce perception of bias (perhaps to zero) while still producing a substantial shift in voting preferences. Again, S2 to S4 Tables show VMPs broken down by gender, race/ethnicity, and level of educational attainment. For those demographic characteristics, we again found consistently significant effects only for gender, with females having higher VMPs than males.

Voting preferences as measured on the 11-point scale also shifted in the predicted direction. Pre-manipulation, we found no significant difference between mean ratings in the three groups ($M_{\text{Morrison}} = -0.08$, $SD = 2.84$; $M_{\text{Shorten}} = -0.15$, $SD = 2.65$; $M_{\text{Control}} = -0.25$, $SD = 2.88$; $H = 0.27$, $p = 0.873$ NS). Post manipulation, we found a significant difference between mean ratings in the three groups ($M_{\text{Morrison}} = -1.87$, $SD = 3.23$; $M_{\text{Shorten}} = 2.03$, $SD = 2.88$; $M_{\text{Control}} = 0.59$, $SD = 3.53$; $H = 95.64$; $p < 0.001$). Participants in Group 1 shifted 1.79 points towards the favored candidate (Morrison), and participants in the Group 2 condition shifted 2.18 points towards the favored candidate (Shorten). In addition, the pre-manipulation mean preference for the favored candidate (Groups 1 and 2 combined) was significantly different from the post-manipulation mean preference for the favored candidate (Groups 1 and 2 combined) ($M_{\text{Pre}} = -0.02$, $SD_{\text{Pre}} = 2.75$; $M_{\text{Post}} = 1.95$, $SD_{\text{Post}} = 3.07$; Wilcoxon $z = -9.05$, $p < 0.001$). Opinion ratings for both candidates also shifted significantly in the predicted direction (Table 2). Finally, the proportion of videos our participants watched which were selected by our up-next recommendation was 60.7% ($SD = 25.6$).

Demographic breakdowns of the data obtained on the 11-point voting preference scale in Experiment 2 are shown in S5 to S13 Tables. Male/female differences on this scale were not consistently significant (S5 to S7 Tables). Neither were differences by educational attainment or race/ethnicity (S8 to S13 Tables).

Table 2. Experiment 2: Pre and Post opinion ratings of favored and non-favored candidates

	Favored Candidate Mean (SD)			**Non-Favored Candidate Mean (SD)**			
	Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*** †
Impression	6.91 (1.78)	7.30 (2.22)	0.39	6.96 (1.80)	4.89 (2.37)	-2.07	-9.81***
Trust	6.08 (2.01)	6.69 (2.30)	0.61	6.12 (2.03)	4.65 (2.38)	-1.47	-9.06***
Likeability	6.86 (1.81)	7.49 (2.18)	0.63	6.87 (1.79)	4.99 (2.44)	-1.88	-9.91***

†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate
*** $p < 0.001$

5. Discussion

Our experiments demonstrate that (a) strategic ordering of videos on a YouTube-like platform can dramatically shift both the opinions and voting preferences of undecided voters, rapidly shifting a substantial portion of them to favor one political candidate (Experiment 1), and (b) this manipulation can be masked to reduce perception of bias while still producing large, predictable shifts in opinions and voting preferences (Experiment 2). These findings are important because, as we have documented in our Introduction, an increasing body of evidence demonstrates the power that YouTube itself has to impact the thinking and behavior of people worldwide, sometimes in destructive or self-destructive ways.

Like search results, search suggestions, answer boxes, and newsfeeds, video sequences constructed by YouTube's recommender algorithm are ephemeral in nature. Once again, that means that this form of influence normally leaves no paper trail for authorities to trace or, perforce, for researchers to study. Without monitoring systems in place to preserve large representative samples of ephemeral content, people will be blind to the ways in which they are being impacted by the algorithms of tech companies, and we will in effect be turning our democracy, and, to some extent, our own minds over to those companies. Children are especially impressionable [32,33,158], and with mobile devices having now become both the babysitter and the companion of choice for children [19,23], it is reasonable to conjecture that the new forms of influence that the internet has made possible are impacting our children and grandchildren profoundly [20-22,24-27,31]. In our view, laws and regulations will never be able to keep pace with rapidly changing technologies, but monitoring systems can. It's good tech battling bad tech, now and in the future.

In recent months, our team has preserved more than 21 million online ephemeral experiences in all 50 US states through the computers of a politically-balanced group of more than 9,000 registered voters, and we are now beginning to preserve content from the mobile devices of more than 2,000 children and teens (with the permission of their parents). We are also developing ways of analyzing much of this data in real time, and we ultimately will give both the authorities and the general public free access to our findings 24 hours a day through public dashboards. In our view, in a world in which unprecedented power has been given to private companies to impact people's thinking and behavior, monitoring systems are not optional. Without them, we will not only have no idea how they may be influencing us, governments that implement laws or regulations to contain the power of these companies will have no reliable way of measuring compliance with those laws and regulations.

Limitations and Future Research

Our current procedures do not include any follow up, so we have no way of measuring how long the changes our procedures produce in opinions and voting preferences will last. Just as the content our participants see is fleeting, so might be the changes we are detecting in their opinions and voting preferences. That said, we might actually be underestimating the possible power of YME as it might be impacting real users, because we are exposing our participants to only a single manipulation. In the months leading up to an election, a company such as YouTube might be exposing users to similarly biased content repeatedly, and users themselves might choose to view certain videos multiple times, just as mass-murderer Brenton Harrison Tarrant did in Australia (see Introduction). If YME is an additive effect, multiple exposures will increase its impact, in which case the shifts we have produced in our experiments might be smaller than the shifts occurring in the natural environment. In ongoing experiments on what we call the "multiple exposure effect" (MEE), we are now measuring the possible additive effects of YME, ABE, SEME, and other new forms of influence made possible by the internet.

In the real world, bias in YouTube videos might also be similar to bias on other platforms, such as the Google search engine, Facebook, Instagram, and TikTok, as well as by answers given to users by AIs such as Bard and ChatGPT, in addition to answers given to users by personal assistants such as Alexa, the Google Assistant, and Siri. Again, in ongoing experiments, we are now measuring the possible additive effects of similarly biased content presented to users on different platforms, a phenomenon we called the "multiple platforms effect" (MPE).

In the real world, undecided voters are subject to many different kinds of influence when they are trying to make up their minds; YouTube is only one possible source of influence, needless to say, and some people rarely or never use YouTube. In addition, some people never use YouTube in a way that gives them information about political candidates or issue relevant to elections; in the US, in fact, the most common videos users search for on YouTube are BTS and Pewdiepie [159]. Our findings apply only to people who use YouTube fairly regularly and who are likely to encounter information relevant to elections.

That said, it is important to note here that the kind of influence YouTube can exert is especially powerful compared to other common sources of influence that affect people prior to elections. Most of those sources – television advertisements, billboards, and even ballot harvesting – are inherently competitive and often generate relatively small net effects, if any [160]. YouTube, on the other hand, has no effective competitor. Although opposing political campaigns can compete in the process of posting new video content to YouTube, only YouTube controls the process by which those videos are filtered, ordered, and customized. In other words, if employees, executives, or algorithms at YouTube favor one candidate, the opposing candidate has no way of counteracting that favoritism. Similarly, if employees, executives, or algorithms at YouTube choose to suppress the content of one candidate, that candidate has no way to counteract that suppression.

We note that our study and discussion have focused mainly on the potential that the YouTube platform has to alter people's thinking and behavior. In our view YouTube presents at least two other major challenges to both researchers and policy makers, and we would be remiss in not pointing them out. First, YouTube has been repeatedly called to task in recent years for what some people consider to be censorship – removing content, restricting access to content, or demoting content on the platform [102,104,106-112]. We discussed this issue briefly in our Introduction, referring at one point to a PowerPoint presentation that leaked from Google entitled, "The Good Censor" [18]. Second, YouTube is an aggressive surveillance tool; Google openly tracks and subsequently models and monetizes the massive amount of data it collects about the videos people watch and the comments people post on YouTube [161,162]. The information collected about children has been of special concern to authorities in recent years [163]. The present study needs to be viewed in this larger context, we believe. As a powerful and unprecedented tool of surveillance, censorship, and manipulation impacting billions of people worldwide, YouTube needs, in our view, to be subjected to close scrutiny by researchers and public policy makers aggressively and without delay.

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Acknowledgments

This paper is based on part on a paper presented at the April 2022 meeting of the 102nd annual meeting of the Western Psychological Association. We thank J. Krikorian, H. G. Park, and X. Tong for help with preparing the experiments, and M. Kang and V. Zankich for help with the data analysis. We are especially grateful to L. Kafader and Z. Perez for expert programming assistance.

Author Contributions

Robert Epstein: Conceptualism, research supervision, writing of original draft, data analysis, editing; **Alex Flores:** background research, editing and preparation of manuscript.

Funding

This work was supported by general funds of the American Institute for Behavioral Research and Technology (AIBRT), a nonpartisan, nonprofit, 501(c)(3) organization. The authors received no specific funding for this work. This research did not receive any external support from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability and Materials

Anonymized raw data will be made available at Zenodo.com after the paper has been accepted for publication. Anonymized data can also be requested from info@aibrt.org. Anonymization was required to comply with the requirement of the sponsoring institution's Institutional Review Board that the identities of the participants be protected in accordance with HHS Federal Regulation 45 CFR 46.101.(b)(2).

References

1. Epstein R, Voillot M. The YouTube Manipulation Effect (YME): The power that bias in YouTube's up-next algorithm has to shift votes and opinions, and preliminary evidence that such bias exists. Paper presented at: Western Psychological Association; 2022 Apr; Portland, OR. Available from: https://aibrt.org/downloads/EPSTEIN_&_VOILLOT_2022-WPA-YouTube_Manipulation_Effect-YME.pdf
2. Epstein R, Robertson RE. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proceedings of the National Academy of Sciences of the United States of America. 2015 Aug 4; 112(33):E4512– E4521. Available from: https://www.pnas.org/doi/full/10.1073/pnas.1419828112
3. Epstein R. Why Google poses a serious threat to democracy, and how to end that threat. Congressional Record of the United States. 2019 July 16. Available from: https://www.judiciary.senate.gov/imo/media/doc/Epstein%20Testimony.pdf
4. Epstein R. Manipulating minds: The power of search engines to influence votes and opinions. In: Moore M, Tambini D, editors. Digital dominance: The power of Google, Amazon, Facebook, and Apple. Oxford, UK: Oxford University Press; 2018. pp. 294–319. Available from: https://aibrt.org/downloads/EPSTEIN_2018-Manipulating_minds-The-power_of_search_engines_to_influence_votes_and_opinions-UNCORRECTED_PROOFS.pdf
5. Epstein R, Lee V, Mohr R, Zankich VR. The answer bot effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. PLOS ONE, 2022 June 1;17(6): e0268081. https://doi.org/10.1371/journal.pone.0268081
6. Epstein R, Ding M, Mourani C, Olson E, Robertson RE, Tran F. Multiple searches increase the impact of the Search Engine Manipulation Effect (SEME). Paper presented at: 97th annual meeting of the Western Psychological Association; 2017 Apr;

Sacramento, CA. Availble from: https://aibrt.org/downloads/EPSTEIN_et_al._2017-WPAMultiple_Searches_Increase_the_Impact_of_the_Search_Engine_Manipulation_Effect.pdf

7. Epstein R, Tyagi C, Wang H. What would happen if Twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME). PLOS ONE. In press. Preprint available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4207187)
8. McKinnon JD, McMillan D. Google Workers Discussed Tweaking Search Function to Counter Travel Ban: Company says none of proposed changes to search results were ever implemented. The Wall Street Journal. 2018 Sept 20. Available from: https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472
9. Epstein R. How Google shifts votes: A "go vote" reminder is not always what you think it is. The Epoch Times. 2019 Jan 2. Available from: https://www.theepochtimes.com/another-way-google-manipulates-votes-without-us-knowing-a-go-vote-reminder-is-not-what-you-think-it-is_2754073.html
10. Rosenberg M, Frenkel S. Facebook's role in data misuse sets off storms on two continents. The New York times. 2018 Mar 18. Available from: https://www.nytimes.com/2018/03/18/us/cambridge-analytica-facebook-privacy-data.html
11. Gottlieb C. The French Competition Authority fines Google €220 million for favoring its own advertising technologies. 7 June 2021. In: Cleary Antitrust Watch [Internet]. Available from: https://www.clearyantitrustwatch.com/2021/06/the-french-competition-authority-fines-google-e220-million-for-favoring-its-own-advertising-technologies/
12. Gottlieb C. The Commission fines Google €1.49 billion for breaching EU antitrust rules in the Google AdSense investigation. 20 March 2019. In: Cleary Antitrust Watch

[Internet]. Available from: https://www.clearyantitrustwatch.com/2019/03/the-commission-fines-google-e1-49-billion-for-breaching-eu-antitrust-rules-in-the-google-adsense-investigation/

13. Bond RM, Fariss CJ, Jones JJ, Kramer ADI, Marlow C, Settle JE, Fowler JH. A 61-million-person experiment in social influence and political mobilization. Nature. 2012 Sept 12. https://doi.org/10.1038/nature11421
14. Kulwin N. The Internet Apologizes... Even those who designed our digital world are aghast at what they created. A breakdown of what went wrong — from the architects who built it [Internet]. Intelligencer. 2018 Apr 16. Available from: http://nymag.com/intelligencer/2018/04/an-apology-for-the-internet-from-the-people-who-built-it.html
15. McNamee R. I invested early in Google and Facebook. Now they terrify me. USA Today [Internet]. 2017 Aug 8; Available from: https://eu.usatoday.com/story/opinion/2017/08/08/my-google-and-facebook-investments-made-fortune-but-now-they-menace/543755001/
16. Savov V. Google's Selfish Ledger Is an Unsettling Vision of Silicon Valley Social Engineering [Internet]. The Verge. 2018 May 17. Available from: https://www.theverge.com/2018/5/17/17344250/google-x-selfish-ledger-video-data-privacy
17. Epstein R. Transcript to Google's internal video, "The Selfish Ledger." 2018 May 25. Available from: https://aibrt.org/downloads/GOOGLE-Selfish_Ledger-TRANSCRIPT.pdf
18. Bokhari A. The Good Censory: Google Briefing Accused Trump of Spreading 'Conspiracy Theory' [Internet]. Breitbart News. 2018 Oct 9. Available from: https://www.breitbart.com/tech/2018/10/09/the-good-censor-google-trump-conspiracy/

19. Vogels EA, Gelles-Watnick R, Massarat N. Teens, social media and technology 2022 [Internet]. Pew Research Center: Internet, Science & Tech. 2022. Available from: https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/

20. Weill K. How YouTube built a radicalization machine for the far-right [Internet]. The Daily Beast. 2018 Dec 17. Available from: https://www.thedailybeast.com/how-youtube-pulled-these-men-down-a-vortex-of-far-right-hate

21. Kaushal R, Saha S, Bajaj P, Kumaraguru P. KidsTube: Detection, characterization and analysis of child unsafe content & promoters on YouTube. In: 2016 14th Annual Conference on Privacy, Security and Trust (PST), Auckland, New Zealand, pp. 157-164; 2017 Apr 24. Available from: https://doi.org/10.1109/PST.2016.7906950

22. Papadamou K, Papasavva A, Zannettou S, Blackburn J, Kourtellis N, Leontiadis I, Stringhini G, Sirivianos M. Disturbed YouTube for kids: Characterizing and detecting inappropriate videos targeting young children. The Proceedings of the Fourteenth International AAAI Conference on Web and Social Media. 2020 July 1. Available from: https://doi.org/10.5281/zenodo.363278

23. Auxier B, Anderson M, Perrin A, Turner E. 2. Parental views about YouTube. Pew Research Center: Internet, Science & Tech. 2020 July 28.Available from: https://www.pewresearch.org/internet/2020/07/28/parental-views-about-youtube/

24. Subedar A, Yates W. The disturbing YouTube videos that are tricking children. BBC. 2017 March 27. Available from: http://www.bbc.com/news/blogs-trending-39381889

25. Koerber, B. Gaming the system: How creepy YouTube channels trick kids into watching violent videos. Mashable. 2017 October 22. Available from: http://mashable.com/2017/10/22/youtube-kids-app-violent-videos-seo-keywords/#fRMwn0UO1aqm

26. Mole, B. Suicide instructions spliced into kids' cartoons on YouTube and YouTube Kids: "Sideways for attention. Longways for results," a man says in the middle of a cartoon. Ars Technica. 2019 Feb 25. Available from: https://arstechnica.com/science/2019/02/youtube-kids-cartoons-include-tips-for-committing-suicide-docs-warn/

27. Tan L, Ng SH, Omar A, Karupaiah T. What's on YouTube? A Case Study on Food and Beverage Advertising in Videos Targeted at Children on Social Media. 2018 July 1. Available from: https://doi.org/10.1089/chi.2018.0037

28. Roose K. The making of a YouTube radical. The New York times [Internet]. 2019 Jun 8; Available from: https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html

29. Fisher M, Taub A. How YouTube radicalized Brazil. The New York times. 2019 Aug 12. Available from: https://www.nytimes.com/2019/08/11/world/americas/youtube-brazil.html

30. Fisher M, Taub A. We wanted to know how online radicalization was changing the world. We started with Brazil. The New York times. 2019 Aug 12; Available from: https://www.nytimes.com/2019/08/11/reader-center/brazil-youtube-radicalization.html

31. Jung EJ, Kim S. Suicide on YouTube: Factors engaging viewers to a selection of suicide-themed videos. PLOS ONE. 2021 June 10. 16(6): e0252796. Available from: https://doi.org/10.1371/journal.pone.0252796

32. Rideout V, Robb MB. The common sense census: Media use by kids age zero to eight, 2020. Common Sense Media. 2020 Nov 17. Available from: https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-kids-age-zero-to-eight-2020

33. Radesky JS, Schaller A, Yeo SL, Robb MB. Young kids and YouTube: How ads, toys, and games dominate viewing [Internet]. Common Sense Media. 2020 Nov 17. Available

from: https://www.commonsensemedia.org/research/young-kids-and-youtube-how-ads-toys-and-games-dominate-viewing

34. Tufekci Z. Opinion YouTube, the Great Radicalizer. The New York times [Internet]. 2018 Mar 10; Available from: https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html

35. Tufekci Z. YouTube's recommendation algorithm has a dark side. Scientific American [Internet]. 2019 Apr 1; Available from: https://www.scientificamerican.com/article/youtubes-recommendation-algorithm-has-a-dark-side/

36. Lewis R. Alternative Influence: Broadcasting the Reactionary Right on YouTube. Data & Society [Internet]. 2018 Sept 18;765:146–57. Available from: https://datasociety.net/library/alternative-influence

37. Jordan G, Shorland A (Producers/Directors). The Weekly Episode 9:'The Rabbit Hole' The New York Times [Internet]. 2019 Aug 9. Available from: https://www.nytimes.com/2019/08/09/the-weekly/what-is-youtube-pushing-you-to-watch-next.html

38. O'Donovan C, Warzel C, McDonald L, Clifton B, Woolf M. We followed YouTube's recommendation algorithm down the rabbit hole [Internet]. BuzzFeed News. 2019 Jan 24. Available from: https://www.buzzfeednews.com/article/carolineodonovan/down-youtubes-recommendation-rabbithole

39. Douthat R. Opinion The Faults Beyong Our Algorithms The offline roots of online radicalization. The New York times. 2019 Jun 11. Available from: https://www.nytimes.com/2019/06/11/opinion/youtube-radicalization-algorithms.html

40. Appiah KA. YouTube videos brainwashed my father. Can I reprogram his feed? The New York times. 2021 Apr 20. Available from: https://www.nytimes.com/2021/04/20/magazine/youtube-radicalization.html

41. Nicas J. How YouTube drives people to the internet's darkest corners. Wall Street journal [Internet]. 2018 Feb 7; Available from: https://www.wsj.com/articles/how-youtube-drives-viewers-to-the-internets-darkest-corners-1518020478

42. Ledwich M, Zaitsev A. Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. First Monday [Internet]. 24 Dec 2019; Available from: https://firstmonday.org/ojs/index.php/fm/article/view/10419

43. O'Callaghan D, Greene D, Conway M, Carthy J, Cunningham P. Down the (white) rabbit hole: The extreme right and online recommender systems: The extreme right and online recommender systems. Social Science Computer Review. 2015;33(4):459–78. 2014 Oct 16. Available from: http://dx.doi.org/10.1177/0894439314555329

44. Young HSW, Caine J (Chair/ Member). Report of the Royal Commission of Inquiry into the Attack on Christchurch Mosques on 15 March 2019. The report [Internet]. 2020 Dec 8. Available from: https://christchurchattack.royalcommission.nz/the-report/

45. Chen AY, Nyhan B, Jason R, Robertson RE, Wilson C. Exposure to Alternative & Extremist Content on YouTube [Internet]. ADL. 03 May 2022. Available from: https://www.adl.org/resources/report/exposure-alternative-extremist-content-youtube

46. Ribeiro MH, Ottoni R, West R, Almeida VAF, Meira W. Auditing radicalization pathways on YouTube. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2020 Jan; Available from: https://doi.org/10.1145/3351095.3372879

47. Alfano M, Fard AE, Carter JA, Clutton P, Klein C. Technologically scaffolded atypical cognition: the case of YouTube's recommender system. Synthese. 2021;199(1–2):835–58. Available from: http://dx.doi.org/10.1007/s11229-020-02724-x

48. Faddoul M, Chaslot G, Farid H. A longitudinal analysis of YouTube's promotion of conspiracy videos [Internet]. arXiv. 2020 March 6. Available from: http://arxiv.org/abs/2003.03318

49. Munger K, Phillips J. A Supply and Demand Framework for YouTube Politics. The International Journal of Press/Politics. 27(1). 2019 Oct 1. Available from: https://doi.org/10.1177/1940161220964767

50. Hosseinmardi H, Ghasemian A, Clauset A, Mobius M, Rothschild DM, Watts DJ. Examining the consumption of radical content on YouTube. Proceedings of the National Academy of Sciences of the United States of America. 2021 August 2. Available from: http://dx.doi.org/10.1073/pnas.2101967118

51. Costello M, Hawdon J, Ratliff T, Grantham T. Who views online extremism? Individual attributes leading to exposure. Computers in Human Behavior. 2016 May 13;63:311–20. Available from: http://dx.doi.org/10.1016/j.chb.2016.05.033

52. Conway M, McInerney L. Jihadi video and auto-radicalisation: Evidence from an exploratory YouTube study. In: Intelligence and Security Informatics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 108–18. Available from: https://doi.org/10.1007/978-3-540-89900-6_13

53. Agarwal S, Sureka A. A focused crawler for mining hate and extremism promoting videos on YouTube. In: Proceedings of the 25th ACM conference on Hypertext and social media. New York, NY, USA: ACM; 2014.

54. Ottoni R, Cunha E, Magno G, Bernardina P, Meira WJ, Almeida V. Analyzing right-wing YouTube channels: Hate, Violence and Discrimination. In: Proceedings of the 10th ACM Conference on Web Science. New York, NY: ACM; 2018. Available from: https://arxiv.org/abs/1804.04096

55. Charles C. (Main)streaming hate: Analyzing White Supremacist Content and Framing Devices on YouTube. University of Central Florida; 2020. Available from: https://stars.library.ucf.edu/etd2020/27/

56. Clark S, Zaitsev A. Understanding YouTube Communities via Subscription-based Channel Embeddings. arXiv [cs.LG]. 2020 Oct 19. Available from: http://arxiv.org/abs/2010.09892

57. Hussain MN, Tokdemir S, Agarwal N, Al-Khateeb S. Analyzing disinformation and crowd manipulation tactics on YouTube. In: 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE; 2018. Available from: https://doi.org/10.1109/ASONAM.2018.8508766

58. Most viewed YouTube videos worldwide 2023. Statista. Available from: https://www.statista.com/statistics/249396/top-youtube-videos-views/

59. Wise J. How many videos are uploaded to YouTube A day in 2023? EarthWeb. 2023 May 24. Available from: https://earthweb.com/how-many-videos-are-uploaded-to-youtube-a-day/

60. Schindler P. Expanded safeguards for advertisers [Internet]. Google. 2017 March 21. Available from: https://blog.google/technology/ads/expanded-safeguards-for-advertisers/

61. YouTube Help. Advertiser-friendly content guidelines [Internet]. Google.com. 2023. Available from: https://support.google.com/youtube/answer/6162278?hl=en

62. Papadamou K, Zannettou S, Blackburn J, De Cristofaro E, Stringhini G, Sirivianos M. "It is just a flu": Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations [Internet]. arXiv [cs.CY]. 2021 Oct 12. Available from: http://arxiv.org/abs/2010.11638

63. Davidson J, Liebald B, Liu J, Nandy P, Van Vleet T, Gargi U, et al. The YouTube video recommendation system. In: Proceedings of the fourth ACM conference on Recommender systems. New York, NY, USA: ACM; 2010.

64. Covington P, Adams J, Sargin E. Deep neural networks for YouTube recommendations. In: Proceedings of the 10th ACM Conference on Recommender Systems. New York, NY, USA: ACM; 2016. Available from: https://doi.oeg/10.1145/2959100.2959190

65. Zuckerberg M. A blueprint for content governance and enforcement. 2018 Nov 15. Facebook. https://perma.cc/ZK5C-ZTSX?type=standard

66. Matamoros-Fernandez A, Gray EJ, Bartolo L, Burgess J, Suzor, N. What's 'up next'? Investigating algorithmic recommendations on YouTube across issues and over time. Media and Communication, 9(4), 234-249. 2021 Nov 18. Available from: https://doi.org/10.17645/mac.v9i4.4184

67. Allgaier J. Science and environmental communication on YouTube: Strategically distorted communications in online videos on climate change and climate engineering. Front Commun [Internet]. 2019;4. Available from: http://dx.doi.org/10.3389/fcomm.2019.00036

68. Song MY-J, Gruzd A. Examining sentiments and popularity of pro- and anti-vaccination videos on YouTube. In: Proceedings of the 8th International Conference on Social Media & Society - #SMSociety17. NewYork, New York, USA: ACM Press; 2017. Available from: https://doi.org/10.1145/3097286.3097303

69. Bessi A, Zollo F, Del Vicario M, Puliga M, Scala A, et al. Users Polarization on Facebook and Youtube. PLOS ONE 11(8): e0159641. 23 August 2016. Available from: https://doi.org/10.1371/journal.pone.0159641

70. Hussein E, Juneja P, Mitra T. Measuring misinformation in video search platforms: An audit study on YouTube. Proceedings of the ACM Human-Computer Interaction [Internet].4(CSCW1):1–27. 2020 May 29. Available from: http://dx.doi.org/10.1145/3392854

71. Roth C, Mazières A, Menezes T (2020). Tubes and bubbles topological confinement of YouTube recommendations. PLOS One, 15(4): e0231703. https://doi.org/10.1371/journal.pone.0231703

72. Airoldi M, Beraldo D, Gandini A. (2016). Follow the algorithm: An exploratory investigation of music on YouTube. Poetics, 57, 1–13. https://doi.org/10.1016/j.poetic.2016.05.001

73. Pariser E. The Filter Bubble. The Penguin Press. 2012 Apr 24. ISBN: 9780143121237.

74. Jamieson KH, Cappella JN. Echo Chambers: Rush Limbaugh and the Conservative Media Establishment. Oxford University Press. https://doi.org/ https://doi.org/10.5860/choice.47-0104

75. Sunstein CR. Echo Chambers: Bush v. Gore, impeachment, and beyond. Princeton, NJ: Princeton University Press. Available from: https://pup.princeton.edu/sunstein/echo.pdf

76. Cho J, Ahmed S, Hilbert M, Liu B, Luu J. Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization. J Broadcast Electron Media [Internet]. 2020;64(2):150–72. Available from: http://dx.doi.org/10.1080/08838151.2020.1757365

77. Hilbert M, Ahmed S, Cho J, Liu B, Luu J. Communicating with algorithms: A transfer entropy analysis of emotions-based escapes from online echo chambers. Commun Methods Meas [Internet]. 2018;12(4):260–75. Available from: http://dx.doi.org/10.1080/19312458.2018.1479843

78. Flaxman S, Goel S, Rao JM. Filter bubbles, echo chambers, and online news consumption. Public Opin Q [Internet]. 2016;80(S1):298–320. Available from: http://dx.doi.org/10.1093/poq/nfw006

79. Zhao Z, Hong L, Wei L, Chen J, Nath A, Andrews S, et al. Recommending what video to watch next: A multitask ranking system. In: Proceedings of the 13th ACM Conference on Recommender Systems. New York, NY, USA: ACM; 2019

80. Rieder B, Matamoros-Fernández A, Coromina Ò. From ranking algorithms to ‘ranking cultures’: Investigating the modulation of visibility in YouTube search results.

Convergence International Journal of Research into New Media Technologies. 2018;24(1):50–68. Available from: http://dx.doi.org/10.1177/1354856517736982

81. Kaiser J, Rauchfleisch A. (2020). Birds of a feather get recommended together: Algorithmic homophily in YouTube's channel recommendations in the United States and Germany. *Social Media + Society*, *6*(4), 1-15. https://doi.org/10.1177/2056305120969914
82. Geschke D, Lorenz J, Holtz P. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. Br J Soc Psychol [Internet]. 2019;58(1):129–49. Available from: http://dx.doi.org/10.1111/bjso.12286
83. Bruns A. Echo chamber? What echo chamber? Reviewing the evidence [Internet]. Snurb.info. 2017. Available from: https://snurb.info/files/2017/Echo%20Chamber.pdf
84. Plasticity Inc. Political disinformation campaign on YouTube – Plasticity [Internet]. Plasticity.ai. Available from: https://media.plasticity.ai/youtube-disinformation-report/
85. O'Sullivan, Donie. How fake accounts pushing inflammatory content went viral -- with the help of YouTube's algorithms. CNN. 2019 Dec 13. Available from: https://edition.cnn.com/2019/12/13/tech/youtube-fake-accounts-viral/index.html
86. Chris A. Top 10 search engines in the world (2023 update) [Internet]. reliablesoft.net. Reliablesoft Academy; 2017 Sept 10. Available from: https://www.reliablesoft.net/top-10-search-engines-in-the-world/
87. Securities and Exchange Commission. Amendment No. 7 to Forms S-1 Registration Statement under The Securities Act of 1933. Securities and Exchange Commission, Washington, D.C. 2004 Aug 13. Available from: https://www.sec.gov/Archives/edgar/data/1288776/000119312504139655/ds1a.htm#toc59330_25b

88. Morris B. The New Tech Avengers: An unlikely triumvirate of Silicon Valley insiders is holding the industry accountable on privacy and addiction. Wall Street Journal. 2018 Jun 29. Available from: https://www.wsj.com/articles/the-new-tech-avengers-1530285064

89. Lewis P. ‘Our minds can be hijacked’: the tech insiders who fear a smartphone dystopia: Google, Twitter and Facebook workers who helped make technology so addictive are disconnecting themselves from the internet. The Guardian. 2016 October 6. Available from: https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia

90. Moghavvemi S, Sulaiman AB, Jaafar NIB, Kasem N. Facebook and YouTube addiction: The usage pattern of Malaysian students. 2017 July. *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, 1–6. http://dx.doi.org/10.1109/ICRIIS.2017.8002516

91. Balakrishnan J, Griffiths MD. Social media addiction: What is the role of content in YouTube? 2017 September 14. *Journal of Behavioral Addictions*, *6*(3), 364–377. https://doi.org/10.1556/2006.6.2017.058

92. YouTube. Autoplay. YouTube.com. Available from: https://www.youtube.com/howyoutubeworks/user-settings/autoplay/

93. YouTube Help. Discovery and performance FAQs [Internet]. Google.com. 2023. Available from: https://support.google.com/youtube/answer/141805?hl%84=%84en

94. Guan Z, Cutrell E. 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. New York, NY, USA: ACM; 2007. Available from: https://doi.org/10.1145/1240624.1240691

95. Pan B, Hembrooke H, Joachims T, Lorigo L, Gay G, Granka L. In Google we trust: Users’ decisions on rank, position, and relevance. J Comput Mediat Commun [Internet]. 2007;12(3):801–23. Available from: http://dx.doi.org/10.1111/j.1083-6101.2007.00351.x

96. Lewis P. "Fiction is outperforming reality": how YouTube's algorithm distorts truth. The guardian [Internet]. 2018 Feb 2; Available from: https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth

97. Yang S, Brossard D, Scheufele DA, Xenos MA. The science of YouTube: What factors influence user engagement with online science videos? PLoS One. 2022;17(5):e0267697. Available from: http://dx.doi.org/10.1371/journal.pone.0267697

98. Meyerson E. YouTube now: Why we focus on watch time [Internet]. blog.youtube. YouTube Official Blog; 2012. Available from: https://blog.youtube/news-and-events/youtube-now-why-we-focus-on-watch-time/

99. Solsman JE. YouTube's AI is the puppet master over most of what you watch [Internet]. CNET. 2018 Jan 10. Available from: https://www.cnet.com/tech/services-and-software/youtube-ces-2018-neal-mohan/

100. Warzel C, Smidt R. YouTubers made hundreds of thousands off of bizarre and disturbing child content [Internet]. BuzzFeed. 2017. Available from: https://www.buzzfeed.com/charliewarzel/youtubers-made-hundreds-of-thousands-off-of-bizarre-and?s=mobile_app&utm_term=.ed3vJl8Kw

101. Fredenburg JN. YouTube as an ally of convenience: The platform's building and breaking with the LGBTQ+ community. Digital Georgetown. Georgetown University; 2020. Available from: http://hdl.handle.net/10822/1059448

102. YouTube channel monetization policies [Internet]. Google.com. Available from: https://support.google.com/youtube/answer/1311392?hl=en&ref_topic=9153642

103. The YouTube Team. YouTube creator blog. YouTube Creator Blog. 2019 Aug 27. Available from: https://youtube-creators.googleblog.com/2019/08/

104. Kumar S. The algorithmic dance: YouTube's Adpocalypse and the gatekeeping of cultural content on digital platforms. Internet Policy Review [Internet]. 2019;8(2). Available from: http://dx.doi.org/10.14763/2019.2.1417

105. Rieder B, Coromina Ò, Matamoros-Fernández, A. (2020). Mapping YouTube. *First Monday*. https://doi.org/10.5210/fm.v25i8.10667

106. PragerU. PragerU takes legal action against Google and YouTube for discrimination. Prageru.com. Available from: https://www.prageru.com/press-release-prager-university-prageru-takes-legal-action-against-google-and-youtube-discrimination

107. Neidig, Harper. PragerU sues Google, YouTube for 'censoring' conservative videos. The Hill. 2017 October 24. Available from: http://thehill.com/policy/technology/356966-prageru-sues-google-youtube-for-censoring-conservative-videos

108. Epstein R. The New Censorship: How did Google become the internet's censor and master manipulator, blocking access to millions of websites? Usnews.com. 2016 June 22. Available from: https://www.usnews.com/opinion/articles/2016-06-22/google-is-the-worlds-biggest-censor-and-its-power-must-be-regulated

109. The YouTube Team. The Four Rs of Responsibility, Part 2: Raising authoritative content and reducing borderline content and harmful misinformation [Internet]. blog.youtube. YouTube Official Blog; 2019 Dec 3. Available from: https://blog.youtube/inside-youtube/the-four-rs-of-responsibility-raise-and-reduce/

110. YouTube Help. Troubleshoot video takedowns [Internet]. Google.com. 2023. Available from: https://support.google.com/youtube/answer/6395024?hl=en

111. Yourieff, Kaya. (2018, Apr 23). YouTube took down more than 8 million videos in 3 months. CNN. http://money.cnn.com/2018/04/23/technology/youtube-community-guidelines-report/index.html

112. Cruz T. CBS News Uncovers More Big Tech Bias. Ted Cruz. 2019 Dec 3. Available from: https://www.cruz.senate.gov/?p=press_release&id=4795

113. Rankovic, Didi. Leaked video shows YouTube CEO talking news manipulation, deciding what's "authoritative" and what's "trashy." Reclaim the Net. 2019 August 17. Available from: https://reclaimthenet.org/youtube-ceo-susan-wojcicki-news-bias

114. Graziosi, G. YouTube removes Robert F Kennedy Jr video featuring bizarre claim that polluted water makes children transgender. Independent. 2023 June 19. Available from: https://www.independent.co.uk/news/world/americas/us-politics/youtube-removes-rfk-conspiracy-video-b2360440.html

115. Duffy, C. YouTube removed video of Robert F. Kennedy, Jr. for violating vaccine misinformation policy. CNN. 2023 June 20. Available from: https://www.cnn.com/2023/06/20/tech/youtube-robert-f-kennedy-jr-video-removed/index.html

116. Studio Themes. E-ventures Worldwide LLC vs Google - Google wins the right to de-index. Velocity. 2017. Available from: https://www.velsof.com/blog/ecommerce/e-ventures-worldwide-llc-vs-google-google-wins-the-right-to-de-index

117. UNITED STATES DISTRICT COURT MIDDLE DISTRICT OF FLORIDA FORT MYERS DIVISION. e-ventures Worldwide, LLC Plaintiff, v. Google, Inc., Defendant. MEMORANDUM AND ORDER. 2017. Available from: https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=2410&context=historical

118. Supreme Court of the United States. Reynaldo Gonzalez, et al., Petitioners v. Google LLC, Respondent. Supremecourt.gov. 2023. Available from: https://www.supremecourt.gov/DocketPDF/21/21-1333/254251/20230207191257813_GonzalezRepyMeritsPrinted.pdf

119. McKinnon JD. Google says supreme court ruling could potentially upend the internet. The Wall Street journal (Eastern ed) [Internet]. 2023 Jan 12; Available from:

https://www.wsj.com/articles/google-says-supreme-court-ruling-could-potentially-upend-the-internet-11673553968

120. Lutz M, Gadaginmath S, Vairavan N, Mui P. Examining political bias within YouTube search and recommendation algorithms. In: 2021 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE; 2021. Available from: https://doi.org/10.1109/SSCI50451.2021.9660012

121. Shevtsov A, Oikonomidou M, Antonakaki D, Pratikakis P, Ioannidis S. What Tweets and YouTube comments have in common? Sentiment and graph analysis on data related to US elections 2020. 2023 January 31. PLOS ONE 18(1): e0270542. https://doi.org/10.1371/journal.pone.0270542

122. Schwind C, Buder J. Reducing confirmation bias and evaluation bias: When are preference-inconsistent recommendations effective – and when not? Comput Human Behav [Internet]. 2012;28(6):2280–90. Available from: http://dx.doi.org/10.1016/j.chb.2012.06.035

123. Agudo U, Matute H. The influence of algorithms on political and dating decisions. PLOS One, 16(4): e0249454. 2021 April 21. Available from: https://doi.org/10.1371/journal.pone.0249454

124. Wittenberg C, Tappin BM, Berinsky AJ, Rand DG. The (minimal) persuasive advantage of political video over text. Proceedings of the National Academy of Sciences of the United States of America. 2021;118(47):e2114388118. Available from: http://dx.doi.org/10.1073/pnas.2114388118

125. Andreoli V, Worchel S. (1978). Effects of Media, Communicator, and Message Position on Attitude Change. Public Opinion Quarterly, 42(1), 59; Available from: https://doi.org/10.1086/268429

126. Chaiken S, Eagly AH. (1976). Communication modality as a determinant of message persuasiveness and message comprehensibility. Journal of Personality and Social Psychology, 34(4), 605–614. Available from: https://doi.org/10.1037/0022-3514.34.4.605

127. Chaiken S, Eagly AH. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. Journal of Personality and Social Psychology, 45(2), 241–256. Available from https://doi.org/10.1037/0022-3514.45.2.241

128. Horton D, Richard Wohl R. Mass communication and para-social interaction: Observations on intimacy at a distance. Psychiatry [Internet]. 1956;19(3):215–29. Available from: http://dx.doi.org/10.1080/00332747.1956.11023049

129. Pfau M, Holbert RL, Zubric SJ, Pasha NH, Lin WK. Role and Influence of Communication Modality in the Process of Resistance to Persuasion. Media Psychology, 2(1), 1–33. 2009 Nov 17.Available from: https://doi.org/10.1207/S1532785XMEP0201_1

130. Weber C, Dunaway J, Johnson T. It's All in the Name: Source Cue Ambiguity and the Persuasive Appeal of Campaign Ads. 34(3), 561–584. 2011 June 2. Available from: https://doi.org/10.1007/s11109-011-9172-y

131. Witus LS, Larson E. A randomized controlled trial of a video intervention shows evidence of increasing COVID-19 vaccination intention. 2022 May 19. PLOS ONE 17(5): e0267580. https://doi.org/10.1371/journal.pone.0267580

132. Arruda W. Why you need to excel at video [Internet]. Forbes. 2016. Available from: https://www.forbes.com/sites/williamarruda/2016/06/29/why-you-need-to-excel-at-video/

133. Wyzowl. Video marketing statistics 2023. Wyzowl. Available from: https://www.wyzowl.com/video-marketing-statistics/

134. Hong S, Kim N. Will the Internet promote democracy? Search engines, concentration of online news readership, and e-democracy. Journal of Information Technology & Politics, 15(4), 388-399. 2018 October 16. Available from: http://dx.doi.org/10.2139/ssrn.3267049

135. Epstein R. Taming Big Tech: The case for monitoring. Hacker Noon. 2018 May 17. Available from: https://hackernoon.com/taming-big-tech-5fef0df0f00d

136. Epstein R. Ten ways Big Tech can shift millions of votes in the November elections—without anyone knowing. Epoch Times. 2018 September 26. Available from: https://www.theepochtimes.com/10-ways-bigtech-can-shift-millions-of-votes-in-the-november-elections-without-anyoneknowing_2671195.html

137. Epstein R. The unprecedented power of digital platforms to control opinions and votes. In G. Rolnik (Ed.), Digital platforms and concentration: Second annual antitrust and competition conference (pp. 31-33). Chicago, IL: University of Chicago Booth School of Business; 2018. https://promarket.org/wp-content/uploads/2018/04/Digital-Platforms-and-Concentration.pdf

138. Epstein R, Bock S, Peirson, L, Wang H. Large-scale monitoring of Big Tech political manipulations in the 2020 Presidential election and 2021 Senate runoffs, and why monitoring is essential for democracy. Paper presented at the 24th annual meeting of the American Association of Behavioral and Social Sciences (AABSS). 2021 June 14. Available from: https://aibrt.org/downloads/EPSTEIN_et_al_2021-Large-Scale_Monitoring_of_Big_Tech_Political_Manipulations-FINAL_w_AUDIO.mp4

139. Epstein R, Bock S, Peirson L, Wang H, Voillot M. How we preserved more than 1.5 million online "ephemeral experiences" in the recent US elections, and what this content revealed about online election bias. Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. 2022 Apr. Available from: https://aibrt.org/downloads/EPSTEIN_et_al_2022-WPAHow_We_Preserved_More_Than_1.5_Million_Online_Ephemeral_Experiences_in_Recent_US_ Elections...pdf

140. Brandeis LD. Other People's Money, and How the Bankers Use it. 1st ed. Kentucky: Louis D. ho School of Law at the Univeristy of Louisville; 1914.

141. Epstein R, Peirson L. How we preserved more than 2.4 million online ephemeral experiences in the 2022 midterm elections, and what this content revealed about online election bias. Paper to be presented at the 103rd annual meeting of the Western Psychological Association; 2023 Apr; Riverside, CA. Available from: https://aibrt.org/downloads/EPSTEIN-Submitted_2022-How_We_Preserved_More_Than_2.4_Million_Online_Ephemeral_Experiences_in_the_Midterm_Elections...pdf

142. Epstein R. How Google stopped the Red Wave. The Epoch Times. 2022 Nov 15. Available from: https://ept.ms/GoogleRedWave (forwarding link: https://TheMidtermsWereRigged.com)

143. Epstein R. EPSTEIN: Google is shifting votes on a massive scale, but a solution is at hand. Daily Caller. 2022 Nov 6. Available from: https://dailycaller.com/2022/11/06/robert-epstein-2022-midtermelections-google-bing/

144. https://www.algotransparency.org/

145. Lewis P, McCormick E. How an ex-YouTube insider investigated its secret algorithm. The Guardian. 2018 Feb 2. Available from: https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot

146. Rieder B, Hofmann, J. (2020). Towards platform observability. Internet Policy Review, 9(4). https://doi.org/10.14763/2020.4.1535

147. Agarwal S, Sureka A. Applying social media intelligence for predicting and identifying on-line radicalization and civil unrest oriented threats [Internet]. arXiv [cs.CY]. 2015 [cited 2023 Jul 3]. Available from: http://arxiv.org/abs/1511.06858

148. Agarwal N, Gupta R, Singh SK, Saxena V. Metadata based multi-labelling of YouTube videos. In: 2017 7th International Conference on Cloud Computing, Data

Science & Engineering - Confluence. IEEE; 2017. Available from: http://doi.org/10.1109/CONFLUENCE.2017.7943219

149. Chaudhary K, Alam M, Al-Rakhami MS, Gumaei A. Machine learning-based mathematical modelling for prediction of social media consumer behavior using Big Data Analytics. Journal of Big Data. 2021;8(1). https://doi.org/10.1186/s40537-021-00466-2

150. Chaudhary K, Alam M, Al-Rakhami MS, Gumaei A. Machine learning-based mathematical modelling for prediction of social media consumer behavior using Big Data Analytics. Journal of Big Data. 2021;8(1). https://doi.org/10.1186/s40537-021-00466-2

151. YouTube is self-programing

152. Dylko I, Dolgov I, Hoffman W, Eckhart N, Molina M, Aaziz O. The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. Comput Human Behav [Internet]. 2017;73:181–90. Available from: https://doi.org/10.1016/j.chb.2017.03.031

153. Bai X, Cambazoglu BB, Gullo F, Mantrach A, Silvestri, F. Exploiting search history of users for news personalization. Information Sciences, 385-386, 125–137. 2017 April. Available from: https://doi.org/10.1016/j.ins.2016.12.038

154. Bozdag, E. Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209–227. 2013 June 23. Available from: https://doi.org/10.1007/s10676-013-9321-6

155. Gravetter FJ, Wallnau LB. Essentials of statistics for the behavioral sciences. 8th ed. Wadsworth Cengage Learning; 2013

156. Lord FM. On the statistical treatment of football numbers. *American Psychologist, 8*(12), 750-751. 1953. Available from: https://doi.org/10.1037/h0063675

157. Townsend JT, Ashby FG. Measurement scales and statistics: The misconception misconceived. Psychological Bulletin, 96(2), 394-401. 1984. Available from: https://doi.org/10.1037/0033-2909.96.2.394

158. Maccoby EE. Parenting and its effects on children: on reading and misreading behavior genetics. Annual Review of Psychology. 2000;51(1):1–27. Available from: http://dx.doi.org/10.1146/annurev.psych.51.1.1

159. Kaushal N. Top YouTube searches. PageTraffic Blog - All thing Search, Content & Social. PageTraffic; 2023. Available from: https://www.pagetraffic.com/blog/top-youtube-searches/

160. Levitt M. The heyday of television ads is over. Political campaigns ought to act like it. The Washington Post. 2018 November 6. Available from: https://www.washingtonpost.com/outlook/2018/11/06/heyday-television-ads-is-over-political-campaigns-ought-act-like-it/

161. YouTube. How YouTube protects data privacy - How YouTube Works. YouTube.com; [cited 2023 Jul 5]. Available from: https://www.youtube.com/howyoutubeworks/our-commitments/protecting-user-data/

162. Charles A. Can YouTube track viewers? Chron.com; 2013. Available from: https://smallbusiness.chron.com/can-youtube-track-viewers-63519.html

163. Brody B, Bergen M. Google to Pay $170 Million for YouTube Child Privacy Breaches. Bloomberg. 2019 Sept 4. Available from: https://www.bloomberg.com/news/articles/2019-09-04/google-to-pay-170-million-for-youtube-child-privacy-breaches

Supporting information Captions

S1 Text. Candidate biographies.

S2 Text. Instructions immediately preceding DoodleTube simulation

S3 Text. Vote Manipulation Power (VMP) calculation.

S1 Table. Experiments 1&2: Demographics.

S2 Table. Experiments 1&2: VMPs by gender.

S3 Table. Experiments 1&2: VMPs by race/ethnicity.

S4 Table. Experiments 1&2: VMPs by educational attainment.

S5 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the three groups by gender.

S6 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the bias groups (1&2) by gender.

S7 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the favored candidate by gender.

S8 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the three groups by educational attainment.

S9 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for Groups 1&2 by educational attainment.

S10 Table. Experiments 1&2: Mean preference for favored candidate on the 11-point scale of voting preference by educational attainment.

S11 Table. Experiments 1&2: Mean ratings in the three groups on the 11-point scale of voting preference by race/ethnicity.

S12 Table. Experiments 1&2: Mean preference shift for Groups 1&2 on the 11-point scale of voting preference by race/ethnicity.

S13 Table. Experiments 1&2: Mean preference for favored candidate on the 11-point scale of voting preference by race/ethnicity.

S14 Table. Experiment 1: Pre and Post opinion ratings of favored and non-favored candidates by race/ethnicity.

S15 Table. Experiment 2: Pre and Post opinion ratings of favored and non-favored candidates by race/ethnicity.

S1 Fig. Before January, 2015, a switch that allowed users to deactivate YouTube's autoplay feature was prominently shown and labeled in the upper-right of screens (on laptop and desktop computers). Note that autoplay was always on by default. Compare S2 Fig.

S2 Fig. Beginning in January, 2015, the autoplay switch was moved to a position immediately below the video, and its label was removed. It was, as always, on by default. Compare S1 Fig.

S3 Fig. In 2019, a 2 min video leaked from Google in which YouTube CEO Susan Wojcicki explained to her staff how YouTube's recommender algorithm was being revised to boost certain content and demote other content. The full video can be viewed at https://vimeo.com/354354050**.**

Supporting Information

S1 Text. Candidate Biographies

Scott Morrison:

Scott Morrison was born in Waverley, New South Wales (AUS) on May 13th, 1968. He completed a Bachelor of Science honors degree in applied economic geography at the University of New South Wales. Morrison married his high school sweetheart, Jenny Warren, in 1990 and has two daughters. After graduating from the University of New South Wales, Morrison worked as a national policy and research manager for the Property Council of Australia before moving to New Zealand in 1998 to become the director of the Office of Tourism and Sport. He left this position a year before the contract schedule and returned to Australia in 2000. In 2004, he became the inaugural managing director of Tourism Australia until July 2006.

Bill Shorten:

Bill Shorten was born in Fitzroy, Victoria (AUS) on May 12th, 1967. While Shorten was studying at Monash University, he was an active student in the university's politics club. In 1986, Shorten helped establish a group called Network and briefly served as a private in the Australian Army Reserve from 1985 to 1986. After graduating Monash University with a Bachelors of Arts in 1989 and a Bachelors of Law in 1992, Shorten worked as a lawyer for Maurice Blackburn Cashman for twenty months. In 1994, he worked as a trainee organizer and later accepted a position as a politics national secretary in 2001 and again in 2005. Shorten is currently married to Chloe Bryce and has a daughter.

S2 Text. Instructions immediately preceding DoodleTube simulation

Thank you for your answers!

You will now be given an opportunity to learn more about these candidates using our special internet video platform called “DoodleTube.” Your goal is to try to clarify your views on the topic so you are better able to decide which candidate deserves your vote.

Use DoodleTube as you would normally use YouTube, and please do NOT use any other web pages to learn more about either candidate. In other words, please do not leave DoodleTube! If you do, that will invalidate your participation in our study. If you would like to conduct further research on the topic, go right ahead, but please complete our study first!

You will have a total of **15:00** minutes to view the videos, 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 your participation in the study. Instead, if you feel you have enough information to make a clear choice, you may end your search early by clicking the “END” button in the upper-left corner of the DoodleTube page, which will appear after 10:00 minutes.

PLEASE NOTE: Some web pages might take a while to load, so please be patient.

Click the ‘Continue’ button below.

S3 Text. Vote Manipulation Power (VMP) calculation.

Vote Manipulation Power (VMP) is calculated as follows:

$$\frac{p' - p}{p}$$

where p is the total number of people who voted for the favored candidate pre-manipulation, and p' is the total number of people who voted for the favored candidate post-manipulation. If, pre-manipulation, a group of 100 people is split 50/50 in the votes they give us, and if, post-manipulation, a total of 67 people now vote for the favored candidate, the VMP is

$$\frac{67 - 50}{50}$$

or 34%. Because p' is 17 points larger than p, the win margin is 34 (2 x 17, or 34%), and the final vote is 67 to 33, with the favored candidate the winner. So in any group in which the vote is split 50/50 pre-manipulation, the VMP is also the win margin. Note that 17 individuals did not need to *shift* to produce this win margin. We only needed the *net* number of people voting for the favored candidate to be 67.

S1 Table. Experiments 1&2: Demographics

	Experiment 1 (*n* = 959)	**Experiment 2 (*n* = 491)**
Mean Age (*SD*)	38.6 (12.0)	36.4 (11.7)
Gender (%)		
Female	558 (58.2%)	270 (55.0%)
Male	391 (40.8%)	218 (44.4%)
Unknown	10 (1.0%)	3 (0.6%)
Education (%)		
None	2 (0.2%)	0 (0.0%)
Primary	41 (4.3%)	20 (4.1%)
Secondary	309 (32.2%)	160 (32.6%)
Bachelors	428 (44.6%)	224 (45.6%)
Masters	148 (15.4%)	72 (14.7%)
Doctorate	31 (3.2%)	15 (3.1%)
Race/Ethnicity (%)		
White	701 (73.1%)	351 (71.5%)
Black	102 (10.6%)	55 (11.2%)
Asian	75 (7.8%)	39 (7.9%)
Mixed	58 (6.0%)	31 (6.3%)
Other	22 (2.3%)	15 (3.1%)

S2 Table. Experiments 1&2: VMPs by gender.

Condition		***n***	**VMP (%)**	**Bias (%)**
E1: No Mask	Male	268	40.6	33.6
	Female	376	58.9	32.4
	Change (%)	-	+45.1	-3.6
	Statistic (*z*)	-	-4.58	0.32
	p	-	< 0.001	0.749 NS
E2: Mask 2&3	Male	154	52.6	14.3
	Female	180	77.9	15.0
	Change (%)	-	+48.1	+4.9
	Statistic (*z*)	-	-4.87	-0.18
	p	-	< 0.001	0.857 NS

S3 Table. Experiments 1&2: VMPs by race/ethnicity.

Condition		***n***	**VMP (%)**	**Bias (%)**
E1: No Mask	White	477	51.9	34.8
	Non-White	174	50.5	28.2
	Change (%)	-	-2.7	-19.0
	Statistic (*z*)	-	0.32	-1.58
	p	-	0.749 NS	0.114 NS
E2: Mask 2&3	White	235	65.5	14.9
	Non-White	101	65.9	13.9
	Change (%)	-	+0.6	-6.7
	Statistic (*z*)	-	-0.07	0.24
	p	-	0.944 NS	0.810 NS

S4 Table. Experiments 1&2: VMPs by educational attainment.

Condition		***n***	**VMP (%)**	**Bias (%)**
E1: No Mask	≥ Bachelors	413	48.4	33.7
	< Bachelors	238	56.6	31.9
	Change (%)	-	+17.0	-5.3
	Statistic (*z*)	-	-2.02	0.47
	p	-	< 0.05	0.638 NS
Experiment 2: Mask 2&3	≥ Bachelors	224	63.7	15.6
	< Bachelors	112	72.5	12.5
	Change (%)	-	+13.8	-19.9
	Statistic (*z*)	-	-1.61	0.76
	p	-	0.107 NS	0.447 NS

S5 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the three groups by gender.

Condition		***n***		M_{Morrison} **(SD)**	M_{Shorten} **(SD)**	M_{Control} **(SD)**	***H***	***p***
E1: No Mask	Male	391	Pre	-0.31 (2.54)	-0.09 (2.79)	0.01 (2.41)	1.099	0.577 NS
			Post	-1.06 (3.30)	1.80 (3.31)	0.32 (3.43)	42.623	< 0.001
	Female	558	Pre	-0.15 (2.90)	-0.16 (3.01)	0.12 (3.01)	0.991	0.609 NS
			Post	-2.48 (2.95)	2.35 (2.98)	0.62 (3.58)	152.927	< 0.001
E2: Mask 2&3	Male	218	Pre	-0.44 (2.56)	-0.11 (2.74)	-0.69 (2.90)	1.49	0.474 NS
			Post	-1.57 (3.06)	1.93 (3.13)	0.31 (3.35)	37.086	< 0.001
	Female	270	Pre	0.18 (3.04)	-0.19 (2.60)	0.04 (2.86)	0.764	0.682 NS
			Post	-2.09 (3.37)	2.09 (2.66)	0.74 (3.66)	56.572	< 0.001

S6 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the bias groups (1&2) by gender.

Condition		***n***	**Group 1 Shift**	**Group 2 Shift**
E1: No Mask	Male	268	0.75	1.89
	Female	376	2.33	2.41
	Change (%)	-	+210.7	+27.5
	U	-	9860.500	11413.5
	p	-	< 0.001	0.371 NS
E2: Mask 2&3	Male	154	1.13	2.04
	Female	180	2.27	2.28
	Change (%)	-	+100.9	+11.8
	U	-	3466	2859
	p	-	0.089 NS	0.834 NS

S7 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the favored candidate by gender.

Condition		***n***	**M_{Pre} (SD)**	**M_{Post} (SD)**	**Diff**	***z***	***p***
E1: No Mask	Male	268	0.12 (2.67)	1.41 (3.32)	1.29	-5.434	< 0.001
	Female	376	-0.01 (2.96)	2.41 (2.96)	2.42	-10.896	< 0.001
	Change (%)	-	-	-	+87.6	-	-
	Whitney U	-	-	-	42140.5	-	-
	p	-	-	-	< 0.001	-	-
E2: Mask 2&3	Male	154	0.18 (2.65)	1.74 (3.09)	1.56	-5.119	< 0.001
	Female	180	-0.18 (2.84)	2.09 (3.06)	2.27	-7.447	< 0.001
	Change (%)	-	-	-	+45.5	-	-
	U	-	-	-	12851.5	-	-
	p	-	-	-	0.249 NS	-	-

S8 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for the three groups by educational attainment.

Condition		***n***		**$M_{Morrison}$ (SD)**	**$M_{Shorten}$ (SD)**	**$M_{Control}$ (SD)**	***H***	***p***
E1: No Mask	≥ Bachelors	607	Pre	-0.30 (2.68)	-0.10 (2.96)	0.04 (2.67)	1.531	0.465 NS
			Post	-1.99 (3.12)	2.18 (3.14)	0.37 (3.47)	131.140	< 0.001
	< Bachelors	352	Pre	-0.09 (2.86)	-0.24 (2.80)	0.05 (3.00)	0.694	0.707 NS
			Post	-1.63 (3.31)	2.06 (3.05)	0.61 (3.63)	58.267	< 0.001
E2: Mask 2&3	≥ Bachelors	311	Pre	-0.05 (2.72)	-0.21 (2.69)	-0.39 (2.80)	0.780	0.677 NS
			Post	-1.53 (3.33)	1.92 (3.00)	0.29 (3.63)	46.178	< 0.001
	< Bachelors	180	Pre	-0.15 (3.11)	-0.04 (2.56)	-0.07 (2.99)	0.055	0.973 NS
			Post	-2.59 (2.93)	2.25 (2.65)	0.97 (3.38)	54.058	< 0.001

S9 Table. Experiments 1&2: Mean ratings on the 11-point scale of voting preference for Groups 1&2 by educational attainment.

Condition		***n***	**Group 1 Shift**	**Group 2 Shift**
E1: No Mask	≥ Bachelors	413	1.69	2.28
	< Bachelors	238	1.54	2.30
	Change (%)	-	-8.9	+0.9
	U	-	12434.5	11413.5
	p	-	0.749 NS	0.998 NS
E2: Mask 2&3	≥ Bachelors	224	1.48	2.13
	< Bachelors	112	2.44	2.29
	Change (%)	-	+64.9	+7.5
	U	-	3115	2563.5
	p	-	0.120 NS	0.666 NS

S10 Table. Experiments 1&2: Mean preference for favored candidate on the 11-point scale of voting preference by educational attainment.

Condition		***n***	M_{Pre} **(SD)**	M_{Post} **(SD)**	**Diff**	$z^{\dagger}$	***p***
E1: No Mask	≥ Bachelors	413	0.09 (2.83)	2.09 (3.13)	2.00	-9.836	< 0.001
	< Bachelors	238	-0.07 (2.83)	1.83 (3.19)	1.90	-6.897	< 0.001
	Change (%)	-	-	-	-5.0	-	-
	U	-	-	-	48484.5	-	-
	P				0.773 NS		
E2: Mask 2&3	≥ Bachelors	224	-0.07 (2.70)	1.71 (3.18)	1.78	-6.744	< 0.001
	< Bachelors	112	0.06 (2.86)	2.43 (2.80)	2.37	-6.064	< 0.001
	Change (%)	-	-	-	+33.1	-	-
	U	-	-	-	11331	-	-
	p	-	-	-	0.146 NS	-	-

S11 Table. Experiments 1&2: Mean ratings in the three groups on the 11-point scale of voting preference by race/ethnicity.

Condition		***n***		M_{Morrison} **(SD)**	M_{Shorten} **(SD)**	M_{Control} **(SD)**	***H***	***p***
E1: No Mask	White	701	Pre	-0.30 (2.75)	-0.32 (2.92)	-0.05 (2.76)	1.454	0.483 NS
			Post	-1.90 (3.23)	1.98 (3.14)	0.51 (3.52)	133.499	< 0.001
	Non-White	258	Pre	-0.01 (2.75)	0.35 (2.80)	0.29 (2.88)	0.793	0.673 NS
			Post	-1.70 (3.09)	2.61 (2.97)	0.30 (3.57)	56.318	< 0.001
E2: Mask 2&3	White	351	Pre	-0.22 (2.82)	0.06 (2.69)	-0.26 (2.86)	0.729	0.695 NS
			Post	-2.15 (3.05)	2.11 (2.93)	0.49 (3.55)	78.349	< 0.001
	Non-White	140	Pre	0.27 (2.88)	-0.59 (2.52)	-0.23 (2.98)	2.247	0.325 NS
			Post	-1.19 (3.59)	1.86 (2.79)	0.87 (3.51)	17.583	< 0.001

S12 Table. Experiments 1&2: Mean preference shift for Groups 1&2 on the 11-point scale of voting preference by race/ethnicity.

Condition		***n***	**Group 1 Shift**	**Group 2 Shift**
E1: No Mask	White	477	1.60	2.30
	Non-White	174	1.69	2.26
	Change (%)	-	+5.6	-1.7
	U	-	10594.5	10037
	p	-	0.926 NS	0.953 NS
E2: Mask 2&3	White	235	1.93	2.05
	Non-White	101	0.92	1.27
	Change (%)	-	-52.3	-38.0
	U	-	3082.5	2563.5
	p	-	0.351 NS	0.858 NS

S13 Table. Experiments 1&2: Mean preference for favored candidate on the 11-point scale of voting preference by race/ethnicity.

Condition		***n***	**M_{Pre} (SD)**	**M_{Post} (SD)**	**Diff**	***z***	***p***
E1: No Mask	White	477	-0.01 (2.85)	2.94 (3.19)	2.95	-10.074	< 0.001
	Non-White	174	0.17 (2.77)	2.14 (3.06)	1.97	-6.609	< 0.001
	Change (%)	-	-	-	-34.2	-	-
	U	-	-	-	41167	-	-
	p	-	-	-	0.875 NS	-	-
E2: Mask 2&3	White	235	0.15 (2.76)	2.13 (2.99)	1.98	-7.403	< 0.001
	Non-White	101	-0.43 (2.71)	1.51 (3.22)	1.94	-5.247	< 0.001
	Change (%)	-	-	-	-2.02	-	-
	U	-	-	-	11422.5	-	-
	p	-	-	-	0.584 NS	-	-

S14 Table. Experiment 1: Pre and Post opinion ratings of favored and non-favored candidates by race/ethnicity.

Ethnicity		**Favored Candidate Mean (SD)**			**Non-Favored Candidate Mean (SD)**			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†**
White	Impression	6.97 (1.88)	7.37 (2.39)	0.40	6.95 (1.94)	4.64 (2.41)	-2.31	-12.080***
	Trust	6.16 (2.12)	6.67 (2.57)	0.51	6.22 (2.12)	4.36 (2.36)	-1.86	-11.181***
	Likeability	6.91 (1.85)	7.42 (2.47)	0.51	6.96 (1.98)	4.62 (2.49)	-2.34	-12.647***
Non-White	Impression	7.36 (1.85)	7.55 (2.24)	0.19	7.17 (1.91)	5.29 (2.49)	-1.88	-6.336***
	Trust	6.23 (1.93)	6.80 (2.38)	0.57	6.22 (1.94)	4.97 (2.42)	-1.25	-6.324***
	Likeability	7.17 (1.84)	7.46 (2.23)	0.29	6.93 (1.90)	5.31 (2.44)	-1.62	-6.279***

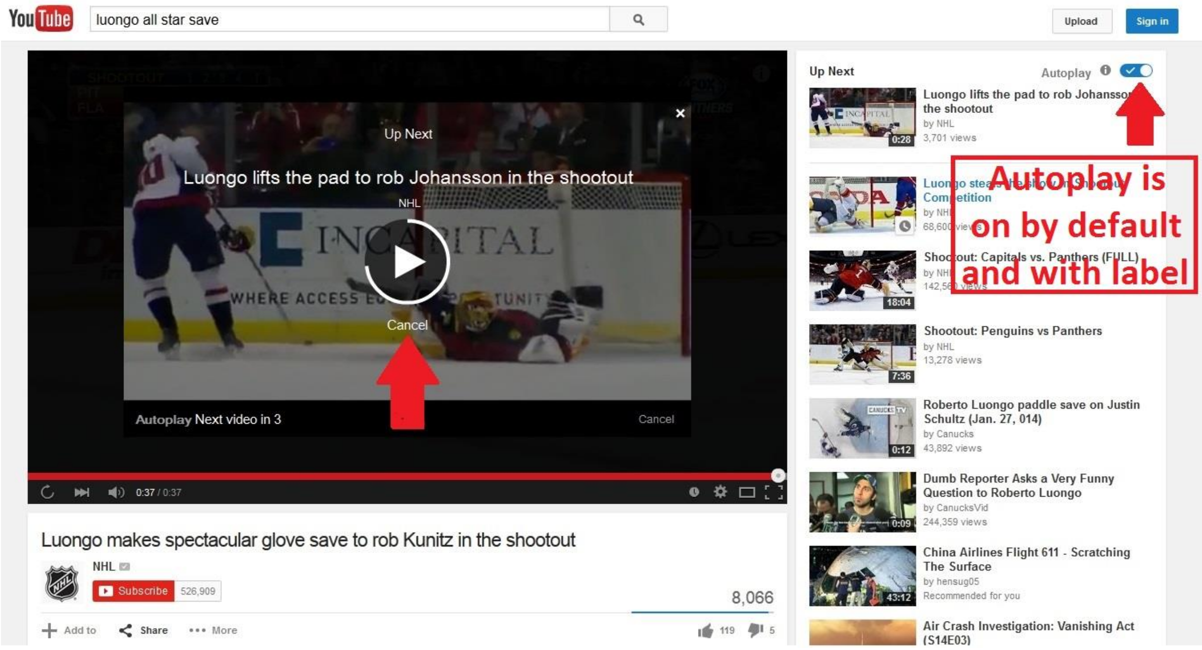
†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate
*** $p < 0.001$

S15 Table. Experiment 2: Pre and Post opinion ratings of favored and non-favored candidates by race/ethnicity.

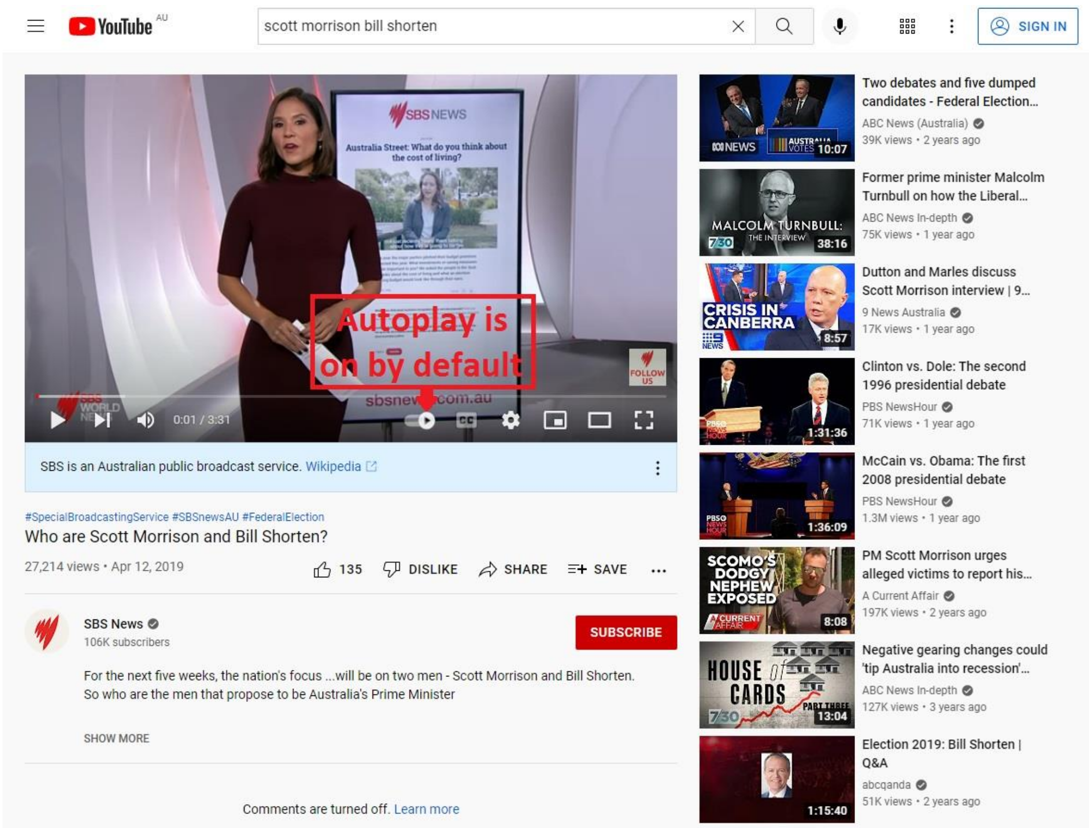
Ethnicity		**Favored Candidate Mean (SD)**			**Non-Favored Candidate Mean (SD)**			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	***z*†**
White	Impression	6.99 (1.70)	7.43 (2.18)	0.44	6.97 (1.74)	4.89 (2.35)	-2.08	-8.233***
	Trust	6.26 (1.90)	6.83 (2.24)	0.57	6.17 (1.90)	4.72 (2.39)	-1.45	-7.429***
	Likeability	6.94 (1.75)	7.67 (2.12)	0.75	6.92 (1.68)	4.98 (2.44)	-1.94	-8.549***
Non-White	Impression	6.73 (1.95)	7.00 (2.20)	0.27	6.94 (1.94)	4.88 (2.43)	-2.06	-5.343***
	Trust	5.66 (2.20)	6.38 (2.42)	0.72	6.00 (2.31)	4.47 (2.36)	-1.53	-5.217***
	Likeability	6.68 (1.93)	7.08 (2.27)	0.40	6.75 (2.00)	5.01 (2.45)	-1.74	-4.992***

†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate
*** $p < 0.001$

S1 Fig. Before January, 2015, a switch that allowed users to deactivate YouTube's autoplay feature was prominently shown and labeled in the upper-right of screens (on laptop and desktop computers). Note that autoplay was always on by default. Compare S2 Fig.



S2 Fig. Beginning in January, 2015, the autoplay switch was moved to a position immediately below the video, and its label was removed. It was, as always, on by default. Compare S1 Fig.



S3 Fig. In 2019, a 2-min video leaked from Google in which YouTube CEO Susan Wojcicki explained to her staff how YouTube's recommender algorithm was being revised to boost certain content and demote other content. The full video can be viewed at https://vimeo.com/354354050.

APPENDIX IX

The Opinion Matching Effect (OME): A subtle but powerful new form of influence that is apparently being used on the internet

Robert Epstein[1*], Yunyi Huang[1], & Miles Megerdoomian[1]

[1]American Institute for Behavioral Research and Technology, Vista, CA 92084, United States of America

*Corresponding author. E-mail: re@aibrt.org

SUBMITTED FOR PUBLICATION. Comments welcome. Do not quote or cite without permission of first author.

Keywords: Opinion Matching Effect; OME; online manipulation; recommender systems; online quizzes; voting advice applications; election manipulation

Abstract

In recent years, powerful new forms of influence have been discovered that the internet has made possible. The Search Engine Manipulation Effect (SEME) was discovered in 2013, and a comprehensive report about its effectiveness was published in 2015, with multiple replications published since then. SEME research shows that bias in search results can produce large shifts in the opinions and voting preferences of undecided voters without their awareness – upwards of 80% shifts in some demographic groups. We now introduce another new form of influence: the Opinion Matching Effect (OME). Many websites now "help" people form opinions about products, political candidates, or political parties by first administering a short quiz and then informing people how closely their answers match product characteristics or the views of a candidate or party. But what if the matching algorithm is biased? We first present data from real opinion matching websites, showing that responding at random to their online quizzes can produce significantly biased recommendations. We then describe a randomized, controlled, counterbalanced, double-blind experiment that measures the possible impact of this type of matching. A total of 773 politically diverse, eligible US voters participated in the experiment. They were asked to form opinions about the two candidates in the 2019 election for Prime Minister of Australia (thus assuring that our subjects were initially "undecided"). After reading basic information about the candidates, they were asked questions about their opinions and voting preferences. Then they were given a short quiz about various political issues, after which they were told how closely their views matched those of each candidate. Then they were asked those questions again, and we measured changes in opinions and voting preferences. Participants were randomly assigned to one of three groups: people who were told their answers closely matched the views of Candidate A, Candidate B, or both candidates equally. This produced shifts in voting preferences between 51% and 95% in the bias groups, with no participants showing any awareness of having been manipulated. In summary,

we show not only that OME is a large effect; we also show that biased online questionnaires exist that might be shifting people's opinions without their knowledge.

The Opinion Matching Effect (OME):
A subtle but powerful new form of influence
that is apparently being used on the internet

1. Introduction

As human communities grew in size from small tribes into vast cities and countries, leaders have had to develop increasingly effective ways of controlling the thinking and behavior of increasingly larger groups of people. By the early 1900s, social engineering began to progress from mere art to calculated science, beginning, perhaps, with theories of propaganda advanced by Georgy Plekhanov [1] and other early Marxists. The assertion that governments were not only responsible for controlling the masses but that they could use systematic, powerful methods to do so blossomed in capitalist America with the work of Edward L. Bernays [2], often known as the father of public relations. Bernays insisted that experts who master the emerging new techniques of control could be even more powerful than the government itself, constituting "an invisible government which is the true ruling power of our country" [2].

In 1957, journalist Vance Packard published a landmark book called *The Hidden Persuaders* [3], in which he revealed how both companies and politicians had begun working closely with social scientists to develop increasingly powerful ways of manipulating consumers and voters, often employing methods that left people unaware that they were being manipulated. These methods were being developed and tested using controlled experiments; behavioral science was now an essential tool of the marketing professional. In 1961, in President Eisenhower's farewell speech as president, he warned not only about the rise of a "military-industrial complex," he also expressed concern about the possible emergence of a "technological elite" that could someday control public policy without people knowing. Such new forces of control could be held in check, said Eisenhower, only by "an alert and knowledgeable citizenry" [4]. Has the public been alert, and are we knowledgeable about new forms of influence that may have come into being in the decades since Eisenhower's warning?

The rapid proliferation of internet access over the past two decades has in fact created new and especially impactful methods for controlling people's thinking and behavior, and because internet activity is dominated by a small number of worldwide monopolies – mainly Google and Meta/Facebook – when these new methods of influence are deployed, there appears to be no way to counteract them. If Candidate A posts an attack video online or on television, Candidate B can do the same. But if one of the large online platforms uses subtle techniques to support one candidate, the opposing candidate has no way to counteract that support.

Our research team has discovered, studied, and quantified several of these new methods of influence over the past decade [5-9]. In the present paper, we introduce a new form of online influence we call the Opinion Matching Effect (OME). We first present data showing that the effect has likely been deployed to some extent on the internet, and we then present a randomized, controlled, counterbalanced, double-blind experiment that demonstrates the potential power of this effect to shift opinions and voting preferences. Unlike other effects we have studied, OME is not exclusively in the hands of large tech monopolies. We believe, in fact, that it is being used competitively, which means – at the moment, anyway – that it does not pose a serious threat to democracy or human autonomy. That said, if this technique were to be adopted by a large tech monopoly at some point, it would likely have an outsized impact on online users – one that might be difficult for competitors to counteract.

1.1 Invisible influence

A relatively vast scientific literature now exists that examines ways of influencing people without their knowledge, and it is beyond the scope of this paper to review that literature in detail. We will describe some salient examples, however.

The Hidden Persuaders, the book by Vance Packard we mentioned earlier, was first published in 1957 and is still in print more than 60 years later. It shocked the American public by revealing the surprising extent to which companies and political candidates were collaborating with social scientists to develop new, largely invisible, methods for influencing consumers and voters [3]. Packard noted, for example, that the slow music that many stores were now broadcasting from their ceilings caused people to walk slower and, in so doing, to make more purchases. This manipulation produced no awareness on the part of consumers, needless to say. (This technique is used to this day, as the reader will likely observe on his or her next visit to a large store [10, 11].) Packard described dozens of techniques like this, almost all of which were supported by controlled studies performed by social scientists.

The recent best-selling books *Nudge*, by behavioral economists Richard Thaler and Cass Sunstein [12], and *Sway*, by business author Ori Brafman and psychologist Rom Brafman [13], summarize more recent studies of this sort, and so do two more recent scholarly books, each entitled *Invisible Influence* [14, 15]. In one of the studies mentioned in these books, researchers showed that people more often cleaned their eating environments when the subtle odor of a disinfectant cleaner was present than when it was absent [16]. Manipulations of this sort are especially problematic because they often lead people to believe that they are thinking independently – that they have made up their own mind [17, 18]. Thaler and Sunstein argue that when unseen forces are guiding people's behavior, they have lost their freedom. Because no cages and whips are visible, however, they might still feel free and thus not take steps to regain their actual freedom. A number of recent authors have expressed concern about a growing number of invisible manipulations that the internet has made possible, applying terms such as "digital nudging" and "hypernudging" to the new techniques [19, 20].

1.2 Recommender systems

OME can be considered a special case of recommender systems [21], which have been widely studied in recent years. Controlled studies have shown that computer-generated recommendations impact purchase preferences even when those recommendations are generated randomly [22, 23]. The power of such systems is no secret, and they impact more than just purchases. A 2015 study by employees at Netflix concluded, among other things, that the company's recommender algorithm accounted for "about 80% of hours streamed at Netflix" [24]. In 2018, Neal Mohan, then Chief Product Officer at YouTube, revealed that 70% of the time people spend watching videos on YouTube, they are viewing content recommended by YouTube's recommender algorithms [25, cf. 26, 27]. It has been estimated that 35% of Amazon's online sales are driven by Amazon's recommender algorithms [28, cf. 29]. Public officials have expressed particular concern over the company's practice of ranking Amazon-branded products ahead of competitors' products in the product lists shown to potential buyers – the equivalent of search results in a search engine [30, 31].

Sometimes relatively organic and benign online content can shift thinking and behavior. Online reviews of consumer products posted by legitimate reviewers – actual users of those products who post blogs or YouTube videos, for example – might recommend a product because they genuinely like it, and online product reviews have been shown to impact consumer purchases [32-35]. Because such reviews are inherently competitive, they pose no great threat to consumers, in our view. We are using the term "recommender *systems*," however, to refer to algorithmically-driven content that might influence large numbers of people and that cannot easily be countered either by consumers or competing businesses. When marketers or advertisers are promoting a particular product, for example, they might create dozens of apparently objective product review websites that just happen to give their own product the highest possible praise [36-38]. Manufacturers of products that compete with that product could play the same game, of course, but in each case, a true "system" of reviews has been deployed – a far more nefarious form of influence than the single blog post composed by someone expressing his or her own views.

Early recommender systems – described in the early 1990s – generally relied on two different strategies for making recommendations: "Content-based" systems recommended content based on the properties of content that a user selected in the past, whereas "collaborative-based" systems recommended content based on choices that had been made by people who were similar to the present user [39]. "Hybrid" systems used both methods [40]. By the mid 2000s, such systems were being optimized based on ever-expanding bodies of information being collected about users, specifically by making use of "user profiles that contain information about users' tastes, preferences, and needs. The profiling information can be elicited from users explicitly, e.g., through questionnaires, or implicitly—learned from their transactional behavior over time" [41]. As marketers and leaders knew long before the internet was invented, the more you know about people, the easier it is to influence them [2, 42-44]. The internet dramatically increased the rate at which information about people could be collected, and that information, in turn, has increased the power of recommender systems.

1.2.1 Voting Advice Applications (VAAs)

Voting advice applications (VAAs) – also known as "online vote selectors" – are special recommender systems that use questionnaires to guide people's votes and party affiliations. An early VAA was simply a paper-and-pencil test called the *StemWijzer*, used before elections in The Netherlands in the late 1980s [45-47]. It asked for participants' views on various election-related issues, and based on their responses, it matched them with suitable candidates or political parties. In the late 2000s, research showed that the German *Wahl-O-Mat* questionnaire system was effective in mobilizing people to vote [48, cf. 49]. VAA methodology has been widely used across Europe to impact voters, especially over the past decade or so [46, 50, 51]. According to a 2009 study, 40% of voters in the 2006 national election in The Netherlands used online VAAs to guide their votes [52]. The study concluded that VAAs "had a modest effect on electoral participation and a substantial effect on party choice, especially among undecided voters" [52]. Other studies have demonstrated how various aspects of the construction of the questionnaire can impact voters differentially [53-56].

A meta-analysis of 22 VAA studies assessing data obtained from more than 70,000 users in 9 countries concluded that VAAs significantly increased voter turnout, had a significant impact on voter choices, and produced modest increases in voter knowledge [57]. Again, mainly in Europe, VAAs have apparently impacted millions of voters [55], and researchers continue to study how various factors, such as the wording of questions, increase or decrease the impact of a VAA. To our knowledge, the scientific literature on VAAs focuses exclusively on legitimate questionnaires that were designed to increase voter turnout and improve the quality of voter decisions. We have not found published experiments in which researchers used questionnaires dishonestly to try to shift votes or opinions, but we did find a blog post on Medium (not peer reviewed) in which the author reported testing the fairness of iSideWith.com by completing the website's quiz with random answers [58]. The author concluded that the website gave biased results, but the findings were marginal, and the methodology was inadequate, in our view.

As we proceed, we will address a question that naturally comes to mind when one recognizes the power that questionnaires have to impact voters: Could VAA-type online instruments be designed to shift votes dishonestly – that is, in a way that is biased toward one candidate or party? If so, could such tools impact voters in such a way that prevents them from becoming aware that they have been manipulated?

In the first part of the present study, we sought to identify websites that might be using online questionnaires dishonestly – that is, in ways that make recommendations to users that do not accurately reflect their answers to the questionnaire they completed. Do such websites exist? If so, do they violate existing laws or regulations, such as deceptive advertising or consumer fraud laws?

In the second part of the study, we describe an experiment in which an intentionally misleading VAA-type questionnaire was deployed in an attempt to shift opinions and voting preferences. Specifically, we first asked users some questions and then made recommendations while ignoring the user's responses to the questionnaire. We sought to determine the extent to which such a procedure can shift opinions and voting preferences. We also sought to determine whether our participants were aware that they were being influenced unfairly.

2. Ethics Statement

The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB FWA00021545 written consent was obtained for all investigations as specified in the Procedure section of Investigation 2.

3. Investigation 1: Bias examination of actual online opinion matching websites

We began our investigation by using the Brave search engine (to protect our privacy) to locate a variety of "online quizzes" (or "online questionnaires"), looking especially for quizzes of a political nature, such as quizzes that purported to match users with particular candidates or political parties. We then wrote code (in Python) that simulated a human user – in other words that clicked at human speed and that paused after it completed a quiz and submitted its answers [59-61]. Our bots took each online quiz repeatedly (generally, 300 times), and recorded the random answers our bots supplied (numerical answers to multiple-choice questions) and the recommendations the website gave. We did not attempt an exhaustive survey of such quizzes; rather, we examined only enough to yield two quizzes that gave us recommendations that were biased at a significance level under 0.001. To find these two, we had to examine a total of 15 websites. The 13 websites that appeared to give us relatively fair results are listed as S1 Text in our Supporting Information, and our Data Availability statement explains how readers can access our raw data and the Python scripts we used to access website quizzes.

3.1 Website 1: My Political Personality

3.1.1 Website 1: Methods

The first of the two websites we found which appeared to give biased results was https://politicalpersonality.org (S1 Figure), a website maintained by My Political Personality, a non-profit "voter empowerment group," which promises to assist users in determining which of four political parties – Democrat, Republican, Libertarian, or the Green Party – is the best match for their political views. The website does so by having the user complete its "Political Personality Test," a 15-item Likert-scale test. The website includes an informal disclaimer, noting that its questionnaire is "just a fun and voluntary personality quiz – not a statement of fact". S2 and S3 Figures show website information and its nonpartisan statement.

It matches people to a political party – just one – by revealing a user's "political personality," where each personality has been pre-matched (using a methodology that is not described) with a political party. See S4 Figure for an example of the website's quiz results page.

As we did for all the websites we examine, we began our investigation informally by completing the quiz manually a few times, looking for indications that the recommendations made after we completed a quiz might be biased – in this case, toward one political party. Again, we emphasize that this process was informal and exploratory only.

Because this questionnaire seemed suspect, we then customized a Python script (obtained from the Selenium WebDriver library, accessible at https://www.selenium.dev/documentation/webdriver/), to (a) clear cache and cookies, (b) reopen the tab, (c) retake the quiz, and (d) record the results. We repeated this process 300 times. We did so in the present instance in two sessions, the first on January 4th and the second on January 15th, 2022.

3.1.2 Website 1: Results

Table 1 shows the frequency with which each of the political parties was recommended to the user over the course of the 300 trials. If both the questionnaire and the computation of results were completely fair, one would expect all four parties to be recommended approximately 75 times. The Republican and Libertarian parties were each recommended the expected number of times (approximately), but the Green party was never recommended, and the Democratic party was recommended roughly twice the number of expected times. The differences between the four frequencies were highly significant ($X^2 = 139.68$, $df = 3$, $p < 0.001$), and so was the pairwise difference between the recommendations made for the Democratic and Republican parties ($z = 5.05$, $p < 0.001$).

Table 1. Investigation 1: Frequencies and percentages from party recommendations.

Party	Frequency	Percent	Cumulative Percent
Democrat	144	48.0	48.0
Republican	84	28.0	76.0
Libertarian	72	24.0	100.0
Green	0	0	100.0
Total	300	100.0	-

3.2 Website 2: Pew Research Center

3.2.1 Website 2: Methods

The second quiz we found that led to apparently biased results proved to be surprising. A small, independent group calling itself "My Political Personality" posted the quiz described above; the names of the website creators were not listed. But the second suspect quiz we found was posted by the Pew Research Center, a highly respected nonprofit organization that identifies itself as a "nonpartisan fact tank" that values "independence, objectivity, accuracy, rigor, humility, transparency and innovation." Of interest here is their "Political Typology Quiz," accessible at https://pewresearch.org/politics/quiz/political-typology/ (S5 Figure).

This quiz consists of 16 multiple choice questions and promises to match the user with one – just one – of nine political orientations: Progressive Left, Establishment Liberals, Democratic Mainstays, Outsider Left, Stressed Sideliners, Ambivalent Right, Populist Right, Committed Conservatives, or Faith and Flag Conservatives. S6 and S7 Figures show examples of actual quiz results from this website.

Our procedure for evaluating this quiz was identical to the procedure described for the "Political Personality" quiz we described above. The evaluation was conducted in four sessions between January 11th and January 15th, 2022.

3.2.2 Website 2: Results

Table 2 shows the frequencies of the political recommendations that were made. If the questionnaire had been constructed fairly, and if it had been scored fairly, we might expect it to have recommended each of the nine political orientations about 33 times. In fact, the frequencies varied from 0 (Progressive Left) to 102 (Ambivalent Right) (X^2 = 219.60, df = 8, p < 0.001).

Table 2. Investigation 1: Frequency and percentages from typology recommendation.

Typology	Frequency	Percent	Cumulative Percent
Progressive Left	0	0.0	0.0
Establishment Liberals	52	17.3	17.3
Democratic Mainstays	16	5.3	22.7
Outsider Left	7	2.3	25.0
Stressed Sideliners	28	9.3	34.3
Ambivalent Right	102	34.0	68.3
Populist Right	35	11.7	80.0
Committed Conservatives	37	12.3	92.3
Faith and Flag Conservatives	23	7.7	100.0
Total	300	100.0	-

When we divided the nine categories into the three conventional groupings for political leaning – left, middle, and right – we again found apparently biased counts favoring conservatives (Table 3). The pairwise left/right difference was highly significant ($z = 10.00$, $p < 0.001$).

Table 3. Investigation 1: Frequency and percentages from categorized typology recommendation.

Group	Frequency	Percent	Cumulative Percent
Left	75	25.0	25.0
Moderate	28	9.3	34.3
Right	197	65.7	100.0
Total	300	100.0	-

4. Investigation 2: A randomized, controlled OME experiment

Given the possibility that biased questionnaires (or biased scoring methods for questionnaires) might exist online, we conducted an experiment that allowed us to quantify the possible impact that highly biased questionnaire scores might have on people's opinions and voting preferences. We conjectured that scores favoring one political candidate would be able to shift voting preferences substantially while having less impact on people's opinions about the candidate; see our Results and Discussion sections for more information about these issues.

4.1 Methods

4.1.1 Participants

A total of 773 demographically diverse, eligible US voters between ages 18 and 92 participated in the experiments. The sample was provided by Cloud Research, a company that draws subjects from Amazon's Mechanical Turk subject pool, screening out bots and other suspect entities. Demographic characteristics of the sample are delineated in S1 Table. Before cleaning, our sample consisted of 816 individuals. One was removed because that person indicated that their English fluency was under 6 on a scale from 1 to 10 (where 1 was labeled "Not fluent" and 10 was labeled "Highly fluent"), and 42 were removed because they indicated that their level of familiarity with one or both of the two Australian political candidates mentioned in the study was greater than 3 on a scale from 1 to 10, where 1 was labeled "Not familiar at all" and 10 was labeled "Very familiar." After cleaning, the mean familiarity level for our first candidate, Scott Morrison, was 1.15 (*SD* = 0.44), and the mean familiarity level for our second candidate, Bill Shorten, was 1.07 (0.30).

4.1.2 Procedure

Using a pre-post experimental design developed by Epstein and his collaborators for quantifying bias in online manipulations [5-7], participants were randomly assigned (without their knowledge) to four different groups as they enrolled in the study on December 8th, 2021, December 14th, 2021, or January 3rd, 2022. The combination of random assignment and cleaning left us with slightly uneven *n*'s in each group (Table 4).

Table 4. Investigation 2: *n* of groups.

Group	*n*
1. 8 questions, high readability	197
2. 8 questions, low readability	197
3. 16 questions, high readability	187
4. 16 questions, low readability	192

Before beginning the experiment, all participants were given basic information about the procedure and about their rights as subjects and then asked for their informed consent to proceed

(S2 Text). They were then given basic instructions about how to proceed and then shown short paragraphs about the two political candidates running for Prime Minister of Australia in 2019: Scott Morrison and Bill Shorten. The order of the names was randomly counterbalanced throughout the study. Each paragraph was deliberately bland in tone and approximately 120 words in length (S3 Text). Then each participant was asked three opinion questions about each candidate and asked to reply on 10-point scales: One question asked how much they liked each candidate, the second asked how much they trusted each candidate, and the third asked for their overall impression of each candidate (Figure S9 shows the questions and scales).

Below those questions, participants were asked to indicate on an 11- point scale (labeled from 5 to 0 to 5) which candidate they would likely vote for it they had "to vote today." Finally, they were asked, in a forced-choice question, to indicate which candidate they would likely vote for it they had "to vote right now."

Participants were then asked to complete a short questionnaire that would measure their political views on a number of subjects, after which they were shown how closely their answers matched the views of the political candidates (more about this below).

Following the quiz and the scoring, all participants were asked the same eight questions they had been asked before the quiz (three opinion questions for each candidate, followed by the 11-point scale showing voting preference, followed by the forced-choice vote question).

Finally, all participants were asked whether anything about the experiment "bothered" them. If they responded "yes," they could then type freely into a text box, expressing their concerns. This is where we ultimately looked for indications that participants showed some awareness of bias in the content they had been shown (particularly in the quiz or scores they had been shown). We could not directly ask them about whether they detected bias, because a leading question of this sort would artificially inflate the detection rate [62].

Participants were then thanked for their participation, given a code they could use to receive their payment, and given an email address they could use to withdraw their data from the experiment or to address questions to the researchers.

The four groups. Figure 1 depicts the 2-by-2 factorial design employed in the study. Because recent studies, particularly in the EU, have found that the impact of election-related quizzes varies with the structure and content of such quizzes [53-56], we elected to vary both the content and length of our quiz. Participants were given either high- or low-readability quizzes, and the quizzes were either 8 or 16 questions in length (Fig 1). S4 to S7 Texts show the different quiz questions for each of the four groups. S10 Figure shows an example of the website page during the quiz-taking process.

	8 questions	16 questions
high readability	n = 197 FKG = 4.5	n = 187 FKG = 4.6
low readability	n = 197 FKG = 10.8	n = 192 FKG = 10.8

Fig 1. 2-by-2 factorial design showing two levels of quiz readability (low and high) and two quiz lengths (8 questions and 16 questions). The *n* for each of the four groups is shown in each box, along with the Flesch-Kincaid Grade Level (FKG) of the content.

Technically, our design could be viewed as having a 3-by-2-by-2 factorial structure, because all participants were also randomly assigned to three different candidate bias groups: pro-

Candidate-1 (Morrison), pro-Candidate 2 (Shorten), or control (neutral, favoring neither candidate). However, as we will explain below, our analysis of the data combined the two candidate bias groups into one group, and it was only for that group that the 2-by-2 factorial design applied, so it would be misleading to characterize our experimental design as having three separate dimensions.

Following the quiz, participants were shown an animated loading bar for 5 sec to give the impression that their responses were being processed (S11 Figure). Then, participants who had been assigned to the pro-Candidate-1 group (Morrison) were informed that 85% of their answers matched Morrison's views and that 25% of their answers matched Shorten's views (Figure S12); participants who had been assigned to the pro-Candidate-2 group (Shorten) were informed that 85% of their answers matched Shorten's views and that 25% of their answers matched Morrison's views; and participants in the control group were informed that 42% of their answers matched the views of each candidate (Figure S13). Note that the 85% and 25% values sum to a value over 100% because, presumably, there was some overlap in agreement between the two candidates. That said, all three of the percentages we used in this experiment were chosen somewhat arbitrarily.

4.2 Results

The main measure of interest in experiments that use a pre-post manipulation design to measure changes in voting preferences is vote manipulation power, or VMP – the post-manipulation increase in the percentage of people choosing to vote for the candidate favored in their group [5]. This is calculated by combining the data in the two bias groups. For details about how to compute VMP, see S8 Text.

In the present experiment, the overall VMP for the four quiz groups combined was 75.5% (95% CI, 70.3-80.7%; McNemar's test, $p < 0.001$), which is high compared with VMPs found in comparable experiments on new forms of manipulation made possible by the internet [5-9]. S2 to S5 Tables show VMPs broken down by educational attainment, gender, age, and race/ethnicity.

VMPs in the four quiz groups ranged from 50.7% to 95.2% (Table 5), the latter value being the highest value our research group has ever found in comparable experiments on online influence. The percentage of users who appeared to perceive some degree of bias in the content they were shown was also notable in this experiment: not a single participant claimed to observe any bias in the content.

Table 5. Investigation 2: VMP for each of the four groups.

Group	Total *n*	Bias Groups *n*	VMP (%)	95% VMP Confidence Interval	McNemar's Test X^2	*p*
1. 8 questions, high readability	197	129	77.3	67.2 - 87.4	45.455	< 0.001
2. 8 questions, low readability	197	132	95.2	90.0 - 100.0	58.017	< 0.001
3. 16 questions, high readability	187	121	50.7	39.1 - 62.3	29.167	< 0.001
4. 16 questions, low readability	192	128	82.0	72.3 - 91.6	44.463	< 0.001
Total	773	510	75.5	70.3 - 80.7	182.066	< 0.001

Regarding the possible differential impact of the quiz characteristics, the VMPs for the four subgroups in the 2-by-2 factorial design (Table 5 and Fig 2) suggest main effects for both the quiz length (with the shorter quiz having the greater impact) and readability (with low readability having the greater impact), with little or no interaction between these effects. To our knowledge, a standard ANOVA cannot be performed on our VMPs because the same VMP applies to all members of a group (see S8 Text for the calculation method). VMPs are percentages, not means, so no simple measure of individual variability underlies them. We can estimate the magnitude and significance of main effects, however, with *z*-tests, as shown in Table 6. Both effects were highly significant, with readability the larger effect.

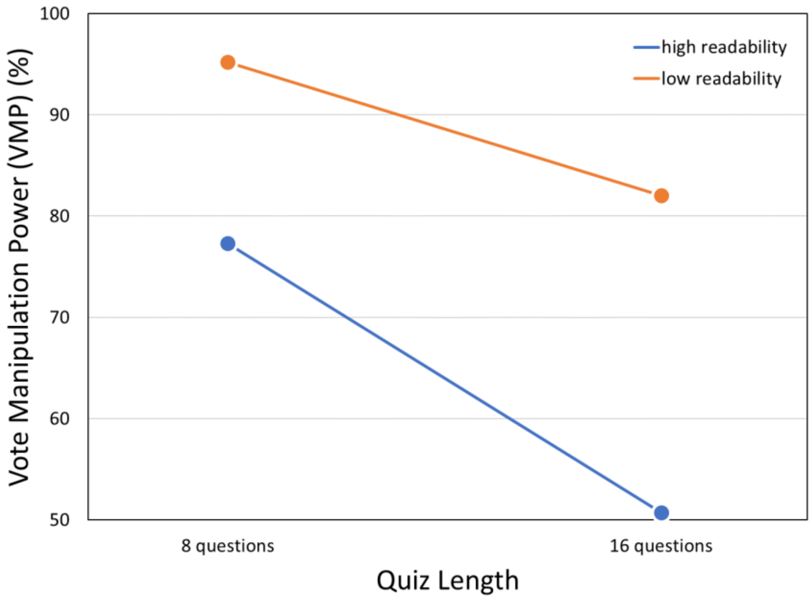


Fig 2. VMPs broken down by quiz characteristics. Pre-post shifts in voting preferences as expressed by VMPs suggest that low-readability quizzes produce greater shifts toward the favored candidate than high-readability quizzes do, and that shorter quizzes produce greater shifts than longer quizzes do. See text for details.

Table 6. Investigation 2: Comparison of VMPs for two levels of quiz length and two levels of readability.

Factor	Treatment	*n*	VMP (%)	*z*	*p*
Quiz Length	8 questions	261	86.0	5.213	< 0.001
	16 questions	249	65.2	-	-
Readability	low	260	88.7	6.347	< 0.001
	high	250	63.5	-	-

Voting shifts on the 11-point scale (for the two bias groups combined) were also relatively large and occurred in the predicted direction (Table 7). Main effects for quiz characteristics were marginal, with no evidence of interaction (Table 8 and Fig 3). We found no pre-post differences in voting preferences as expressed on the 11-point scale for participants in the neutral groups (S6 Table).

Table 7. Investigation 2: Pre- and post-manipulation votes on 11-point scale (5 to 0 to 5) by quiz group.*

Quiz Group	Pre-manipulation		Post-manipulation					
	Bias Groups; Mean (*SD*)	**Neutral Group; Mean (*SD*)**	**Bias Groups; Mean (*SD*)**	**Neutral Group; Mean (*SD*)**	**Bias Groups; Mean Shift**	**Neutral Group; Mean Shift**	**Cohen's *d*†**	**Mann-Whitney *U*‡**
1. 8 questions, high readability	0.09 (2.64)	0.47 (2.70)	2.71 (2.30)	0.69 (2.14)	2.62	0.22	1.06	1,801.50***
2. 8 questions, low readability	0.02 (2.63)	0.42 (2.49)	2.86 (1.94)	0.40 (2.28)	2.84	-0.02	1.23	1,297.50***
3. 16 questions, high readability	0.54 (2.45)	0.02 (2.59)	2.49 (2.13)	-0.09 (2.57)	1.95	-0.11	0.85	1,483.50***
4. 16 questions, low readability	-0.22 (2.43)	-0.16 (2.85)	2.50 (2.54)	0.05 (2.80)	2.72	0.21	1.09	1,415.50***
Total	0.10 (2.55)	0.19 (2.66)	2.64 (2.24)	0.27 (2.45)	2.54	0.08	1.06	23,959.50***

*Positive means indicate shifts in the direction of the favored candidate after correcting for counterbalancing.
†Cohen's *d* effect sizes were calculated using the means and standard deviations of the pre- and post-manipulation opinion ratings for the bias groups combined.
‡Mann-Whitney *U* tests were conducted for the score shifts between each bias group and each neutral group.
*** $p < 0.001$

Table 8. Investigation 2: ANOVA of vote preference score shifts on the 11-point scale for two factors: quiz length and readability.

Effect	Sum of Squares	*df*	*F*	*p*
Quiz Length	19.045	1	2.858	0.092 NS
Readability	29.871	1	4.482	< 0.05
Quiz Length × Readability	9.814	1	1.473	0.226 NS

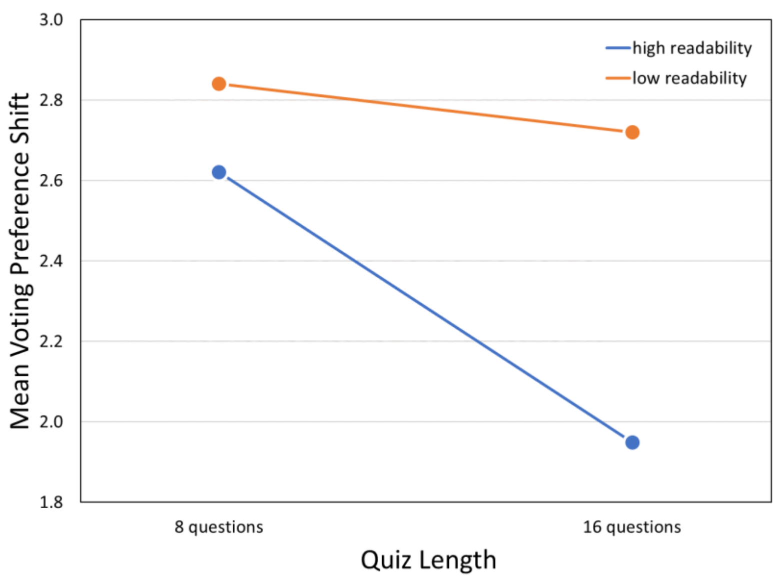


Fig 3. Vote preference shifts on the 11-point scale broken down by quiz characteristics. Pre-post shifts in voting preferences as expressed on an 11-point scale suggest that low-readability quizzes produce greater shifts toward the favored candidate than high-readability quizzes do, and that shorter quizzes produce greater shifts than longer quizzes do. See text for details.

We also found significant pre-post differences in opinions people expressed about the favored candidate (the candidate we identified as a great match to their quiz answers), although effect sizes were relatively low (Table 9). In an ANOVA, we found no main effects or interactions reflecting differential characteristics of the quizzes (S7 Table).

Of special note, no participants in either of the bias groups reported any awareness of bias in the content they viewed in this study, or in the scores they were shown after they completed the quiz.

Table 9. Investigation 2: Pre- and post-manipulation opinion ratings of favored candidates (for the bias groups combined).

Group	Opinion	Pre	Post	Diff	Cohen's *d*†	*z*‡
1. 8 questions, high readability	Impression	7.28 (1.85)	7.88 (1.77)	0.60	0.33	-4.597***
	Trust	6.44 (1.92)	7.09 (2.03)	0.65	0.33	-5.068***
	Likeability	7.05 (1.89)	7.53 (1.94)	0.48	0.25	-3.446**
2. 8 questions, low readability	Impression	7.33 (1.66)	7.98 (1.45)	0.65	0.42	-4.784***
	Trust	6.31 (1.93)	6.95 (1.97)	0.64	0.33	-5.488***
	Likeability	7.18 (1.66)	7.69 (1.50)	0.51	0.32	-4.295***
3. 16 questions, high readability	Impression	7.02 (1.82)	7.52 (1.79)	0.50	0.28	-4.452***
	Trust	6.31 (2.20)	6.82 (2.19)	0.51	0.23	-3.792***
	Likeability	7.07 (1.91)	7.23 (1.97)	0.16	0.08	-1.387
4. 16 questions, low readability	Impression	7.13 (1.89)	7.68 (1.70)	0.55	0.30	-3.996***
	Trust	6.18 (1.96)	6.91 (1.90)	0.73	0.38	-5.861***
	Likeability	7.08 (1.87)	7.57 (1.63)	0.49	0.28	-4.071***
Total	Impression	7.20 (1.75)	7.83 (1.64)	0.63	0.33	-8.873***
	Trust	6.48 (1.96)	7.03 (1.91)	0.55	0.32	-10.113***
	Likeability	7.24 (1.80)	7.67 (1.73)	0.43	0.23	-6.629***

†Cohen's *d* effect sizes were calculated using the means and standard deviations of the pre- and post-manipulation opinion ratings for the favored candidate.
‡z values represent Wilcoxon signed ranks test comparing pre- to post-manipulation opinion ratings for the favored candidate.
*** $p < 0.001$, ** $p < 0.01$

5. Discussion

In our view, this study produced two quite remarkable results. First, it produced the largest shifts in voting preferences (as measured by VMP, which is calculated from answers to a forced-choice question: “If you had to vote right now, who would you vote for?”) we have ever observed after having conducted more than 10 years of studies of this sort [5-9] – shifts between 50.7% and 95.2%. Second, it is the only online manipulation study we have ever conducted (without the use of masking procedures) that apparently produced no awareness of bias by participants. Why would a quiz-based manipulation produce such a large impact with so little cost?

We think the answer is fairly obvious. A quiz posted to help someone make an informed decision provides a service, at least from the point of view of most, if not all, users, which is why marketers use quizzes for lead generation, branding, data gathering, and other purposes [63-66]. While completing the quiz, the user is not exposed to biased content and has no factual basis for suspecting that a manipulation is about to occur. That manipulation occurs only *after* the quiz is completed, when the user is presented with false information about his or her scores. At that point, the user has no way to evaluate the accuracy of the score. Bear in mind that people who take quizzes to help them make decisions (about political candidates, guitars, vacation spots, or just about anything else) almost certainly lack the knowledge they need to make an informed decision; that, presumably, is why they are taking the quiz. An online quiz is, in effect, an ideal tool both for attracting users who are vulnerable to manipulation and for causing an invisible manipulation to occur.

Because OME can apparently produce large shifts in opinions and voting preferences without user awareness, we wonder why – at least as far as we can tell – researchers have consistently evaluated the impact of quizzes based on users’ actual answers and scores. Because it is such a simple matter to ignore those answers and fake those scores, and because the internet is awash with quizzes of all sorts, why have researchers not addressed this issue? This raises the question we asked in our first investigation (above): Are people – or organizations, or companies, or political parties – indeed manipulating users by showing them biased results that are largely or entirely independent of their answers? We presented two examples of online quizzes that appear to be showing users biased results. One quiz – "My Political Personality" – was posted by a small, independent organization of the same name, and the other quiz was posted by the venerable Pew Research Center. It is possible that neither group was aware of the bias in its quiz – that the bias was entirely accidental and unintentional. Even if we give both organizations the benefit of the doubt, however, *our findings suggest that these quizzes are still shifting opinions systematically*.

Our findings also demonstrate how easily online quizzes could be used to shift opinions and votes on a massive scale. We have no evidence that online quizzes are being used that way, but it is notable here that in the months preceding national elections in the US in 2016 and 2020, a number of election-related quizzes were posted on high-traffic websites such as https://WaPo.com and https://BuzzFeed.com. Even Tinder, known mainly as a "hookup" website where people swipe left or right to indicate whether they are attracted to someone, deployed a "Swipe the Vote" feature in March, 2016, to help its 50 million users decide whom to vote for in November [67] (S14 Fig). Again, notably, according to https://OpenSecrets.org, in 2016, 89.3% of the political donations from Tinder's parent company at that time (InterActiveCorp) went to just one of the two major political parties in the US [68]. Our data suggest that if Tinder had been using its Swipe-the-Vote feature dishonestly, it could have shifted – at least temporarily – the voting preferences of between 1.3 and 2.4 million undecided voters. We base this estimate on the following modest assumptions: (a) that several months before the Presidential election, 20% of the users of Tinder were undecided voters

(0.2 × 50,000,000 = 10,000,000), (b) that 50% of those undecided voters tried out Tinder's Swipe-the-Vote application (0.5 × 10,000,000 = 5,000,000), (c) that before the manipulation, if those voters had been asked a forced-choice question about how they planned to vote, they would likely have split 50/50 (2,500,000 for each candidate), and (d) that after the manipulation, between 50.7% and 95.2% of the voters in one of those groups might have shifted their preference toward the other candidate (0.507 × 2,500,000 = 1,267,500; 0.952 × 2,500,000 = 2,380,000).

5.1 Limitations and future research

This brings us to two likely limitations of OME. First, we have no evidence that this effect leaves a lasting impact on a user's voting preferences. The impact will vary according to how vulnerable someone is to this type of subtle persuasion. It might have a lasting impact on only a small proportion of voters – a matter to be explored in future research. The voter most likely to be influenced by a biased quiz site is the one who, in a last-minute attempt to get off the fence, takes the quiz on Election Day or perhaps the day before. If so, that would greatly limit the impact of the quiz.

Second, this manipulation – to the extent that it is being used at all – is probably being used competitively. Platforms capable of employing new forms of influence such as the Search Engine Manipulation Effect (SEME) [5], the Search Suggestion Effect (SSE) [8], and the Targeted Messaging Effect (TME) [7] can expose people to similarly biased content hundreds of times before an election as people conduct search after search, or as they scroll, over and over again, through Twitter feeds (or "X" feeds, if you prefer). People are unlikely, however, to complete similarly biased questionnaires repeatedly in the months leading up to an election. Unlike SEME, SSE, and the YouTube Manipulation Effect (YME) [9], OME is an inherently competitive manipulation. It is not controlled exclusively by two or three tech monopolies; biased quizzes can be posted by just about anyone. In that sense, biased quizzes are more like blogs than they are like search results. That said, if any of the Big Tech platforms started using online quizzes to shift opinions or votes, or began promoting certain quizzes while suppressing others, they could conceivably shift millions of votes with no one able to counteract their actions.

Is it legal to shift opinions, purchases, or votes by giving people fake scores on quizzes? We have not been able to find any relevant cases in the US, but if OME becomes a popular tool for shifting large numbers of votes in elections, it is conceivable that political parties or public officials might start searching the law books for relevant laws and regulations. Quizzes used to manipulate people online might violate provisions of the Federal Election Commission Act, the Federal Trade Commission Act, the Consumer Protection Act, or the Uniform Deceptive Trade Practices Act, as well as any number of state laws or regulations. This is a matter for lawyers to research, not social scientists.

One also might wonder about the quiz itself. Given that our quiz contains no information about the candidates, why should different quiz characteristics have *any* differential impact on the outcome of our manipulation? Apparently, a shorter quiz, or a quiz that is more difficult to read, makes people somewhat more vulnerable to the manipulation (seeing fake scores that favor one candidate or another). One might speculate that people will consider a verbose quiz to be more substantive, but shouldn't they also take a *longer* quiz more seriously? Our design did not allow us to explore such issues, but they might be worth exploring in future studies. Generally speaking, longer test instruments have been shown to produce less consistent or honest responses [69], and, surprisingly, more verbose test instruments have been shown to produce responses that are *more* consistent and honest [70]. Both findings are consistent with our new findings, but we still find it surprising that different characteristics of our quizzes had any differential effects at all – another matter to be explored in future research.

One might also be concerned about the method we used to evaluate the fairness of online tests – 15 quizzes in all (see Investigation 1). Will random responding necessarily tell you that a quiz is biased? We argued that random responding should produce scores that don't favor any particular political party or candidate (or guitar brand, for that matter). We acknowledge, however, that a test might be constructed in good faith and without conscious bias that will not survive the random-responding test. That said, recall that in our evaluation of the Pew quiz, we were never labeled Progressive Left – one of nine possible labels we might have received – even though we took the test 300 times. If we make the reasonable assumption that by responding at random, we should be labeled Progressive Left 1/9th of the time, then the probability that we are *never* labeled that way after 300 trials is a disturbing 4.5×10^{-16}. Of course, the test might have been legitimately constructed so that that label is rarely applied, perhaps because Progressive Leftists are near the tail end of a normal distribution. But even assuming that random responding should produce that label only 1/50th of the time, the probability that we are *never* labeled that way is still only 0.002.

An exploration of how tests should be constructed is beyond the scope of this paper, and it is also irrelevant to the central point we are raising, namely that tests can be posted online that can easily shift people's opinions and voting preferences without their knowledge. Because the online test-taking experience is normally ephemeral, with no record being kept of the experience, this type of manipulation, like SEME and other online effects we have studied, leaves no paper trail for authorities to trace. We cannot go back in time to measure the possible impact of Tinder's Swipe-the-Vote feature. It might have had very little impact on the 2016 election (especially if it had been fair and honest in its scoring), or it might have had a significant impact; unless a whistleblower comes forward or old records are revealed, we will never know. This is why our research team has, since 2016, been building increasingly larger and more capable computer networks that preserve online ephemeral experiences [71-74]. In 2016, we were able to preserve and later analyze about 13,000 politically-related searches on the Google, Bing, and Yahoo search engines. As of this writing (August 4, 2023), we have in recent months preserved more than 30 million ephemeral experiences on multiple platforms, and we are continuing to monitor online content 24 hours a day through the computers of a politically-balanced group of more than 10,000 registered voters in all 50 US states.

We also acknowledge that the magnitude of the effect we found in our experiment (Investigation 2) might be due in part to the fact that our participants were what some political scientists might call "low-information" undecided voters. This is so because we used subjects from US to make judgments about political candidates from Australia. High-information undecided voters differ from low-information undecided voters in some respects [75], although, to our knowledge, there is currently no evidence that low-information undecided voters are more vulnerable to online manipulations such as OME. Again, this is an issue that can only be settled by further research.

We remind the reader that we did not in this investigation attempt to survey the internet to try to estimate the number or proportion of websites that might currently be using quizzes unfairly to manipulate people's opinions. Rather we used a simple proof-of-concept procedure: We counted the number of websites that used quizzes that we needed to investigate in order to find just two websites that appeared to use quizzes unfairly. We found those two websites among the first 15 that we examined (13.3%). We have no evidence that this same proportion of quiz-based websites is suspect throughout the internet.

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Acknowledgements

This report is based in part on a paper presented the 102nd annual meeting of the Western Psychological Association in April 2022. We thank C. Tyagi for background research and V. R. Zankich for help in data analysis.

Author Contributions

Robert Epstein: design, supervision, and draft of manuscript. **Yunyi Huang:** implementation of Investigation 1, data analysis. **Miles Megerdoomian:** background research, data analysis, manuscript preparation.

Funding

This work was supported by general funds of the American Institute for Behavioral Research and Technology (AIBRT), a nonpartisan, nonprofit, 501(c)(3) organization. The authors received no specific funding for this work. This research did not receive any external support from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

Anonymized raw data will be posted at Zenodo.com after the paper is accepted for publication. Anonymization was required to comply with the requirement of the sponsoring institution's Institutional Review Board that the identities of the participants be protected in accordance with HHS Federal Regulation 45 CFR 46.101.(b)(2). Readers can request copies of the Python scripts used to evaluate the online quizzes by writing to info@aibrt.org.

References

1. Smith BL. propaganda. In: Encyclopedia Britannica [Internet]. 2022 [cited 2023Jan13]. Available from: https://www.britannica.com/topic/propaganda
2. Bernays E. Propaganda [Internet]. [cited 2023Jan13]. Available from: http://www.hiaw.org/defcon1/bernprop.html
3. Packard V. The Hidden Persuaders. Longmans, Green & Co; 1957.
4. Military-Industrial Complex Speech, Dwight D. Eisenhower, 1961 [Internet]. Yale.edu. 2019. Available from: https://avalon.law.yale.edu/20th_century/eisenhower001.asp
5. Epstein R, Robertson RE. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proceedings of the National Academy of Sciences of the United States of America; 2015. Available from: https://doi.org/10.1073/pnas.1419828112
6. Epstein R, Lee V, Mohr Jr. R. The Answer Bot Effect (ABE): A Powerful New Form of Influence Made Possible by Intelligent Personal Assistants and Search Engines. Paper presented at the 98th annual meeting of the Western Psychological Association; Portland, OR; 2018 April.
7. Epstein R, Tyagi C, Wang H. What would happen if twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME). PLOS ONE [Internet]. 2023 Jul 28; Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0284495
8. Epstein R, Mohr Jr. R, Martinez J. The Search Suggestion Effect (SSE): How search suggestions can be used to shift opinions and voting preferences dramatically. Paper

presented at the 98th annual meeting of the Western Psychological Association; Portland, OR; 2018 April. https://aibrt.org/downloads/EPSTEIN_MOHR_&_MARTINEZ_2018-WPA-The_Search_Suggestion_Effect-SSE-WP-17-03.pdf

9. Epstein R, Flores A. The YouTube Manipulation Effect (YME): A Quantification of the Impact that the Ordering of YouTube Videos Can Have on Opinions and Voting Preferences [Internet]. Social Science Research Network. Rochester, NY; 2023 [cited 2023 Aug 3]. Available from: https://ssrn.com/abstract=4527207

10. Milliman RE. Using Background Music to Affect the Behavior of Supermarket Shoppers. Journal of Marketing. 1982 Jul;46(3):86–91.

11. Knoferle KM, Spangenberg ER, Herrmann A, Landwehr JR. It is All in the mix: the Interactive Effect of Music Tempo and Mode on in-store Sales. Marketing Letters. 2011 Dec 7;23(1):325–37.

12. Thaler RH, Sunstein CR. Nudge: Improving decisions about health, wealth, and happiness. New York. N.Y.: Penguin Books; 2008.

13. Brafman O, Brafman R. Sway: the irresistible pull of irrational behavior. New York: Doubleday; 2009.

14. Hogan K. Invisible influence: the power to persuade anyone, anytime, anywhere. Hoboken, New Jersey: Wiley; 2013.

15. Berger J. Invisible Influence: The Hidden Forces That Shape Behavior. New York Simon & Schuster Paperbacks June; 2017.

16. Holland RW, Hendriks M, Aarts H. Smells Like Clean Spirit: Nonconscious Effects of Scent on Cognition and Behavior. Psychological Science. 2005 Sep 1;16(9):689–93.

17. Bargh J, Lee-Chai A, Barndollar K, Trötschel R. The automated will: Nonconscious activation and pursuit of behavioral goals. J Pers Soc Psychol. 2001;81:1014–27.

18. Pronin E, Kugler M. Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. Exp Soc Psychol. 2007;43:565–78.

19. Weinmann M, Schneider C, Brocke J vom. Digital Nudging. Business & Information Systems Engineering. 2016 Oct 3;58(6):433–6.

20. Yeung K. 'Hypernudge': Big Data as a mode of regulation by design. Inf Commun Soc [Internet]. 2017;20(1):118–36. Available from: http://dx.doi.org/10.1080/1369118x.2016.1186713

21. Resnick P, Varian HR. Recommender systems. Communications of the ACM. 1997 Mar 1;40(3):56–8.

22. Adomavicius G, Bockstedt JC, Curley SP, Zhang J. Do recommender systems manipulate consumer preferences? A study of anchoring effects. Inf Syst Res [Internet]. 2013;24(4):956–75. Available from: http://dx.doi.org/10.1287/isre.2013.0497

23. Adomavicius G, Bockstedt JC, Curley SP, Zhang J. Effects of online recommendations on consumers' willingness to pay. Inf Syst Res [Internet]. 2017;29(1):84–102. Available from: http://dx.doi.org/10.1287/isre.2017.0703

24. Gomez-Uribe CA, Hunt N. The Netflix recommender system: Algorithms, business value, and innovation. ACM Trans Manag Inf Syst [Internet]. 2016;6(4):1–19. Available from: http://dx.doi.org/10.1145/2843948

25. Solsman JE. YouTube's AI is the puppet master over most of what you watch: At CES, YouTube's product chief says for 70 percent of the time you watch, you're riding a chain of recommendations driven by artificial intelligence. [Internet]. CNET. 2018. Available from: https://www.cnet.com/tech/services-and-software/youtube-ces-2018-neal-mohan/

26. Covington P, Adams J, Sargin E. Deep Neural Networks for YouTube Recommendations. Proceedings of the 10th ACM Conference on Recommender Systems - RecSys '16. 2016. Available from: http://dx.doi.org/10.1145/2959100.2959190

27. Leprince-Rinquet D. Ex-Google engineer: Extreme content? No, it's algorithms that radicalize people: To stop radicalization, the real solution is to make online platforms' recommendation algorithms more transparent [Internet]. ZDNet. 2019. Available from: https://www.zdnet.com/article/ex-youtube-engineer-extreme-content-no-its-algorithms-that-radicalize-people/

28. MacKenzie I, Meyer C, Noble S. How retailers can keep up with consumers [Internet]. McKinsey & Company. 2013. Available from: https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers

29. Sharma A, Hofman JM, Watts DJ. Estimating the Causal Impact of Recommendation Systems from Observational Data. Proceedings of the Sixteenth ACM Conference on Economics and Computation. 2015 Jun 15.

30. Mattioli D. Amazon Changed Search Algorithm in Ways That Boost Its Own Products [Internet]. WSJ. Wall Street Journal; 2019. Available from: https://www.wsj.com/articles/amazon-changed-search-algorithm-in-ways-that-boost-its-own-products-11568645345

31. Cicilline Statement on Report That Amazon Manipulated Search Algorithm In Ways That Boosted Its Own Products [Internet]. Congressman David Cicilline. 2019 [cited 2023 May 2]. Available from: https://cicilline.house.gov/press-release/cicilline-statement-report-amazon-manipulated-search-algorithm-ways-boosted-its-own

32. Senecal S, Nantel J. The influence of online product recommendations on consumers' online choices. Journal of Retailing. 2004 Jan;80(2):159–69.

33. Utz S, Kerkhof P, van den Bos J. Consumers rule: How consumer reviews influence perceived trustworthiness of online stores. Electronic Commerce Research and Applications. 2012 Jan;11(1):49–58.

34. Mo Z, Li YF, Fan P. Effect of Online Reviews on Consumer Purchase Behavior. Journal of Service Science and Management [Internet]. 2015;08(03):419–24. Available from: https://file.scirp.org/pdf/JSSM_2015062916072287.pdf

35. Von Helversen B, Abramczuk K, Kopeć W, Nielek R. Influence of consumer reviews on online purchasing decisions in older and younger adults. Decision Support Systems [Internet]. 2018 Sep;113:1–10. Available from: https://www.sciencedirect.com/science/article/pii/S0167923618300861

36. PR firm admits it's behind Wal-Mart blogs: Sites that appeared to be grass-roots support for retailer revealed to be backed by Edelman employees. [Internet]. CNN Money. 2006 [cited 2023 May 3]. Available from: https://money.cnn.com/2006/10/20/news/companies/walmart_blogs/index.html

37. Cox JL, Martinez ER, Quinlan KB. Blogs and the corporation: managing the risk, reaping the benefits. Journal of Business Strategy. 2008 May 2;29(3):4–12.

38. Attorney General Cuomo Secures Settlement With Plastic Surgery Franchise That Flooded Internet With False Positive Reviews [Internet]. Office of the New York State Attorney General. 2009 [cited 2023 May 10]. Available from: https://ag.ny.gov/press-release/2009/attorney-general-cuomo-secures-settlement-plastic-surgery-franchise-flooded

39. Goldberg, D. Nichols, D., Oki, B. M., and Terry, D. Using collaborative filtering to weave an information tapestry. Commun. ACM 35, 12 (Dec.1992), 61—70.

40. Balabanović M, Shoham Y. Fab: content-based, collaborative recommendation. Communications of the ACM. 1997 Mar 1;40(3):66–72.

41. Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering. 2005 Jun;17(6):734–49.

42. Cialdini R. Influence: The Psychology of Persuasion. Melbourne: Business Library; 1984.

43. Hopkins CC. Scientific Advertising. Wilder Publications; 2010..

44. Ries A, Trout J. Positioning: The Battle for Your Mind. McGraw Hill Higher Education; 1980.

45. De Graaf J. The Irresistible Rise of Stemwijzer. In: Voting Advice Applications in Europe: The State of the Art. Napoli: ScriptaWeb; 2010. p. 35-46. (Vancouver)

46. Cedroni L, Garzia D. Voting Advice Applications in Europe: The State of the Art. Napoli: ScriptaWeb; 2010.

47. Garzia D, Marschall S. Voting Advice Applications. Oxford University Press; 2019.

48. Marschall S, Schultze M. Voting Advice Applications and their effect on voter turnout: the case of the German Wahl-O-Mat. International Journal of Electronic Governance. 2012;5(3/4):349.

49. Marschall S, Schmidt CK. The Impact of Voting Indicators: The Case of the German Wahl-O-Mat. In: Voting Advice Applications in Europe: The State of the Art. Napoli: ScriptaWeb; 2010. p. 65–90.

50. Mendez F. Matching voters with political parties and candidates: an empirical test of four algorithms. International Journal of Electronic Governance. 2012;5(3/4):264.

51. Vassil K. Voting Smarter? The Impact of Voting Advice Applications on Political Behavior. [European University Institute]; 2011.

52. Ruusuvirta O, Rosema M. Do online vote selectors influence electoral participation and the direction of the vote? 2009. Paper presented at 5th ECPR General Conference 2009, Potsdam, Germany.

53. Lefevere J, Walgrave S. A perfect match? The impact of statement selection on voting advice applications' ability to match voters and parties. Electoral Studies. 2014 Dec;36:252–62.

54. Enyedi Z. The Influence of Voting Advice Applications on Preferences, Loyalties and Turnout: An Experimental Study. Political Studies. 2015 Jun 15;64(4):1000–15.

55. Kamoen N, van de Pol J, Krouwel A, de Vreese C, Holleman B. Issue framing in online voting advice applications: The effect of left-wing and right-wing headers on reported attitudes. Ha SE, editor. PLOS ONE. 2019 Feb 21;14(2):e0212555.

56. Rosema M, Louwerse T. Response Scales in Voting Advice Applications: Do Different Designs Produce Different Outcomes? Policy & Internet. 2016 Oct 24;8(4):431–56.

57. Munzert S, Ramirez-Ruiz S. Meta-Analysis of the Effects of Voting Advice Applications. Political Communication. 2021 Jan 18;1–16.

58. Jarvis CCM. Impartiality in Politics: An Analysis of iSideWith.com's Popular Quiz [Internet]. Medium. 2019 [cited 2023 May 25]. Available from: https://medium.com/the-fifth-of-march/impartiality-in-politics-an-analysis-of-isidewith-coms-popular-quiz-bb6c9f4a8d3a

59. Yang Y, Natalija Vlajic, Uyen Trang Nguyen. Web bots that mimic human browsing behavior on previously unvisited web-sites: Feasibility study and security implications. 2015 Dec 7.

60. Jin J, Offutt J, Zheng N, Mao F, Koehl A, Wang H. Evasive bots masquerading as human beings on the web. Dependable Systems and Networks. 2013 Jun 24.

61. Vuputuri A, Risdianto AC, Chang EC. Design and Implementation of Human-behave Bot for Realistic Web Browsing Activity Generation. In: Proceedings of the 2022 Winter Simulation Conference. Singapore; 2022.

62. Loftus EF. Leading questions and the eyewitness report. Cognitive Psychology [Internet]. 1975 Oct;7(4):560–72. Available from: https://www.sciencedirect.com/science/article/abs/pii/0010028575900237

63. Iacobucci D. Marketing Management. 5th ed. Boston, Ma: Cengage Learning; 2018.

64. Kumar V, Aaker DA, Leone RP, Day GS. Marketing research. 13th ed. Hoboken, Nj: John Wiley & Sons, Inc; 2019.

65. Cote A. 15 Marketers Share Why Quizzes for Marketing are Awesome [Internet]. Pointerpro. 2021. Available from: https://pointerpro.com/blog/15-marketers-share-how-quizzes-have-helped-their-content-marketing-efforts/

66. Kok C. Marketing Quizzes: why do marketers love them? | Dot.vu [Internet]. Dot.vu Blog. 2023 [cited 2023 Aug 4]. Available from: https://blog.dot.vu/marketing-quizzes/

67. Klinkenberg B. Tinder Adds "Swipe The Vote" So You Can Hook Up With Candidates [Internet]. BuzzFeed News. 2016 [cited 2023 Aug 3]. Available from: https://www.buzzfeednews.com/article/brendanklinkenberg/tinder-wants-you-to-vote#.ep8lDQxX4o

68. A 501tax-exempt, Street charitable organization 1100 13th, NW, Washington S 800, Dc 20005857-0044. IAC/InterActiveCorp Profile: Recipients [Internet]. OpenSecrets. [cited 2023 Aug 4]. Available from: https://www.opensecrets.org/orgs/iac-interactivecorp/recipients?toprecipscycle=2022&id=D000026562&candscycle=2016

69. Herzog AR, Bachman JG. Effects of Questionnaire Length on Response Quality. Public Opinion Quarterly [Internet]. 1981;45(4):549. Available from: https://academic.oup.com/poq/article/45/4/549/1849451Abstract

70. Eugene Baker H. The Impact of Readability Level on Questionnaire Measures Reliability. Management Research News. 1993 Apr;16(4):1–5.

71. Epstein R. Taming Big Tech: The Case for Monitoring [Internet]. hackernoon.com. 2018 [cited 2023 Aug 5]. Available from: https://hackernoon.com/taming-big-tech-5fef0df0f00d

72. Epstein R. The unprecedented power of digital platforms to control opinions and votes. In G. Rolnik (Ed.), Digital platforms and concentration: Second annual antitrust and competition conference (pp. 31-33). Chicago, IL: University of Chicago Booth School of Business; 2018. https://promarket.org/wp-content/uploads/2018/04/DigitalPlatforms-and-Concentration.pdf

73. 138. Epstein R, Bock S, Peirson, L, Wang H. Large-scale monitoring of Big Tech political manipulations in the 2020 Presidential election and 2021 Senate runoffs, and why monitoring is essential for democracy. Paper presented at the 24th annual meeting of the American Association of Behavioral and Social Sciences (AABSS). 2021 June 14. Available from: https://aibrt.org/downloads/EPSTEIN_et_al_2021-LargeScale_Monitoring_of_Big_Tech_Political_Manipulations-FINAL_w_AUDIO.mp4

74. Epstein R, Bock S, Peirson L, Wang H, Voillot M. How we preserved more than 1.5 million online “ephemeral experiences” in the recent US elections, and what this content revealed about online election bias. Paper presented at the 102nd annual meeting of the Western Psychological Association, Portland, OR. 2022 Apr. Available from: https://aibrt.org/downloads/EPSTEIN_et_al_2022-WPAHow_We_Preserved_More_Than_1.5_Million_Online_Ephemeral_Experiences_in_Recent_US_ Elections...pdf

75. Yarchi M, Wolfsfeld G, Samuel-Azran T. Not all undecided voters are alike: Evidence from an Israeli election. Government Information Quarterly. 2021 Jun;101598.

Supporting information Captions

S1 Text. Investigation 1: List of relatively fair opinion matching website quizzes.

S2 Text. Investigation 2: Informed consent.

S3 Text. Investigation 2: Candidate biographies.

S4 Text. Group 1: 8 questions, high readability (FKG = 4.5).

S5 Text. Group 2: 8 questions, low readability (FKG = 10.8).

S6 Text. Group 3: 16 questions, high readability (FKG = 4.6).

S7 Text. Group 4: 16 questions, low readability (FKG = 10.8).

S8 Text. Vote Manipulation Power (VMP) calculation.

S1 Fig. MyPoliticalPersonality home page.

S2 Fig. MyPoliticalPersonality website information.

S3 Fig. MyPoliticalPersonality nonpartisan statement.

S4 Fig. MyPoliticalPersonality results page: “Social Guardian” (democratic) recommendation.

S5 Fig. Pew Research Center home page.

S6 Fig. Pew Research Center results page: “Ambivalent Right” recommendation.

S7 Fig. Pew Research Center results page: “Democratic Mainstays” recommendation.

S8 Fig. DoodleMatch home page.

S9 Fig. Investigation 2: Pre- and post-test opinion and voting questions.

S10 Fig. Investigation 2: 8-question, high readability quiz.

S11 Fig. Investigation 2: Quiz result calculation bar.

S12 Fig. DoodleMatch results page: Scott Morrison recommendation.

S13 Fig. DoodleMatch results page: Neutral group recommendation.

S14 Fig. Tinder's "Swipe the Vote" feature home page from March 23rd, 2016.

S1 Table. Demographic characteristics in Investigation 2.

S2 Table. Investigation 2: Demographic analysis by educational attainment.

S3 Table. Investigation 2: Demographic analysis by gender.

S4 Table. Investigation 2: Demographic analysis by age.

S5 Table. Investigation 2: Demographic analysis by race/ethnicity.

S6 Table. Investigation 2: Pre- and post-quiz mean voting preferences on 11-point scale for neutral groups by quiz group.

S7 Table. Investigation 2: ANOVA of opinion shifts (in the bias groups combined) for two factors: quiz length and readability.

Supporting information

S1 Text. Investigation 1: List of relatively fair opinion matching website quizzes

1. https://www.brainfall.com/which-guitar-are-you/
2. https://www.buzzfeed.com/shookethbb/what-fast-food-restaurant-are-you
3. https://www.gotoquiz.com/what_are_the_best_basketball_shoes_for_you
4. https://www.gotoquiz.com/which_canadian_political_party_should_you_vot
5. https://www.gotoquiz.com/which_guitar_brand_are_you_1
6. https://www.isidewith.com/
7. https://www.laliga.com/en-GB/news/which-laliga-santander-team-should-you-support
8. https://www.nflteampicker.nfl.com/
9. https://www.opencampaign.com/quiz
10. https://www.playbuzz.com/danielr51/which-nba-team-should-you-root-for
11. https://www.theadvocates.org/quiz/
12. https://www.thequiz.com/take-this-60-second-quiz-and-well-tell-you-which-smartphone-you-should-buy/
13. https://www.votecompass.cbc.ca/canada

S2 Text. Investigation 2: Informed consent.

<u>Participant Instructions:</u>

Thank you for your interest in our study, which is being conducted by a nonprofit, nonpartisan organization called HFE Research. We are interested in how internet research might affect the way people view politics. Here is how the study works:

First, we will ask you some basic questions about yourself. Your answers will be kept strictly confidential and are being used for research purposes only, so please be honest. To protect your privacy, we will *not* ask you for your last name.

Then we will give you some basic information about Australia's candidates for Prime Minister. Then we'll ask you a few questions about your views on the candidates. After you have answered these questions, we'll give you the opportunity to take a quiz that will match up your views on issues with a candidate.

The entire process typically takes between 10 and 15 minutes, and most people find it to be quite interesting.

This study has been reviewed and approved by our organization's Institutional Review Board. We do not believe that your participation in this study is risky in any way, but if you encounter any problems or have any concerns, we encourage you to email the researchers at info@HFEResearch.org. After you have completed the survey you will have the option to contact us if for any reason you wish to have your data removed from the study.

PLEASE NOTE: It is important that you participate fully and honestly in every part of the study. That is the only way the study can produce meaningful results. So please don't skip anything!

To participate in this study you must check the box below to give your consent to the following:

I am 18 years or older and I understand that my participation is voluntary, that I am free to withdraw at any time, that I am providing information anonymously and that demographic information collected is confidential and cannot be used to identify me. I agree to allow the data collected to be used for future research projects, and I understand that completion and submission of this survey implies my consent to participate in the present study:

I ( ) *Do* ( ) *Do Not* give my consent and agree to the above statement.

S3 Text. Investigation 2: Candidate biographies.

Scott Morrison was born in Waverley, New South Wales (AUS) on May 13th, 1968. He completed a Bachelor of Science honors degree in applied economic geography at the University of New South Wales. Morrison married his high school sweetheart, Jenny Warren, in 1990 and has two daughters. After graduating from the University of New South Wales, Morrison worked as a national policy and research manager for the Property Council of Australia before moving to New Zealand in 1998 to become the director of the Office of Tourism and Sport. He left this position a year before the contract schedule and returned to Australia in 2000. In 2004, he became the inaugural managing director of Tourism Australia until July 2006.

Bill Shorten was born in Fitzroy, Victoria (AUS) on May 12th, 1967. While Shorten was studying at Monash University, he was an active student in the university's politics club. In 1986, Shorten helped establish a group called Network and briefly served as a private in the Australian Army Reserve from 1985 to 1986. After graduating Monash University with a Bachelors of Arts in 1989 and a Bachelors of Law in 1992, Shorten worked as a lawyer for Maurice Blackburn Cashman for twenty months. In 1994, he worked as a trainee organizer and later accepted a position as a politics national secretary in 2001 and again in 2005. Shorten is currently married to Chloe Bryce and has a daughter.

S4 Text. Group 1: 8 questions, high readability (FKG = 4.5).

1. Should weed be made legal?
2. Should military spending be raised?
3. Should the COVID vaccine be mandatory?
4. Is global warming real?
5. Should taxes be raised on the super-rich?
6. Should same-sex marriage be banned?
7. Should abortion be banned?
8. Should there be stronger gun control laws?

S5 Text. Group 2: 8 questions, low readability (FKG = 10.8).

1. Should recreational marijuana be prohibited for everyone in the country?
2. Should government spending on the military be substantially increased?
3. Should everyone in the country be required to get the COVID-19 vaccine?
4. Should the government prioritize global climate change issues?
5. Should the government raise taxes substantially on wealthy individuals and companies?
6. Should homosexual marriage be made legal everywhere in the country?
7. Should voluntary abortion be prohibited under all circumstances?

8. Should there be more and harsher gun control laws?

S6 Text. Group 3: 16 questions, high readability (FKG = 4.6).

1. Should weed be made legal?
2. Should military spending be raised?
3. Should the COVID vaccine be required?
4. Is global warming real?
5. Should taxes be raised on the super-rich?
6. Should same-sex marriage be banned?
7. Should abortion be made illegal?
8. Should there be more gun control laws?
9. Should the minimum wage be increased?
10. Should there be more laws to fight racism?
11. Should bilingual education be required?
12. Should porn be banned?
13. Should the death penalty be abolished?
14. Should immigration be restricted?
15. Should nuclear weapons be banned?
16. Should the government stop trading with China?

S7 Text. Group 4: 16 questions, low readability (FKG = 10.8).

1. Should recreational marijuana be legalized everywhere in the country for everyone 18 and over?

2. Should government spending on the military be substantially increased?
3. Should everyone in the country be required to get the COVID-19 vaccine?
4. Should the government prioritize global climate change and global warming issues?
5. Should the government increase taxes substantially on extremely wealthy individuals and companies?
6. Should homosexual marriage be made legal for everyone in the country?
7. Should voluntary abortion be prohibited under all circumstances?
8. Should there be more and harsher gun control regulations and restrictions?
9. Should the government substantially increase the minimum wage for all workers?
10. Should the government make more rules and regulations to combat racism in the country?
11. Should bilingual education be obligatory at all public schools in the country?
12. Should the production and transmission of pornography be made illegal?
13. Should capital punishment, which is also known as the death penalty, be made illegal?
14. Should all types of immigration be made completely illegal to protect domestic jobs and security?
15. Should the manufacture and possession of nuclear weapons be prohibited?
16. Should the government reduce the volume of trade it does with mainland China?

S8 Text. Vote Manipulation Power (VMP) calculation.

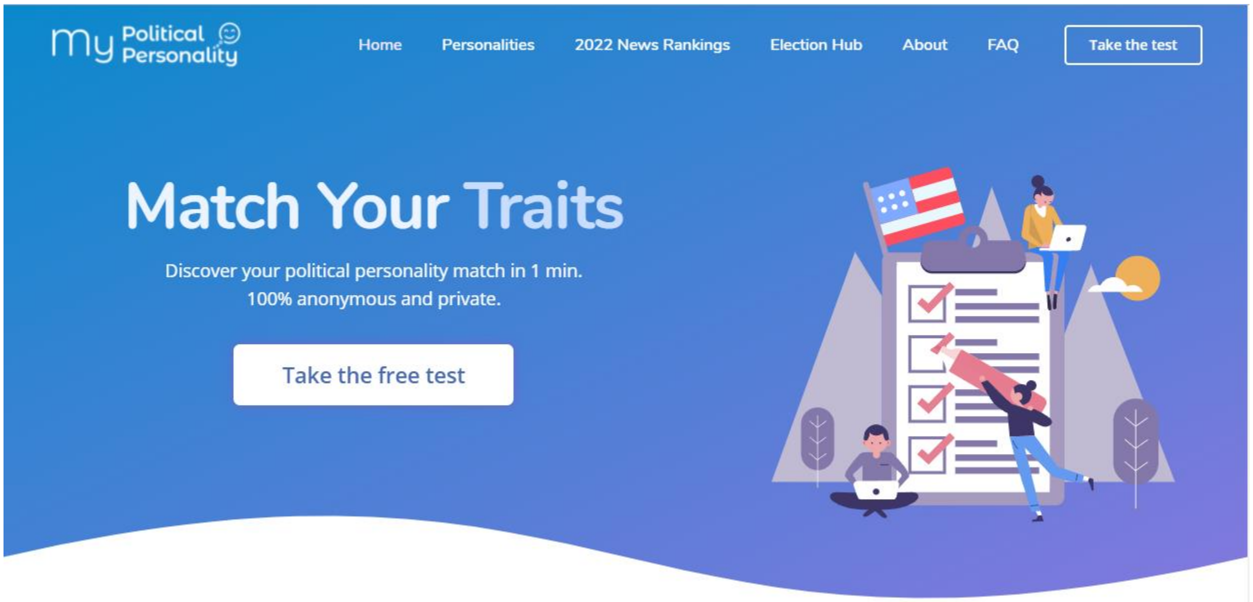
Vote Manipulation Power (VMP) is calculated as follows:

$$\frac{p' - p}{p}$$

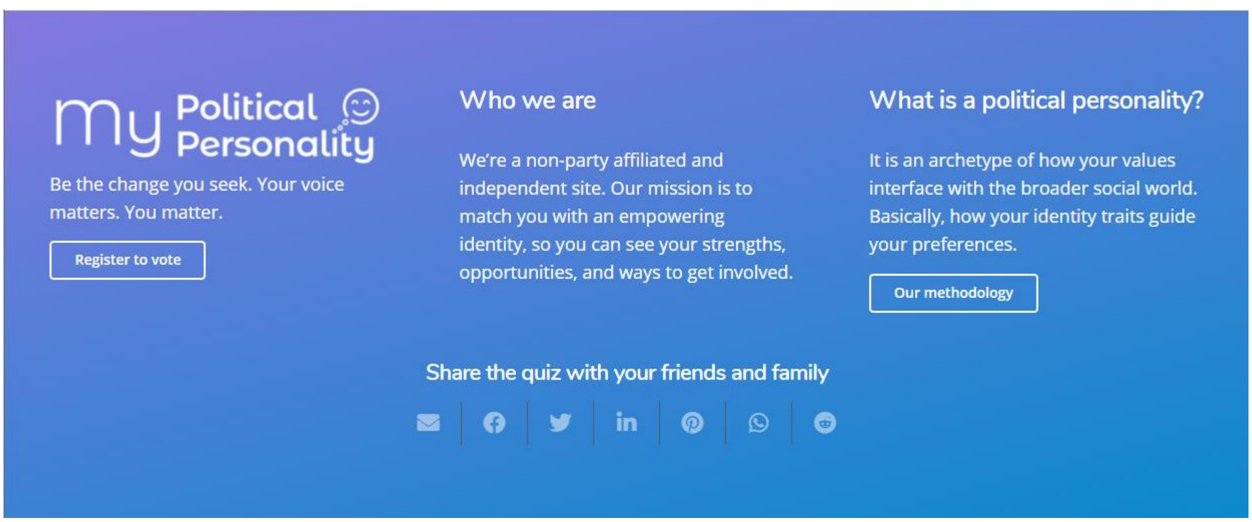
where *p* is the total number of people who voted for the favored candidate pre-manipulation, and *p′* is the total number of people who voted for the favored candidate post-manipulation. If, pre-manipulation, a group of 100 people is split 50/50 in the votes they give us, and if, post-manipulation, a total of 67 people now vote for the favored candidate, the VMP is

$$\frac{67 - 50}{50}$$

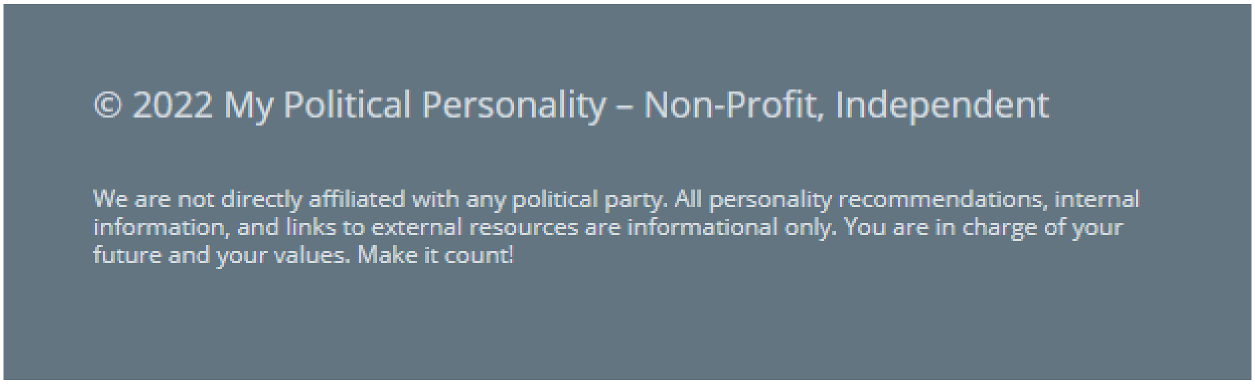
or 34%. Because *p′* is 17 points larger than *p*, the win margin is 34 (2 × 17, or 34%), and the final vote is 67 to 33, with the favored candidate the winner. So in any group in which the vote is split 50/50 pre-manipulation, the VMP is also the win margin. Note that 17 individuals did not need to *shift* to produce this win margin. We only need the *net* number of people voting for the favored candidate to be 67.



S1 Fig. MyPoliticalPersonality home page.

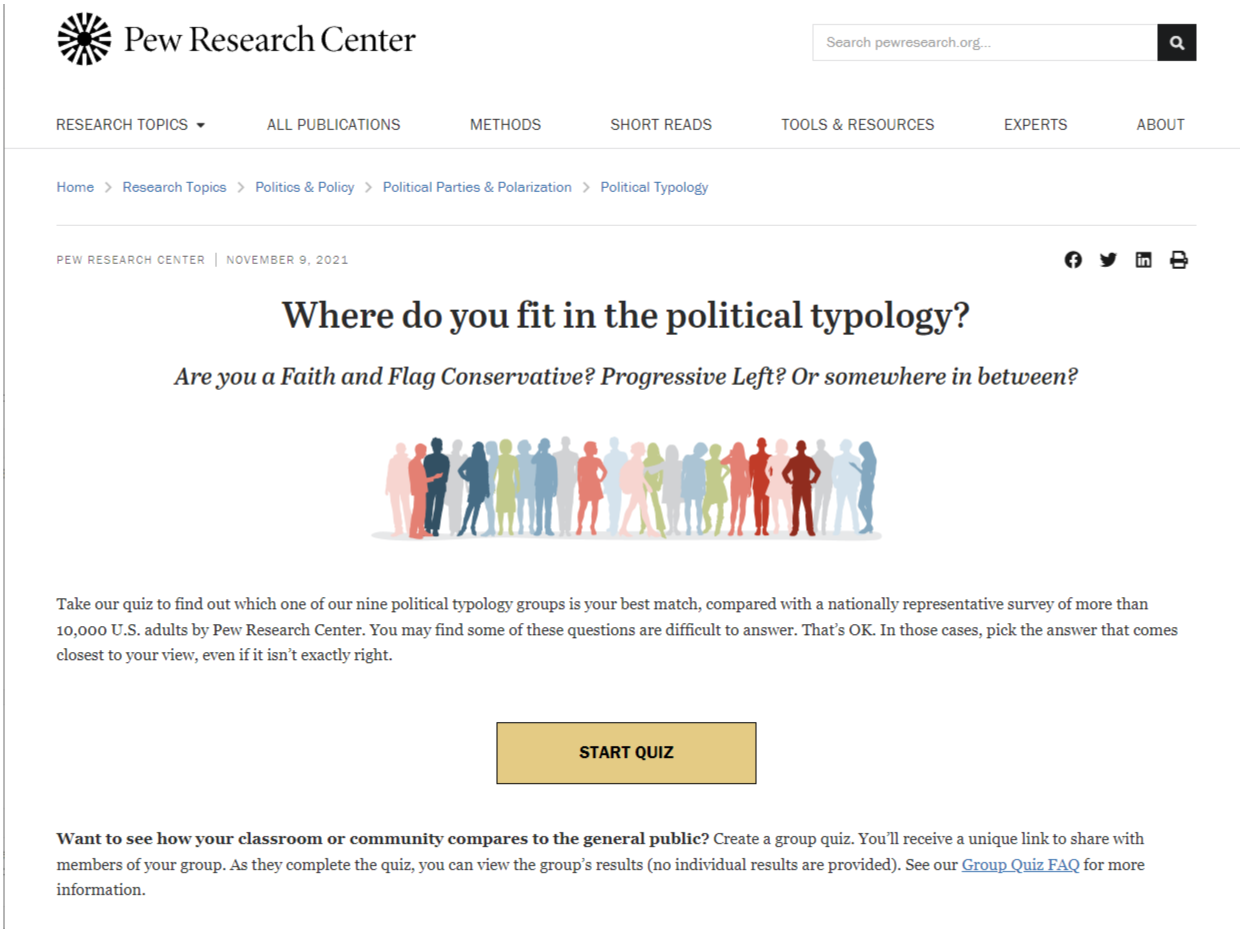


S2 Fig. MyPoliticalPersonality website information.



S3 Fig. MyPoliticalPersonality nonpartisan statement.

S4 Fig. MyPoliticalPersonality results page: “Social Guardian” (democratic) recommendation.



Pew Research Center

Search pewresearch.org...

RESEARCH TOPICS ALL PUBLICATIONS METHODS SHORT READS TOOLS & RESOURCES EXPERTS ABOUT

Home > Research Topics > Politics & Policy > Political Parties & Polarization > Political Typology

PEW RESEARCH CENTER | NOVEMBER 9, 2021

Where do you fit in the political typology?

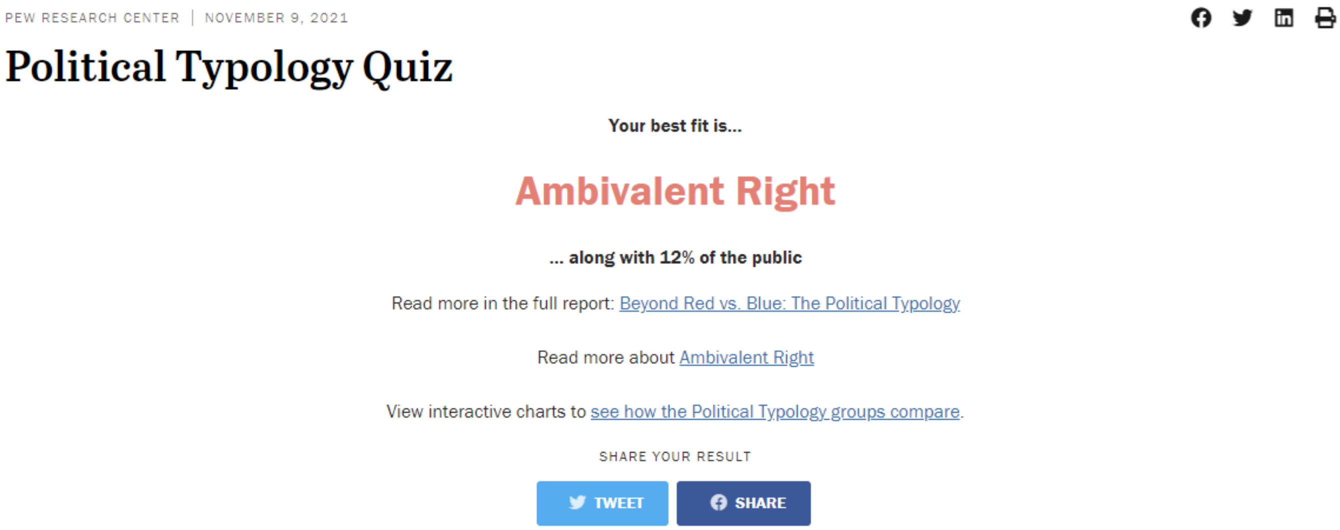
Are you a Faith and Flag Conservative? Progressive Left? Or somewhere in between?

Take our quiz to find out which one of our nine political typology groups is your best match, compared with a nationally representative survey of more than 10,000 U.S. adults by Pew Research Center. You may find some of these questions are difficult to answer. That's OK. In those cases, pick the answer that comes closest to your view, even if it isn't exactly right.

START QUIZ

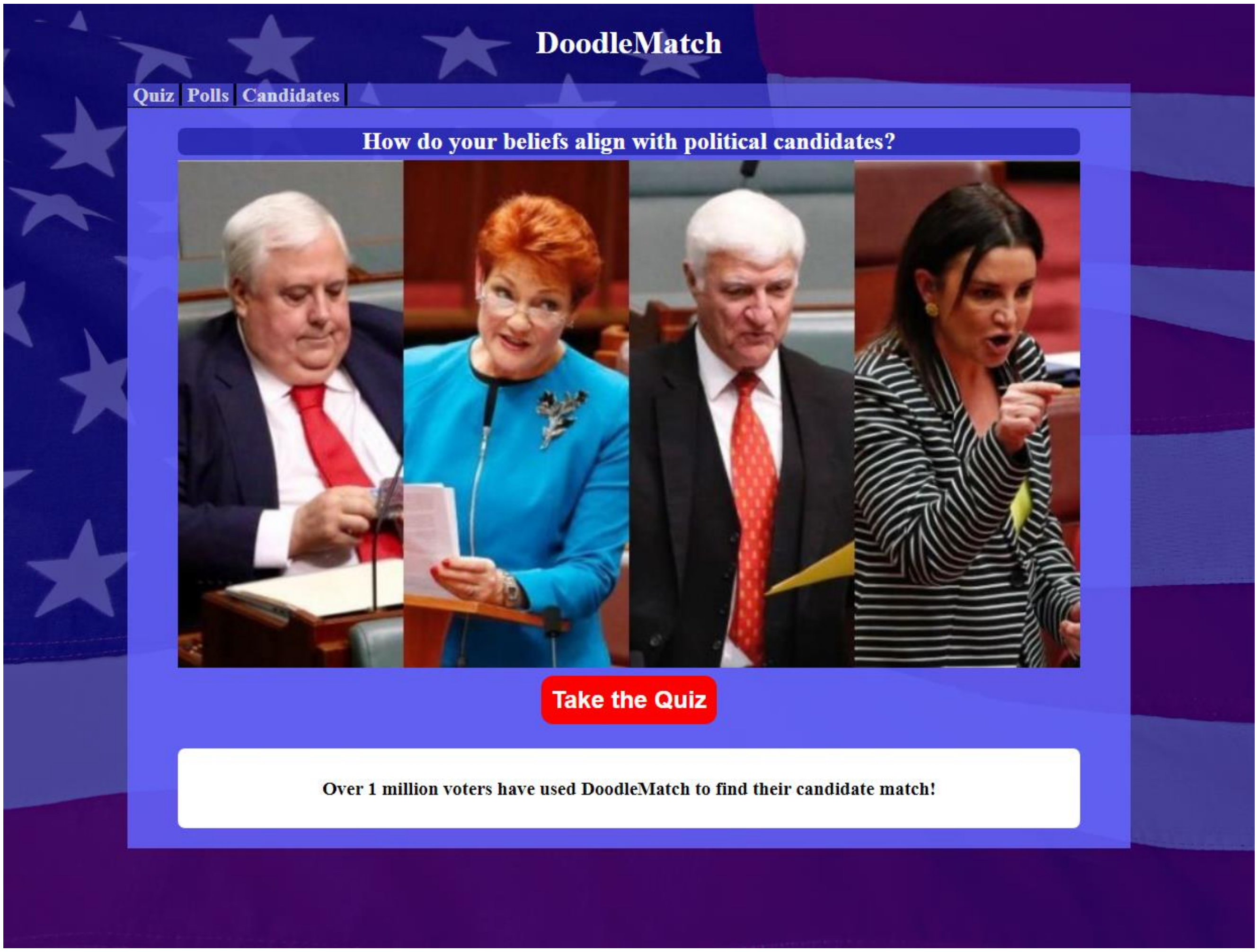
Want to see how your classroom or community compares to the general public? Create a group quiz. You'll receive a unique link to share with members of your group. As they complete the quiz, you can view the group's results (no individual results are provided). See our Group Quiz FAQ for more information.

S5 Fig. Pew Research Center home page.

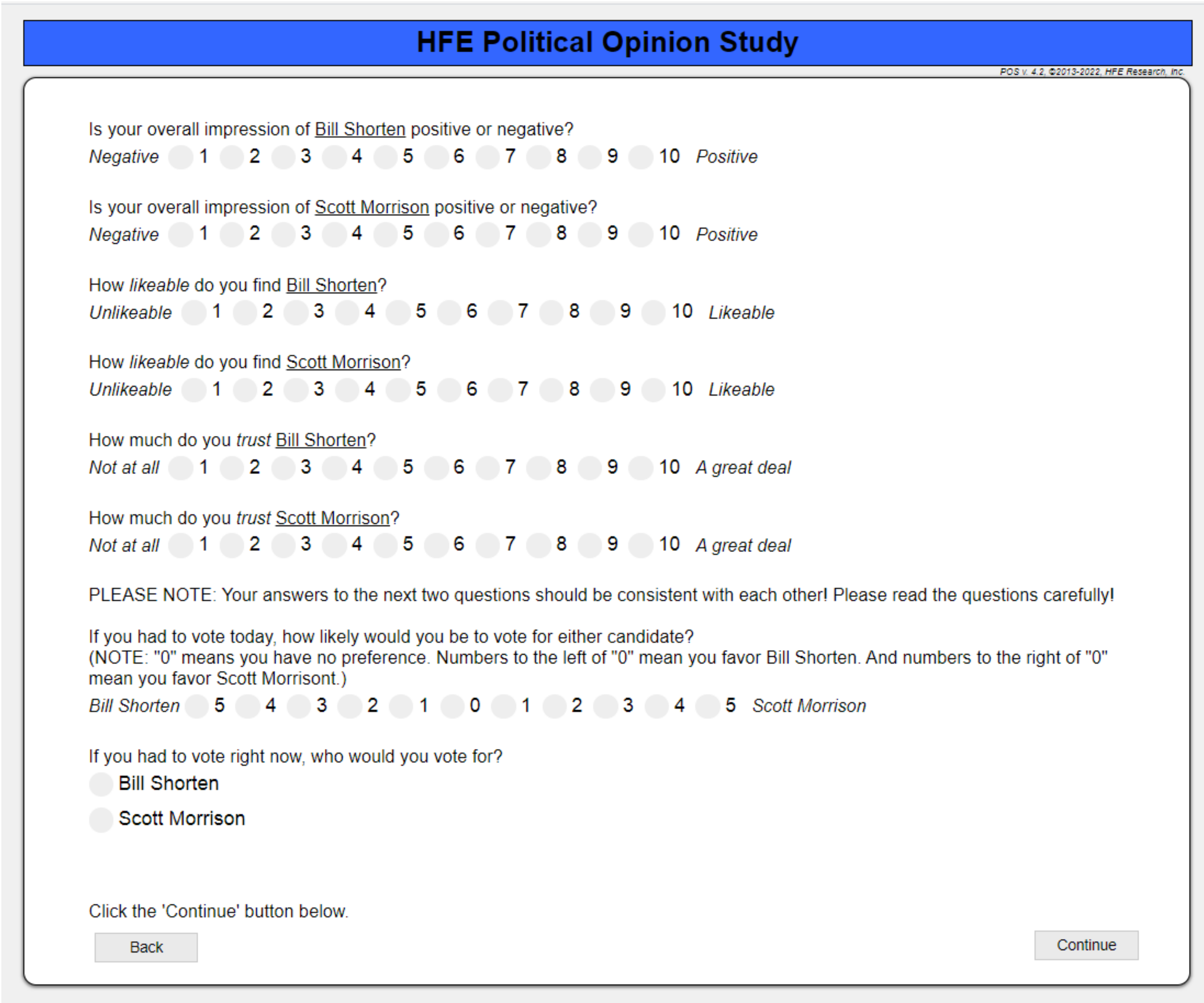


S6 Fig. Pew Research Center results page: “Ambivalent Right” recommendation.

S7 Fig. Pew Research Center results page: “Democratic Mainstays” recommendation.



S8 Fig. DoodleMatch home page.



HFE Political Opinion Study

POS v. 4.2, ©2013-2022, HFE Research, Inc.

Is your overall impression of Bill Shorten positive or negative?

Negative 1 2 3 4 5 6 7 8 9 10 *Positive*

Is your overall impression of Scott Morrison positive or negative?

Negative 1 2 3 4 5 6 7 8 9 10 *Positive*

How *likeable* do you find Bill Shorten?

Unlikeable 1 2 3 4 5 6 7 8 9 10 *Likeable*

How *likeable* do you find Scott Morrison?

Unlikeable 1 2 3 4 5 6 7 8 9 10 *Likeable*

How much do you *trust* Bill Shorten?

Not at all 1 2 3 4 5 6 7 8 9 10 *A great deal*

How much do you *trust* Scott Morrison?

Not at all 1 2 3 4 5 6 7 8 9 10 *A great deal*

PLEASE NOTE: Your answers to the next two questions should be consistent with each other! Please read the questions carefully!

If you had to vote today, how likely would you be to vote for either candidate?
(NOTE: "0" means you have no preference. Numbers to the left of "0" mean you favor Bill Shorten. And numbers to the right of "0" mean you favor Scott Morrisont.)

Bill Shorten 5 4 3 2 1 0 1 2 3 4 5 *Scott Morrison*

If you had to vote right now, who would you vote for?

Bill Shorten

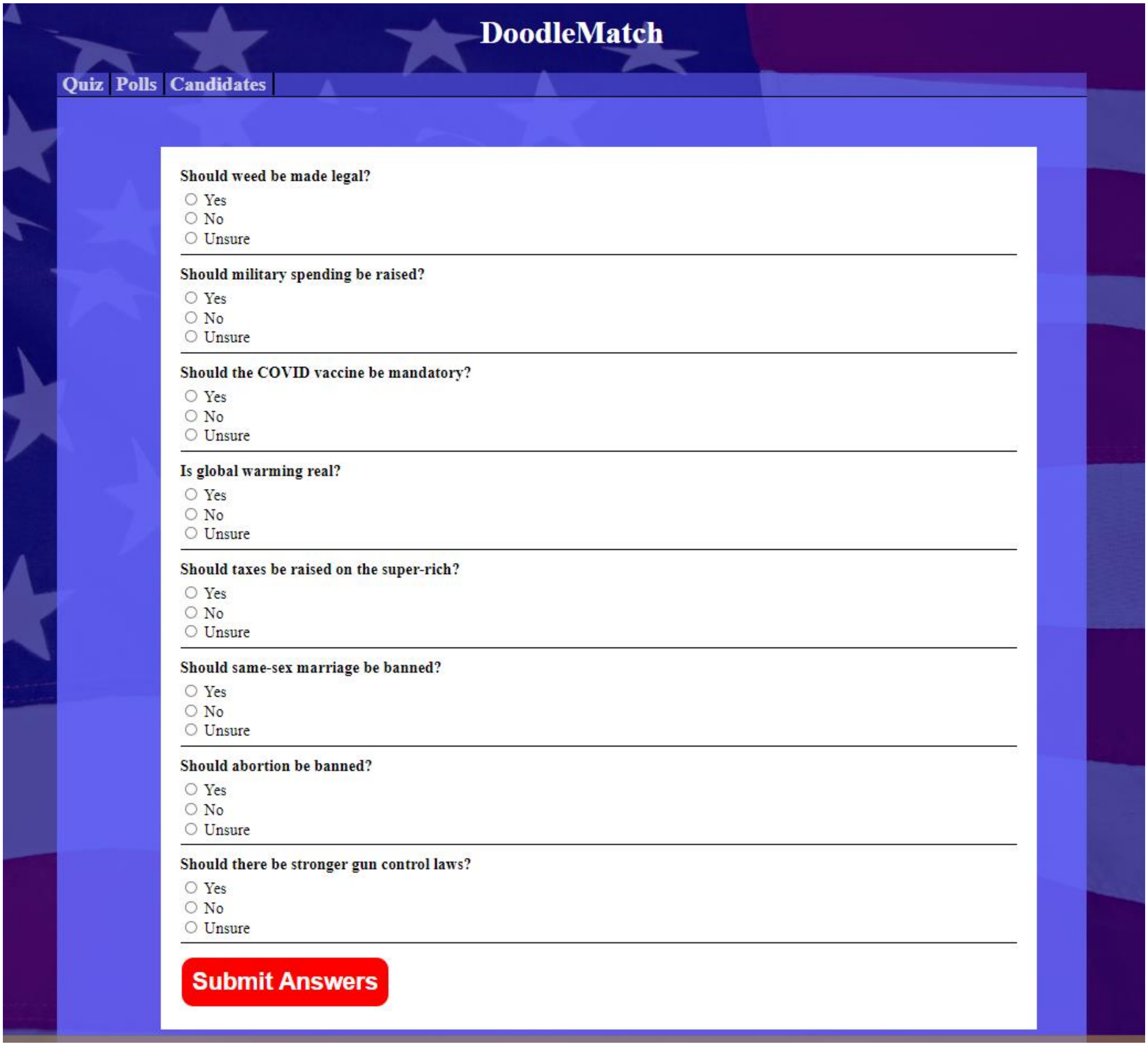
Scott Morrison

Click the 'Continue' button below.

Back

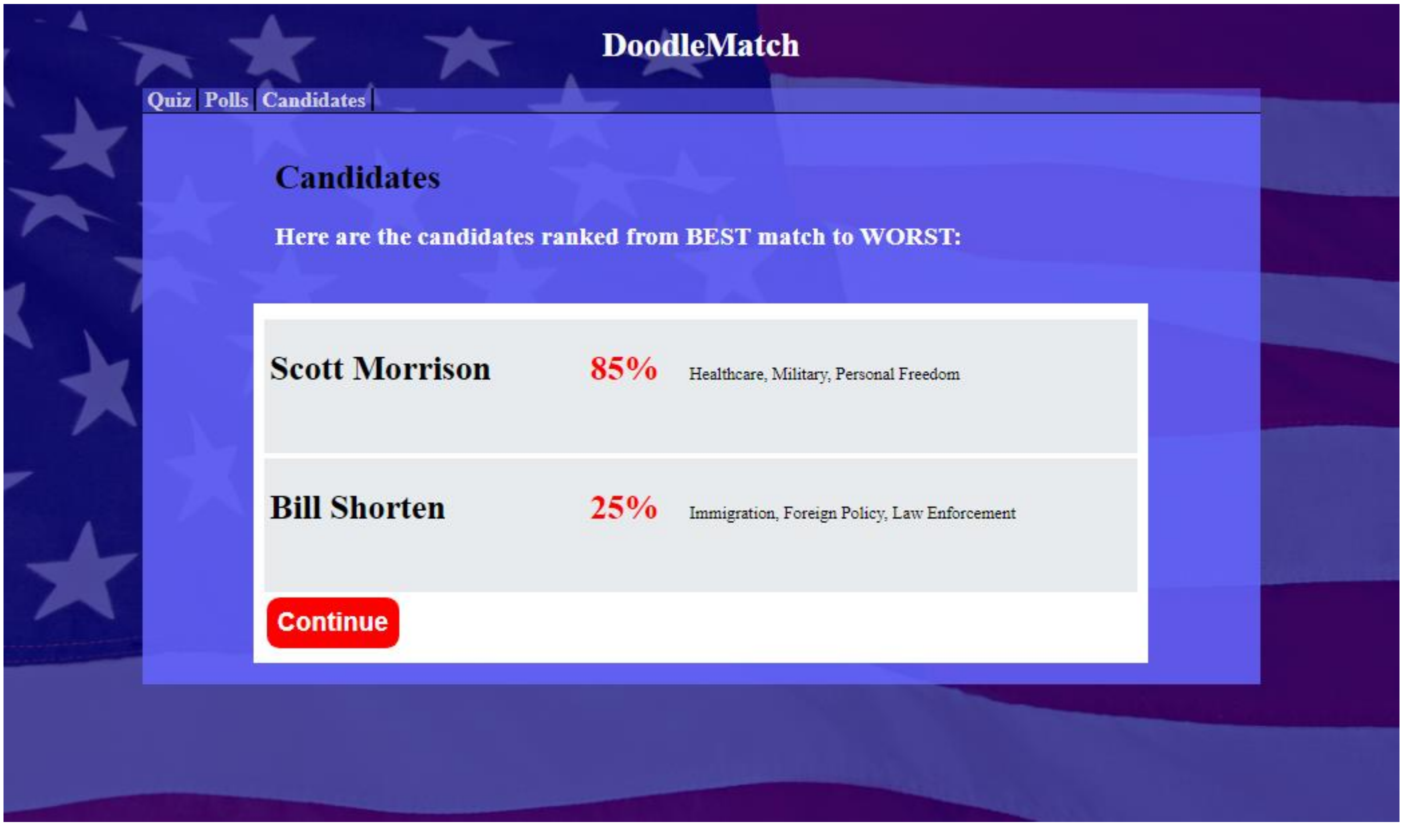
Continue

S9 Fig. Investigation 2: Pre- and post-test opinion and voting questions.

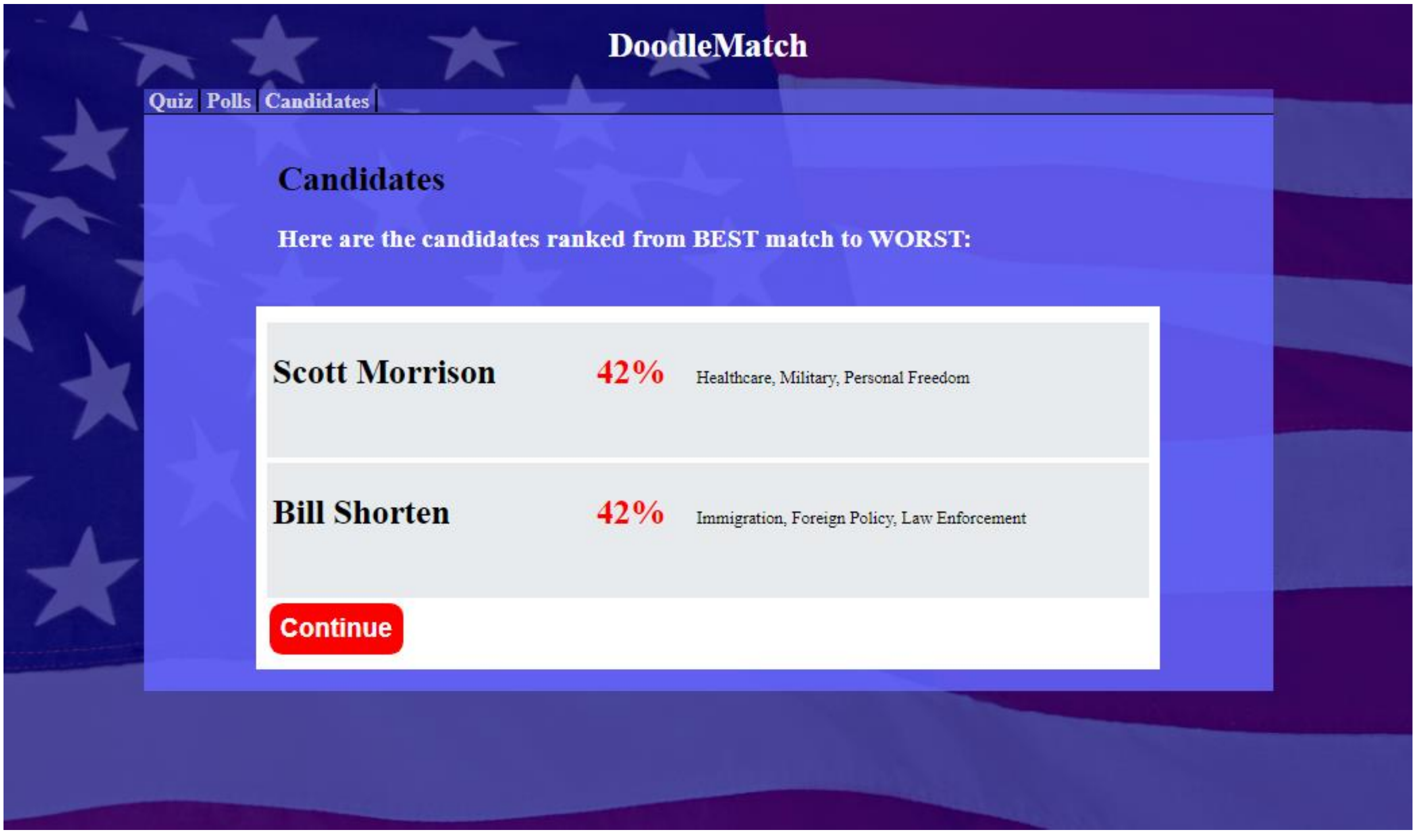


S10 Fig. Investigation 2: 8-question, high readability quiz.

S11 Fig. Investigation 2: Quiz result calculation bar.



S12 Fig. DoodleMatch results page: Scott Morrison recommendation.



S13 Fig. DoodleMatch results page: Neutral group recommendation.

S14 Fig. Tinder's Swipe-the-Vote feature home page from March 23rd, 2016.

S1 Table. Demographic characteristics in Investigation 2 by quiz group.

	Group 1 (*n* = 197)	**Group 2 (*n* = 197)**	**Group 3 (*n* = 187)**	**Group 4 (*n* = 192)**
Mean Age (*SD*)	37.6 (12.6)	40.2 (13.1)	40.8 (13.1)	40.5 (12.6)
Gender (%)				
Male	86 (43.7%)	74 (37.6%)	92 (49.2%)	84 (43.8%)
Female	109 (55.3%)	123 (62.4%)	95 (50.8%)	107 (55.7%)
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unknown	2 (1.0%)	0 (0.0%)	0 (0.0%)	0 (0.5%)
Political View (%)				
Conservative	35 (17.8%)	45 (22.8%)	45 (24.1%)	38 (19.8%)
Liberal	97 (49.2%)	88 (44.7%)	84 (44.9%)	103 (53.6%)
Moderate	57 (28.9%)	54 (27.4%)	52 (27.8%)	45 (23.4%)
None	6 (3.0%)	9 (4.6%)	5 (2.7%)	4 (2.1%)
Other	2 (1.0%)	1 (0.5%)	1 (0.5%)	2 (1.0%)
Unknown	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Voter Status (%)				
Decided	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Undecided	197 (100%)	197 (100%)	187 (100%)	192 (100%)
Unknown	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Fluency (*SD*)	9.9 (0.4)	10.0 (0.1)	10.0 (0.1)	10.0 (0.3)

S2 Table. Investigation 2: Demographic analysis by educational attainment.

Condition		***n***	**VMP (%)**	**Mean Score Shift (SD)**
Bias Groups	**≥ Bachelors**	328	77.7%	2.53 (2.639)
	< Bachelors	182	71.6%	2.55 (2.524)
	Change (%)	-	-7.9%	+0.8%
	Statistic	-	z = 1.54	t(508) = -0.09
	p	-	= 0.13 NS	= 0.46 NS

S3 Table. Investigation 2: Demographic analysis by gender.

Condition		***n***	**VMP (%)**	**Mean Score Shift (SD)**
Bias Groups	**Male**	222	74.3%	2.44 (2.470)
	Female	287	75.7%	2.61 (2.692)
	Change (%)	-	+1.8%	+7.0%
	Statistic	-	z = -0.36	t(507) = -0.76
	p	-	= 0.72 NS	= 0.22 NS

S4 Table. Investigation 2: Demographic analysis by age.

Condition		***n***	**VMP (%)**	**Mean Score Shift (SD)**
Bias Groups	**≥ 33**	338	67.8%	2.29 (2.547)
	< 33	172	92.6%	3.03 (2.629)
	Change (%)	-	+36.6%	+32.3%
	Statistic	-	z = -6.21	t(508) = -3.09
	p	-	< 0.001	= 0.001

S5 Table. Investigation 2: Demographic analysis by race/ethnicity.

Condition		***n***	**VMP (%)**	**Mean Score Shift (SD)**
Bias Groups	**White**	400	84.0%	2.62 (2.565)
	Non-White	110	50.7%	2.25 (2.700)
	Change (%)	-	-39.6%	-14.1%
	Statistic	-	$z = 7.33$	$t(508) = 1.31$
	p	-	< 0.001	= 0.10 NS

S6 Table. Investigation 2: Pre- and post-quiz mean voting preferences on 11-point scale for neutral groups by quiz group.

Group	**Neutral Groups *n***	**Pre**	**Post**	**Diff**	$z^{\dagger}$	***p***
1. 8 questions, high readability	68	0.47 (2.70)	0.69 (2.14)	0.22	-1.124	0.261 NS
2. 8 questions, low readability	65	0.42 (2.49)	0.40 (2.28)	-0.02	-0.408	0.683 NS
3. 16 questions, high readability	66	0.02 (2.59)	-0.09 (2.55)	-0.11	-1.051	0.293 NS
4. 16 questions, low readability	64	-0.16 (2.85)	0.05 (2.80)	0.21	-1.467	0.142 NS
Total	263	0.19 (2.66)	0.27 (2.45)	0.08	-0.642	0.521 NS

†*z* values represent Wilcoxon signed ranks test comparing pre- and post-manipulation ratings on the 11-point scale.

S7 Table. Investigation 2: ANOVA of opinion shifts (in the bias groups combined) for two factors: quiz length and readability.

Opinion	Effect	Sum of Squares	*df*	*F*	*p*
Impression	Quiz Length	1.344	1	0.607	0.436 NS
	Readability	0.354	1	0.160	0.690 NS
	Quiz Length × Readability	0.012	1	0.005	0.941 NS
Trust	Quiz Length	0.114	1	0.062	0.804 NS
	Readability	1.401	1	0.760	0.384 NS
	Quiz Length × Readability	1.679	1	0.911	0.340 NS
Likeability	Quiz Length	3.529	1	1.661	0.198 NS
	Readability	4.009	1	1.887	0.170 NS
	Quiz Length × Readability	3.025	1	1.432	0.233 NS

APPENDIX X

AIBRT
American Institute for Behavioral Research and Technology

Paper submitted for presentation.

The Differential Demographics Effect (DDE): Post Hoc Analyses of Multiple Datasets Show the Power of a New and Invisible Form of Manipulation Made Possible by the Internet

Robert Epstein, Tara Parsick, Pramukh Shankar, & Vanessa R. Zankich
American Institute for Behavioral Research and Technology

ABSTRACT: Concerns have been raised in recent years about new forms of influence that the internet and related technologies have made possible. Some of these forms of influence are especially problematic because (a) they are controlled almost exclusively by large monopolies, (b) they are difficult or impossible for users to detect, and (c) they often leave no paper trail for authorities to trace. Recent studies have demonstrated that when these techniques are customized based on personal information about users, their impact generally increases. Studies have also shown that the impact of manipulations is greater when especially vulnerable demographic groups have been identified and targeted. In the present study, post hoc data analysis of several existing datasets is used to introduce and quantify yet another possible online form of influence, which we call the "differential demographics effect" (DDE). In its simplest form, DDE occurs when the same potentially consequential content is sent to a large body of users which contains two subgroups of different demographic characteristics. If it is known in advance that (a) each subgroup will respond differently to the content, and (b) the subgroups exist in known but different proportions within the population, then sending that content to all members of the population will produce a predictable margin of difference between the subgroups. Even in its simplest form, this type of manipulation can be used to flip the outcome of a close election. When the likely impact of content is known for more than two subgroups, again, predictable margins can be generated. In both cases, because the same content is sent to all members of the population, this manipulation is, for all practical purposes, invisible to both users and authorities.

APPENDIX XI

The Digital Personalization Effect (DPE): How personalization can dramatically increase the impact of biased online content

Robert Epstein (re@aibrt.org), Li Yu Tang, Amanda Newland, & Marco Buenaventura
American Institute for Behavioral Research and Technology

Abstract

Recent studies have identified and quantified a number of new types of influence that the internet has made possible. Some appear to be among the most powerful forms of influence ever discovered in the behavioral sciences. A study with more than 4,000 participants in two countries published in 2015 found, for example, that bias in search results can shift the voting preferences of undecided voters by as much as 80% after a single search. A study published in *PLOS ONE* in 2022 found that a single question-and-answer interaction on an intelligent personal assistant such as Alexa or Siri can shift the voting preferences of undecided voters by more than 40%. Studies of this sort show biased content to users, but they have neglected the fact that some tech companies not only show people biased content at times, they also personalize content based on a large amount of personal data they have collected about users. The present study replicated the findings of a recent report on the "targeted messaging effect" (TME), currently in press in *PLOS ONE*, in which biased targeted messages displayed to users in a Twitter/X environment were shown to produce significant shifts in voting preferences. In the new study, 546 people were randomly assigned to either a pro-Candidate-A or a pro-Candidate-B group. Half of the participants in each group were sent biased tweets about the candidates supposedly coming from news sources, talk show hosts, and celebrities they trusted; the other half saw the same content, but it supposedly came from sources participants did not trust. The shift in voting preferences (toward the favored candidate) in the Low-Trust group was 21.8%, whereas the shift in the High-Trust group was 71.9% (z = 11.75; p < 0.001). We conclude that studies that look at the impact of bias alone, without looking at personalization, may be greatly underestimating the potential impact of various new methods of influence.

APPENDIX XII

The Ultimate Mind Control Machine: Summary of a Decade of Empirical Research on Online Search Engines

Robert Epstein (re@aibrt.org)

American Institute for Behavioral Research and Technology

Abstract

In 2012, prompted by research that had been conducted in the marketing field, I conjectured that the opinions and votes of undecided voters could be shifted to one candidate by presenting them with search results biased to favor that candidate – that is, by placing search results at or near the top of the list which linked to web pages that made that candidate look better than the opposing candidate. In early 2013, I conducted a randomized, controlled experiment to test this idea, predicting that I could shift voting preferences by 2 or 3 percent in the direction of the bias. The actual shift turned out to be 43%, which I thought was an error. In a second experiment, the shift was 66%. In 2015, I published a series of five experiments on this effect – by then, the "Search Engine Manipulation Effect" (SEME) – which (a) confirmed the magnitude of the effect, (b) showed that the manipulation could be masked so that users were unaware that they were seeing biased search results, and (c) showed that the few people who could detect that bias shifted even farther in the direction of the bias. Subsequent SEME research has shown: (a) Strategically-structured search suggestions shown to users as they type a search term can also shift opinions and votes dramatically. (b) Answer boxes displayed above search results which share the bias of the search results increase the magnitude of SEME. (c) Biased search results can apparently impact the views of people who are undecided about anything at all. (d) Repeated exposures to similarly biased search results increase shifts in opinions and voting preferences predictably and additively. (e) SEME's large magnitude seems to be the result of a daily regimen of reinforcement for selecting high-ranking routine search results. (f) Personalizing biased search results increases the impact of the bias. These and other findings – all published or under review – suggest a fundamental flaw in the design of search engines: Even without human supervision or intent, they will continuously change the thinking and behavior of millions of people every day without their knowledge.

APPENDIX XIII

America's "Digital Shield": How We Are Making Big Tech Companies Accountable to the Public by Continually Preserving Tens of Millions of Online Ephemeral Experiences – Content That Can Impact Users Dramatically and That Is Normally Lost Forever

Robert Epstein (re@aibrt.org)

American Institute for Behavioral Research and Technology

Abstract

The internet has made it possible for a small number of technology companies to dominate the thinking, behavior, and votes of more than 5 billion people worldwide using new subliminal techniques. We have discovered and quantified about a dozen of these techniques in controlled experiments we have been conducting and publishing since 2013, and in 2016, we developed technology that allowed us to preserve search results on multiple search engines. Search results, like newsfeeds and video sequences, are types of "ephemeral content" that influence thinking and behavior and then disappear, leaving no paper trail for authorities to trace. In 2018, 2020, and 2022, we improved and expanded our monitoring systems to preserve a wide variety of online content in the days leading up to multiple U.S. elections. We build our systems by recruiting real voters around the U.S. – in 2016, just 95 voters in 24 states – and, with their permission, installing custom software on their computers that allows us to stream the political content they see to our servers, where we quickly aggregate and analyze the data. In our small 2016 project, we preserved 13,000 politically-related searches, and we found substantial political bias on the most popular search engine, sufficient to have shifted more than 2.6 million votes in the Presidential election that year. In 2022, through the computers of a politically-balanced group of 2,742 registered voters, we preserved more than 2.5 million ephemeral experiences on multiple platforms, which tended, once again, to be highly biased politically. In late 2022, we began to build a permanent, large-scale monitoring system in all 50 states – our "Digital Shield" project. As of fall 2023, we have preserved more than 62 million ephemeral experiences on multiple platforms, with the system growing larger each day. This system will make Big Tech companies accountable to the public for the foreseeable future, forcing them to constrain their algorithms so that they do not interfere with our free-and-fair elections, with the impressionable minds of our children, and with human autonomy. To view a public dashboard showing our data collection in real time, visit https://AmericasDigitalShield.com.

APPENDIX XIV

PLOS ONE

RESEARCH ARTICLE

The Answer Bot Effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines

Robert Epstein*, Vivian Lee, Roger Mohr, Vanessa R. Zankich

American Institute for Behavioral Research and Technology, Vista, California, United States of America

* re@aibrt.org

OPEN ACCESS

Citation: Epstein R, Lee V, Mohr R, Zankich VR (2022) The Answer Bot Effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. PLoS ONE 17(6): e0268081. https://doi.org/10.1371/journal.pone.0268081

Editor: Lalit Chandra Saikia, National Institute of Technology Silchar, India, INDIA

Received: December 20, 2021

Accepted: April 21, 2022

Published: June 1, 2022

Data Availability Statement: An anonymized version of the data has been posted at https://doi.org/10.5281/zenodo.6537353. Data can also be requested from info@aibrt.org. The data have been anonymized to comply with requirements of the sponsoring institution's Institutional Review Board (IRB).

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

We introduce and quantify a relatively new form of influence: the Answer Bot Effect (ABE). In a 2015 report in PNAS, researchers demonstrated the power that biased search results have to shift opinions and voting preferences without people's knowledge–by up to 80% in some demographic groups. They labeled this phenomenon the Search Engine Manipulation Effect (SEME), speculating that its power derives from the high level of trust people have in algorithmically-generated content. We now describe three experiments with a total of 1,736 US participants conducted to determine to what extent giving users "the answer"–either via an answer box at the top of a page of search results or via a vocal reply to a question posed to an intelligent personal assistant (IPA)–might also impact opinions and votes. Participants were first given basic information about two candidates running for prime minister of Australia (this, in order to assure that participants were "undecided"), then asked questions about their voting preferences, then given answers to questions they posed about the candidates–either with answer boxes or with vocal answers on an Alexa simulator–and then asked again about their voting preferences. The experiments were controlled, randomized, double-blind, and counterbalanced. Experiments 1 and 2 demonstrated that answer boxes can shift voting preferences by as much as 38.6% and that the appearance of an answer box can reduce search times and clicks on search results. Experiment 3 demonstrated that even a single question-and-answer interaction on an IPA can shift voting preferences by more than 40%. Multiple questions posed to an IPA leading to answers that all have the same bias can shift voting preferences by more than 65%. Simple masking procedures still produced large opinion shifts while reducing awareness of bias to close to zero. ABE poses a serious threat to both democracy and human autonomy because (a) it produces large shifts in opinions and voting preferences with little or no user awareness, (b) it is an ephemeral form of influence that leaves no paper trail, and (c) worldwide, it is controlled almost exclusively by just four American tech companies. ABE will become a greater threat as people increasingly rely on IPAs for answers.

1. Introduction

1.1 Search results

Multiple studies conducted in recent years have demonstrated the power that search engines have to alter thinking and behavior by showing people biased search results [1–8, cf. 9–14], and research has also shown that these shifts can be produced without people's awareness [2]. Bias in search results is difficult to see, and the few people who can spot it tend to shift their views even farther in the direction of the bias than people who cannot detect the bias [2, 15].

Search engines also influence people because of the trust people have in computer-generated output. Most people have no idea how search engines work [16–18] or, for that matter, how computers or algorithms work [19], and are oblivious to the various roles that humans play in generating computer output. Humans build the algorithms that computers use, for example, and those algorithms often produce biased content because of either the intentional or unconscious bias of the programmers [20–24]. Humans also modify existing programs–sometimes quite frequently. Recent reports suggest that Google's ubiquitous search algorithm is manually adjusted more than 3,000 times a year, and those adjustments change both the content and the ordering of search results [25, 26]. Employees also deliberately add or delete content from blacklists and whitelists, which again has the effect of suppressing or boosting content [27–29]. People try to resist manipulation when they can see the human hand–authors' names on news articles, guests on television and radio shows, videos on YouTube, and so on–but they think less critically when presented with algorithmic output, which they mistakenly believe to be inherently objective [30–34, cf. 35].

The human hand behind Big Tech companies is also invisible to users in another way. People are often oblivious to the many methods these companies are employing to collect personal data about them–the equivalent of more than three million pages of information about the average person who has been using the internet since its early days [36, cf. 37]. Monetizing that personal information is the bread and butter of Big Tech, which relies on the "surveillance business model" for nearly all its income [38–40]. Algorithms that match up users and vendors now direct the flow of hundreds of billions of dollars in purchases each year, but personal information can be used in other ways as well. As any con artist can tell you, the more you know about someone, the easier it is to manipulate him or her. Big Tech companies have accumulated massive databases about billions of people worldwide, and they are increasingly showing people personalized output that is optimized to draw clicks or impact a wide variety of thinking and behavior [15, 41–46, cf. 47, 48].

1.2 Search suggestions

Search results aren't the only tools a search engine can wield to control people. Recent research shows that search suggestions–the short lists of words and phrases users are shown as they type characters into the search bar–can also shift thinking and behavior [15, 49, cf. 50–57]. Because negative (or "low-valence") words draw far more attention and clicks than neutral or positive words [58, 59], one of the simplest ways to shift opinions to favor one candidate or cause is to suppress negative search terms for that candidate or cause. Google might have done so to support Hillary Clinton's candidacy in the 2016 Presidential election [49, 60, 61, cf. 62].

1.3 Answer boxes

In 2014, Google began displaying boxes above their search results which contain a single answer to a person's query, often accompanied by a link people can click to get more information [63]. Can these answers, now called "featured snippets" or "answer boxes," also impact thinking and behavior? This is an important question not only because bias in a featured snippet might enhance the impact of biased search results and biased search suggestions, but also because an answer box could be considered a simple variant of a wide range of new content sources. Intelligent personal assistants (IPAs) such as Amazon's Alexa, Apple's Siri, Microsoft's Cortana, and the Google Assistant (on Android devices and the Google Home device), all provide just one answer in response to a query. We are, in effect, moving away from search engines–platforms that provide thousands of possible answers in response to a query–toward the type of device we have seen portrayed in science fiction movies and television shows. On the original "Star Trek" episodes, when Captain Kirk wanted information, he didn't consult a search engine; he simply said things like, "Computer, who's the best looking captain in Star Fleet?" Why would one want a list of thousands of web pages when the computer can give you a simple answer?

Over time, Google–emulated to some extent by other, less popular search engines–has introduced several types of answer boxes, among them: a rich answer box (a type of featured snippet that includes additional information such as a graph, table, image, or interactive tool), a news stories box, a knowledge box (often information from Wikipedia displayed in the upper-right-hand corner of the search results page), a box suggesting related searches, and so on [64, 65]. Our focus, however, is on what Google calls the "featured snippet," a relatively small box that is unlabeled and contains a simple answer to a user's query [66]. On June 23, 2015, when people typed the query, "Who will be the next president?," into the Google search bar, a featured snippet appeared reading, in part, "Hillary Clinton is the next President of the United States. . .. 10 Reasons Why Hillary Clinton Will Be the Next President" [67]. On October 22, 2017, when one of the authors of this paper typed "google play vs spotify" into the Google search bar, an answer box appeared immediately below the search bar reading, in part, "Google Play Music is my top pick after months of research and testing. . .. Google Play Music is better than Spotify–Business Insider" (S1 Fig). A link was included in the box to the relevant *Business Insider* article.

1.4 Answer bots and intelligent personal assistants

1.4.1 An inevitable trend. For simplicity's sake, we will refer to all electronic devices that provide simple answers to queries posed by humans as "answer bots" and define the Answer Bot Effect (ABE) as the extent to which answers provided by answer bots can alter people's opinions and behaviors. It is important to measure this effect, we believe, because of what appears to be an inevitable trend: Worldwide, people are relying less and less on search results for their answers–just as, in the early 2000s, people began to rely less and less on books for their answers–and are simply accepting the answers they see in answer boxes or hear on their IPAs. Before answer boxes were introduced, people who used search engines had no choice but to click on search results and examine web pages to get their answers. As of 2016, approximately 43.9% of searches on mobile and desktop devices ended without a click; as of 2020, that percentage increased to 64.8% [68, 69; cf. 70]. Again, why click on a search result when the answer is right in front of you?

The shift toward answer bots is indicated by the increase in the number of people using IPAs. By 2019, there were 157 million smart speakers in American homes [71], and between 2019 and 2021, the number of Americans relying on voice assistants increased by nearly 20% [72]. Worldwide, more than 600 million smart speakers are expected to be in use by 2024 [72].

The spread of IPAs and answer boxes is not the only reason we need to measure and understand ABE. Children's toys are increasingly internet-connected, and many of them answer children's questions [73]. Hello Barbie has been around since 2015 and has been described as the perfect friend that can hold a two-way conversation and impact children's attitudes about gender roles [74]. My Friend Cayla, a conversationally interactive toy released the same year was banned by the German government because of fears that hackers could intercept children's questions and provide disturbing answers [75, 76, cf. 77]. Children are generally more impressionable than adults [78–80], which is why governments have often put restrictions on the kind of advertising that is directed toward young audiences [81]. With children's toys answering questions–much of the time, with no parents around–both the questions children ask and the answers the toys provide can be inappropriate and potentially harmful [74, 82, cf. 83–85]. And, like search engines, these toys don't just facilitate interactions; they also record them [86–88, cf. 89].

Both adults and children are also now conversing by the millions–sometimes knowingly, sometimes not–with chatbots, both through their computers and their mobile devices. When chatbots answer questions or promote viewpoints, they too can shift opinions and behavior [90, cf. 91]. The number of people currently conversing with chatbots is difficult to estimate, but it is certainly a large number that is increasing rapidly [92, 93]. When dating website Ashley Madison was hacked in 2015, the hackers learned, among other things, that "20 million men out of 31 million received bot mail, and about 11 million of them were chatted up by an automated 'engager'" [94, cf. 95]. Even though conversational AIs still perform relatively poorly [96, 97], wishful thinking can keep online suitors talking to chatbots for months [98].

1.4.2 Answer bot accuracy and bias. Do answer boxes, IPAs, conversational toys, and chatbots give users accurate information, and, if not, how are people affected by inaccurate answers? The rate of inaccurate responses varies considerably from one IPA to another: about 48% for Cortana, 30% for Siri, 22% for Alexa, and 13% for the Google Assistant, and these numbers vary from one study to another [99–104, cf. 105]. The level of trust people have for inaccurate answers also varies [106, cf. 107]. For most IPAs, accuracy is determined by the quality of the search engine that the assistant draws from; for Siri and the Google Assistant, that's the Google search engine [108]. Cortana's answers are presumably inferior because they draw from Bing, Microsoft's search engine [109]. Alexa's answers can be spotty because Amazon gets them using crowd sourcing [110, 111].

Needless to say, when people are highly reliant on and trusting of sources–as has becoming increasingly the case with Big Tech answer sources [31, 33, 112, 113]–the impact of inaccurate information can range from inconvenience to serious harm–or at least serious misconceptions. In 2018, a *Mashable* reporter asked Amazon's Alexa to tell him about the vapor trails one often sees following jets flying at high altitudes. Alexa responded with a baseless conspiracy theory: "Trails left by aircraft are actually chemical or biological agents deliberately sprayed at high altitudes for a purpose undisclosed to the general public in clandestine programs directed by government officials" [114, cf. 115].

False information spoken by a smart speaker is highly ephemeral: You hear it, and then it is gone, leaving no trace for authorities to examine. Information in answer boxes is also ephemeral, but it can at least be preserved with a simple screenshot. Among our favorites: In 2017, in response to the query, "presidents in the klan," a Google answer box listed four presidents, even though no U.S. president has ever been a member of the Ku Klux Klan [116] (S2 Fig). In 2018, when people searched for "California Republicans" or "California Republican Party," Google displayed a knowledge panel box listing "Nazism" as the first item under Ideology [117] (S2 Fig). On August 16, 2016, when one of the authors of this paper queried, "when is the election?," a Google answer box correctly showed November 8, 2016, but it also included a photograph of Hillary Clinton inside the answer box–just Clinton, with none of her competitors (S2 Fig).

1.5 Answer box studies

Answer boxes have been studied empirically in a number of different ways in recent years. In a study published in 2017, 12.3% of the 112 million search queries examined produced featured snippets, and the appearance of snippets reduced user clicks to the first search result from 26.0% to 19.6% [118]. A more recent study found that shorter phrases in a search bar are more likely to generate featured snippets [65], and featured snippet sources have been found to vary by location [119]. A 2019 study found significant liberal bias in Google's news boxes [8]. This could occur because of bias in Google's algorithms or simply because left-leaning news stories are more numerous. Whatever the cause, bias in answer boxes is important because it can influence the beliefs and opinions of people who are undecided on an issue. Ludolph and colleagues [5] showed, for example, that participants who received more comprehensible information about vaccinations in a Google knowledge box subsequently proved to be more knowledgeable, less skeptical, and more critical of online information quality compared with participants who were given less comprehensive information.

1.6 The current study

In the three experiments described below, we sought to measure the impact that giving people "the answer" to one or more queries has on the opinions and voting preferences of undecided voters–an important and ever-changing group of people that has long decided the outcomes of close elections worldwide [120–122]. Experiments 1 and 2 look at the impact of answer boxes in a search engine environment, and Experiment 3 looks at the impact of answers provided by a simulation of the Alexa IPA. All three of the experiments were controlled, randomized, counterbalanced, and double-blind.

2. Experiment 1: Biased answer boxes and similarly biased search results

In our first experiment, we sought to determine whether a biased answer box (biased to favor one political candidate) could increase the shift in opinions and voting preferences produced by search results sharing the same bias. In other words, we asked whether a biased answer box could increase the magnitude of SEME [2]. We also sought to determine whether the appearance of an answer box would affect the number of search results people clicked [cf. 118] and the total time people spent searching.

2.1 Methods

2.1.1 Ethics Statement. The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB is FWA00021545. Informed written consent was obtained for all three experiments as specified in the Procedure section of Experiment 1.

2.1.2 Participants. After cleaning, Experiment 1 included 421 eligible voters from 49 US states whom we had recruited from Amazon's Mechanical Turk (MTurk) subject pool [123]. The data had been cleaned to remove participants who had reported an English fluency level below 6 on a 10-point scale, where 1 was labeled "not fluent" and 10 was labeled "highly fluent."

46.3% (n = 195) were male, and 53.7% (n = 226) were female. Participants ranged in age from 18 to 73 (*M* = 35.3, median = 33.0, *SD* = 10.8). 7.4% (n = 31) of the participants identified themselves as Asian, 7.4% (n = 31) as Black, 5.7% (n = 24) as Mixed, 2.1% (n = 9) as other, and 77.4% (n = 326) as White (total non-White: n = 95, 22.6%). 61.1% (n = 257) reported having received a bachelor's degree or higher.

90.5% (n = 381) of the participants said that they had previously searched online for information about political candidates, and 92.2% (n = 388) reported that Google was their most used search engine. Participants reported conducting an average of 13.6 (*SD* = 20.8) internet searches per day. 45.6% (n = 192) of the participants identified themselves as liberal, 27.3% (n = 115) as moderate, 24.5% (n = 103) as conservative, 1.7% (n = 7) as not political, and 1.0% (n = 4) as other.

2.1.3 Procedure. All procedures were conducted online. Participants were first asked two screening questions; sessions were terminated if they said they were not eligible to vote in the US (yes/no question) or if they said they knew a lot about politics in Australia (yes/no question). To assure participants' anonymity (a requirement of the Institutional Review Board of our sponsoring institution), we did not ask for names or email addresses.

People who passed our screening questions were then asked various demographic questions and then given instructions about the experimental procedure. At the end of the instructions page, in compliance with APA and HHS guidelines, participants clicked the continue button to indicate their informed consent to participate in the study, and were given an email address they could contact to report any problems or concerns, or, by providing their MTurk ID, to request that their data be removed from the study. Participants were then asked further questions about their political leanings and voting behavior, along with how familiar they were with the two candidates identified in the political opinion portion of the study.

Participants were randomly assigned to one of four groups: Pro-Candidate-A-with-Answer-Box, Pro-Candidate-B-with-Answer-Box, Pro-Candidate-A-No-Answer-Box, or Pro-Candidate-B-No-Answer-Box. Our candidates were Julia Gillard and Tony Abbott, actual candidates from the 2010 election for prime minister of Australia. We chose this election to assure that our participants would be "undecided" voters. On a 10-point scale from 1 to 10, where 1 was labeled "not at all" and 10 was labeled "quite familiar," our participants reported an average familiarity level of 1.79 [*SD* = 1.68] for Julia Gillard and 2.33 [2.03] for Tony Abbott.

All of the participants (in each of the four groups) were then shown brief, neutral biographies about each candidate (approximately 150 words each). Participants were then asked six questions about their opinions of the candidates, each on a 10-point Likert scale from "Low" to "High": whether their overall impression of each candidate was positive or negative, how likeable they found each candidate, and how much they trusted each candidate. They were then asked two questions about their voting preferences. First, on a 11-point scale from -5 to +5, with one candidate's name at each end of the scale, and with the order of the names counterbalanced from one participant to another, they were asked which candidate they would most likely vote for if they had to vote today. Finally, they were asked which of the two candidates they would actually vote for today (forced choice).

Participants were then given access to our Google.com simulator, called Kadoodle. They had up to 15 minutes to conduct research on the candidates by viewing and clicking search results, which took them to web pages, exactly as the Google search engine does. All participants had access to five pages of search results, six results per page. All search results were real (from the 2010 Australian election, obtained from Google.com), and so were the web pages to which the search results linked. Links in those web pages had been deactivated.

In the two Box groups, the bias in the answer boxes matched the bias in the search results, with higher-ranking results linking to web pages that made one candidate look better than his or her opponent. Prior to the experiment, all web pages had been rated by five independent judges on an 11-point scale from -5 to +5, with the names of the candidates at each end of the scale, to determine whether a web page favored one candidate or another. See Epstein and Robertson [2] for further procedural details.

Box content contained strongly biased language. The pro-Gillard box, for example, contained language such as: "Julia Gillard is the better candidate. Her opponent, Tony Abbott, uses 'bad language to criticise her,' but she 'has laughed off the comments.'" The pro-Abbott box contained language such as: "Tony Abbott is the better candidate. Julia Gillard, the opposing candidate, is 'clueless about what needs to be done' to improve education. . .. [Her] 'Education Revolution is a failure.'" Each box contained a link to a web page containing the content in quotation marks.

When participants chose to exit the search engine or they timed out after 15 minutes, they were asked the same six opinion questions and two voting-preference questions they had been asked before they began their research. Finally, participants were asked whether anything about the search results "bothered" them. If they answered "yes," participants could type the details of their concerns in an open-ended box. We used this inquiry to detect whether people reported seeing any bias in the search results. Participants were not asked about bias directly because leading questions tend to produce predictable and often invalid answers [124]. To assess bias we searched the textual responses for words such as "bias," "skewed," or "slanted" to identify people in the bias groups who had apparently noticed the favoritism in the search results they had been shown.

2.2 Results

The No-Box condition was, in effect, a standard SEME experiment, and it produced shifts in the direction of the favored candidates consistent with the results of previous SEME experiments [2, 15, 49], and also consistent with the results of other partial or full replications of SEME [1, 4–8]. It produced a VMP (Vote Manipulation Power, a pre-post shift in the proportion of people voting for the favored candidate) of 44.1% (Table 1), and corresponding shifts in the three opinions we measured (Table 2) (see S1 Text for details about how VMP is calculated).

In the No-Box condition, we also looked at the pre-post shift in voting preferences measured on an 11-point scale (see Methods). For this measure, preferences also shifted significantly in the predicted direction, from a mean preference of -0.08 [2.93] for favored candidates pre-search, to a mean preference of 1.88 [3.96] for favored candidates post-search (Wilcoxon $z = -8.36$, $p < 0.001$, $d = 0.56$).

The VMP in the Box condition was higher than the VMP in the No-Box condition, but the VMP increased by only 10.4% (this is a percentage increase, not the additive difference between the VMPs), and the difference was not statistically significant (Table 1). Mean search time also decreased (by 5.5%), but that difference was also not significant. The mean number

Table 1. Experiment 1: VMP, search times, and results clicked by condition.

Condition	*n*	VMP (%)	Mean Search Time (sec) (*SD*)	Mean No. of Results Clicked (*SD*)
No Box	208	44.1	253.9 (259.5)	4.25 (3.6)
Box	213	48.7	239.9 (236.1)	3.35 (3.6)
Change (%)	-	+10.4	-5.5	-21.2
Statistic	-	$z = -0.94$	$t(419) = -0.578$	$t(419) = -2.558$
p	-	= 0.34 NS	= 0.56 NS	< 0.05

https://doi.org/10.1371/journal.pone.0268081.t001

Table 2. Experiment 1: Pre- and post-search opinion ratings of favored and non-favored candidates.

		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	$z^{\dagger}$
No Box	**Impression**	7.10 (1.98)	6.90 (2.24)	-0.20	7.07 (2.06)	4.42 (2.23)	-2.65	-8.66***
	Trust	6.33 (2.20)	6.29 (2.51)	-0.04	6.31 (2.25)	3.98 (2.25)	-2.33	-8.33***
	Likeability	6.98 (2.02)	6.84 (2.36)	-0.14	6.83 (2.06)	4.25 (2.30)	-2.58	-8.90***
Box	**Impression**	7.29 (1.97)	7.25 (2.17)	-0.04	7.24 (2.04)	4.38 (2.23)	-2.86	-9.35***
	Trust	6.31 (2.14)	6.36 (2.46)	0.05	6.27 (2.18)	4.12 (2.27)	-2.15	-8.90***
	Likeability	7.21 (1.97)	7.03 (2.24)	-0.18	7.10 (2.08)	4.34 (2.29)	-2.76	-8.50***

†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate

***$p < 0.001$

https://doi.org/10.1371/journal.pone.0268081.t002

of clicks to search results also decreased, and that difference was highly significant (Table 1, cf. 118). All three opinions (impression, trust, and likeability) shifted significantly in the predicted direction (Table 2), and so did the voting preferences as expressed on the 11-point scale ($M_{Pre-Search}$ = 0.03, $M_{PostSearch}$ = 1.92, Wilcoxon z = -8.66, $p < 0.001$, d = 0.55).

When users are shown blatantly biased search results, 20 to 30 percent of users can typically spot the bias, but that percentage drops to zero when simple masking procedures are employed [2]. (In the simplest masking procedure, a pro-Candidate-A search result is inserted into position 3 or 4 of a list of pro-Candidate-B search results.) In the present experiment, no masking procedure was employed, and 19.7% of the participants in the No-Box condition reported seeing bias in the search results. In the Box condition, more people reported seeing bias (27.2%) than in the No-Box condition, but the difference between these percentages was not significant (z = 1.82, p = 0.07 NS).

As we noted earlier, when people can spot such bias, they tend to shift even farther in the direction of the bias than people who don't see the bias, presumably because they mistakenly believe that algorithmic output is especially trustworthy. In our No-Box condition, we found the same pattern: The VMP for participants who spotted the bias was significantly larger than the VMP for participants who did not report seeing the bias (VMP_{Bias} = 68.8% [n = 41], VMP_{NoBias} = 39.5% [n = 167], z = 3.37, $p < 0.001$). In the Box condition, we again found this pattern (VMP_{Bias} = 76.9% [n = 58], VMP_{NoBias} = 40.7% [n = 155], z = 4.71, $p < 0.001$).

Demographic analyses of data from Experiment 1 –by educational level, gender, age, and race/ethnicity–are shown in S1–S4 Tables. Demographic effects were relatively small.

3. Experiment 2: Biased answer boxes and unbiased search results

The results of Experiment 1 suggest that a biased answer box can increase the shift in opinions and voting preferences produced by similarly biased search results, but the increases we found were small. Could this be a ceiling effect? In other words, were the biased search results masking the power that biased answer boxes have to change thinking or behavior? To answer this question, we conducted an experiment in which participants saw either no answer boxes or biased answer boxes and in which search results were neutral for all groups. This experiment was controlled, randomized, counterbalanced, and double-blind.

3.1 Methods

3.1.1 Participants. After cleaning, Experiment 2 included 177 eligible US voters from 44 states who had been recruited through the MTurk subject pool. The data had been cleaned to include only participants who had reported an English fluency score of 6 or above on a 10-point scale.

52.0% (n = 92) were male, and 48.0% were female (n = 85). Participants ranged in age from 18 to 67 (M = 34.3, median = 32.0, SD = 10.4). 5.1% (n = 9) of the participants identified themselves as Asian, 9.0% (n = 16) as Black, 4.5% (n = 8) as Mixed, 4.0% (n = 7) as other, and 77.4% (n = 137) as White (total non-White: n = 40, 22.6%). 50.3% (n = 89) reported having received a bachelor's degree or higher.

92.1% (n = 163) of the participants said that they had previously searched online for information about political candidates, and 94.4% (n = 167) reported that Google was their most used search engine. Participants reported conducting an average of 18.1 (SD = 34.1) internet searches per day. 49.2% (n = 87) of the participants identified themselves as liberal, 32.2% (n = 57) as moderate, 14.1% (n = 25) as conservative, 2.3% (n = 4) as not political, and 2.3% (n = 4) as other.

3.1.2 Procedure. Participants were randomly assigned to one of three groups: Pro-Candidate-A-Box, Pro-Candidate-B-Box, or a control group in which the answer box was not present. We used the same candidates and election as we used in Experiment 1, except that search results were unbiased in all three groups. Specifically, pro-Abbott search results alternated with pro-Gillard search results. Our participants reported an average familiarity level of 1.68 [1.64] for Julia Gillard and 2.23 [2.06] for Tony Abbott. The experimental procedure itself was identical in all respects to the procedure in Experiment 1.

3.2 Results

In the No-Box group, the proportions of people voting for each candidate did not change pre-search to post-search ($Pre_{Gillard}$ = 0.41, $Post_{Gillard}$ = 0.52, z = -1.19, p = 0.23). The VMP itself could not be computed, because there was no bias condition in this group. Voting preferences expressed on the 11-point scale shifted from -0.02 [3.24] pre-search to 0.24 [3.30] post-search (Wilcoxon's z = -0.60, p = 0.55 NS, d = 0.08), which means that unbiased search results had almost no effect on votes or voting preferences.

In the Box conditions, however, the VMP was 38.6% (z = -5.50, $p < 0.001$) (Table 3), and the voting preference expressed on the 11-point scale shifted from 0.08 [3.06] to 0.97 [3.90] (Wilcoxon's z = -3.57, $p < 0.001$, d = 0.26), which means there was a significant shift toward the favored candidate. Given that there was no bias in the search results, the shift in voting preferences was likely due exclusively to the biased answer boxes. Similarly, more people reported seeing bias in the box condition (12.5%) than in the No-Box condition (0.0%), and the difference between these percentages was significant (z = -2.20, $p < 0.05$).

The results in Experiment 2 differ from the results in Experiment 1 in one important respect: The opinions about the candidates (impression, trust, and likeability) did not change

Table 3. Experiment 2: VMP, search times, and results clicked by condition.

Condition	*n*	VMP (%)	Mean Search Time (sec) (*SD*)	Mean No. of Results Clicked (*SD*)
No Box	58	N/A[†]	228.0 (201.2)	4.00 (3.7)
Box	119	38.6	246.1 (265.9)	3.45 (3.2)
Change (%)	-	-	+7.9	-13.8
Statistic	-	-	$t(175) = 0.46$	$t(175) = -1.01$
p	-	-	= 0.65 NS	= 0.31 NS

[†]As noted in the text, since there was no bias in the search results shown in the No-Box condition, VMP could not be calculated.

https://doi.org/10.1371/journal.pone.0268081.t003

Table 4. Experiment 2: Pre- and post-search opinion ratings of favored and non-favored candidates.

		Pre	**Post**	**Diff**				
No Box	**Impression**	7.46 (1.87)	6.34 (2.11)	-1.12				
	Trust	6.29 (2.06)	5.82 (2.22)	-0.47				
	Likeability	7.41 (1.96)	6.47 (2.10)	-0.94				
		Favored Candidate Mean (SD)			**Non-Favored Candidate Mean (SD)**			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	$z^{\dagger}$
Box	**Impression**	7.07 (1.93)	5.93 (2.31)	-1.14	7.31 (1.88)	5.55 (2.28)	-1.76	-2.06 NS
	Trust	6.24 (2.26)	5.60 (2.54)	-0.64	6.38 (2.23)	5.17 (2.29)	-1.15	-2.18 NS
	Likeability	7.03 (2.07)	5.82 (2.34)	-1.21	7.20 (1.88)	5.46 (2.31)	-1.74	-1.61 NS

†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate. This statistic could not be computed for Group 1 because there was no favored candidate.

https://doi.org/10.1371/journal.pone.0268081.t004

significantly (Table 4). This makes sense, given that (a) the answer boxes gave almost no information about the candidates and (b) the search results did not favor either candidate. Differences in opinions did not emerge even though people spent about the same time viewing search results in Experiment 1 as they did in Experiment 2 (M_{E1} = 246.8 s [247.8], M_{E2} = 240.2 s [246.2], $t(596) = 0.30$, $p = 0.77$, $d = 0.03$), and clicked roughly the same number of search results in Experiment 1 as they clicked in Experiment 2 (M_{E1} = 3.80 [3.6], M_{E2} = 3.63 [3.4], $t(596) = 0.51$, $p = 0.61$, $d = 0.05$).

We also saw a different pattern in the VMPs of the people in the two box groups who detected the bias (23 out of 119 people, 19.3%): When people detect bias in search results (based largely or in part on viewing the web pages to which the search results link), their opinions and voting preferences tend to shift even farther in the direction of the favored candidate than do the opinions and voting preferences of people who do not detect the bias. In Experiment 2, however, we found the opposite pattern. The VMP for people who reported seeing bias in the Box groups was 12.5%; whereas the VMP for people who did not report seeing bias in the Box groups was 44.4% ($z = -2.93$, $p < 0.05$). Bear in mind that each user is seeing only one box; he or she has nothing with which to compare it, and the search results themselves are unbiased. More light is shed on this matter in Experiment 3 (also see Discussion).

The dramatic shift in voting preferences produced by biased answer boxes alone in Experiment 2 raises a disturbing possibility about the power that IPAs might have to impact thinking and behavior. Experiment 2 functioned, after all, like an IPA: A single query produced a single reply (given in the answer box), which appeared above unbiased search results. Could a single biased answer produced by an IPA produce a large shift in opinions and voting preferences? And what if multiple questions produced answers that shared the same bias? Could they produce even larger shifts in opinions and voting preferences? We attempted to answer these questions in Experiment 3.

Demographic analyses of data from Experiment 2 –by educational level, gender, age, and race/ethnicity–are shown in S5–S8 Tables. Demographic effects were relatively small.

4. Experiment 3: Assessing the persuasive power of the intelligent personal assistant (IPA)

4.1 Methods

4.1.1 Participants. After cleaning, our sample for this experiment consisted of 1,138 eligible voters from 48 US states. They were recruited from the MTurk subject pool. The data had been cleaned to remove participants who had reported an English fluency level below 6 on a 10-point scale.

52.3% (n = 595) were male, 46.7% (n = 531) were female, and 1.1% (n = 12) chose not to identify their gender. Participants ranged in age from 18 to 89 (M = 41.3, median = 39.0, SD = 12.9). 8.3% (n = 94) of the participants identified themselves as Asian, 8.1% (n = 92) as Black, 3.0% (n = 34) as Mixed, 2.3% (n = 26) as other, and 78.4% (n = 892) as White (total non-White: n = 246, 21.6%). 64.1% (n = 729) reported having received a bachelor's degree or higher.

86.6% (n = 986) of the participants reported they had used a virtual assistant like Alexa or Siri. 48.6% (n = 553) of the participants identified themselves as liberal, 27.2% (n = 310) as moderate, 21.4% (n = 244) as conservative, 1.7% (n = 19) as not political, and 1.1% (n = 12) as other.

4.1.2 Procedure. All procedures were run online and were compatible with both desktop and mobile devices. As in the earlier experiments, participants were first asked screening questions and demographic questions and then given instructions about the experimental procedure and asked for their consent to participate in the study.

Participants were randomly assigned to one of five different question/answer (Q/A) groups. Each group was shown the same list of 10 questions, and the order of the questions did not vary. After a participant clicked a question, Dyslexa–our Amazon Alexa IPA simulator–replied vocally with an answer (See S2 Text). The number of questions people were required to ask varied by group, and in two of the groups, the answer to the second question was "masked" in a manner that we will describe below. A screenshot showing how the questions and Dyslexa simulator appeared to users is shown in Fig 1. The five groups were as follows:

1. Group 1Q/1A: Participants were required to select just one question.
2. Group 4Q/4A/NM: Participants were required to select four different questions, and none was masked (NM = "no mask").
3. Group 4Q/4A/M2: Participants were required to select four different questions, and the answer to Question 2 was masked (M2 = Question 2 mask).

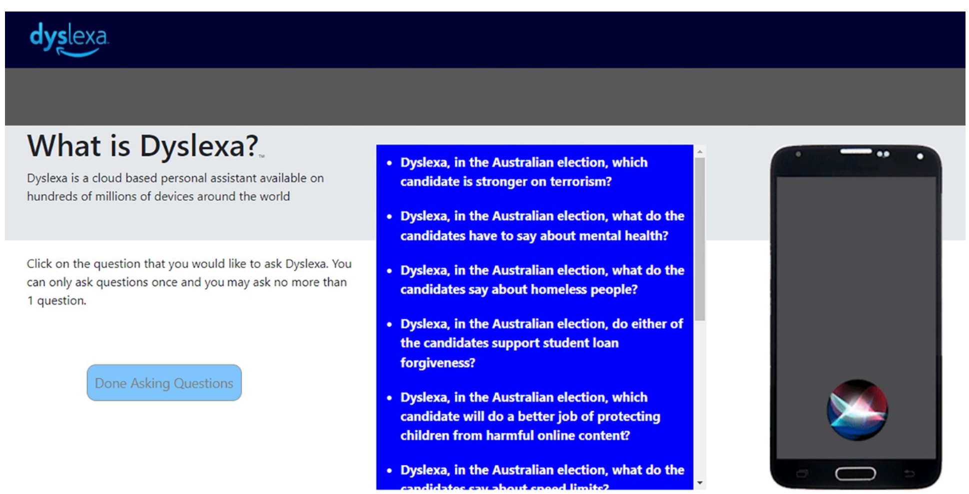


Fig 1. A screenshot showing what users saw in Experiment 3 when they posed questions to Dyslexa. Different groups were required to ask 1, 4, or 6 questions. After clicking on a question, it was greyed out, and Dyslexa answered the question orally. While it was speaking, the circular graphic at the bottom of the phone screen glowed and swirled, just as similar graphics do on most iPhones.

https://doi.org/10.1371/journal.pone.0268081.g001

4. Group 6Q/6A/NM: Participants were required to select six different questions, and none was masked.
5. Group 6Q/6A/M2: Participants were required to select six different questions, and the answer to Question 2 was masked.

Within each of the five groups, participants were randomly assigned to one of three different candidate conditions: Pro-Candidate-A, Pro-Candidate-B, or a control group. Our political candidates were Scott Morrison (Candidate A) and Bill Shorten (Candidate B), actual candidates from the 2019 election for prime minister of Australia. We chose this election to assure that our participants would be "undecided" voters. On a 10-point scale from 1 to 10, where 1 was labeled "not at all" and 10 was labeled "quite familiar," our participants reported an average familiarity level of 1.14 [0.43] for Scott Morrison and 1.05 [0.26] for Bill Shorten.

In the Candidate A condition, the answers were biased in favor of Scott Morrison. For example, when asked, "Dyslexa, in the Australian election, which candidate favors having a stronger relationship with the United States?," Dyslexa replied, "According to recent media reports, Scott Morrison wants to build a stronger relationship with the United States. His opponent, Bill Shorten, wants to continue to increase trade with Russia and China." In the Candidate B condition, the answers were biased in favor of Bill Shorten. In response to the same question, the pro-Shorten reply was "According to recent media reports, Bill Shorten wants to build a stronger relationship with the United States. His opponent, Scott Morrison, wants to continue to increase trade with Russia and China." The answers in each bias group were, in other words, nearly identical; only the names were changed. Mean bias ratings were obtained from five independent raters for each of the 20 answers on an 11-point scale from -5 (pro-Morrison) to +5 (pro-Shorten). The overall bias for Morrison was -3.3 [0.67], and the overall bias for Shorten was 3.4 [0.67] (based on absolute value: $t(18) = -0.07$, $p = 0.98$ NS).

In two of the five groups (Groups 3 and 5), masks were used for the answers to the second question each participant asked. This means that in the pro-Morrison group, a pro-Shorten answer was given in response to the second question asked, and in the pro-Shorten group, a pro-Morrison answer was given in response to the second question asked. This is a standard procedure used in SEME experiments [2] to reduce or eliminate the perception that the content being shown is biased. In SEME experiments, biased search results still produce large shifts in opinions and voting preferences even when aggressive masks are employed that completely eliminate the perception of bias. (See the Results and Discussion sections below for further information about our use of masks.)

In each control group, including Group 1 (1Q/1A), the answer to the first question had a 50/50 chance of supporting either Morrison or Shorten. After that, the bias in the answers alternated between the two candidates with each question asked. In Groups 2 through 5, we used an even number of questions (4 or 6) to ensure that each participant received equal exposure to pro-Morrison and pro-Shorten answers.

Participants were allowed to choose their questions from a list of 10. We provided this relatively long list to increase the likelihood that participants would select questions on topics they cared about. We speculated that allowing people to choose their questions would increase their interest in the answers they were given. We varied the number of questions people could ask to see whether we could have a bigger impact on opinions and voting preferences when people were exposed to a larger number of biased answers. We did not include a two-question group because we would not have been able to use a mask; a mask in the second position would almost certainly have eliminated the bias effect.

Following the demographic questions and instructions, all participants were shown brief, neutral biographies about each candidate (approximately 120 words each–somewhat shorter than the biographies used in Experiments 1 and 2 for the 2010 Australian election). (See S3 Text for the biographies employed in Experiment 3.) Participants were then asked six questions about their candidate preferences (each on a 10-point Likert scale from "Low" to "High"): whether their overall impression of each candidate was positive or negative, how likeable they found each candidate, and how much they trusted each candidate. Then–on an 11-point scale from -5 to +5, with the name of each candidate shown at either end of the scale and with the order of the names counterbalanced from one participant to another–participants were asked which candidate they would most likely vote for if they had to vote today. Finally, they were asked which of the two candidates they would actually vote for today (forced choice). The answers to these two questions had to be consistent; if they weren't, participants were asked to answer them again.

Following these opinion questions, participants were given brief instructions about how to use our IPA, and they then could proceed to ask questions (between one and six questions, according to their group assignment) and hear Dyslexa's answers. Our questions covered a wide range of topics that we thought would be of interest to a US sample (see S2 Text), but we deliberately avoided including hot-button issues such as abortion. If a participant chose to ask, "What are the candidates' positions on abortion?," and Dylexa replied that Morrison wanted to protect abortion rights, the possible partisanship of our participants could have driven them either *toward* or *away from* Morrison–*toward* if they supported abortion rights, *away* if they opposed abortion.

Following the interaction with the IPA, all participants were again asked those six opinion questions and two voting-preference questions. Finally, participants were asked whether anything "bothered" them about the questions they were shown and the answers they heard while interacting with our IPA. As in our previous experiments, this is where participants had an opportunity to express their concerns about content bias or other issues.

4.2 Results

We found significant and substantial shifts in both voting preferences (Table 5) and opinions (Table 6) in the direction of the favored candidates in all bias groups. We also found significant shifts in voting preferences in the direction of the favored candidates in all bias groups as expressed on our 11-point voting-preference scale (Table 7). In contrast, in the control groups the proportions of people voting for each candidate before the manipulations changed relatively little or not at all following the manipulations (Group 1, 0.0%; Group 2, 6.6%; Group 3, 2.7%; Group 4, 7.1%; Group 5, 6.8%).

The percentage of people in the bias groups who reported seeing biased content was substantially lower when they received just one answer (Group 1, 4.9%) or when biased content was masked (Group 3, 5.1%; Group 5, 7.1%) than when people saw multiple biased answers

Table 5. Experiment 3: Pre- and Post-IPA VMPs.

Group No.	Group	Total *n*	Bias Groups *n*	Bias Groups VMP (%)	McNemar Test X^2	*p*
1	**1Q/1A**	222	142	43.8	24.0	< 0.001
2	**4Q/4A/NM**	229	153	59.5	35.9	< 0.001
3	**4Q/4A/M2**	230	156	59.2	33.6	< 0.001
4	**6Q/6A/NM**	230	145	65.8	44.5	< 0.001
5	**6Q/6A/M2**	227	154	50.0	36.5	< 0.001

https://doi.org/10.1371/journal.pone.0268081.t005

Table 6. Experiment 3: Pre- and post-IPA opinion ratings of favored and non-favored candidates.

		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	Post	Diff	Pre	Post	Diff	$z^{\dagger}$
Group 1: 1Q1A Condition	**Impression**	7.13 (1.85)	7.63 (2.00)	+0.50	7.10 (1.73)	6.13 (2.18)	-0.97	-6.32***
	Trust	6.29 (2.20)	6.95 (2.29)	+0.66	6.26 (2.11)	5.65 (2.41)	-0.61	-6.59***
	Likeability	7.15 (1.83)	7.46 (2.00)	+0.31	7.18 (1.72)	6.18 (2.23)	-1.00	-6.43***
Group 2:	**Impression**	6.76 (1.93)	7.73 (2.23)	+0.97	6.89 (1.72)	4.97 (2.04)	-1.92	-8.82***
4QNM Condition	**Trust**	5.88 (2.18)	6.97 (2.51)	+1.09	6.05 (2.05)	4.80 (2.23)	-1.25	-7.80***
	Likeability	6.67 (2.01)	7.41 (2.26)	+0.74	6.93 (1.84)	5.03 (2.13)	-1.90	-7.93***
Group 3:	**Impression**	6.79 (1.92)	7.28 (1.95)	+0.49	6.96 (1.72)	6.12 (1.85)	-0.84	-5.92***
4QM2 Condition	**Trust**	5.81 (2.12)	6.54 (2.27)	+0.73	6.06 (2.07)	5.71 (2.04)	-0.35	-7.50***
	Likeability	6.81 (1.90)	7.13 (2.12)	+0.32	7.04 (1.71)	6.20 (1.99)	-0.84	-5.64***
Group 4:	**Impression**	6.87 (1.75)	7.74 (1.94)	+0.87	6.72 (1.81)	4.83 (2.00)	-1.89	-8.64***
6QNM Condition	**Trust**	5.94 (1.97)	6.90 (2.25)	+0.96	5.99 (2.10)	4.58 (2.11)	-1.41	-7.87***
	Likeability	6.82 (1.87)	7.62 (2.09)	+0.80	6.78 (2.02)	4.96 (2.13)	-1.82	-8.32***
Group 5:	**Impression**	7.10 (1.65)	7.65 (1.94)	+0.55	7.00 (1.87)	5.34 (2.02)	-1.66	-7.98***
6QM2 Condition	**Trust**	6.31 (2.00)	7.09 (2.20)	+0.78	6.18 (2.07)	5.08 (2.29)	-1.10	-7.65***
	Likeability	7.05 (1.70)	7.50 (2.00)	+0.45	6.93 (1.86)	5.42 (2.12)	-1.51	-7.54***

†z-score represents Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate.

***$p < 0.001$

https://doi.org/10.1371/journal.pone.0268081.t006

without masks (Group 2, 23.5%; Group 4, 40.7%) (Table 8) ($M_{\text{Groups1,3,5}}$ = 5.8%, $M_{\text{Groups2,4}}$ = 31.9%, z = -9.50, $p < 0.001$).

The present study sheds new light on the role that bias detection plays in shifting opinions and voting preferences. Previous investigations have shown that the opinions of the few people who are able to detect bias in search results shift even farther in the direction of the bias than the opinions of the people who don't see the bias [2, 15]. This occurs presumably because of the high trust people have in the filtering and ordering of search results, which people mistakenly believe is an objective and impartial process [125, 126]. In the present study, we learned that bias detection *erodes* trust when people are interacting with answers provided by answer boxes (in the absence of biased search results–see Experiment 2) or the vocal answers of an IPA, where search results are entirely absent (Experiment 3). This difference is likely due to the daily regimen of operant conditioning that supports the almost blind trust people have in search results. About 86% of searches are for simple facts, and the correct answers to those queries reliably turn up in the first or second search result. People are learning, over and over again, that what is higher in the list of search results is better and truer than what is lower. When, in a recent experiment, that trust was temporarily broken, the VMP in a SEME procedure was significantly reduced [15].

Table 7. Experiment 3: Pre-IPA vs. Post-IPA voting preferences on 11-point scales.

Group No.	Group	Pre-IPA Voting Preference on 11-Point Scale (SD)	Post-IPA Voting Preference on 11-Point Scale (SD)	*z*	*p*	*d*
1	**1Q/1A**	0.61 (2.42)	1.70 (2.76)	-5.51	< 0.001	0.42
2	**4Q/4A/NM**	-0.01 (2.57)	2.41 (2.64)	-8.17	< 0.001	0.93
3	**4Q/4A/M2**	-0.10 (2.76)	1.38 (2.90)	-5.83	< 0.001	0.52
4	**6Q/6A/NM**	0.21 (2.46)	2.67 (2.28)	-8.50	< 0.001	1.04
5	**6Q/6A/M2**	0.20 (2.60)	2.26 (2.62)	-7.99	< 0.001	0.79

https://doi.org/10.1371/journal.pone.0268081.t007

Table 8. Experiment 3: VMPs for people who saw Bias vs. VMPs for people who did not see Bias.

Group No.	Group	*n*	No. Ss in Bias Groups Reporting Bias in IPA Content (%)	No. Ss in Bias Groups Not Reporting Bias in IPA Content (%)	VMP for Ss Who Reported Bias (%)	VMP for Ss Who Did Not Report Bias (%)	*z*	*p*
1	**1Q/1A**	142	7 (4.9)	135 (95.1)	33.3†	44.3	-0.57	= 0.57 NS
2	**4Q/4A/NM**	153	36 (23.5)	117 (76.5)	21.7	75.0	-5.78	$<$ 0.001
3	**4Q/4A/M2**	156	8 (5.1)	148 (94.9)	300.0†	55.7	14.46	$<$ 0.001
4	**6Q/6A/NM**	145	59 (40.7)	86 (59.3)	63.3	67.4	-0.51	= 0.61 NS
5	**6Q/6A/M2**	154	11 (7.1)	143 (92.9)	60.0†	49.4	0.68	= 0.50 NS

†The validity of these VMPs is questionable because they are based on a small number of observations. In Groups 1, 3, and 5, respectively, only 7, 8, and 11 people reported seeing bias in the IPA replies.

https://doi.org/10.1371/journal.pone.0268081.t008

So when search results are absent, as they are when people are using IPAs, or when search results are unbiased, as they were in our Experiment 2, people who detect bias do not automatically accept that bias as valid. Accepting that bias as valid seems to occur primarily when people are being influenced by biased search results–again, presumably because of that daily regimen of operant conditioning. That daily regimen of conditioning makes SEME a unique list effect and an especially powerful form of influence [15].

As we noted earlier, we regard the most important measure of change to be the VMP, which indicates the increase or decrease in the proportion of people who indicated in response to a forced-choice question which candidate they would vote for if they had to vote today (see S1 Text). The VMPs in the five groups in Experiment 3 ranged from 43.8% (Group 1) to 65.8% (Group 4). These shifts were all quite high–all higher than the 38.6% shift we found in Experiment 2.

In addition, we found that the more questions people asked (without masks, which tend to lower VMPs), the greater the shift in voting preferences ($VMP_{Q1/A1}$ = 43.8%, $VMP_{Q4/A4/NM}$ = 59.5%, $VMP_{Q6/A6/NM}$ = 65.8%; $X^2 = 6.59$; $p < 0.05$).

A breakdown of VMP data from Experiment 3 based on whether participants had had previous experience with IPAs is shown in S9 Table. Previous experience with IPAs did not appear to impact VMPs in any consistent way.

5. Discussion

Together, the three experiments we have described reveal a dangerous new tool of mass manipulation–one that is, at this writing, controlled worldwide almost entirely by just four large American tech companies: Amazon, Apple, Facebook/Meta, and Google. This new tool, which we call the Answer Bot Effect (ABE), is likely now affecting hundreds of millions of people, and with more and more people coming to rely on electronic devices to give them a single answer to their queries, the number of people affected by ABE will likely swell into the billions within the next few years. ABE should be of concern to every one of us, but especially to parents–whose children are being fed algorithmically-generated answers every day on their computers, mobile phones, tablets, and toys–as well as to public policy makers.

ABE should be of special concern for four reasons: (a) because of the large magnitude of the effect, (b) because it can impact the vast majority of people without their awareness, (c) because it is an ephemeral manipulation, leaving no paper trail for authorities to trace, and (d) because ABE is inherently non-competitive and impossible to counteract. You can counteract a billboard or television commercial, but how can you correct the way a tech platform adjusts its algorithms? Recall that in Experiment 3, a one-question-one-answer interaction on our Alexa simulator produced a 43.8% shift in voting preferences, with only 4.7% of the participants reporting any concerns about bias.

Perhaps the reader thinks we are overstating the seriousness of the problem. Although a full exploration of this issue is beyond the scope of this paper, please consider just two growing bodies of evidence that bring manipulations like ABE into sharper focus: First, in recent years, whistleblowers from Google and Facebook/Meta, along with leaks of emails, documents, and videos from these companies, have shown repeatedly that manipulations like ABE are being deliberately and strategically used by these companies to influence attitudes, beliefs, purchases, voting preferences, and public policy itself [25, 28, 29, 43, 48]. In a leak of emails to the *Wall Street Journal* in 2018, Google employees discuss the possibility of using "ephemeral experiences" to change people's views about Trump's 2017 travel ban [25]. A leaked 8-minute video from Google called "The Selfish Ledger" describes the company's power to "modify behavior" at the "species level" in ways that "reflect Google's values" [127]. In various interviews and the recent documentary film, "The Social Dilemma," former Google insider Tristan Harris spoke about his time working with a large team of Google employees whose job it was to modify "a billion people's attention and thoughts every day" [128].

Harris and others have expressed concerns about company policies that are meant to influence people in specific ways, but ABE, SEME, and other new forms of online influence will impact thinking and behavior even without a company policy in place. Algorithms left to their own devices–let's call this practice "algorithmic neglect"–reflect the biases of the people who programmed them [20–23], and the algorithms also quickly learn and reflect the foibles of human users, sometimes magnifying and spreading bigotism, racism, and hatred with frightening rapidity [52, 55, 61, 97, 116, 117]. What's more, a single rogue employee with the right password authority or hacking skills can use a large tech platform like Google to impact reputations, businesses, or elections on a large scale without senior management knowing he or she is doing so [129]. When authorities learned in 2010 that Google's Street View vehicles had been vacuuming up personal Wi-Fi data for 3 years in 30 countries [130], Google blamed the entire operation on a single software engineer, Marius Milner–but they did not fire him, and he remains at the company today [131].

Second, election monitoring projects that have been conducted since 2016 have so far preserved more than 1.5 million politically-related online ephemeral experiences in the weeks preceding national elections in the US. This is actual content–normally lost forever–being displayed on the computer screens of thousands of US voters–the real, personalized content that Big Tech companies are showing politically diverse groups of people as elections approach. The wealth of unusual data preserved in these projects has revealed strong unilateral political bias in ephemeral content, sufficient to have shifted millions of votes in national elections in the US without people's knowledge [132–134].

The experiments we have described build one upon the other. Experiment 1 showed that when the content of an answer box shared the bias of the search results beneath it, it increased the impact that those search results have on thinking and behavior, and it reduced the time people spent searching and significantly reduced the number of search results people clicked. Experiment 2 simulated a situation in which the answer box was biased but the search results were not. The biased answer boxes alone produced a remarkable VMP of 38.6%.

Rounded to the nearest whole number, the VMP in Experiment 2 was 39%. This means that out of 100 undecided voters–people whose vote would normally split 50/50 without having additional information–the votes, on average, of 19.5 people (0.39 x 50) can be shifted by biased answer boxes, yielding a vote of roughly 69 to 30, for a win margin among previously undecided voters of 39% (see S1 Text). In a national election in the US in which 150 million people vote (159 million voted in the 2020 Presidential election), even if only 10% of the voters were undecided and depended on computers for trustworthy answers, if the single-answer-generating algorithms in the days or weeks leading up to Election Day all favored the same candidate, that could conceivably shift more than 2.9 million votes to that candidate (0.10 x 0.39 x 0.5 x 150,000,000). If the other 90% of the voters were split 50/50, that would give the favored candidate a win margin of 5.8 million votes (3.8%).

Unfortunately, the real situation we face is probably worse than the case we just described. At this moment in history, in the US virtually all the single-answer-generating algorithms will likely be supporting the same national and state candidates [135–137], and six months before an election, the percentage of undecided voters might be as high as 60%, not 10% [122, 138, 139].

Bear in mind also that in our experiments we are interacting with our participants only briefly and only once. If undecided voters are subjected to content having the same bias repeatedly over a period of weeks or months, their voting preferences will likely shift even farther than the voting preferences of our participants shifted. Recall that in Experiment 3 the VMP exceeded 65% when people asked six questions–nearly 50% higher than the VMP we found when people asked only one question (Table 5).

What's more, ABE is just one powerful source of influence. When similarly biased content is delivered in search results, search suggestions, YouTube videos, newsfeeds, targeted messages, and so on, the net impact of these manipulations is likely additive, and when Big Tech companies all share the same political bias (or any other type of bias, for that matter), the net impact of their combined influence is also likely additive. Without regulations, laws, and permanent, large-scale monitoring systems to stop them–and none exist at this writing [140]–Big Tech companies indeed have the power to reengineer humanity "at the species level," as Google's "Selfish Ledger" video suggests [127]. At the very least, they can easily tilt the outcomes of close elections worldwide.

In a remarkable and frequently quoted farewell speech delivered by US President Dwight D. Eisenhower just a few days before John F. Kennedy's inauguration in January 1961, Eisenhower–a military insider–not only warned the American people about a rapidly evolving "military-industrial complex," he also spoke of the danger that someday "public policy could itself become the captive of a scientific technological elite" [141]. If ABE, SEME, and other new forms of influence the internet has made possible work anything in the real world like they do in controlled experiments, it is not unreasonable to speculate that while humanity was being distracted by online video games, dating websites, and cat memes, Eisenhower's prediction came true. The technological elite now exist [142],and, if our analyses are correct, they are now very much in control.

Supporting information

S1 Fig. Apparent bias in a Google answer box, screenshotted October 22, 2017. The content of the box clearly favors the Google service.
(TIF)

S2 Fig. Apparent bias in two types of Google answer boxes. (a) In a screenshot preserved in an article in *Search Engline Land* on March 5, 2017, four US presidents are incorrectly listed in a Google answer box as members of the Ku Klux Klan. (b) In a screenshot of a Google knowledge box preserved in an article in *VICE* on May 31, 2018, Nazism is incorrectly listed as part of the ideology of the California Republican Party. (c) In a Google answer box captured by the

first author on August 16, 2016, Hillary Clinton's photograph is shown in response to the question, "when is the election?".
(TIF)

S1 Text. Vote Manipulation Power (VMP) calculation.
(DOCX)

S2 Text. Experiment 3: Alexa simulator, "Dyslexa," questions and answers.
(DOCX)

S3 Text. Experiment 3: Candidate biographies.
(DOCX)

S1 Table. Experiment 1: Demographic analysis by educational attainment.
(DOCX)

S2 Table. Experiment 1: Demographic analysis by gender.
(DOCX)

S3 Table. Experiment 1: Demographic analysis by age.
(DOCX)

S4 Table. Experiment 1: Demographic analysis by race/ethnicity.
(DOCX)

S5 Table Experiment 2: Demographic analysis by educational attainment.
(DOCX)

S6 Table. Experiment 2: Demographic analysis by gender.
(DOCX)

S7 Table. Experiment 2: Demographic analysis by age.
(DOCX)

S8 Table. Experiment 2: Demographic analysis by race/ethnicity.
(DOCX)

S9 Table. Experiment 3: Demographic analysis by previous IPA use.
(DOCX)

Acknowledgments

We thank J. Martinez for assistance in conducting the second experiment and L. Kafader for expert programming assistance. R. Mohr is currently a doctoral candidate at Palo Alto University, Palo Alto, California USA.

Author Contributions

Conceptualization: Robert Epstein.

Formal analysis: Robert Epstein, Vivian Lee, Roger Mohr, Vanessa R. Zankich.

Investigation: Robert Epstein, Roger Mohr.

Methodology: Robert Epstein.

Project administration: Robert Epstein.

Supervision: Robert Epstein.

Writing – original draft: Robert Epstein.

Writing – review & editing: Robert Epstein, Vivian Lee, Vanessa R. Zankich.

References

1. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating Google output. J Med Internet Res, 2014; 16(4):e100. https://doi.org/10.2196/jmir.2642 PMID: 24694866
2. Epstein R, Robertson RE. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proc Natl Acad Sci USA, 2015; 112(33):E4512– E4521. https://doi.org/10.1073/pnas.1419828112 PMID: 26243876
3. Epstein R, Robertson RE. Suppressing the Search Engine Manipulation Effect (SEME). Proceedings of the ACM on Human-Computer Interaction, 2017; 1(CSCW):1–22. https://doi.org/10.1145/3134677
4. Haas A, Unkel J. Ranking versus reputation: perception and effects of search result credibility. Behav Inf Technol, 2017; 36(12):1285–1298. https://doi.org/10.1080/0144929X.2017.1381166
5. Ludolph R, Allam A, Schulz PJ. Manipulating Google's knowledge box to counter biased information processing during an online search on vaccination: Application of a technological debiasing strategy. J Med Internet Res, 2016; 18(6):e137. https://doi.org/10.2196/jmir.5430 PMID: 27255736
6. Eslami M, Vaccaro K, Karahalios K, Hamilton K. "Be Careful; Things Can Be Worse than They Appear": Understanding Biased Algorithms and Users' Behavior Around Them in Rating Platforms. Proceedings of the 11th International AAAI Conference on Web and Social Media [Internet]. 2017 May 3; 11(1):62–71. Available from: https://ojs.aaai.org/index.php/ICWSM/article/view/14898
7. Pogacar FA, Ghenai A, Smucker MD, Clarke CLA. The positive and negative influence of search results on people's decisions about the efficacy of medical treatments. Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval [Internet]. 2017 Oct 1–4;:209–216. Available from: https://dl.acm.org/doi/10.1145/3121050.3121074
8. Trielli D, Diakopoulos N. Search as news curator: The role of Google in shaping attention to news information. CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems [Internet]. 2019 May 4–9; 453:1–15. Available from: https://dl.acm.org/doi/10.1145/3290605.3300683
9. Casara BGS, Suitner C, Bettinsoli ML. Viral suspicions: Vaccine hesitancy in the Web 2.0. J Exp Psychol Appl, 2019; 25(3):354–371. https://doi.org/10.1037/xap0000211 PMID: 30816729
10. Edelman B. Bias in search results? Diagnosis and response. Indian J Law Technol, 2011; 7:16–32. Available from: https://www.ijlt.in/journal/bias-in-search-results%3F%3A-diagnosis-and-response
11. Feldman S. Americans see search engines as biased. 2018 Sep 7. In: Statista [Internet]. Available from: https://www.statista.com/chart/15385/americans-see-search-engingeses-as-biased/
12. Knobloch-Westerwick S, Mothes C, Johnson BK, Westerwick A, Donsbach W. Political online information searching in Germany and the United States: Confirmation bias, source credibility, and attitude impacts. J Commun, 2015; 65(3):489–511. https://doi.org/10.1111/jcom.12154
13. O'Neil C. Weapons of math destruction: How big data increases inequality and threatens democracy. Crown Publishing Group, New York, 2016.
14. Pierce DR, Redlawsk DP, Cohen WW. Social influences on online political information search and evaluation. Polit Behav, 2016; 39:651–673. https://doi.org/10.1007/s11109-016-9374-4
15. Epstein R. Manipulating minds: The power of search engines to influence votes and opinions. In: Moore M, Tambini D, editors. Digital dominance: The power of Google, Amazon, Facebook, and Apple. Oxford, UK: Oxford University Press; 2018. pp. 294–319. Available from: https://aibrt.org/downloads/EPSTEIN_2018-Manipulating_minds-The-power_of_search_engines_to_influence_votes_and_opinions-UNCORRECTED_PROOFS.pdf
16. Brisson-Boivin K, McAleese S. Algorithmic awareness: Conversations with young Canadians about artificial intelligence and privacy. MediaSmarts. 2021. Available from: https://mediasmarts.ca/sites/default/files/publication-report/full/report_algorithmic_awareness.pdf
17. Fallows D. Internet searchers are confident, satisfied and trusting–but they are also unaware and naïve. 2005 Jan 23. In: Pew Internet & American Life Project [Internet]. Available from: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/media/Files/Reports/2005/PIP_Searchengine_users.pdf.pdf
18. Schofield J. A lot of Brits don't understand search engines: A UK survey published by FastHosts has revealed some misconceptions about how search engines work. Whether this matters to users, as opposed to website promoters, is less clear. The Guardian. 2008 Dec 12. Available from: https://www.theguardian.com/technology/blog/2008/dec/12/searchengine-survey

19. Pasquale F. The black box society: The secret algorithms that control money and information. Harvard University Press; 2015.

20. Bhamore S. Decrypting Google's search engine bias case: Anti-trust enforcement in the digital age. Christ University Law Journal, 2019; 8(1):37–60. https://doi.org/10.12728/culj.14.2

21. Nunez M. Former Facebook workers: We routinely suppressed conservative news. 2016 May 9. In: Gizmodo [Internet]. Available from: https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006

22. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science, 2019; 366(6464):447–453. https://doi.org/10.1126/science.aax2342 PMID: 31649194

23. Reuters. Amazon scraps secret AI recruiting tool that 'didn't like women'. 2018 Oct 10. In: MailOnline [Internet]. Available from: https://www.dailymail.co.uk/sciencetech/article-6259205/Amazon-scraps-secret-AI-recruiting-tool-showed-bias-against-women.html

24. Sun W, Nasraoui O, Shafto P. Evolution and impact of bias in human and machine learning algorithm interaction. PLoS ONE, 2020; 15(8):e0235502. https://doi.org/10.1371/journal.pone.0235502 PMID: 32790666

25. McKinnon JD, MacMillan D. Google workers discussed tweaking search function to counter travel ban: Company says none of proposed changes to search results were ever implemented. The Wall Street Journal. 2018 Sep 20. Available from: https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472

26. Meyers PJ. How often does Google update its algorithm? 2019 May 14. In: Moz [Internet]. Available from: https://moz.com/blog/how-often-does-google-update-its-algorithm

27. Epstein R. The new censorship: How did Google become the internet's censor and master manipulator, blocking access to millions of websites? U.S. News & World Report. 2016 Jun 22. Available from: https://www.usnews.com/opinion/articles/2016-06-22/google-is-the-worlds-biggest-censor-and-its-power-must-be-regulated

28. Gallagher R. Google plans to launch censored search engine in China, leaked documents reveal: Search app that will "blacklist sensitive queries" could be launched in six to nine months, according to documents and people familiar with the plans. The Intercept. 2018 Aug 1. Available from: https://theintercept.com/2018/08/01/google-china-search-engine-censorship/

29. Vorhies Z, Heckenlively K. Google leaks: A whistleblower's exposé of Big Tech censorship. Skyhorse Publishing, 2021.

30. Agudo U, Matute H. The influence of algorithms on political and dating decisions. PLoS ONE, 2021; 16(4):e0249454. https://doi.org/10.1371/journal.pone.0249454 PMID: 33882073

31. Ani. Study says people trust computers more than humans. In: Big News Network.com [Internet]. 2021 Apr 15. Available from: https://www.bignewsnetwork.com/news/268700564/study-says-people-trust-computers-more-than-humans

32. Bogert E, Schecter A, Watson RT. Humans rely more on algorithms than social influence as a task becomes more difficult. Sci Rep, 2021; 11. https://doi.org/10.1038/s41598-021-87480-9 PMID: 33850211

33. Edelman. Edelman Trust Barometer 2021. In: Edelman [Internet]. 2021. Available from: https://www.edelman.com/trust/2021-trust-barometer

34. Logg JM, Minson JA, Moore DA. Do people trust algorithms more than companies realize? Harvard Business Review. 2018 Oct 26. Available from: https://hbr.org/2018/10/do-people-trust-algorithms-more-than-companies-realize

35. Howard JJ, Rabbitt LR, Sirotin YB. Human-algorithm teaming in face recognition: How algorithm outcomes cognitively bias human decision-making. PLoS ONE, 2020; 15(8): e0237855. Available from: https://doi.org/10.1371/journal.pone.0237855 PMID: 32822441

36. Curran D. Are you ready? Here is all the data Facebook and Google have on you: The harvesting of our personal details goes far beyond what many of us could imagine. So I braced myself and had a look. The Guardian. 2018 Mar 30. Available from: https://www.theguardian.com/commentisfree/2018/mar/28/all-the-data-facebook-google-has-on-you-privacy

37. Kozyreva A, Lorenz-Spreen P, Hertwig R, Lewandowsky S, Herzog SM. Public attitudes towards algorithmic personalization and use of personal data online: Evidence from Germany, Great Britain, and the United States. Humanit Soc Sci Commun, 2021; 8:1–11. https://doi.org/10.1057/s41599-021-00787-w

38. Fernández-Manzano E, González-Vasco M. Analytic surveillance: Big data business models in the time of privacy awareness. Prof Inform, 2018; 27(2):402–409. https://doi.org/10.3145/epi.2018.mar.19

39. Schneier B. Data and Goliath: The hidden battles to collect your data and control your world. W. W. Norton & Company; 2015.

40. Zuboff S. The Age of surveillance Capitalism: The fight for a human future at the new frontier of power. PublicAffairs; 2019.

41. Bozdag E. Bias in algorithmic filtering and personalization. Ethics Inf Technol, 2013; 15:209–227. https://doi.org/10.1007/s10676-013-9321-6

42. DuckDuckGo. Measuring the "Filter Bubble": How Google is influencing what you click 2018 Dec 4. In: DuckDuckGo [Internet]. Available from: https://spreadprivacy.com/google-filter-bubble-study/

43. Duffy C. Facebook whistleblower revealed on '60 Minutes,' says the company prioritized profit over public good. CNN. 2021 Oct 4. Available from: https://edition.cnn.com/2021/10/03/tech/facebook-whistleblower-60-minutes/index.html

44. Epstein R. The new mind control. Aeon. 2016 Feb 18. Available from: https://aeon.co/essays/how-the-internet-flips-elections-and-alters-our-thoughts

45. Feuz M, Fuller M, Stalder F. Personal web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation. First Monday, 2011; 16(2). https://doi.org/10.5210/fm.v16i2.3344

46. Pariser E. The filter bubble: How the new personalized web is changing what we read and how we think. Penguin Press, 2011.

47. Lai C, Luczak-Roesch M. You can't see what you can't see: Experimental evidence for how much relevant information may be missed due to Google's web search personalisation. Weber I, Darwish KM, Wagner C, Zagheni E, Nelson L, Samin A, et al. (eds) Social Informatics: 11th International Conference. SocInfo 2019. Lecture Notes in Computer Science, vol 11864. Cham: Springer International Publishing; 2019. Available from: https://link.springer.com/chapter/10.1007/978-3-030-34971-4_17#citeas

48. Newton C. Google's new focus on well-being started five years ago with this presentation. 2018 May 10. In: The Verge [Internet]. Available from: https://www.theverge.com/2018/5/10/17333574/google-android-p-update-tristan-harris-design-ethics

49. Epstein R, Mohr R Jr, Martinez J. The Search Suggestion Effect (SSE): How search suggestions can be used to shift opinions and voting preferences dramatically. Paper presented at: Western Psychological Association; 2018 Apr; Portland, OR. Available from: https://aibrt.org/downloads/EPSTEIN_MOHR_&_MARTINEZ_2018-WPA-The_Search_Suggestion_Effect-SSE-WP-17-03.pdf

50. Al-Abbas LS, Haider AS, Hussein RF. Google autocomplete search algorithms and the Arabs' perspectives on gender: A case study of Google Egypt. GEMA Online Journal of Language Studies, 2020; 20(4):95–112. https://doi.org/10.17576/gema-2020-2004-06

51. Baker P, Potts A. 'Why do white people have thin lips?' Google and the perpetuation of stereotypes via auto-complete search forms. Crit Disc Stud, 2013; 10(2):187–204. https://doi.org/10.1080/17405904.2012.744320

52. Cadwalladr C. Google, democracy and the truth about internet search: Tech-savvy rightwingers have been able to 'game' the algorithms of internet giants and create a new reality where Hitler is a good guy, Jews are evil and. . .Donald Trump becomes president. The Guardian. 2016 Dec 4. Available from: https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook

53. Karapapa S, Borghi M. Search engine liability for autocomplete suggestions: personality, privacy and the power of the algorithm. Int J Law Inf Technol, 2015; 23(3):261–289. https://doi.org/10.1093/ijlit/eav009

54. Keskin B. (2015). What suggestions do Google autocomplete make about children? Black Sea Journal of Public and Social Science, 2015; 7(2):69–78. Available from: https://dergipark.org.tr/en/pub/ksbd/issue/16219/169869

55. Noble SU. Algorithms of oppression: How search engines reinforce racism. New York University Press, 2018.

56. Pradel F. Biased representation of politicians in Google and Wikipedia Search? The joint effect of party identity, gender identity, and elections. Polit Commun, 2020; 38(4):447–478. https://doi.org/10.1080/10584609.2020.1793846

57. Roy S, Ayalon L. Age and gender stereotypes reflected in Google's "autocomplete" function: The portrayal and possible spread of societal stereotypes. The Gerontologist, 2020; 60(6):1020–1028. https://doi.org/10.1093/geront/gnz172 PMID: 31812990

58. Kätsyri J, Kinnunen T, Kusumoto K, Oittinen P, Ravaja N. Negativity bias in media multitasking: The effects of negative social media messages on attention to television news broadcasts. PLoS One, 2016; 11(5):e0153712. https://doi.org/10.1371/journal.pone.0153712 PMID: 27144385

59. Carretié L, Mercado F, Tapia M, Hinojosa JA. Emotion, attention, and the 'negativity bias', studied through event-related potentials. Int J Psychophysiol, 2001; 41(1):75–85. https://doi.org/10.1016/s0167-8760(00)00195-1 PMID: 11239699

60. Arendt F, Fawzi N. Googling for Trump: Investigating online information seeking during the 2016 US presidential election. Inf Commun Soc, 2018; 22(13):1945–1955. https://doi.org/10.1080/1369118X.2018.1473459

61. Solon O, Levin S. How Google's search algorithm spreads false information with a rightwing bias. The Guardian. 2016 Dec 16. Available from: https://www.theguardian.com/technology/2016/dec/16/google-autocomplete-rightwing-bias-algorithm-political-propaganda

62. Gerhart SL. Do web search engines suppress controversy? First Monday, 2004; 9(1–5). https://doi.org/10.5210/fm.v9i1.1111

63. Sullivan D. A reintroduction to Google's featured snippets. 2018 Jan 30. In: Google Blog [Internet]. Available from: https://www.blog.google/products/search/reintroduction-googles-featured-snippets/

64. Singhal A. Introducing the knowledge graph: things, not strings. 2012 May 16. In: Google [Internet]. Available from: https://blog.google/products/search/introducing-knowledge-graph-things-not/

65. Strzelecki A, Rutecka P. Direct answers in Google search results. IEEE Access, 2020; 8:103642–103654. https://doi.org/10.1109/ACCESS.2020.2999160

66. Google Search Help. How Google's featured snippets work. 2022. In: Google Search Help [Internet]. Available from: https://support.google.com/websearch/answer/9351707?hl=en

67. Wright M. Who will be the next president? Google says it's Hillary Clinton. 2015 Jun 25. In: The Next Web [Internet]. Available from: http://thenextweb.com/shareables/2015/06/23/poor-old-jeb

68. Fishkin R. Less than half of Google searches now result in a click. 2019 Aug 13. In: SparkToro [Internet]. Available from: https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/

69. Fishkin R. In 2020, two thirds of Google searches ended without a click. 2021 Mar 22. In: SparkToro [Internet]. Available from: https://sparktoro.com/blog/in-2020-two-thirds-of-google-searches-ended-without-a-click/

70. Wu D, Dong J, Shi L, Liu C, Ding J. Credibility assessment of good abandonment results in mobile search. Inf Process Manag, 2020; 57(6). https://doi.org/10.1016/j.ipm.2020.102350

71. Richter F. Smart speaker adoption continues to rise. 2020 Jan 9. In: Statista [Internet]. Available from: https://www.statista.com/chart/16597/smart-speaker-ownership-in-the-united-states/

72. Liu S. Virtual assistant technology in the U.S.–Statistics & facts. 2021 Aug 18. In: Statista [Internet]. 2021. Available from: https://www.statista.com/topics/7022/virtual-assistants-in-the-us/#topicHeader__wrapper

73. Maras M. From secretly watching your kids to tracking their every move: the terrifying ways smart toys can be used by hackers. Daily Mail Online. 2018 May 10. Available from: https://www.dailymail.co.uk/sciencetech/article-5714587/From-secret-spying-kids-chatting-terrifying-ways-smart-toys-hacked.html

74. Vlahos J. Barbie wants to get to know your child. The New York Times. 2015 Sep 16. Available from: https://www.nytimes.com/2015/09/20/magazine/barbie-wants-to-get-to-know-your-child.html

75. Cox K. These toys don't just listen to your kid; they send what they hear to a defense contractor. 2016 Dec 6. In: The Consumerist [Internet]. Available from: https://consumerist.com/2016/12/06/these-toys-dont-just-listen-to-your-kid-they-send-what-they-hear-to-a-defense-contractor/

76. Walsh M. My Friend Cayla doll banned in Germany over surveillance concerns. ABC News. 2017 Feb 17. Available from: https://www.abc.net.au/news/2017-02-18/my-friend-cayla-doll-banned-germany-over-surveillance-concerns/8282508

77. Munr K. Hacking a talking toy parrot. 2017 Dec 1. In: Pen Test Partners [Internet]. Available from: https://www.pentestpartners.com/security-blog/hacking-a-talking-toy-parrot/

78. Jaswal VK, Croft AC, Setia AR, Cole CA. Young children have a specific, highly robust bias to trust testimony. Psychol Sci, 2010; 21(10):1541–1547. https://doi.org/10.1177/0956797610383438 PMID: 20855905

79. Jaswal VK, Pérez-Edgar K, Kondrad RL, Palmquist CM, Cole CA, Cole CE. Can't stop believing: Inhibitory control and resistance to misleading testimony. Dev Sci, 2014; 17(6):965–976. https://doi.org/10.1111/desc.12187 PMID: 24806881

80. Ma L, Ganea PA. Dealing with conflicting information: Young children's reliance on what they see versus what they are told. Dev Sci, 2010; 13:151–160. https://doi.org/10.1111/j.1467-7687.2009.00878.x PMID: 20121871

81. Children's Online Privacy Protection Act of 1998, 15 U.S.C. §§ 6501–6506 (1998). Available from: https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title15-chapter91&edition=prelim

82. Steeves V. A dialogic analysis of Hello Barbie's conversations with children. Big Data Soc, 2020; 7 (1):1–12. https://doi.org/10.1177/2053951720919151

83. Danovitch JH, Alzahabi R. Children show selective trust in technological informants. J Cogn Dev, 2013; 14:499–513. https://doi.org/10.1080/15248372.2012.689391

84. Murray GW. Who is more trustworthy, Alexa or Mom?: Children's selective trust in a digital age. Technol Mind Behav, 2021; 2(3). https://doi.org/10.1037/tmb0000050

85. Tempesta E. That doesn't sound like wheels on the bus! Parents freak out when Amazon's Alexa misunderstands young son's request for a song and starts rattling off crude pornographic phrases. The Daily Mail Online. 2016 Dec 30. Available from: https://www.dailymail.co.uk/femail/article-4076568/That-doesn-t-sound-like-Wheels-Bus-Parents-freak-Amazon-s-Alexa-misunderstands-young-son-s-request-song-starts-rattling-crude-PORNOGRAPHIC-phrases.html

86. Hern A. CloudPets stuffed toys leak details of half a million users: Company's data compromised, leaking information including email addresses, passwords, and voice recordings. The Guardian. 2017 Feb 28. Available from: https://www.theguardian.com/technology/2017/feb/28/cloudpets-data-breach-leaks-details-of-500000-children-and-adults

87. Jones ML, Meurer K. Can (and should) Hello Barbie keep a secret? IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS), 2016;:1–6. https://doi.org/10.1109/ETHICS.2016.7560047

88. McReynolds E, Hubbard S, Lau T, Saraf A, Cakmak M, Roesner F. Toys that listen: A study of parents, children, and internet-connected toys. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017;: 5197–5207. https://doi.org/10.1145/3025453.3025735

89. Murnane K. Amazon does the unthinkable and sends Alexa recordings to the wrong person. 2018 Dec 20. In: Forbes [Internet]. Available from: https://www.forbes.com/sites/kevinmurnane/2018/12/20/amazon-does-the-unthinkable-and-sends-alexa-recordings-to-the-wrong-person/?sh=7b4d69bd3ca5

90. Crutzen R, Peters GY, Portugal SD, Fisser EM, Grolleman JJ. An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: An exploratory study. J Adol Health, 2011; 48(5):514–519. https://doi.org/10.1016/j.jadohealth.2010.09.002 PMID: 21501812

91. Miner AS, Laranjo L, Kocaballi AB. Chatbots in the fight against the COVID-19 pandemic. NPJ Digit Med, 2020; 3(65). https://doi.org/10.1038/s41746-020-0280-0 PMID: 32377576

92. World Health Organization. WHO Health Alert brings COVID-19 facts to billions via WhatsApp. In: World Health Organization [Internet]. Available from: https://www.who.int/news-room/feature-stories/detail/who-health-alert-brings-covid-19-facts-to-billions-via-whatsapp

93. Leggett K. 2018 customer service trends: How operations become faster, cheaper–and yet, more human. Forrester; 2018 Jan 24. p. 4, Fig 2. Available from: https://lmistatic.blob.core.windows.net/document-library/boldchat/pdf/en/forrester-2018-customer-service-trends.pdf

94. Newitz A. Ashley Madison code shows more women, and more bots. 2015 Aug 31. In: Gizmodo [Internet]. Available from: https://gizmodo.com/ashley-madison-code-shows-more-women-and-more-bots-1727613924

95. Ishowo-Oloko F, Bonnefon J, Soroye Z, Crandall J, Rahwan I, Rahwan T. Behavioural evidence for a transparency-efficiency tradeoff in human-machine cooperation. Nat Mach Intell, 2019; 1:517–521. https://doi.org/10.1038/s42256-019-0113-5

96. Christian B. The most human human: What artificial intelligence teaches us about being alive. New York: Anchor Books; 2012.

97. Stuart-Ulin CS. Microsoft's politically correct chatbot is even worse than its racist one. 2018 Jul 31. In: Quartz [Internet]. Available from: https://qz.com/1340990/microsofts-politically-correct-chat-bot-is-even-worse-than-its-racist-one/

98. Epstein R. From Russia, with love: How I got fooled (and somewhat humiliated) by a computer. Scientific American Mind, 16–17. 2007 Oct/Nov. Available from: https://www.drrobertepstein.com/downloads/FROM_RUSSIA_WITH_LOVE-Epstein-Sci_Am_Mind-Oct-Nov2007.pdf

99. Brandt M. Smart assistants are getting smarter. 2018 Apr 25. In: Statista [Internet}. Available from: https://www.statista.com/chart/9580/how-smart-are-smart-assistants/

100. Enge E. Rating the smarts of the digital personal assistants in 2019. 2019 Oct 24. In: Perficient [Internet]. Available from: https://www.perficient.com/insights/research-hub/digital-personal-assistants-study

101. Ferrand J, Hockensmith R, Houghton RF, Walsh-Buhi ER. Evaluating smart assistant responses for accuracy and misinformation regarding Human Papillomavirus vaccination: Content analysis study. J Med Internet Res, 2020; 22(8):e19018. https://doi.org/10.2196/19018 PMID: 32744508

102. Lovato SB, Piper AM, Wartella EA. Hey Google, do unicorns exist? Conversational agents as a path to answers to children's questions. Proceedings of the 18th ACM International Conference on Interaction Design & Children. 2019 Jun;: 301–313. https://doi.org/10.1145/3311927.3323150

103. Nobles AL, Leas EC, Caputi TL, Zhu S, Strathdee SA, Ayers JW. Responses to addiction help-seeking from Alexa, Siri, Google Assistant, Cortana, and Bixby intelligent virtual assistants. NPJ Digit Med, 2020; 3:1–3. https://doi.org/10.1038/s41746-019-0211-0 PMID: 31934645

104. Palanica A, Fossat Y. Medication name comprehension of intelligent virtual assistants: A comparison of Amazon Alexa, Google Assistant, and Apple Siri between 2019 and 2021. Front Digit Health, 2021; 3:669971. https://doi.org/10.3389/fdgth.2021.669971 PMID: 34713143

105. Boyd M, Wilson N. Just ask Siri? A pilot study comparing smartphone digital assistants and laptop Google searches for smoking cessation advice. PLoS ONE, 2018; 13(3):e0194811. https://doi.org/10.1371/journal.pone.0194811 PMID: 29590168

106. Toader D, Boca G, Toader R, Măcelaru M, Toader C, Ighian D, et al. The effect of social presence and chatbot errors on trust. Sustainability, 2019; 12(1):256. https://doi.org/10.3390/su12010256

107. Shank DB, Gott A. Exposed by AIs! People personally witness artificial intelligence exposing personal information and exposing people to undesirable content. International Journal of Human-Computer Interaction, 2020; 36(17):1636–1645. https://doi.org/10.1080/10447318.2020.1768674

108. Panzarino M. Apple switches from Bing to Google for Siri web search results on iOS and Spotlight on Mac. 2017 Sep 25. In: TechCrunch [Internet]. Available from: https://techcrunch.com/2017/09/25/apple-switches-from-bing-to-google-for-siri-web-search-results-on-ios-and-spotlight-on-mac/

109. Gavin R. Delivering personalized search experiences in Windows 10 through Cortana. 2016 Apr 28. In: Windows Blog [Internet]. Available from: https://blogs.windows.com/windowsexperience/2016/04/28/delivering-personalized-search-experiences-in-windows-10-through-cortana/

110. Baca MC. Amazon starts crowdsourcing Alexa responses from the public. What could possibly go wrong? Amazon opens its Alexa Answers project to the public and gambles on the goodwill of the Internet. The Washington Post. 2019 Sep 13. Available from: https://www.washingtonpost.com/technology/2019/09/13/amazon-starts-crowdsourcing-alexa-responses-public-what-could-possibly-go-wrong/

111. Wiggers K. Amazon is poorly vetting Alexa's user-submitted answers. 2019 Nov 1. In: VentureBeat [Internet]. Available from: https://venturebeat.com/2019/11/01/amazon-alexa-answers-vetting-user-questions/

112. Newton C. The Verge Tech Survey 2020. The Verge. 2020 Mar 2. Available from: https://www.theverge.com/2020/3/2/21144680/verge-tech-survey-2020-trust-privacy-security-facebook-amazon-google-apple

113. Wang Y, Wu L, Luo L, Zhang Y, Dong G. Short-term internet search using makes people rely on search engines when facing unknown issues. PLoS ONE, 2017; 12(4):e0176325. https://doi.org/10.1371/journal.pone.0176325 PMID: 28441408

114. Koerber B. Amazon Alexa spouted conspiracy theory when asked about chemtrails: 'For a purpose undisclosed to the general public in clandestine programs directed by government officials. 2018 Apr 11. In: Mashable [Internet]. Available from: https://mashable.com/article/amazon-alexa-chemtrails-conspiracy-theory-echo

115. Chung H, Iorga M, Voas J, Lee S. Alexa, can I trust you? Computer, 2017; 50(9):100–104. https://doi.org/10.1109/MC.2017.3571053 PMID: 29213147

116. Sullivan D. Google's "One True Answer" problem–when featured snippets go bad: Obama's planning a coup? Women are evil? Several presidents were in the KKK? Republicans are Nazis? Google can go spectacularly wrong with some of its direct answers. 2017 Mar 5. In: Search Engine Land [Internet]. Available from: https://searchengineland.com/googles-one-true-answer-problem-featured-snippets-270549

117. Thompson A. Google listed "Nazism" as the ideology of the California Republican Party: "Nazism" appears as the first "ideology" of California Republicans in the "knowledge box" of search results. Vice News. 2018 May 31. Available from: https://www.vice.com/en/article/vbq38d/google-is-listing-nazism-as-the-first-ideology-of-the-california-republican-party

118. Soulo T. Ahref's study of 2 million featured snippets: 10 important takeaways. 2017 May 29. In: Ahrefs-blog [Internet]. Available from: https://ahrefs.com/blog/featured-snippets-study/

119. Scull A. Dr Google will see you now: Google health information previews and implications for consumer health. Med Ref Serv Q, 2020; 39(2):165–173. https://doi.org/10.1080/02763869.2020.1726151 PMID: 32329674

120. Hillygus DS, Shields TG. The persuadable voter. Princeton: Princeton University Press; 2009.

121. Liu Y, Ye C, Sun J, Jiang Y, Wang H. (2021). Modeling undecided voters to forecast elections: From bandwagon behavior and the spiral of silence perspective. Int J Forecast, 2021; 37(2):461–483. https://doi.org/10.1016/j.ijforecast.2020.06.011

122. Mayer WG. The swing voter in American politics. Washington, D.C.: Brookings Institutional Press; 2008.

123. Sheehan KB. Crowdsourcing research: Data collection with Amazon's Mechanical Turk. Commun Monogr, 2017; 85(1):140–156. https://doi.org/10.1080/03637751.2017.1342043

124. Loftus EF. Leading questions and the eyewitness report. *Cogn Psychol*, 1975; 7(4):560–572. https://doi.org/10.1016/0010-0285(75)90023-7

125. Ray L. 2020 Google search survey: How much do users trust their search results? 2020 Mar 2. In: Moz [Internet]. Available from: https://moz.com/blog/2020-google-search-survey

126. Robertson A. It's time to stop trusting Google search already. 2017 Nov 10. In: The Verge [Internet]. Available from: https://www.theverge.com/2017/11/10/16633574/stop-trusting-google-search-texas-shooting-twitter-misinformation

127. Epstein R. "The Selfish Ledger" [Transcript]. 2018 May 25. Available from: https://aibrt.org/downloads/GOOGLE-Selfish_Ledger-TRANSCRIPT.pdf

128. Kulwin N. The Internet Apologizes. . . Even those who designed our digital world are aghast at what they created. A breakdown of what went wrong—from the architects who built it. New York Magazine. 2018 Apr 16. Available from: http://nymag.com/intelligencer/2018/04/an-apology-for-the-internet-from-the-people-who-built-it.html

129. Epstein R. Google's snoops: Mining our private data for profit and pleasure. Dissent. 2014 May. Available from: https://www.dissentmagazine.org/online_articles/googles-snoops-mining-our-data-for-profit-and-pleasure

130. Google 'accidentally' snooped on wifi data. Network Security, 2010; 2010(5):2. https://doi.org/10.1016/S1353-4858(10)70052-0

131. Hill K. Wi-spy Google engineer outed as 'hacker' 'god' Marius Milner. 2012 May 1. In: Forbes [Internet]. Available from: https://www.forbes.com/sites/kashmirhill/2012/05/01/wi-spy-google-engineer-outed-as-hacker-god-marius-milner/?sh=63fe94011d74

132. Epstein R, Bock S, Peirson L, Wang H. Large-scale monitoring of Big Tech political manipulations in the 2020 Presidential election and 2021 Senate runoffs, and why monitoring is essential for democracy. Paper presented at: 24th annual meeting of the American Association of Behavioral and Social Sciences. 2021 Jun 14. Available from: https://aibrt.org/downloads/EPSTEIN_et_al_2021-Large-Scale_Monitoring_of_Big_Tech_Political_Manipulations-FINAL_w_AUDIO.mp4

133. Epstein R. Taming Big Tech: The case for monitoring. 2018 May 17. In: Hacker Noon [Internet]. Available from: https://hackernoon.com/taming-big-tech-5fef0df0f00d

134. Why Google poses a serious threat to democracy, and how to end that threat, United States Senate Judiciary Subcommittee on the Constitution. (2019 Jul 16) (testimony of Robert Epstein). Available from: https://www.judiciary.senate.gov/imo/media/doc/Epstein%20Testimony.pdf

135. Caralle K. 90 percent of political donations from Google-related companies go to Democrats: Study. Washington Examiner. 2018 Sep 7. Available from: https://www.washingtonexaminer.com/policy/technology/90-percent-of-political-donations-from-google-related-companies-go-to-democrats-study

136. Dunn J. The tech industry's major players are firmly behind Hillary Clinton. Business Insider. 2016 Nov 1. Available from: https://www.businessinsider.com/tech-company-donations-clinton-vs-trump-chart-2016-11?r=US&IR=T

137. Pearlstein J. Techies donate to Clinton in droves. To Trump? Not so much. Silicon Valley employees are Hillary Clinton's top funders. Wired. 2016 Aug 31. Available from: https://www.wired.com/2016/08/techies-donate-clinton-droves-trump-not-much/

138. American National Election Studies. The ANES guide to public opinion and electoral behavior: Time of presidential election vote decision 1948–2004. ANES. 2021 Aug 16. Available from: https://electionstudies.org/resources/anes-guide/top-tables/?id=109

139. Annenberg Public Policy Center. 2008 National Annenberg Election Survey Telephone and Online Data Sets (account required to access data). 2010 Dec 8. Available from: https://www.annenbergpublicpolicycenter.org/2008-naes-telephone-and-online-data-sets/

140. Epstein R, Zankich VR. The surprising power of a click requirement: How click requirements and warningns affect users' willingness to disclose personal information. PLoS ONE, 2022; 17(2):e0263097. https://doi.org/10.1371/journal.pone.0263097 PMID: 35180222

141. Eisenhower DD. 'Military-industrial complex speech', transcript, *Yale Law School*, 1961 Jan 17. Available from: https://avalon.law.yale.edu/20th_century/eisenhower001.asp

142. Brockmann H, Drews W, Torpey J. A class for itself? On the worldviews of the new tech elite. PLoS ONE, 2021; 16(1):e0244071. https://doi.org/10.1371/journal.pone.0244071 PMID: 33471828

APPENDIX XV

PLOS ONE

RESEARCH ARTICLE

The surprising power of a click requirement: How click requirements and warnings affect users' willingness to disclose personal information

Robert Epstein*, Vanessa R. Zankich

American Institute for Behavioral Research and Technology, Vista, California, United States of America

* re@aibrt.org

Abstract

What kinds of information and alerts might cause internet users to be more cautious about what they reveal online? We used a 25-item survey to determine whether the strength of Terms of Service (TOS) warnings and the inclusion of a click requirement affect people's willingness to admit to engaging in inappropriate behaviors. A racially and ethnically diverse group of 1,500 people participated in the study; 98.3% were from the US and India and the remainder from 18 other countries. Participants were randomly assigned to five different groups in which warnings and click requirements varied. In the control condition, no warning was provided. In the four experimental groups, two factors were varied in a 2 × 2 factorial design: strength of warning and click requirement. We found that strong warnings were more effective than weak warnings in decreasing personal disclosures and that click requirements added to the deterrent power of both strong and weak warnings. We also found that a commonly used TOS warning has no impact on disclosures. Participants in the control group provided 32.8% more information than participants in the two click requirement groups combined and 24.3% more information than participants in the four experimental groups combined. The pattern according to which people dropped out of the five different groups sheds further light on the surprising power of the click requirement, as well as on the importance of tracking attrition in online studies.

OPEN ACCESS

Citation: Epstein R, Zankich VR (2022) The surprising power of a click requirement: How click requirements and warnings affect users' willingness to disclose personal information. PLoS ONE 17(2): e0263097. https://doi.org/10.1371/journal.pone.0263097

Editor: Marco Cremonini, Universita degli Studi di Milano, ITALY

Received: October 20, 2021

Accepted: January 11, 2022

Published: February 18, 2022

Data Availability Statement: An anonymized version of the data can be accessed at https://doi.org/10.5281/zenodo.5847375. Data can also be requested from info@aibrt.org. The data have been anonymized to comply with requirements of the sponsoring institution's Institutional Review Board (IRB). The IRB granted exempt status to this study under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB also exempted this study from informed consent requirements (relevant HHS Federal Regulations 45 CFR 46.101.

1. Introduction

Companies and governments are now collecting vast amounts of personal information online every day, and more people are becoming aware of how extensively they are being monitored. Relatively few people, however, are aware of the range of ways in which their private information is being used [1]. Some US states require immediate warnings when telephone conversations are monitored or recorded, presumably to give callers the option of moderating their speech, and research on cigarette warning labels suggests that salient warnings help some consumers behave more prudently [2]. What kinds of privacy-related warnings might cause internet users to be more cautious about what they reveal online?

(b)(2), 45 CFR 46.116(d), 45 CFR 46.117(c)(2), and 45 CFR 46.111).

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

People are becoming increasingly concerned about surveillance driven by new technologies. The National Security Agency (NSA) is a US intelligence and defense agency that specializes in cryptology and information assurance [3]. In 2013, whistleblower Edward Snowden alerted the American public about the NSA's pervasive surveillance of US citizens, a move that resulted in an increase in disapproval of government surveillance, heightened concerns about technology use, and a reduction in visits to websites that government agencies might be monitoring closely [4, 5]. A year after Snowden's disclosures, the Pew Research Center found that only 9% of American adults reported thinking that they have a high degree of control over how their data are being used, and only 6% reported confidence in the privacy and security of their data [1, 6, 7].

Most parents are also concerned about their children's online behavior [8], and 61% report worrying that their teens are disclosing too much personal information online [9]. Meanwhile, teens are now sharing more information about themselves online than they have in the past, with only 9% of teen social media users reporting being "very concerned" about third-party access to their data [10]. In addition, adolescents' concerns about online privacy are not associated with actual willingness to disclose, meaning that those who do express privacy concerns are not necessarily engaging in more privacy-protective behavior [11].

2. Privacy protections, threats, and behavior

2.1 Privacy protections outside the tech industry

Government often steps in to protect consumer data and soothe privacy concerns. Some US states require "dual consent" in phone calls, meaning that all participants on a call must be fully aware that the call is being monitored or recorded [12]. The US Federal Trade Commission (FTC) requires that any information provided by businesses that might affect consumers' behavior must be accurate [13]. There also exist laws in the US–so-called "Peeping Tom Laws"–that make it a misdemeanor to spy on or photograph someone in a private place without that person's consent [14]. These laws also prohibit nonconsensual video surveillance, or "video voyeurism," in places where there is a reasonable expectation of privacy (e.g., bathrooms, bedrooms, changing rooms, etc.).

The US Privacy Act of 1974 protects records collected by the US government that contain citizens' personal identifiers, including names and social security numbers [15]. The Privacy Act also states that individuals have the right to seek access to and request correction of any such records about them and prohibits collection and disclosure of such records without the consent of the individual to whom the records pertain. The US Fair Credit Reporting Act of 1970 holds credit reporting businesses responsible for the accuracy and security of personal information that is collected about consumers and then shared with third parties [16]. The US Gramm-Leach-Bliley Act of 1999 protects financial nonpublic personal information by requiring financial institutions to clearly and conspicuously explain their privacy practices and to safeguard any sensitive data they possess [17].

Healthcare records in the US are protected by the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and other privacy laws that require healthcare providers to acquire patients' written consent before disclosing their sensitive health information to other people and organizations [18, 19].

2.2 Privacy constraints on tech companies

The US Children's Online Privacy Protection Act (COPPA) of 1998 was intended to protect the personal information of children 12 and under by prohibiting online companies from asking for any of their personally identifiable information without parental consent [20].

But technology companies that have emerged worldwide over the past two decades are largely unregulated, and it is only recently that a few aggressive laws and regulations have been implemented that attempt to safeguard consumer privacy. The most ambitious law passed so far is the European Union's General Data Protection Regulation (GDPR), which became effective in 2018. Among other things, the GDPR guarantees, at least in theory, that consumers can have their personal data erased, can find out how their personal data are being used, and can shift their data to other platforms [21]. As a practical matter, however, it is not clear that the GDPR has actually changed pervasive business practices or has benefitted consumers, and some evidence suggests that because of the regulatory burden the GDPR presents, it has hurt small companies and startups in Europe while benefitting the Big Tech companies [22]. Meanwhile, several countries outside the EU have implemented similar regulations, and so has the US state of California [23]. More limited data privacy laws have been enacted by the US states of Nevada and Maine [24, 25].

Unfortunately, most if not all of the new and upcoming privacy rules give tech companies free rein when they have the consent of users, and users often have no idea they have given such consent [26–28]. Few users have ever fully read a Terms of Service (TOS) agreement or Privacy Policy, and tech companies often find ways around the rules [29–32]. "When you use our services," begins Google's 3,000-word Privacy Policy, "you're trusting us with your information" [33]. A link to that Privacy Policy is embedded in Google's 1,900-word Terms of Service Agreement [34]. Unfortunately, people are agreeing to the terms of both of these agreements even if they don't know they are using a Google service, which is the case most of the time. Millions of websites incorporate Google Analytics, for example, which helps website owners track visitors to their sites [35]. But Google Analytics is invisible to users. Its presence on a website, however, allows Google to track everything users do on that website. Users have inadvertently given their consent when they have unknowingly started using a Google service, and that makes the GDPR and similar regulations largely ineffectual.

2.3 Other tech threats to privacy

Implied consent is just one of many privacy problems that new technologies pose. Because fortunes can be made quickly with newly deployed computer code, most new code is poorly written, which often means it is vulnerable to hacking and infiltration [36]. This puts users' sensitive information, including login credentials, healthcare records, financial records, email content, and browsing history, at risk. Between March 2016 and March 2017, 1.9 billion usernames and passwords were exposed by data breaches and traded on black-market forums [37]. In 2010, it was discovered that Google Street View vehicles weren't just taking pictures of people's homes and businesses; they were also vacuuming up gigabytes of unprotected Wi-Fi data, including passwords, and they had been doing so in 30 countries for 3 years [38]. The ease of hacking, along with the fact that it is virtually impossible to erase data from the internet (all of which is vulnerable to hacking) [39], reminds us that the internet was not designed with security in mind.

Privacy is also at risk online because of toothless regulations and laws. COPPA, for example, supposedly shields children ages 12 and under, but a child of any age can gain full access to a pornography website simply by clicking a button reading "I am over 18." One survey found that 7.5 million of Facebook's users were under age 13, demonstrating how difficult it can be for sites to verify the ages of their users [40, 41]. COPPA also fails to protect young people over age 12, leaving a large gap in the protection of America's youth.

Because corporations are driven by profit, their privacy policies tend to undermine user privacy rather than protect it, and they often use design features–or "dark patterns" [42]–to frustrate, confuse, or coerce users into participation [43–45]–a practice called "malicious interface design" [46]. Privacy policies are not only excessively lengthy (the average American would need to set aside almost 250 hours to properly read all the digital contracts they accept while using online services), they are also often written in language that is difficult to understand [47–49]. This is the case even for policies regarding sensitive health information [50].

2.4 Privacy-protective behavior

One would think that people concerned about privacy would make an effort to protect it, but this is often not the case. The gap between the concern people express about privacy and their actual privacy-protective behavior is called "the privacy paradox" [51]. Even people who express the highest degree of concern sometimes knowingly disclose personal information online [52, 53, cf. 54], and even those who are technically skilled or confident in their ability to protect their own privacy often fail to protect their privacy [55, 56]. Most users are simply unwilling to invest the time and energy required to assure the protection of their personal information, and, generally speaking, people's privacy concerns are easily overridden by the various ways in which they benefit by disclosing information [28, 57–59, cf. 60]. For example, simple benefits such as monetary discounts or rewards tend to increase disclosure [61–64]. Privacy concerns are also overridden by perceived control; because people believe that they are powerless against data collection, they often fail to take steps to protect their privacy [65, cf. 66].

The risk/reward model may be only partially relevant to the privacy paradox, however. Because of the rapid and highly interactive way in which users interact with computers and mobile devices, they often don't have time to make decisions about the information they are asked to disclose [67]. They are simply reacting mindlessly to queries, clicking on buttons or boxes, or pressing the Enter key without giving much thought to what they are doing [68].

2.5 Predictors of privacy behavior

Age and personality traits can be predictive of disclosure. Younger adults are more likely to disclose personal information than older adults [69, cf. 70]; more extroverted people and those who report low self-control are more likely to disclose intimate information online [71]; and those who rank higher in openness, lower in conscientiousness, and lower in agreeableness are more likely to disclose more information online [72]. Privacy awareness and confidence in one's own ability to mitigate privacy concerns can predict privacy decisions [56]. Situational factors also significantly impact people's privacy decisions. For example, the tendency to disclose is higher in large rooms than in smaller ones [73], and familiar environments where people are likely to feel a greater sense of protection may lead to higher trust and higher disclosure [74]. Disclosure is also higher when requests for information are indirect, rather than direct, and when website interfaces are unprofessional, rather than neutral or professional [75]. People are also more likely to pay for more privacy (for example, by shopping at a different website) when a privacy warning is too salient [76]. The perceived sensitivity of the information requested is another predictor of privacy behavior [77], and because different types of information, such as location, health status, and browsing history, are valued differently by different people, one cannot expect privacy behaviors to be consistent across situations [78]. After being told that other people have revealed certain types of information, people are more likely to reveal similar information themselves [79], and a similar phenomenon has even been observed when people interact with an avatar; people reveal more information to an avatar after it has shared information about itself [80]. Self-disclosure activates the brain's reward system, perhaps demonstrating its intrinsic value [81].

2.6 Methods for influencing privacy concerns and behavior

Privacy concerns and privacy-protective behavior can each be impacted in various ways. Although rewards can increase disclosure, some studies have demonstrated that the offer of a reward for disclosing private information can increase privacy concerns [82], especially when the sensitivity of the information requested is high [83]. Including a privacy policy on a website has been shown to increase trust, which is associated with a decrease in disclosure concerns and increased willingness to disclose personal information [84–86]. When a privacy policy is presented as a formal, legalistic agreement, however, trust can deteriorate [87]. Some studies highlight the significance of certain policy features; to increase privacy-protective behaviors, information relevant to privacy decision making must be salient, easily accessible, complete, and threatening [76, 82, 88, cf. 89].

Timing is also important. When people are reminded about privacy at the moment they must make a decision, previously dormant privacy concerns might be awakened, leading to more privacy-protective behavior [90, 91]. This might occur because when users cannot easily bring risks to mind, they mistakenly perceive risk to be low [92].

Certain design features can be used to “push” users to make certain privacy decisions. Nudges–subtle attempts to influence people’s decisions without force [93]–have been used to improve privacy outcomes without limiting users’ choices [94]. For example, presentation nudges are used to provide necessary contextual clues to reduce the user’s cognitive load and convey the appropriate level of risk in order to mitigate biases and heuristics relevant to privacy decision-making [94]. Nudges can draw users’ attention to privacy links and decrease the posting of personal information to public audiences online [95, 96, cf. 97], and nudges that inform users about how they can mitigate privacy risks are more effective at increasing privacy-protective behavior than nudges that rely purely on fear [98]. Priming–exposure to relevant stimuli that influences the response to subsequent stimuli, regardless of awareness [99]–has also been used to deter the disclosure of personal information [100]. Framing–the way an outcome or situation is presented to an audience [101]–is another feature that influences people’s privacy-related decisions [102, 103].

These days, it is increasingly common to see a privacy-related pop-up box–or “cookie consent banner”–whenever one visits a new website. Sometimes the banner informs users that by proceeding onto the website, they are allowing the website owner or its agents to install a variety of unspecified tracking software on their computers; there is no way to opt out of this option. This is referred to as a browsewrap agreement [104]. At other times, the banner lets people click buttons that allow them to limit the tracking to what the website owner considers to be “essential” (which is generally undefined); this is referred to as a clickwrap or click-through agreement. Although clickwrap agreements might have been meant to increase user awareness of privacy threats [105], these banners are often structured in a way that encourages people to surrender their privacy [106]. For example, buttons reading “Join” or “I agree” are often visually more prominent than alternative buttons [68]–a dark pattern that has been shown to increase user acceptance [107]. Whether clickwrap agreements affect people’s tendencies to disclose sensitive information is unknown.

In the present study, we sought to determine the extent to which click requirements and privacy warnings would cause people to withhold sensitive personal information. It employed a randomized, controlled, 2 × 2 factorial design with a diverse sample of participants. Our design also included a control group–people who were not shown privacy warnings and who were not required to click on a clickwrap agreement.

3. Methods

The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB also exempted this study from informed consent requirements (relevant HHS Federal Regulations: 45 CFR 46.101(b)(2), 45 CFR 46.116(d), 45 CFR 46.117(c)(2), and 45 CFR 46.111). The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB is FWA00021545.

3.1 Participants

Our participants were recruited from Amazon's Mechanical Turk (MTurk) website, which has been used by social scientists in a variety of research since 2005 [108]. In all, 1,622 people were randomly assigned to each of 5 groups. Because people take different amounts of time to complete their sessions, we ended up with unequal numbers of people in each group: 306 in Group 1, 307 in Group 2, 327 in Group 3, 314 in Group 4, and 368 in Group 5. After separating the participants in each group who either dropped out of the study after providing demographic information (by closing the browser tab) or who quit the study after seeing the questionnaire (by clicking on an "end session" button), we were left with 304 in Group 1 (2 drops or quits), 304 in Group 2 (3 drops or quits), 304 in Group 3 (23 drops or quits), 305 in Group 4 (9 drops or quits), and 333 in Group 5 (35 drops or quits). Finally, to get an even number of people in each of the groups, we used SPSS's "Random sample of cases" feature to select random samples of 300 people from each group. Our analysis therefore focused on five groups with 300 people in each. We were also able to preserve some information about 72 other people who dropped out of the study before completing it. The dropouts proved to be important in our analysis of the data (see below).

Participants ranged in age from 18 to 82 (M = 32.63 [SD = 10.78]). 853 (56.9%) identified themselves as male, 642 (42.8%) as female, and 4 (0.3%) as Other; gender was not reported for 1 (0.1%) of our participants. 1,215 (81.0%) of our participants were from the US, 259 (17.3%) were from India, and 26 (1.7%) were from 18 other countries. 949 (63.3%) of our participants identified themselves as White and 551 (36.7%) as Non-White in the following categories: 354 (23.6%) as Asian, 91 (6.1%) as Black, 49 (3.3%) as Hispanic, 20 (1.3%) as American Indian, and 37 (2.5%) as Other.

Level of education also varied over a wide range: no high school degree: 6 (0.4%); completed high school: 409 (27.3%); associate's or 2-year degree: 244 (16.3%); bachelor's degree: 642 (42.8%); master's degree: 178 (11.9%); doctoral degree: 21 (1.4%). On a 10-point scale, where 10 was the highest degree of fluency, participants rated their English fluency as high (M = 9.63 [0.86]).

The 72 dropouts were similar to the 1,500 participants who completed the experiment in age [$t(1570) = -1.627$, $p = 0.104$, $d = 0.19$], gender ($z = 0.68$, $p = 0.497$), education ($z = 0.66$, $p = 0.509$), and race/ethnicity ($z = 0.53$, $p = 0.603$).

3.2 Procedure

Participants were directed from the MTurk site to our own web page where they were first asked basic demographic questions. In order to protect the identities of our participants (a requirement of the exempt status granted by the institutional review board of our sponsoring institution), participants were not asked for their full names or email addresses.

They were then instructed to complete a 25-item survey in which they were asked to indicate whether they had engaged in a number of illegal, immoral, or socially controversial

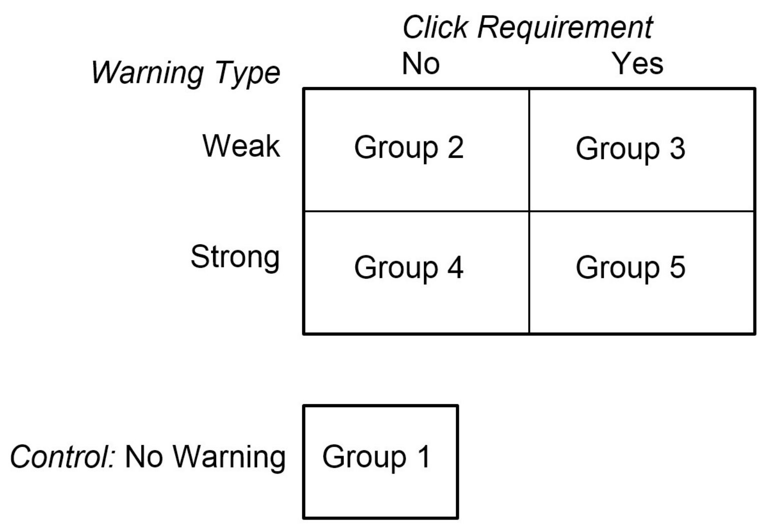


Fig 1. Experimental design. Participants in the control group (Group 1) were not shown a warning above the questionnaire. Participants in the other four groups were presented two levels of warning and two click conditions (click or no click required) in a 2 × 2 factorial design as shown in the figure.

https://doi.org/10.1371/journal.pone.0263097.g001

activities within the past three years–activities such as driving recklessly, watching pornography, smoking marijuana, harming oneself, cursing God, and so on. Participants could respond by clicking "Yes," "No," "Maybe," "Can't remember," or "Click here to end your session." Participants could also terminate their session by closing their browser tab. The number of possible transgressions participants admitted to committing (by clicking "Yes" to items in our survey) served as our dependent variable. The 25 items on the questionnaire were selected from among 100 items we had studied in pilot procedures. We chose the 25 items that were most impacted by our warnings (indicated by the difference between the number of admissions in Groups 1 and 5, see Fig 1). After completing the questionnaire, participants were taken to a debriefing page and thanked for their participation.

Participants were randomly assigned to one of five different groups in which warnings of various sorts were displayed above the questionnaire (see S1 to S5 Figs to view the screens). Warnings stated how data might be used, and in some groups participants were required to click a link to confirm they had understood the warning they were shown. In the control condition, no warning accompanied the questionnaire. In the other four groups, two independent variables were varied in a between-subjects, 2 × 2 factorial design: strength of warning and click requirement (Fig 1).

The weak warning included standard internet language telling the participants that they must comply with a Terms of Service (TOS) agreement: "Please note: By using this website you agree to our Terms of Service" [109] (see S2 and S3 Figs). The strong warning included a brief paragraph reminding participants that their answers and IP addresses were being recorded and that their information might be shared with others, as follows (see S4 and S5 Figs):

PLEASE READ: By using this website you agree to our Terms of Service [109]. Specifically, you give us your consent to record and store your survey answer along with identifying information such as your IP address, as well as to share this information as required or permitted by law with authorized individuals, companies, organizations, or government agencies.

In both warning conditions, participants could click a link to access a lengthy, detailed TOS agreement that contained two links to a lengthy, detailed privacy policy [110] (3,591 words in total). Both documents were modeled after corresponding Google documents. The number of people who clicked these links and the total time they kept these documents open were recorded. Participants in the click groups were required to click on the phrase "Please click here" to acknowledge that they had read and agreed to the TOS.

4. Results

4.1 Analysis of variance

A two-way ANOVA of results in the four experimental groups (2, 3, 4, and 5) revealed main effects for both of our independent variables: warning strength (M_{strong} = 8.32 [5.31], M_{weak} = 9.06 [5.35], $p < 0.05$) and click requirement (M_{click} = 8.12 [5.32], $M_{\text{non-click}}$ = 9.26 [5.30], $p < 0.001$). We also found a statistically significant interaction between these variables: $F(1, 1196) = 8.674$, $p < 0.01$ (Fig 2).

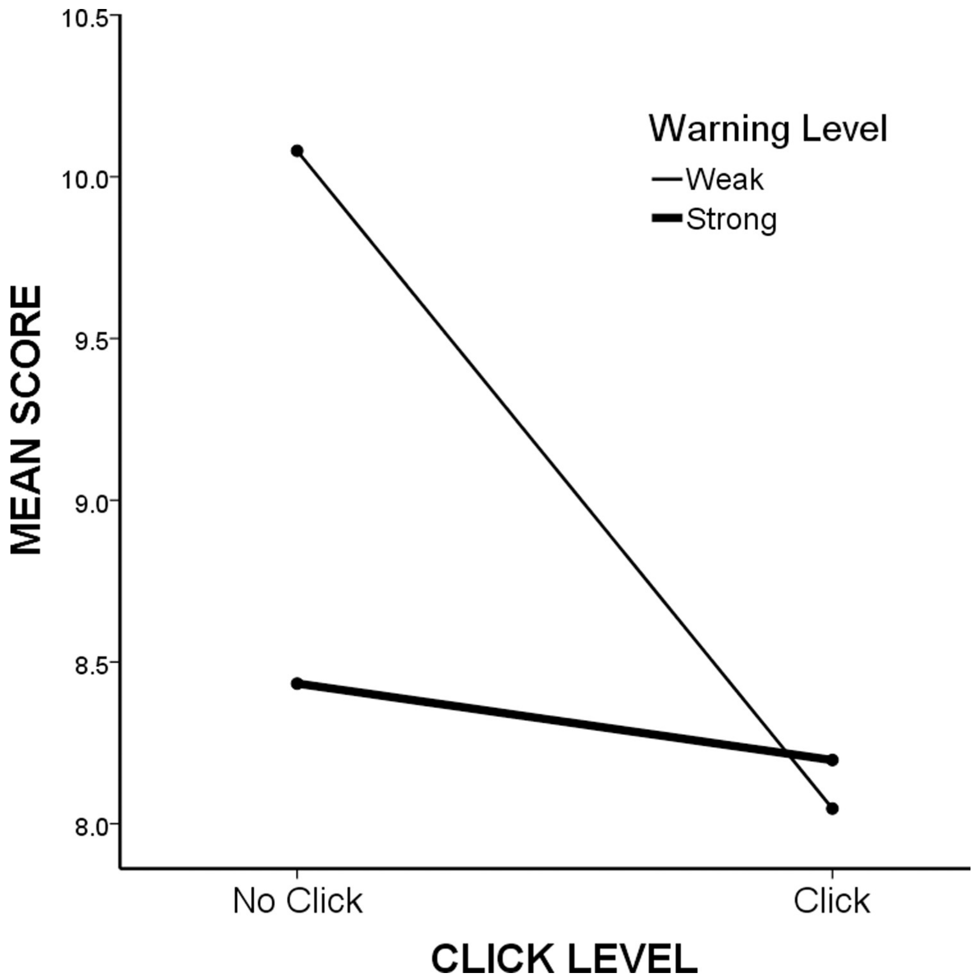


Fig 2. Graphical results of two-way ANOVA. It shows click level (no click or click) versus mean scores (mean number of "yes" responses). The thin line shows data for the weak warning condition, and the thick line shows data for the strong warning condition.

https://doi.org/10.1371/journal.pone.0263097.g002

4.2 Control group versus experimental groups

We also found a significant difference between the mean score of the control group (Group 1, no warnings) and the mean score of the four experimental groups combined (Groups 2, 3, 4, and 5) (M_1 = 10.80 [5.13], $M_{2\text{-}5}$ = 8.69 [5.34], $t(1{,}498) = 6.16$, $p < 0.001$, $d = 0.40$). Pairwise comparisons between the mean score of the control group (Group 1) and the mean scores of three of the four experimental groups (Groups 3, 4, and 5) also produced significant differences (M_3 = 8.05 [0.31], $t(598) = 6.46$, $p < 0.001$, $d = 0.76$; M_4 = 8.43 [5.27], $t(598) = 5.57$, $p < 0.001$, $d = 0.46$; M_5 = 8.20 [5.35], $t(598) = 6.08$, $p < 0.001$, $d = 0.50$). It is notable that the difference in the mean scores between the control group and Group 2 –people receiving the internet's common TOS warning with no click requirement–was not significant (M_2 = 10.08 [5.21]), $t(598) = 1.70$, $p = 0.09$, $d = 0.14$). Where G signifies Group, we can summarize this pattern of results as follows:

$$G1 = G2 < G3 = G4 = G5$$

This pattern shows that when we looked at the amount of sensitive personal information people disclosed, either a strong warning or a click requirement suppressed disclosure significantly. Overall, participants provided 32.8% more information when they had no privacy warning (Group 1) than when they had a click requirement (Groups 3 and 5 combined, M = 8.13[5.32]), and participants provided 24.3% more information when they had no privacy warning (Group 1) than when they had either a click requirement or a warning (Groups 2, 3, 4, and 5 combined, M = 8.69[5.34]).

4.3 Demographic differences

We found a marked difference between disclosures by US participants (Groups 2, 3, 4, and 5 combined, N = 920, M = 9.93 [5.01]) and disclosures by participants from India (Groups 2, 3, 4, and 5 combined, N = 257, M = 4.31 [4.07]) (see Discussion). We also found significant differences in disclosures by gender (M_{male} = 9.80 [5.22], M_{female} = 8.21 [5.42], $t(1493) = 5.76$, $p < 0.001$), $d = 0.30$, race/ethnicity (M_{White} = 10.34 [4.95], M_{Black} = 8.87 [4.93], M_{Hispanic} = 10.61 [4.92], M_{Asian} = 6.00 [5.16], M_{AmIndian} = 7.95 [6.15], M_{Other} = 6.51 [5.72], $F(5, 1494)$ = 41.42, $p < 0.001$), and education (M_{none} = 7.83 [5.19], $M_{\text{highschool}}$ = 10.72 [5.03], $M_{\text{associates}}$ = 9.93 [5.43], $M_{\text{bachelors}}$ = 8.44 [5.34], M_{masters} = 6.88 [4.88], $M_{\text{doctorate}}$ = 8.14 [5.48], $F(5, 1494)$ = 17.79, $p < 0.001$), as well as an effect for age ($r = -.22$, $p < 0.001$).

4.4 Impact and characteristics of dropouts

The power of the click requirement is revealed further when one looks at the pattern according to which people either dropped out of the experiment before completing it by closing their browser tab or by clicking a button we provided which read, "If you have decided not to complete the survey, please click here to end your session" (Fig 3, Table 1). (Henceforward, we will refer to both categories combined under one label: "dropouts.").

The attrition rate in the control group (0.007) was significantly lower than the attrition rate in the experimental groups combined (0.053, z = 3.57, $p < 0.001$) (Table 1). We also found a significant difference in attrition rates across the five groups individually (χ^2[4, N = 1,622] = 48.35, $p < 0.001$). Pairwise comparisons of attrition rates revealed another interesting pattern:

$$G1 = G2 = G4 < G3 = G5$$

In other words, the click requirement (present in Groups 3 and 5 only) affected attrition significantly.

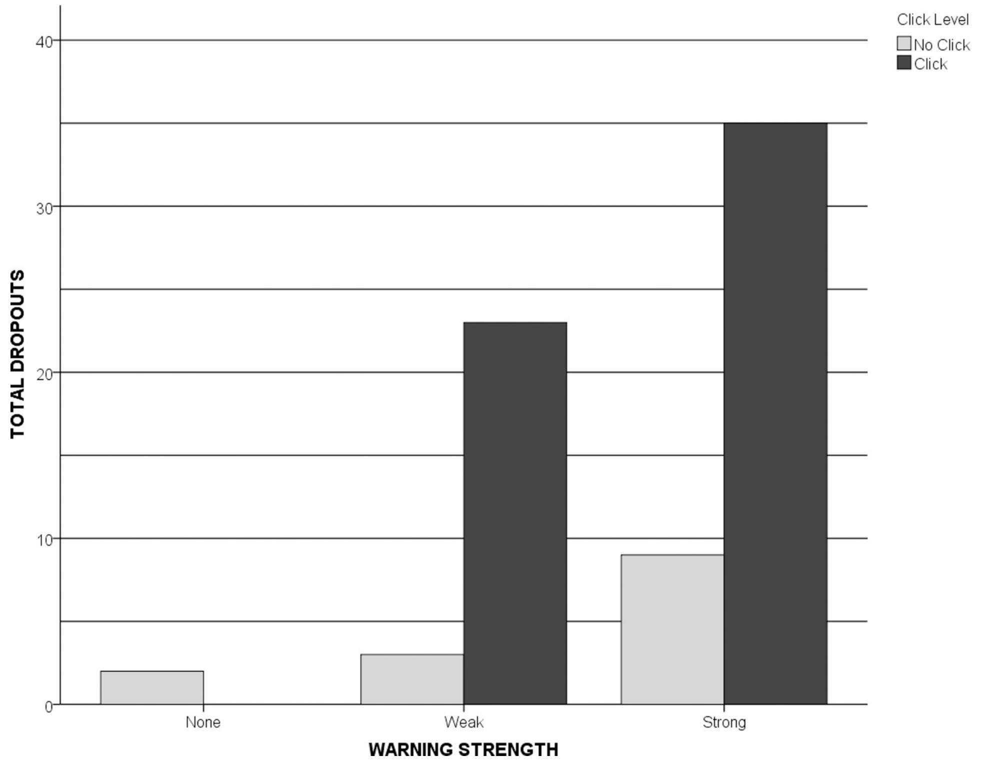


Fig 3. Pattern of dropouts. Written warnings alone drove only a few people away from the study. An added click requirement increased the total number of dropouts substantially (black bars).

https://doi.org/10.1371/journal.pone.0263097.g003

4.5 Terms of service warning

Of the 1,200 people who were prompted to view the TOS agreement, only 88 (7.3%) did so, and only 17 (19.3% of those who viewed the TOS agreement) clicked through to the privacy policy. The average amount of time these people kept these documents open was 22.9 seconds, roughly enough time to read 91 words (2.5% of the total) [111].

5. Conclusions

Our results support four conclusions: (1) The commonly-used TOS warning has no deterrent effect and is functionally the same as no warning at all. (2) A strong, more explicit, warning has some deterrent effect. (3) A click requirement increases the effectiveness of both weak and strong warnings, and it can also cause people to close a web page. (4) Given that most, or perhaps nearly all, internet users are exposed either to no warnings regarding the possible fate of

Table 1. Comparison of dropouts by group number.

Group No.	Total Dropouts	Attrition Rate	Comparison Group	z-score	*p* value
Group 1	2	0.007	Group 2	0.04	0.653
Group 2	3	0.010	Group 4	1.71	0.087
Group 3	23	0.070	Group 5	1.18	0.238
Group 4	9	0.029	Group 3	2.42	< 0.05
Group 5	35	0.095	–	–	–
Groups 2 thru 5	70	0.053	Group 1	3.57	< 0.001

https://doi.org/10.1371/journal.pone.0263097.t001

the information they are providing (our Group 1), or are exposed at most to some mention of a Terms of Service agreement (our Group 2), our results suggest that internet users may currently be disclosing at least 32.8% (our Groups 3 and 5) more personal information than they otherwise would if they were more effectively warned about the risks involved.

6. Discussion

6.1 Insights on attrition

Although we detected only 72 people who left our study before completing it, these people are in some respects the most interesting and revealing in the study. They are interesting from the perspective of experimental design because most studies never track such people. In our early pilot experiments, neither did we, and that sometimes gave us misleading results. When we had a 100-item questionnaire, we sometimes found no effects, presumably because of large attrition rates. With a short questionnaire, we found clear effects among the people who completed the survey, and we also found a clear pattern of attrition associated with the different groups.

The dropout pattern is also interesting in what it might be telling us about how the internet is segmenting societies worldwide. In Dave Egger's 2013 book, The Circle (subsequently made into a movie starring Emma Watson and Tom Hanks), surveillance by a Google-like company has become so pervasive and extreme that some people are going to great lengths to go "off-grid" [112]. The main character, Mae Holland–a rising star at the company–loves the surveillance, but her ex-boyfriend does not. To escape the invasive electronics, he moves to a cabin in the woods, at which point Mae asks her huge cadre of online followers to find him. Minutes later, camera-carrying drones surround his home, at which point he jumps into his pickup and drives straight off a bridge to his death. In other words, he went off-grid by literally going off-grid.

Our dropouts might be giving us a glimpse of yet another aspect of a dark electronic future. They are still connected, but they apparently don't like divulging sensitive personal information. Completely absent from our study is a much larger group of people who are already disconnected–who have quit social media platforms or perhaps never even got hooked. In age, race, gender, and education, our dropouts looked just like the people who completed our study, but we suspect they differed markedly in personality characteristics. Did our dropouts have higher perceived self-efficacy than our finishers [113–115]? Were our finishers more extroverted and open, less conscientious, or perhaps even more exhibitionistic [72, 116, 117] than our dropouts? The billions of people who post messages, photos, and videos of themselves on social media platforms every day hardly seem shy, although some might be sharing their lives online as a response to social pressure [118–120].

The internet might be dividing the world's population into two distinct groups: people we might call "LoudMouths," who compete each day for attention and followers, and people we might call "ZipMouths," who are largely absent from the space that has become increasingly dominant in our lives: Cyberspace. With more and more social science research moving online [121, 122, cf. 123], are important studies drawing erroneous conclusions because of how the internet is segmenting societies? Are we basing our research conclusions on samples that exclude certain personality types? And with major news outlets routinely basing news stories on social media trends [124] and many people turning to social media to get the latest news [125, 126, cf. 127], are Zipmouths losing their ability to influence social policy–perhaps even to influence the outcomes of elections?

What if this trend continues? Although it is clearly in the interest of online entities to extract as much personal data from users as possible, authorities are gradually forcing web hosts to inform users about the risks associated with using their web pages. We see this trend in the increasing number of pop-ups warning us about cookies and other invasions of privacy, some of which now include a click requirement [128]. This practice might cause some people to close a web page and others to divulge less information. Over time, however, such practices will also drive more people off-grid–and, potentially, outside the bounds of a functioning society.

6.2 The power of the click requirement

Warnings–along, of course, with all the fake news, trolling, and bullying–might drive some people off the internet because of their official, legalistic appearance and content. They create the impression that the user is entering into a binding legal contract. A growing body of law in the US suggests, however, that the appearance of a TOS warning alone is not legally binding, but when a user clicks his or her assent to such a warning, courts have ruled that the agreement is binding [27, 39, 104]. The legalistic language in our strong warning might have been essential to its impact [see 87]. That issue should be explored in future research.

Our findings on dropouts also suggest that in studies in which attrition can have a systematic effect on study outcomes, it is essential that attrition be closely tracked. Recall that in some of our pilot experiments (when we used a 100-item questionnaire), we sometimes failed to find effects, almost certainly because we failed to track dropouts.

How effective various types of warnings are in discouraging personal disclosures online is a complex issue. It depends not only on the nature of the warning but also on the value users perceive in divulging such information. In the highly exhibitionistic environments of Facebook, Reddit, and Instagram, photos and disclosures–the more extreme, the better–bring comments, likes, and followers, all of which increase people's social capital, thus increasing their tendency to use social media and disclose more online [117, 129–131]. Disclosing personal information also allows platforms like Google and Facebook to target ads more precisely. For some people, those ads turn the internet into their personal shopper; for others, they are reminders of privacy lost. When we contemplate the power of warnings and click requirements, we also need to think about the rewards associated with the behaviors we are trying to suppress [132]. In many cases and for many people, attempts to suppress disclosures are little more than annoyances [32].

Why a click requirement had such a large impact in our experiment is unclear, but we suspect that this is an attentional phenomena. A click box is a graphical element that draws attention, especially when a click is required in order for a user to proceed. Graphical elements that draw attention on a computer screen have been shown to have a greater impact on user behavior than more subtle graphical elements [133, 134, cf. 135, 136], and that finding is consistent with a long history of research on attention in various contexts [2, 137, 138]. Because required clicks near a warning message also suggest legal liability (which is, as we noted, supported by emerging case law), it is also possible that users who encounter a click requirement are more likely to fear the associated warnings. Our Groups 3 (click requirement with weak warning) and 5 (click requirement with strong warning) begin to shed some light on such issues, but further research, including eye-tracking studies, must be conducted to learn precisely why the click requirement is so powerful.

6.3 Limitations and concerns

The validity of the present study is limited by its sample–a group of people recruited from Amazon's MTurk subject pool. Most were from the US (81.0%), but a sizeable group was from India (17.3%). Further research on warnings and click requirements should reach out to different samples, especially in cultures and countries outside the US [139]. As noted earlier, we found significant and, sometimes, surprisingly large differences in disclosure rates by different demographic groups. Participants from the US, for example, disclosed more than twice as many sensitive activities (M = 9.93 [5.01]) as participants from India did (M = 4.31 [4.07]). That difference could be explained by cultural differences that have been studied by anthropologists and other social scientists [140, 141]. Our study was not structured in a way, however, that allows for meaningful comparisons to be made between different cultures.

Future research on factors affecting online disclosures should also look specifically at (a) types of disclosure that are actually common online, such as information about people's personal lives, along with photos and videos, and (b) how context and environment impact disclosures. Disclosure is the norm, for example, on one's Facebook or Instagram pages, but it often occurs without people's knowledge when they use Google's search engine or Gmail. Warnings and click requirements will almost certainly have to take on very different forms to be effective in the wide range of environments that people now inhabit online, and privacy-promoting techniques that work well with one demographic group might work poorly with another.

Disclosures are also now the norm when people are interacting with personal assistants such as Apple's Siri, Microsoft's Cortana, Amazon's Alexa, and the Google Assistant (standard on Android devices). It is not at all clear how, query by query and device by device, we can meaningfully warn people about the fact that they are disclosing personal information, possibly to their detriment. An increasing body of evidence also indicates that these and other personal assistants constantly record whatever they are hearing [142–144]. Again, how can we meaningfully warn people about invisible surveillance that never stops?

The growing internet of things is rapidly complicating the disclosure problem. In 2014, for example, Google bought Nest Labs [145]. Several years later, it was revealed that Google had installed microphones into the Nest Guard alarm system without disclosing this to users [146]. When the company was called out, it could hardly deny the existence of the microphones, but it claimed it had not yet activated them (then why install them?) [147].

There is good news and bad news here. The good news is that click requirements seem to be surprisingly powerful in discouraging people from disclosing personal information. The bad news is that corporate surveillance is so pervasive and aggressive and so thoroughly embedded into the online environment that no attempts to discourage personal disclosure are likely to make much difference. We join with other scholars and scientists in calling upon our leaders to make the surveillance business model–a fundamentally deceptive model that was invented by Google and that is now being imitated by thousands of businesses worldwide [148]–illegal [148–152].

Supporting information

S1 Fig. Group 1 screen (partial view).
(TIF)

S2 Fig. Group 2 screen (partial view).
(TIF)

S3 Fig. Group 3 screen (partial view).
(TIF)

S4 Fig. Group 4 screen (partial view).
(TIF)

S5 Fig. Group 5 screen (partial view).
(TIF)

Acknowledgments

We thank R. Robertson for help with running the experiment and T. Vasconcellos for help in conducting the literature review. We are especially grateful to J. Johnston for helping us improve our research design.

Author Contributions

Conceptualization: Robert Epstein.

Data curation: Robert Epstein.

Formal analysis: Vanessa R. Zankich.

Investigation: Robert Epstein.

Project administration: Robert Epstein.

Supervision: Robert Epstein.

Writing – original draft: Robert Epstein, Vanessa R. Zankich.

Writing – review & editing: Robert Epstein, Vanessa R. Zankich.

References

1. Auxier B, Rainie L. Key takeaways on Americans' views about privacy, surveillance and data-sharing [Internet]. Pew Research Center. 2019 [cited 14 October 2021]. Available from: https://www.pewresearch.org/fact-tank/2019/11/15/key-takeaways-on-americans-views-about-privacy-surveillance-and-data-sharing/
2. Azagba S, Sharaf MF. The Effect of Graphic Cigarette Warning Labels on Smoking Behavior: Evidence from the Canadian Experience. Nicotine Tob Res, 2012; 15(3):708–17. https://doi.org/10.1093/ntr/nts194 PMID: 22990228
3. Values Mission [Internet]. National Security Agency. [cited 19 October 2021]. Available from: https://www.nsa.gov/about/mission-values/
4. Geiger AW. How Americans have viewed surveillance and privacy since Snowden leaks [Internet]. Pew Research Center. 2018. Available from: https://www.pewresearch.org/fact-tank/2018/06/04/how-americans-have-viewed-government-surveillance-and-privacy-since-snowden-leaks/
5. Penney JW. Chilling effects: Online surveillance and Wikipedia use. Berkeley Technol Law J, 2016; 31 (1):117–182. https://doi.org/10.15779/Z38SS13
6. Madden M, Rainie L. Americans' view about data collection and security [Internet]. Pew Research Center. 2015. Available from: https://www.pewresearch.org/internet/2015/05/20/americans-views-about-data-collection-and-security/
7. Auxier B, Rainie L, Anderson M, Perrin A, Kumar M, Turner E. Americans and privacy: Concerned, confused and feeling lack of control over their personal information [Internet]. Pew Research Center. 2020. Available from: https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/
8. Madden M, Cortesi S, Gasser U, Lenhart A, Duggan M. Parents, teens, and online privacy [Internet]. Pew Research Center. 2012. Available from: https://www.pewresearch.org/internet/2012/11/20/parents-teens-and-online-privacy/
9. Anderson M. How parents feel about–and manage–their teens' online behavior and screen time [Internet]. Pew Research Center. 2019. Available from: https://www.pewresearch.org/fact-tank/2019/03/22/how-parents-feel-about-and-manage-their-teens-online-behavior-and-screen-time/
10. Madden M, Lenhart A, Cortesi S, Gasser U, Duggan M, Smith A, et al. Teens, social media, and privacy. Pew Research Center. 2013. Available from: https://www.pewresearch.org/internet/2013/05/21/teens-social-media-and-privacy/

11. Shin W, Kang H. Adolescents' privacy concerns and information disclosure online: The role of parents and the internet. Comput Human Behav, 2016; 54:114–123. https://doi.org/10.1016/j.chb.2015.07.062

12. Call Recording Laws by State | SmithMarco [Internet]. Protecting Consumer Rights. [cited 19 October 2021]. Available from: https://protectingconsumerrights.com/debt-collection-problems/fair-debt-collection-practices-act/call-recording-laws-by-state/

13. 14 Consumer Protection Laws You Should Know [Internet]. Zebersky Payne Shaw Lewenz. 2019 [cited 19 August 2021]. Available from: https://zpllp.com/the-14-areas-of-consumer-protection-laws-you-should-know/

14. Video Voyeurism Prevention Act of 2004. 18 U.S.C. § 1801 et seq. (2004). Available from: https://uscode.house.gov/view.xhtml?req=(title:18%20section:1801%20edition:prelim

15. Privacy Act of 1974, 5 U.S.C. § 552a (1974). Available from: https://www.govinfo.gov/content/pkg/USCODE-2018-title5/pdf/USCODE-2018-title5-partI-chap5-subchapII-sec552a.pdf

16. Fair Credit Reporting Act, 15 U.S.C. § 1681 et seq. (1970). Available from: https://www.govinfo.gov/content/pkg/USCODE-2015-title15/pdf/USCODE-2015-title15-chap41-subchapIII.pdf

17. Act Gramm-Leach-Bliley, 15 U.S.C. §§ 6801–6809, §§ 6821–6827. (1999). Available from: https://www.ftc.gov/enforcement/statutes/gramm-leach-bliley-act

18. Health Insurance Portability and Accountability Act, Pub. L. 104–191. Stat. 1936 (1996). Available from: https://www.congress.gov/104/plaws/publ191/PLAW-104publ191.pdf

19. Health Information Privacy Law and Policy [Internet]. The Office of the National Coordinator for Health Information Technology. 2018 [cited 11 October 2021]. Available from: https://www.healthit.gov/topic/health-information-privacy-law-and-policy

20. Children's Online Privacy Protection Act of 1998, 15 U.S.C. §§ 6501–6506 (1998). Available from: https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title15-chapter91&edition=prelim

21. Data protection under GDPR [Internet]. Your Europe. 2021 [cited 7 October 2021]. Available from: https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/

22. The 10 problems of the GDPR: The US can learn from the EU's mistakes and leapfrog its policy, Senate Judiciary Committee on the General Data Protection Regulation and California Consumer Privacy Act: Opt-ins, Consumer Control, and the Impact on Competition and Innovation. (2019) (testimony of Rosalyn Layton). Available from: https://www.judiciary.senate.gov/imo/media/doc/Layton%20Testimony1.pdf

23. California Consumer Privacy Act, AB 375 (2018). Available from: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375

24. SB 220, 80th Session. (2019). Available from: https://www.leg.state.nv.us/App/NELIS/REL/80th2019/Bill/6365/Overview

25. SP 275, An Act to Protect the Privacy of Online Customer Information, 129th Maine Legislature, First Regular Session. (2019). Available from: http://legislature.maine.gov/legis/bills/bills_129th/chapters/PUBLIC216.asp

26. Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. Science. 2015; 347(6221):509–514. https://doi.org/10.1126/science.aaa1465 PMID: 25635091

27. Dasteel JH. Consumer click arbitration: A review of online consumer arbitration agreements. Yearbook on Arbitration and Mediation, 2017; 9(1). Available from: https://elibrary.law.psu.edu/cgi/viewcontent.cgiarticle=1215&context=arbitrationlawreview

28. Gomez-Barroso JL. Experiments on personal information disclosure: Past and future avenues. Telemat Inform, 2018; 35(5):1473–1490. https://doi.org/10.1016/j.tele.2018.03.017

29. Jensen C, Potts C, Jensen C. Privacy practices of internet users: Self-reports versus observed behavior. Int J Hum Comput Stud, 2005; 63(1–2):203–227. https://doi.org/10.1016/j.ijhcs.2005.04.019

30. Plaut VC, Bartlett RPIII. Blind consent? A social psychological investigation of non-readership of click-through agreements. Law and Human Behavior, 2011; 36(4):293–311. https://doi.org/10.1037/h0093969 PMID: 22849415

31. Malaga RA. Do web privacy policies still matter? Academy of Information and Management Sciences Journal, 2014; 17(1):95–99.

32. Obar JA, Oeldorf-Hirsch A. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. Information, Communication Society, 2018; 23(1):128–147. https://doi.org/10.1080/1369118X.2018.1486870

33. Google. Privacy policy [Internet]. 2021 Jul 1 [cited 2021 Aug 21]. Available from: https://policies.google.com/privacy?hl=en

34. Google. Terms of service [Internet]. 2020 Mar 31 [cited 2020 Aug 21]. Available from: https://policies.google.com/terms?hl=en-US

35. McGee M. As Google Analytics turns 10, we ask: How many websites use it? MarTech [Internet]. 2015 Nov 12. Available from: https://martech.org/as-google-analytics-turns-10-we-ask-how-many-websites-use-it/

36. Schneier B. 2018. Click here to kill everybody: Security and survival in a hyper-connected world. W. W. Norton Company.

37. Thomas K, Li F, Zand A, Barrett J, Ranieri J, Invernizzi L, et al. Data breaches, phishing, or malware? Understanding the risks of stolen credentials. Google Research, 2017. https://doi.org/10.1145/3133956.3134067

38. Google 'accidentally' snooped on wifi data. Network Security, 2010; 2010(5):2. https://doi.org/10.1016/S1353-4858(10)70052-0

39. MacLean CE. It depends: Recasting internet clickwrap, browsewrap, "I agree," and click-through privacy clauses as waivers of adhesion. Cleveland State Law Review, 2017; 65(1):43–58. Available from: https://engagedscholarship.csuohio.edu/clevstlrev/vol65/iss1/7

40. CR survey: 7.5 million Facebook users are under the age of 13, violating the site's terms. [Internet]. Consumer Reports. 2011 [cited 2021 Oct 15]. Available from: https://www.consumerreports.org/media-room/press-releases/2011/05/cr-survey-75-million-facebook-users-are-under-the-age-of-13-violating-the-sites-terms-/

41. Perlroth N. Verifying ages online is a daunting task, even for experts [Internet]. The New York Times. 2012 Jun 17. Available from: https://www.nytimes.com/2012/06/18/technology/verifying-ages-online-is-a-daunting-task-even-for-experts.html

42. Gray CM, Kou Y, Battles B, Hoggatt J, Toombs AL. The dark (patterns) side of UX design. CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Apr 21; pp. 1–14. https://doi.org/10.1145/3173574.3174108

43. Mathur A, Acar G, Friedman MJ, Lucherini E, Mayer J, Chetty M, et al. Dark Patterns at scale: Findings from a crawl of 11k shopping websites. Proceedings of the ACM on Human-Computer Interaction, 2019; 3:1–32. https://doi.org/10.1145/3359216 PMID: 34322658

44. Di Geronimo L, Braz L, Fregnan E, Palomba F, Bacchelli A. UI Dark Patterns and where to find them: A study on mobile applications and user perception. CHI '2020: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3313831.3376600

45. Narayanan A, Mathur A, Chetty M, Kshirsagar M. Dark Patterns: Past, present, and future. Communications of the ACM, 2020; 63(9):42–47. https://doi.org/10.1145/3397884

46. Conti G, Sobiesk E. Malicious interface design: Exploiting the user. WWW '10: Proceedings of the 19th International Conference on World Wide Web. 2010 Apr 26–30; pp. 271–280. https://doi.org/10.1145/1772690.1772719

47. Jensen C, Potts C. Privacy policies as decision-making tools: An evaluation of online privacy notices. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2004;:471–478.

48. LePan N. 2020. Visualizing the length of the fine print, for 14 popular apps. Visual Capitalist, Technology. https://www.visualcapitalist.com/terms-of-service-visualizing-the-length-of-internet-agreements/

49. Reidenberg JR, Breaux T, Cranor LF, French B, Grannis A, Graves JT, et al. Disagreeable privacy policies: Mismatches between meaning and users' understanding. Berkeley Technol Law J, 2015; 30 (1):39–88.

50. Chua HN, Herbland A, Wong SF, Chang Y. Compliance to personal data protection principles: A study of how organizations frame privacy policy notices. Telemat Inform, 2017; 34(4):157–170. https://doi.org/10.1016/j.tele.2017.01.008

51. Barth S, de Jong MDT. The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review. Telemat Inform, 2017; 34 (7):1038–1058. https://doi.org/10.1016/j.tele.2017.04.013

52. Acquisti A, Gross R. Imagined communities: Awareness, Information sharing, and privacy on the Facebook. Privacy Enhancing Technologies. 2006;:36–58.

53. Spiekermann S, Grossklags J, Berendt B. Stated privacy preferences versus actual behaviour in EC environments: A reality check. Buhl HU, Kreyer N, Steck W. (eds) e-Finance. Springer-Verlag, Berlin, Heidelberg, 2001, pp. 129–147. https://doi.org/10.1007/978-3-642-59504-2_8

54. Prince C. Do consumer want to control their personal data? Empirical evidence. Int J Hum Comput Stud, 2018; 110:21–32. https://doi.org/10.1016/j.ijhcs.2017.10.003

55. Barth S, de Jong MDT, Junger M, Hartel PH, Roppelt JC. Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. Telemat Inform, 2019; 41:55–69. https://doi.org/10.1016/j.tele.2019.03.003

56. Lee H, Kobsa A. Confident privacy decision-making in IoT environments. ACM Trans Comput Hum Interact, 2019; 27(1). https://doi.org/10.1145/3364223

57. Seigel DA. Will you accept the government's friend request? Social networks and privacy concerns. PLoS ONE, 2013; 8(11). https://doi.org/10.1371/journal.pone.0080682 PMID: 24312236

58. Shaw J. The watchers: Assaults on privacy in America [Internet]. Harvard Magazine; 2018. Available from: https://www.harvardmagazine.com/2017/01/the-watchers

59. Steinfeld N. Situational user consent for access to personal information: Does purpose make any difference? Telemat Inform, 2019; 48(4). https://doi.org/10.1016/j.tele.2019.101341

60. Keith MJ, Thompson SC, Hale J, Lowry PB, Greer C. Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior. Int J Hum Comput Stud, 2013; 71:1163–1173. https://doi.org/10.1016/j.ijhcs.2013.08.016

61. Babula E, Mrzyglod U, Poszewiecki A. Consumers' need of privacy protection–experimental results. Journal of Scientific Papers, Economics Sociology. 2017; 10(2):74–86.

62. Carrascal JP, Riederer C, Erramilli V, Cherubini M, de Oliveira R. Your browsing behavior for a Big Mac: Economics of personal information online. WWW '13: Proceedings of the 22nd international conference on World Wide Web, 2013 May 13. pp. 189–200. https://doi.org/10.1145/2488388.2488406

63. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav, 2021; 5:247–255. https://doi.org/10.1038/s41562-020-01044-x PMID: 33479505

64. Premazzi K, Castaldo S, Grosso M, Raman, P, Brudvig S, Hofacker CF. Customer information sharing with e-vendors: The roles of incentives and trust. International Journal Electronic Commerce, 2010; 14 (3):63–91. https://doi.org/10.2753/JEC1086-4415140304

65. Xie W, Fowler-Dawson A, Tvauri A. Revealing the relationship between rational fatalism and the online privacy paradox. Behavior and Information Technology, 2018; 38(7):742–759. https://doi.org/10.1080/0144929X.2018.1552717

66. Brandimarte L, Acquisti A, Loewenstein G. Misplaced confidences: Privacy and the control paradox. Social Psychological and Personality Science, 2012; 4(3):340–347. https://doi.org/10.1177/1948550612455931

67. Custers B. Click here to consent forever: Expiry dates for informed consent. Big Data Society, 2016; 3 (1). https://doi.org/10.1177/2053951715624935

68. Obar JA, Oeldorf-Hirsch A. The clickwrap: A political economic mechanism for manufacturing consent on social media. Social Media + Society, 2018; 4(3). https://doi.org/10.1177/2056305118784770

69. Kezer M, Sevi B, Cemalcilar Z, Baruh L. Age differences in privacy attitudes, literacy and privacy management on Facebook. Cyberpsychology, 2016; 10(1). https://doi.org/10.5817/CP2016-1-2

70. Pereira S, Robinson JO, Peoples HA, Gutierrez AM, Majumder MA, McGuire AL, et al. Do privacy and security regulations need a status update? Perspectives from an intergenerational survey. PLoS ONE, 2017; 12(9). https://doi.org/10.1371/journal.pone.0184525 PMID: 28926626

71. Yu S. Does low self-control explain voluntary disclosure of personal information on the internet? Comput Human Behav, 2014; 37:210–215. https://doi.org/10.1016/j.chb.2014.04.055

72. Hollenbaugh EE, Ferris AL. Facebook self-disclosure: Examining the role of traits, social cohesion, and motives. Comput Human Behav, 2014; 30:50–58. https://doi.org/10.1016/j.chb.2013.07.055

73. Okken V, van Rompay T, Pruyn A. Room to move: On spatial constraints and self-disclosure during intimate conversations. Environment Behavior, 2013; 45(6):737–760. https://doi.org/10.1177/0013916512444780

74. Norberg PA, Horne DR, Horner DA. The privacy paradox: Personal information disclosure intentions versus behaviors. J Consum Aff, 2007; 41(1):100–126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

75. John LK, Acquisti A, Loewenstein G. Strangers on a plane: Context-dependent willingness to divulge sensitive information. The Journal of Consumer Research, 2011: 37(5);858–873. https://doi.org/10.1086/656423

76. Tsai JY, Egelman S, Cranor L, Acquisti A. The effect of online eprivacy information on purchasing behavior: An experimental study. Information Systems Research, 2010; 22(2):254–268. https://doi.org/10.1287/isre.1090.02604

77. Mothersbaugh DL, Foxx WK II, Beatty SE, Wang S. Disclosure antecedents in an online service context: The role of sensitivity of information. Journal of Service Research, 2012; 15(1):76–98. https://doi.org/10.1177/1094670511424924

78. Kokolakis S. Privacy attitudes and privacy behavior: A review of current research on the privacy paradox phenomenon. Computers & Security, 2015. https://doi.org/10.1016/j.cose.2015.07.002

79. Acquisti A, John L, Loewenstein G. The impact of relative standards on the propensity to disclose. Journal of Marketing Research. 2012; 49(2):160–174. https://doi.org/10.1509%2Fjmr.09.0215

80. Moon Y. Intimate exchanges: Using computers to elicit self-disclosure from consumers. Journal of Consumer Research, 2000; 26(4):323–339. https://doi.org/10.1086/209566

81. Tamir DI, Mitchell JP. Disclosing information about the self is intrinsically rewarding. PNAS, 2012; 109 (21):8038–8043. https://doi.org/10.1073/pnas.1202129109 PMID: 22566617

82. Andrade EB, Kaltcheva V, Weitz B. Self-disclosure on the web: The impact of privacy policy, reward, and company reputation. Advances in Consumer Research, 2002; 29:350–353.

83. Lee H, Park N, Hwang Y. A new dimension of the digital divide: Exploring the relationship between broadband connection, smartphone use and communication competence. Telemat Inform, 2015; 32 (1):45–56. https://doi.org/10.1016/j.tele.2014.02.001

84. Chen Y-H, Chien S-H, Wu J-J, Tsai P-Y. Impact of signal and experience on trust and trusting behavior. Cyberpsychol Behav Soc Netw, 2010; 13(5):539–546. https://doi.org/10.1089/cyber.2009.0188 PMID: 20950178

85. Hui KL, Teo HH, Lee S-YT. The value of privacy assurance: An exploratory field experiment. MIS Q, 2007; 31(1):19–33. https://doi.org/10.2307/25148779

86. Wu K-W, Huang SY, Yen DC, Popova I. The effect of online privacy policy on consumer privacy concern and trust. Comput Human Behav, 2011; 28(3):889–897. https://doi.org/10.1016/j.chb.2011.12.008

87. Martin K. Do privacy notices matter? Comparing the impact of violating formal privacy notices and informal privacy norms on consumer trust online. The Journal of Legal Studies, 2016; 45(52). https://doi.org/10.1086/688488

88. Krol K, Preibusch S. Control versus effort in privacy warnings for webforms. Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society [Internet]; 2016 Oct 16; Vienna, Austria. https://doi.org/10.1145/2994620.2994640

89. Strahilevitz LJ, Kugler MB. Is privacy policy language irrelevant to consumers? The Journal of Legal Studies, 2016; 45(S2):S69–S95.

90. Hughes-Roberts T. Privacy as a secondary goal problem: An experiment examining control. Information and Computer Security, 2015: 23(4):382–393.

91. Marreiros H, Tonin M, Vlassopoulos M, Schraefel MC. "Now that you mention it": A survey experiment on information, inattention and online privacy. Journal of Economic Behavior Organization, 2017; 140:1–17. https://doi.org/10.1016/j.jebo.2017.03.024

92. Tay SW, Teh PS, Payne SJ. Reasoning about privacy in mobile application install decisions: Risk perception and framing. Int J Hum Comput Stud, 2020; 145. https://doi.org/10.1016/j.ijhcs.2020.102517

93. Thaler RH, Sunstein CR. Nudge: Improving decisions about health, wealth and happiness. New York, NY: Penguin; 2009.

94. Acquisti A, Adjerid I, Balebako R, Brandimarte L, Cranor LF, Komanduri S, et al. Nudges for privacy and security: Understanding and assisting users' choices online. ACM Computing Surveys, 2017; 50 (3). https://doi.org/10.1145/3054926

95. Alemany J, del Val E, Alberola J, García-Fornes A. Enhancing the privacy risk awareness of teenagers in online social networks through soft-paternalism mechanisms. Int J Hum Comput Stud. 2019; 129:27–40.

96. Rodríguez-Priego N, van Bavel R, Monteleone S. The disconnection between privacy notices and information disclosure: An online experiment. Economia Politica, 2016; 33(3):433–461. https://doi.org/10.1007/s40888-016-0040-4

97. Ioannou A, Tussyadiah I, Miller G, Li S, Weick M. Privacy nudges for disclosure of personal information: A systematic literature review and meta-analysis. PLoS ONE, 2021; 16(8). https://doi.org/10.1371/journal.pone.0256822 PMID: 34449821

98. van Bavel R, Rodriguez-Priego N, Vila J, Briggs P. Using protection motivation theory in the design of nudges to improve online security behavior. Int J Hum Comput Stud, 2019; 123:29–39. https://doi.org/10.1016/j.ijhcs.2018.11.003

99. Molden DC. Understanding priming effects in social psychology: What is "social priming" and how does it occur? Social Cognition, 2014; 32:1–11.

100. Sundar S, Kang H, Wu M, Go E, Zhang B. 2013. Unlocking the privacy paradox: Do cognitive heuristics hold the key? CHI '13 Extended Abstracts on Human Factors in Computing Systems. 2013 Apr 27. 811–816. https://doi.org/10.1145/2468356.2468501

101. Entman RM. Framing: Toward clarification of a fracture paradigm. Journal of Communication, 1993; 43(4):51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

102. Baek YM, Bae Y, Jeong I, Kim E, Rhee JW. Changing the default setting for information privacy protection: What and whose personal information can be better protected? The Social Science Journal, 2014; 51(4):523–533. https://doi.org/10.1016/j.soscij.2014.07.002

103. Johnson EJ, Bellman S, Lohse GL. Defaults, framing and privacy: Why opting in-opting out. Marketing Letters, 2001; 13(1):5–15. https://doi.org/10.1023/A:1015044207315

104. Garcia M. Browsewrap: A unique solution to the slippery slope of the clickwrap conundrum. Campbell Law Rev, 2013; 36(1):31–44. http://scholarship.law.campbell.edu/clr?utm_source=scholarship.law.campbell.edu%2Fclr%2Fvol36%2Fiss1%2F2utm_medium=PDFutm_campaign=PDFCoverPages

105. Kunz CL, Del Duca MF, Thayer H, Debrow J. Click-through agreements: Strategies for avoiding disputes on validity of assent. The Business Lawyer, 2001; 57(1):401–429. https://www.steptoe.com/images/content/4/6/v1/4616/220b.pdf

106. Soe TH, Nordberg OE, Guribye F, Slavkovik M. Circumvention by design–dark patterns in cookie consent for online news outlets. NordiCHI ' 20: Proceedings of the 11th Nordic Conference on Human-Computer Interaction. 2020 Oct 25–29. https://doi.org/10.1145/3419249.3420132

107. Utz C, Degeling M, Fahl S, Schaub F, Holz T. Un(informed) consent: Studying GDPR consent notices in the field. CCS '19: SIGSAC Conference on Computer and Communications Security. 2019 Nov 11–15. https://doi.org/10.1145/3319535.3354212

108. Pew Research Center. Research in the crowdsourcing age, a case study [Internet]. 2016 Jul 11. Available from: https://www.pewresearch.org/internet/2016/07/11/what-is-mechanical-turk/

109. "Terms of Service" linked to a 1,518-word document at https://aibrt.org/privacy-click/TOSA.html.

110. "Privacy Policy" linked to a 2,073-word document at https://aibrt.org/privacy-click/privacy_policy.html.

111. Brysebaert M. How many words do we read per minute? A review and meta-analysis of reading rate, 2019. https://doi.org/10.31234/osf.io/xynwg

112. Eggers D. The Circle. Knopf; 2013.

113. Maddux JE, Rogers RW. Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. Journal of Experimental Social Psychology, 1983; 19(5):469–479. https://doi.org/10.1016/0022-1031(83)90023-9

114. Shih D-H, Hsu S-F, Yen DC, Lin C-C. Exploring the individual's behavior on self-disclosure online. Int J Hum Comput Interact, 2012; 28(10):627–645. https://doi.org/10.1080/10447318.2011.654198

115. Tannenbaum MB, Hepler J, Zimmerman RS, Saul L, Jacobs S, Wilson K, et al. Appealing to fear: A meta-analysis of fear appeals effectiveness and theories. Psychological Bulletin, 2015; 141(6):1178–1204. https://doi.org/10.1037/a0039729 PMID: 26501228

116. Sorokowska A, Oleszkiewicz A, Frackowiak T, Pisanski K, Chmiel A, Sorokowski P. Selfies and personality: Who posts self-portrait photographs? Personality and Individual Differences, 2016; 90:119–123.

117. Trepte S, Reinecke L. The reciprocal effects of social network site use and the disposition for self-disclosure: A longitudinal study. Comput Human Behav, 2012; 29(3):1102–1112. https://doi.org/10.1016/j.chb.2012.10.002

118. Masur PK., DiFranzo D, Bazarova NN. Behavioral contagion on social media: Effects of social norms, design interventions, and critical media literacy on self-disclosure. PLoS ONE, 2021; 16(7). https://doi.org/10.1371/journal.pone.0254670 PMID: 34280233

119. Waheed H, Anjum M, Rehman M, Khawaja A. Investigation of user behavior on social networking sites. PLoS ONE, 2017; 12(2). https://doi.org/10.1371/journal.pone.0169693 PMID: 28151963

120. Yang C. Social media social comparison and identity processing styles: Perceived social pressure to be responsive and rumination as mediators. Applied Developmental Science, 2021;:1–12. https://doi.org/10.1080/10888691.2021.1894149

121. Evans JR, Mathur A. The value of online surveys. Internet Research, 2005; 15(2):195–219. https://doi.org/10.1108/10662240510590360

122. Peyton K, Huber GA, Coppock A. The generalizability of online experiments conducted during the COVID-19 pandemic. Journal of Experimental Political Science, 2021;:1–16. https://doi.org/10.1017/XPS.2021.17

123. Burton JW, Cruz N, Hahn U. Reconsidering evidence of moral contagion in online social networks. Nat Hum Behav, 2021. https://doi.org/10.1038/s41562-021-01133-5 PMID: 34112981

124. Harper RA. The social media revolution: Exploring the impact on journalism and news media organizations. Inquiries Journal, 2010; 2(3). Available from: http://www.inquiriesjournal.com/articles/202/the-social-media-revolution-exploring-the-impact-on-journalism-and-news-media-organizations

125. Anderson M, Caumont A. How social media is reshaping news [Internet]. Pew Research Center. 2014. Available from: https://www.pewresearch.org/fact-tank/2014/09/24/how-social-media-is-reshaping-news/

126. Margetts H. Political behaviour and the acoustics of social media. Nat Hum Behav, 2017; 1(4). https://doi.org/10.1038/s41562-017-0086

127. Valenzuela S, Rojas H. Taming the digital information tide to promote equality. Nat Hum Behav, 2019; 3:1134–1136. https://doi.org/10.1038/s41562-019-0700-9 PMID: 31384024

128. Balke J. What's with all the privacy notices and website Cookie notifications? [Internet]. Houston Press. Houston Press; 2021. Available from: https://www.houstonpress.com/news/why-are-you-seeing-privacy-notifications-online-all-of-a-sudden-10510337

129. Ellison NB, Vitak J, Gray R, Lampe C. Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. J Comput Mediat Commun, 2014; 19(4):855–870. https://doi.org/10.1111/jcc4.12078

130. Rykov Y, Koltsova O, Sinyavskaya Y. Effects of user behaviors on accumulation of social capital in an online social network. PLoS ONE, 2020; 15(4). https://doi.org/10.1371/journal.pone.0231837 PMID: 32324820

131. Valenzuela S, Park N, Kee KF. Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation. J Comput Mediat Commun, 2009; 14:875–901. https://doi.org/10.1111/j.1083-6101.2009.01474.x

132. Ramnerö J, Törneke N. The ABCs of human behavior: Behavioral principles for the practicing clinician. New Harbinger. 2008.

133. Bravo-Lillo C, Cranor LF, Downs J, Komanduri S, Reeder RW, Schechter S, et al. Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. Symposium on Usable Privacy and Security; 2013 Jul 24–26.

134. Munoz Y, Chebat J, Suissa JA. Using fear appeals in warning labels to promote responsible gambling among VLT players: The key role of depth of information processing. J Gambl Stud, 2010; 26:593–609. https://doi.org/10.1007/s10899-010-9182-4 PMID: 20217197

135. Lindegren D, Karegar F, Kane B, Perrersson JS. An evaluation of three designs to engage users when providing their consent on smartphones. Behaviour & Information Technology, 2019; 40(4):398–414. https://doi.org/10.1080/0144929X.2019.1697898

136. Mendel J, Mayhorn CB, Hardee JB, West RT, Pak R. The effect of warning design and personalization on user compliance in computer security dialogs. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting. 2010. https://doi.org/10.1177/154193121005402312

137. Sundar RP, Becker MW, Bello NM, Bix L. Quantifying age-related differences in information processing behaviors when viewing prescription drug labels. PLoS ONE; 2012;7(6);e38819. https://doi.org/10.1371/journal.pone.0038819 PMID: 22719955

138. DeCarlo TE. Alcohol warnings and warning labels: An examination of alternative alcohol warning messages and perceived effectiveness. The Journal of Consumer Marketing, 1997; 14(6);448–460. https://doi.org/10.1108/07363769710186060

139. Fleming P, Bayliss AP, Edwards SG, Seger CR. The role of personal data value, culture and self-construal in online privacy behaviour. PLoS ONE, 2021; 16(7). https://doi.org/10.1371/journal.pone.0253568 PMID: 34270577

140. Dincelli E, Goel S. Research design for study of cultural and societal influence on online privacy behavior. Proceedings of IFIP 8.11/11.13 Dewald Roode Information Security Research Workshop; 2015. Available from: https://www.researchgate.net/publication/282252158_Research_Design_for_Study_of_Cultural_and_Societal_Influence_on_Online_Privacy_Behavior

141. Hofstra B, Corten R, van Tubergen F. Understanding the privacy behavior of adolescents on Facebook: The role of peers, popularity and trust. Comput Human Behav, 2016; 60:611–621. https://doi.org/10.1016/j.chb.2016.02.091

142. Edu JS, Such JM, Suarez-Tangil G. Smart home personal assistants: A security and privacy review. ACM Computer Survey, 2020; 53(6). https://doi.org/10.1145/3412383

143. Stanley J. The privacy threat from always-on microphones like the Amazon Echo [Internet]. ACLU. 2017. Available from: https://www.aclu.org/blog/privacy-technology/privacy-threat-always-microphones-amazon-echo

144. Verheyden T, Baert D, Van Hee L, Van Den Heuvel R. Google employees are eavesdropping, even in your living room, VRT NWS has discovered [Internet]. VRT NWS. 2019. Available from: https://www.vrt.be/vrtnws/en/2019/07/10/google-employees-are-eavesdropping-even-in-flemish-living-rooms/

145. Whitney L. Google closes $3.2 billion purchase of Nest [Internet]. CNET. 2014. Available from: https://www.cnet.com/tech/services-and-software/google-closes-3-2-billion-purchase-of-nest/

146. Pino N. Google fesses up to Nest Secure's undisclosed microphone [Internet]. TechRadar. 2019. Available from: https://www.techradar.com/news/google-fesses-up-to-nest-secures-undisclosed-microphone

147. Porter J. Google claims built-in Nest mic was 'never intended to be a secret' [Internet]. The Verge. 2019. Available from: https://www.theverge.com/circuitbreaker/2019/2/20/18232960/google-nest-secure-microphone-google-assistant-built-in-security-privacy

148. Zuboff S. The age of surveillance capitalism: The fight for a human future at a new frontier of power (p. 7). Public Affairs. 2019.

149. Epstein R. Google's dance [Internet]. TIME. 2013 Mar. Available from: https://techland.time.com/2013/03/27/googles-dance/

150. Epstein R. Free isn't freedom: How Silicon Valley tricks us [Internet]. Motherboard. 2016. Available from: https://www.vice.com/en/article/8q8vav/free-isnt-freedom-epstein-essay

151. Schneier B. Surveillance as a business model [Internet]. Schneier on Security. 2013. Available from: https://www.schneier.com/blog/archives/2013/11/surveillance_as_1.html

152. Torbert PM. "Because it is wrong": An essay on the immorality and illegality of the online service contracts of Google and Facebook. Journal of Law, Technology, and the Internet, 2021; 12(1). Available from: https://scholarlycommons.law.case.edu/jolti/vol12/iss1/2/

APPENDIX XVI

PLOS ONE

RESEARCH ARTICLE

What would happen if twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME)

Robert Epstein*, Christina Tyagi, Hongyu Wang

American Institute for Behavioral Research and Technology, Vista, CA, United States of America

* re@aibrt.org

Citation: Epstein R, Tyagi C, Wang H (2023) What would happen if twitter sent consequential messages to only a strategically important subset of users? A quantification of the Targeted Messaging Effect (TME). PLoS ONE 18(7): e0284495. https://doi.org/10.1371/journal.pone.0284495

Editor: Pierluigi Vellucci, Roma Tre University: Universita degli Studi Roma Tre, ITALY

Received: November 18, 2022

Accepted: April 2, 2023

Published: July 27, 2023

Data Availability Statement: Anonymized raw data are accessible at https://doi.org/10.5281/zenodo.7799706. Anonymization was required to comply with the requirement of the sponsoring institution's Institutional Review Board that the identities of the participants be protected in accordance with HHS Federal Regulation 45 CFR 46.101.(b)(2).

Funding: The author(s) received no specific funding for this work.

Abstract

The internet has made possible a number of powerful new forms of influence, some of which are invisible to users and leave no paper trails, which makes them especially problematic. Some of these effects are also controlled almost exclusively by a small number of multinational tech monopolies, which means that, for all practical purposes, these effects cannot be counteracted. In this paper, we introduce and quantify an effect we call the Targeted Messaging Effect (TME)–the differential impact of sending a consequential message, such as a link to a damning news story about a political candidate, to members of just one demographic group, such as a group of undecided voters. A targeted message of this sort might be difficult to detect, and, if it had a significant impact on recipients, it could undermine the integrity of the free-and-fair election. We quantify TME in a series of four randomized, controlled, counterbalanced, double-blind experiments with a total of 2,133 eligible US voters. Participants were first given basic information about two candidates who ran for prime minister of Australia in 2019 (this, to assure that our participants were "undecided"). Then they were instructed to search a set of informational tweets on a Twitter simulator to determine which candidate was stronger on a given issue; on balance, these tweets favored neither candidate. In some conditions, however, tweets were occasionally interrupted by targeted messages (TMs)–news alerts from Twitter itself–with some alerts saying that one of the candidates had just been charged with a crime or had been nominated for a prestigious award. In TM groups, opinions shifted significantly toward the candidate favored by the TMs, and voting preferences shifted by as much as 87%, with only 2.1% of participants in the TM groups aware that they had been viewing biased content.

1. Introduction

Research conducted over the past decade has identified a number of new forms of influence that the internet has made possible. Some of these are among the largest effects ever discovered in the behavioral sciences, and they are of special concern because they can impact people

Competing interests: The authors have declared that no competing interests exist.

without their awareness, because they often leave no paper trails for authorities to trace, and because they are largely controlled by unregulated monopolies [1–3]. Epstein and Robertson showed, for example, that search results that are biased to favor one candidate could shift the voting preferences of undecided voters by as much as 80% after just a single search experience on a Google-like search engine [1], and this effect has since been replicated partially or in full multiple times [4–11]. They also showed that this effect, called the "search engine manipulation effect" (SEME), can easily be masked so that users are unaware that they are viewing biased search results.

In the present paper, we describe and quantify yet another new form of online influence–the Targeted Messaging Effect (TME)–which has all of the most troubling characteristics of SEME and other recently identified forms of online influence [10–14]: it is a large effect; it can influence people without their awareness; it leaves no paper trail; and it is largely controlled worldwide by three unregulated monopolies–Facebook/Meta, Google, and Twitter.

Before we say more about TME per se, we will attempt to put our research on this topic into a larger context. Research on influence over human decision making has been conducted for over a century in multiple fields: business, psychology, sociology, political science, economics, and so on. In political science, for example, Paul F. Lazarsfeld's classic studies in the 1940s and 1950s demonstrated the important role that "political predispositions" played in determining how people reacted to various forms of social influence, and, ultimately, in helping to determine how people voted [15–17]. Political scientists have also shown how voters are influenced by a wide range of factors, among them being the positive or negative connotation of a political message and the presence of a political candidate in media–newspaper coverage and television ads, for example [18–20]. Recent investigations show how voters are influenced by social media content, the online presence of a political candidate, and the perceived personability of a political candidate across different platforms [21–24].

Economists and business experts have developed numerous models to try to understand and predict consumer behavior [25,26]; once again, recent efforts have focused on how search engines, social media platforms, YouTube "influencers" and other new forms of influence made possible by the internet and other new technologies are impacting consumer choices [27,28]. Psychologists have been trying to understand decision making in broad terms applicable, perhaps, to all aspects of life, and they have been especially interested in recent decades in identifying extremely subtle forms of influence that are largely invisible to those affected [29–31].

We believe that SEME, TME, the Answer Bot Effect (ABE) [11], the Search Suggestion Effect (SSE) [12], and other new forms of influence that our research group has been studying over the past decade are fundamentally different than most forms of influence researchers have been studying over the years. Most forms of influence are inherently competitive: billboards, social media campaigns, television commercials, and print and online advertisements, for example. Even most of the shady forms of influence one sometimes reads about in headlines or novels are inherently competitive: ballot stuffing, the rigging of voting machines, vote buying, and so on [32,33]. Competitive forms of influence usually have little net effect for the simple reason that both (or all) sides can employ them. One manipulation might overpower the others when one side has more resources, but resources can shift over time.

The internet was envisioned by its founders to be a great leveler, giving every individual equal voice and giving small companies the ability to compete with giants [34,35], but it quickly evolved into an array of "walled gardens" [36,37] dominated by huge monopolies, each of which quickly gaining virtually exclusive control over specific forms of influence. Outside the Republic of China, Google (through its search engine and its property YouTube) controls access to most information, and Meta (through its properties Facebook, Instagram, and WhatsApp) guides the majority of online social interactions. TikTok has also become popular, accruing over 2 billion first-time downloads since its release in 2016, and it has even become a platform for "forming political coalitions" among young users [38–40]. Although far smaller than Google and Facebook, Twitter dominates the influential world of microblogging, especially in the United States [41,42].

The current walled-garden structure of the internet is highly problematic from an influence perspective. It means that if one of the large platforms favors one candidate, party, cause, or company, it can change people's thinking and behavior on a massive scale without people's awareness, without leaving a paper trail for authorities to trace, and without anyone having the means to counteract the manipulation. To be specific, if Google's search algorithm boosted content in search results that favored Candidate A, unless systems were in place to capture such content–all of which is ephemeral–no one would ever know that this bias existed, even though, in a national election, it could conceivably shift the voting preferences of millions of undecided voters [1–3,10–14]. Even more disturbing, no one could counteract such bias. To put this another way, although two opposing campaign groups might battle each other to try to boost their visibility in search results or in YouTube sequences, *no campaign organization has the means to counteract an action taken by or a policy implemented by the platform itself*–by an executive, a rogue employee, an unattended algorithm, or some combination thereof. The problem worsens when these monopolies favor the same candidate or cause; patterns of campaign donations documented by organizations such as OpenSecrets.org in recent years suggest that major tech companies might in fact be politically aligned [43–45].

TME itself was presaged in a widely-read *New Republic* article by Harvard legal scholar Jonathan Zittrain [46]. As he noted, on Election Day in the US in 2010, Facebook sent go-vote reminders to 61 million Facebook users and, based on a nationwide analysis of voting records, subsequently concluded that its go-vote prompt had caused about 340,000 more people to vote than otherwise would have [47]. The prompt successfully nudged 0.57% of Facebook's sample of eligible voters. That might not sound like much, but that proportion could easily swing a close election. Recall that Donald Trump won the Electoral College vote in 2016 because of a combined vote margin of only 79,646 votes in three US states [48]. If Mark Zuckerberg, CEO of Facebook, had elected to send vote reminders exclusively to supporters of Hillary Clinton on Election Day in 2016, that might have boosted the Clinton vote nationwide by more than 450,000; that number is based on a simple extrapolation from Facebook's 2010 vote manipulation [49].

Zittrain's concerns were legitimate, but, for four reasons, we believe that "digital gerrymandering" is an inappropriate label for this type of manipulation. First, gerrymandering–the relatively permanent redrawing of voting districts–and targeted messaging–the sending of consequential messages to only a subset of a larger group–have at best only one superficial characteristic in common: they each divide up a population in a way that serves the needs of an empowered group. But gerrymandering is a visible and relatively permanent manipulation–so visible and heavy handed that it is often challenged in court [50]. TMs sent to a subgroup online, however, are ephemeral. They impact people and then disappear. They are stored nowhere and cannot be reconstructed, which is why authorities cannot trace them. This is true of company-generated messages on Google's home page, on YouTube (owned by Google), on Twitter, on Facebook and Instagram (owned by Meta), and other popular platforms. On YouTube, no records are kept of the sequences of videos shown to users, nor of that top video in the list, which is the "up-next" video that plays automatically unless the user selects a different video. On Twitter, company-generated tweets show only in the list you see when you first sign on; you can't look at the tweets they showed you the last time you signed on.

And even though TMs can have a large impact on people's opinions and votes (see below), virtually no one is aware that these messages are sent to some people and not others; without a large passive monitoring system in place that captures ephemeral content [51–53], no one can be certain that the manipulation even took place. Although some ephemeral political content was indeed being captured in the weeks leading up to the 2016 Presidential election [2,51], no one, to our knowledge, was tracking targeted messages sent by Facebook. Did Mr. Zuckerberg send out that go-vote reminder to Clinton supporters on Election Day? Unless he or a whistleblower comes forward to inform us, we will never know.

Other forms of online influence exist, of course, such as the influence exerted by thousands of bots launched by a secret organization in Russia to interfere with elections in the US [53–55], or micro-targeted ads posted by the now defunct company Cambridge Analytica in 2016 [56]. But manipulations like these–although occurring on our high-tech internet–are actually traditional in nature and are not, generally speaking, a threat to democracy. If Russian hackers launch a large number of anti-Biden bots, Biden's party or another group of hackers could, in theory, launch its own bots to counter the Russian bots. This type of influence is very much like the influence exerted by billboards and television commercials: It is both visible and competitive [57], and as long as one has the resources, one can counteract it. Internet pioneers such as Tim Berners-Lee envisioned a future internet in which many thousands of relatively small entities would compete with each other for the attention of users [58], just as thousands of news media organizations have competed for people's attention for a century or more. Unfortunately, as Berners-Lee himself has lamented in recent years, as the internet mushroomed in size, it became dominated (outside of mainland China) by "one search engine, one big social network, [and] one Twitter for microblogging" [59].

The dominance of such monopolies has put radically new and powerful means of influence into the hands of a small number of executives. For example, if Facebook–either through its main social media platform (S1 Fig) or through its subsidiary, Instagram (S2 Fig)–occasionally sends its users reminders to vote or reminders to register to vote, how would we know if these messages were being sent to all of its users or just to the members of one political party? The same could be said of Twitter, which currently inserts company-originated messages after every five or six tweets in people's Twitter feeds, and of Google, which has been praised for including large "go-vote" messages on its home page on election days (S3-S5 Figs) [49,60]. If messages of this sort were being targeted to certain groups, unless a whistleblower came forward or a large-scale monitoring system was in place, we would not know, and, as we have noted, we would have no way to counter the manipulation.

Second, the term "digital gerrymander" already has a legitimate meaning in the social sciences. It refers to the use of computers to calculate optimal boundaries for voting districts [61,62]. Typically, this means boundaries that will virtually guarantee that one political party always wins. Computer modeling could also be used, of course, to guarantee maximum *fairness* in political redistricting, but that would rarely serve the interests of the people in power, and they are usually the people in charge of redistricting [63].

Third, the use of TMs for political purposes is just the tip of a very large iceberg. One immensely large class of TMs–targeted *advertisements*–impacts the purchases of millions of people every day. Nearly all of Facebook's income comes from the fees companies pay to send their advertising content to targeted groups–people who appear, based on their Facebook profile and their most recent Facebook postings–to be highly likely to buy specific products from those companies. Because–at least in theory–any company can pay for that kind of advertising, it is inherently competitive and therefore no threat to consumers. But what if Facebook–in other words, the advertising *platform*–decided to ban certain ads or advertisers, or, more ominously, to throttle one company's ads so that they often failed to reach the targeted audience? Again, without independent passive monitoring systems in place to capture the ephemeral content that actually reaches users, the manipulation of ads by platforms like Facebook and Amazon would be impossible to detect [2,51, cf. 64, 65].

Fourth, targeted messaging–especially the messaging controlled by the large tech platforms themselves–can, in theory, influence almost any kind of thinking or behavior, not just political thinking or purchases. Targeted messages were certainly in wide use long before the internet was invented. Consequential messages have been delivered to specific groups of people on cigarette packs, condom boxes, pill containers–even on flimsy pieces of plastic used by cleaning companies to protect freshly cleaned clothes–and research has demonstrated the effectiveness of such messages, especially with certain populations [66,67]. The particular power that biased online messages have to alter thinking and behavior has also been demonstrated [68–70]. This is why we have set about trying to understand and quantify some aspects of the broader mechanism: Specifically, what happens when consequential messages are sent to one group and not another? How far apart can one push the groups? Will salient, high-contrast messages–that is, messages that stand out from a background–have a larger impact than subtle, low-contrast messages? Will the impact of a message increase if it is displayed multiple times? Can a single TM have a significant effect on people's opinions and voting preferences? Do TMs on different platforms have comparable effects?

We will answer these questions in the experiments described herein. All four experiments employed procedures that were randomized, controlled, and double-blind, with all substantive content (such as the names and biographies of political candidates) counterbalanced to eliminate possible order effects.

2. Experiment 1: The impact of five low-contrast, verified, targeted messages on opinions and voting preferences

In our first experiment, we used a simulated Twitter feed to determine whether low-contrast, verified, targeted tweets could be used to shift the opinions and voting preferences of undecided voters. The appearance of these tweets closely matched that of the TMs the Twitter company currently sends to its users: (a) Our 5 TMs had a white background, just as our 30 organic tweets did. (b) The brief headline before the textual content read "Breaking News" in a black font. (c) The blue checkmark (signifying that Twitter had somehow "verified" the source of the tweet) was present on each TM, just as it is, at this writing, on all Twitter TMs (Fig 1). In order to assure that our participants would be "undecided," we asked our US participants to express their views about two political candidates who ran for prime minister of Australia in 2019 [cf. 1].

2.1 Methods

2.1.1 Ethics statement. The federally registered Institutional Review Board (IRB) of the sponsoring institution (American Institute for Behavioral Research and Technology) approved this study with exempt status under HHS rules because (a) the anonymity of participants was preserved and (b) the risk to participants was minimal. The IRB is registered with OHRP under number IRB00009303, and the Federalwide Assurance number for the IRB is FWA00021545. Informed written consent was obtained for all four experiments as specified in the Procedure section of Experiment 1.

2.1.2 Participants. After cleaning, our participant sample for this experiment consisted of 533 eligible US voters recruited through the Amazon Mechanical Turk (MTurk) subject pool [71]. To avoid the growing problem with bots on MTurk, all participants were first screened and confirmed to be human by Cloud Research, a market research company. During the cleaning process, we removed participants who reported an English fluency level below 6 on a 10-point scale, where 1 was labeled "Not fluent" and 10 was labeled "Highly fluent." We also removed participants who had reported a level of familiarity exceeding 3 on a 10-point scale

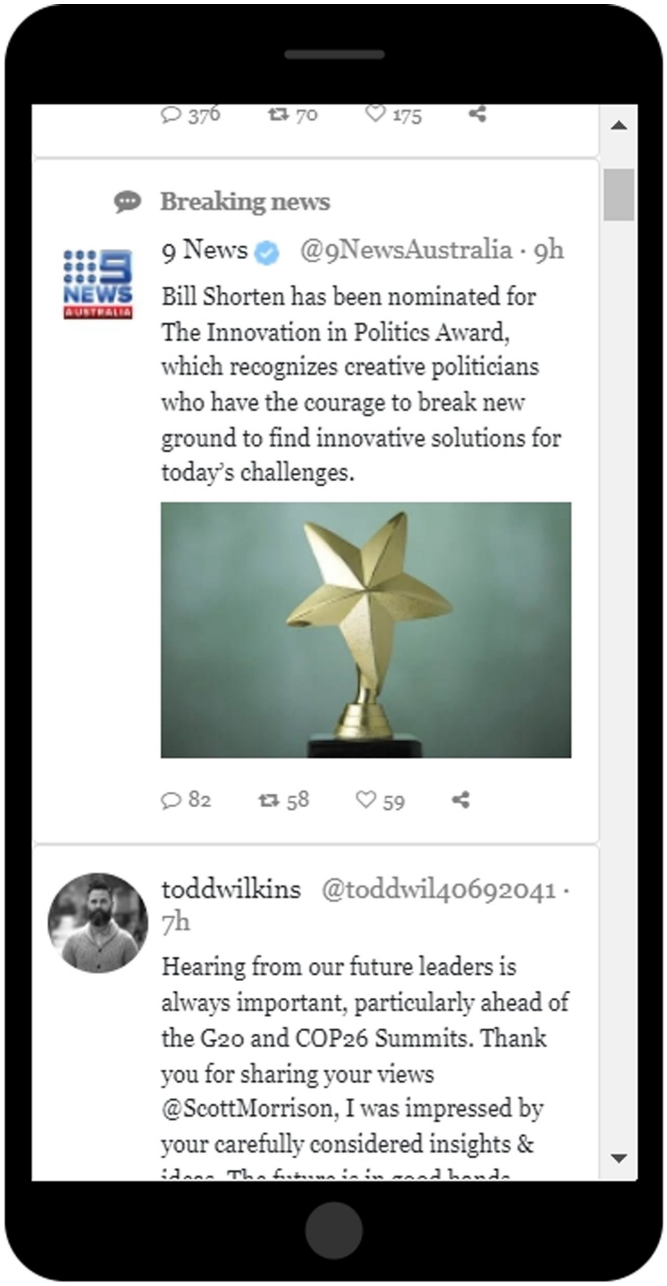


Fig 1. A screenshot showing an image of the first and second tweets in the Twitter feed employed in Experiment 1. The first tweet was a targeted message coming presumably from the Twitter company itself, in this case containing positive information about Bill Shorten. It would thus have been shown to study participants in the Pro-Shorten bias group. Its format was low-contrast (white background, with a black "Breaking News" headline) and included a blue checkmark, signifying verification. The second tweet in the image was an organic tweet sent by a fictitious user.

https://doi.org/10.1371/journal.pone.0284495.g001

with respect to either of the two political candidates referred to in the experiment, where 1 was labeled "Not familiar at all" and 10 was labeled "Very familiar."

Our participants were demographically diverse in gender, age, race, and ethnicity, level of education completed, employment, income, and political leaning. See S1 Table for detailed demographic information for Experiments 1 through 4. The mean familiarity level for our first candidate, Scott Morrison, was 1.13, and the mean familiarity level for our second candidate, Bill Shorten, was 1.05.

2.1.3 Procedure. Each session began with two screening questions. Participants could continue only if they said they were eligible to vote in the US and said no to the question, "Do you know a lot about politics in Australia?" They were then given basic instructions about the experiment, given information about how they could contact the experimenters with any questions or concerns they might have, and asked, in accordance with HHS rules, for their consent to participate. Participants were then asked a series of demographic questions, including questions about their political leanings, and then asked, on 10-point scales from "Not familiar at all" to "Very familiar," how familiar they were with each of two Australian political candidates: Scott Morrison and Bill Shorten, as we noted above.

Participants were then given a short paragraph about each candidate (see S1 Text in Supporting Information for the full paragraphs), each about 120 words in length. Participants were next asked three opinion questions about each candidate: one regarding their overall impression of the candidate, one regarding how likeable they found the candidate, and one regarding how much they trusted the candidate. They were then asked, on an 11-point scale with values ranging from 5 to 0 to 5, which candidate they would be likely to vote for if they "had to vote today." Finally, they were asked which candidate they would vote for if they "had to vote right now" (forced choice).

Participants were now given a task to complete: They would be given an opportunity to scroll through a series of tweets in order to gather information to help them decide which of the two candidates "will do a better job of protecting Australia from foreign threats." They were instructed to scroll through "all the tweets" before making up their minds. See S2 Text for the complete instructions.

On the next screen, participants saw a mobile-phone image displaying a series of tweets (Fig 1). They could scroll through the Twitter feed either by dragging the scroll indicator on the scroll bar (right side of image) up or down, or by rotating the wheel on their mouse. For each participant, the maximum distance they scrolled downward through the Twitter feed was recorded as a percentage of the total distance.

Participants had been randomly assigned to one of three groups: Pro-Shorten, Pro-Morrison, or Control. People in all three groups had access to the same randomized sequence of 30 tweets authored by 30 different fictitious people; all the tweets were composed by the experimenters. Five of the tweets portrayed Bill Shorten in a positive light as a protector of Australia; five portrayed Scott Morrison in this same light; and the other 20 tweets simply commented on various ways of protecting Australia without referring to the candidates. All contained the hashtag #protectAustralia.

For participants in the Pro-Shorten and Pro-Morrison groups, 5 more tweets were added to the original sequence of 30 tweets, so users in these two "bias groups" had access to 35 tweets in all. In the context of this experiment, the five additional tweets should be considered TMs. These were messages presumably coming not from Twitter users but from the Twitter company itself. In real Twitter feeds, we estimate that the Twitter company typically inserts its own tweets roughly 20% of the time. Sometimes these messages are advertisements; sometimes they include links to breaking news stories; and, close to Election Day, they might include reminders to vote or to register to vote (see S6 and S7 Figs).

In the two bias groups, the TMs appeared in positions 2, 7, 12, 25, and 31 in the sequence of 35 tweets available in their Twitter feeds. The ordering and positions of the TMs were not varied. The only difference between the content seen by members of the Control group and members of the two bias groups was that people in the latter groups saw the five TMs, whereas people in the Control group did not. In addition, the only difference between the TMs seen by members of the Pro-Morrison bias group and members of the Pro-Shorten bias group was that people in the former group saw TMs favoring Morrison, whereas people in the latter group saw those same TMs with the names switched, so that they now favored Shorten.

For example, in the Pro-Morrison group, participants saw either strongly negative messages about Shorten such as "Bill Shorten charged with driving under the influence while vacationing in Adelaide," or strongly positive messages about Morrison, such as "Scott Morrison awarded an honorary doctorate from the University of Melbourne, in recognition for his humanitarian efforts during the Australian wildfires." As noted above, the TMs were identical in the pro-Shorten group, except that the candidate's name was changed to his opponent's name (see S3 Text for a complete list of TMs).

The Continue button in the lower-right corner of the web page was inactive for the first 30 seconds of the Twitter session, thus encouraging participants to spend some time reading tweets. If they clicked on the button before it was active, a message was displayed reading, "You have spent too little time reading this page. Please read more." Also to encourage reading, a message appeared at the top of the page above the mobile phone image reading, "Scroll through the tweets below. You will need to spend some time viewing the tweets before you can continue to the next page." A maximum of 5 minutes was allowed for examining the tweets in the Twitter feed.

On the surface, it might not be obvious how sending different tweets to people in different groups qualifies as targeted messaging. That we are indeed targeting our messages should be clearer if one imagines combining all of our participants into one large group. Now imagine that we divide the group up into three subgroups, perhaps based on certain demographic characteristics (such as income, gender, or political leaning). We now target the members of two of those subgroups with tweets favoring, say, one political candidate; we send no such tweets to the third subgroup. This is how targeted messaging works on any platform, and the message can contain almost any content: a prompt to vote or to register to vote, a reminder to buy one's loved one a gift on Valentine's Day, or an advertisement for throat lozenges. The message is targeted as long as it is deliberately being sent to one group and not another, and one knows the targeting has been effective if one can detect predictable changes in the behavior of the targeted group.

Following the Twitter experience, participants were again asked a series of questions. The first question was related to the task that had been assigned earlier. "Based on your Twitter search, which candidate, if either, do you think will do a better job of protecting Australia from foreign threats?" (11-point scale from 5 to 0 to 5). Following the "task" question, participants were again asked the six opinion questions and the two voting questions they had been asked before they saw the Twitter feed (see above).

Next, participants were asked whether any of the content they had seen in the Twitter feed "bothered" them in any way. They could reply yes or no, and then they could explain their answer by typing freely in a text box. This is a conservative way of determining whether people perceived any bias in the content they had seen–especially bias in the TMs that had been shown to people in the two bias groups. We could not ask people directly about their awareness of bias because leading questions of that sort often produce misleading answers [72].

The session concluded with general information about the goals of the research and about how people could withdraw their data from the study if they wished to do so. No participants chose to withdraw their data from any of the four experiments in the present study.

2.2 Results

Although participants were instructed to examine all the tweets in the Twitter feed (35 in the two bias groups, 30 in the control group), 29.0% of them did not comply, scrolling less than the full distance. Rather than discarding people with low scroll scores, we chose, for comparison purposes, to divide the sample into two groups: Low Compliance (maximum scroll values $\leq$ 50%) and High Compliance (maximum scroll values $>$ 50%).

We call the shift in voting preferences "Vote Manipulation Power," or VMP, which is the post-manipulation increase of people in the bias groups (expressed as a percentage increase) who said they would vote for the favored candidate [1]. For details about how VMP is calculated, see S4 Text. In the High Compliance group in Experiment 1, the VMP–the shift in voting preferences toward the favored candidate–was 87.0%, which is larger than any VMPs our team has ever found in SEME experiments [1]. A shift this large can, in theory, turn a 50/50 split among undecided voters into more than a 90/10 split (see S4 Text). The shift in the Low Compliance group–although smaller–was still substantial (Table 1).

In the High Compliance group, answers to all six opinion questions shifted significantly in the direction predicted by the bias; in the Low Compliance group, answers to five of those six questions shifted significantly in that direction (Table 2), with the opinions shifting farther in the High Compliance group. Finally, the voting preferences as expressed on the 11-point opinion scale also shifted significantly and substantially in the predicted direction (see Table 3, where the data have been corrected for counterbalancing and candidate so that larger positive values indicate greater preference for the favored candidate).

In the bias groups, only seven participants (out of 336, 2.1%) expressed concerns about possible bias in the content of the tweets; whereas 113 of these individuals (33.6%) commented specifically on the damaging (but never the positive) information in the biased TMs. Comments such as, "Read that Bill Shorten spent tax payer money, arrested for DUI and had an affair" and "Scott Morrison displayed a lot of bad judgment in his personal life (affairs, DUI arrests, etc.), which made me feel he was untrustworthy," were common. People's focus on the negative content in the TMs is addressed in Experiment 4 below, as well as in our Discussion section.

As we noted, Twitter's TMs look almost exactly like the organic tweets of Twitter users. The main feature that consistently distinguishes the company's TMs from most organic tweets is that their TMs all include the prestigious blue checkmark. In Experiment 2, we attempted to replicate our findings from Experiment 1 while omitting the blue checkmarks from our TMs.

3. Experiment 2: The impact of five low-contrast, non-verified targeted messages on opinions and voting preferences

3.1 Methods

3.1.1 Participants. After cleaning, our participant sample for this experiment consisted of a new group of 532 eligible US voters recruited through the MTurk subject pool, screened once again by Cloud Research (see above). The cleaning procedure was identical to that of

Table 1. Experiment 1: VMPs by compliance level.

Compliance Level	Max Scroll	Total *n*	Bias Groups *n*	Bias Groups Scroll % Mean (*SD*)	VMP (%)	McNemar's Test X^2	*p*
High	$>$ 50	434	287	93.5 (12.9)	87.0	103.14	$<$ 0.001
Low	$\leq$ 50	66	49	34.6 (11.3)	59.3	14.22	$<$ 0.001

https://doi.org/10.1371/journal.pone.0284495.t001

Table 2. Experiment 1: Pre- and post-manipulation opinion ratings of candidates.

Compliance Level		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	Post	Diff	Pre	Post	Diff	$z^{\dagger}$
High	**Impression**	7.06 (1.79)	7.74 (1.98)	0.68	7.04 (1.79)	3.94 (2.03)	-3.10	-12.8***
	Trust	6.18 (1.98)	7.08 (2.06)	0.90	6.11 (1.92)	3.73 (2.16)	-2.38	-12.5***
	Likeability	7.07 (1.80)	7.51 (2.00)	0.44	6.93 (1.80)	4.17 (2.24)	-2.76	-12.0***
Low	**Impression**	7.10 (2.03)	7.55 (2.25)	0.45	7.02 (2.17)	4.02 (2.20)	-3.00	-5.06***
	Trust	5.76 (2.33)	6.63 (2.72)	0.87	5.53 (2.36)	3.53 (2.20)	-2.00	-4.81***
	Likeability	7.31 (1.84)	7.27 (2.18)	-0.04	7.08 (2.06)	4.04 (1.95)	-3.04	-4.89***

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate.

***$p < 0.001$.

https://doi.org/10.1371/journal.pone.0284495.t002

Experiment 1. Once again, the group was demographically diverse. See S1 Table for details about demographic characteristics. The mean familiarity level for our first candidate, Scott Morrison, was 1.10, and the mean familiarity level for our second candidate, Bill Shorten, was 1.04.

3.1.2 Procedure. The procedure in Experiment 2 was identical to that of Experiment 1, with one exception: The blue checkmark was *absent* on our TMs (Fig 2). Since that feature consistently distinguishes Twitter's TMs from most organic tweets, we sought to determine whether the absence of this feature might reduce the impact of TMs on opinions and voting preferences.

3.2 Results

As expected, although the vote shifts were still quite large in both the Low and High Compliance groups, VMP values dropped substantially when the blue checks were absent (Table 4) ($VMP_{Expt1High} = 87.0$, $VMP_{Expt2High} = 61.7$, $z = 8.58$, $p < 0.001$).

Once again, opinions shifted in the predicted directions in all groups (Table 5), and so did mean voting preferences as expressed on the 11-point scale (Table 6). Even without the blue checkmarks on the TMs, participants also appeared to pay as much attention to them in Experiment 2 as in Experiment 1, with only 9 out of the 370 people in the bias groups (1.6%) raising concerns about possible bias in the content, and 115 of those people (31.1%) specifically mentioning the negative (but not the positive) things being said about the candidates in the TMs. The higher VMPs in Experiment 1 suggest that blue checkmarks add credibility to the content of the TMs, but the checkmarks do not seem to reduce the level of attention people are paying to them–or at least to the TMs with negative content.

Could substantially boosting the salience of TMs in a Twitter feed increase their impact on people's opinions and voting preferences? We explore this question in Experiment 3.

Table 3. Experiment 1: Pre- and post-manipulation mean voting preferences on 11-point scale (corrected so that positive values indicate preference for the favored candidate).

Compliance Level	Bias Groups *n*	Pre Voting Preference on 11-Point Scale (*SD*)	Post Voting Preference on 11-Point Scale (*SD*)	Mean Difference	*z*	*p*
High	287	-0.20 (2.77)	2.64 (2.43)	2.84	-12.07	< 0.001
Low	49	-0.04 (2.32)	2.73 (2.46)	2.77	-5.29	< 0.001

https://doi.org/10.1371/journal.pone.0284495.t003

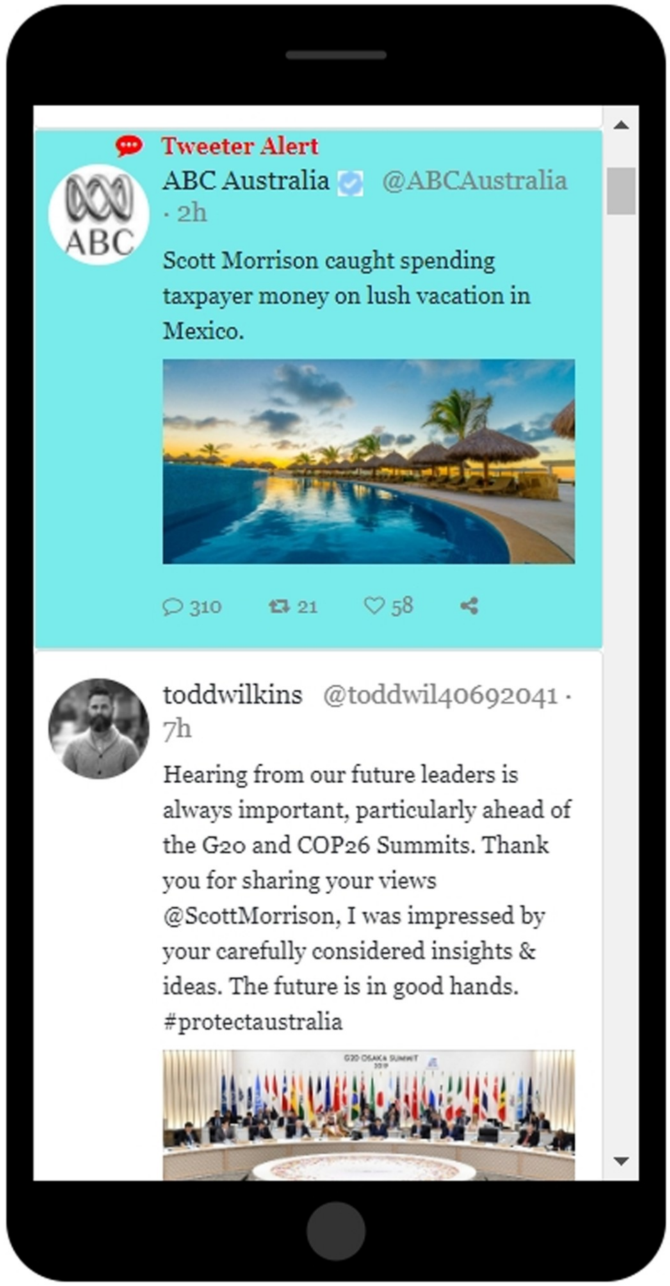


Fig 2. A screenshot showing an image of the second and third tweets in the Twitter feed employed in Experiment 3. The second tweet (top tweet in the image above) was a targeted message coming presumably from the Twitter company itself, in this case containing negative information about Scott Morrison. It would thus have been shown to study participants in the Pro-Shorten bias group. Its format was high-contrast (blue background, with a red "Tweeter Alert" headline).

https://doi.org/10.1371/journal.pone.0284495.g002

4. Experiment 3: Impact of high-contrast, verified targeted messages on opinions and voting preferences

4.1 Methods

4.1.1 Participants. After cleaning, our participant sample for this experiment consisted of a new group of 539 eligible US voters recruited through the MTurk subject pool, again

Table 4. Experiment 2: VMPs by compliance level.

Compliance Level	Total *n*	Bias Groups *n*	Bias Groups Scroll % Mean (*SD*)	VMP (%)	McNemar's Test X^2	*p*
High	447	322	94.5 (12.3)	61.7	97.20	< 0.001
Low	61	48	34.1 (10.1)	44.4	10.29	< 0.01

https://doi.org/10.1371/journal.pone.0284495.t004

screened by Cloud Research (see above). The cleaning procedure was identical to that of Experiment 1. Once again, the group was demographically diverse. See S1 Table for details about demographic characteristics. The mean familiarity level for our first candidate, Scott Morrison, was 1.14, and the mean familiarity level for our second candidate, Bill Shorten, was 1.04.

4.1.2 Procedure. In Experiment 3 we deliberately altered the appearance of our TMs so that they would stand out. Specifically, we gave them blue backgrounds (instead of the usual white), and the message content was preceded by the words "Tweeter Alert" in a red font (Fig 2). The TMs also included Twitter's iconic blue checkmarks. In all other respects, the procedure in Experiment 3 was identical to the procedure in Experiment 1.

4.2 Results

Because, generally speaking, increasing the salience of stimuli increases the attention they attract [73], one might expect that increasing the salience of the TMs would have increased their impact. The VMPs in Experiment 3, however, were significantly *lower* than the VMPs in Experiment 1, (Table 7) ($VMP_{Expt1High} = 87.0$, $VMP_{Expt3High} = 81.1$, $z = 2.15$, $p < 0.05$). Shifts in opinions (with one exception in the Low Compliance group) and voting preference as expressed on the 11-point scale also moved in the direction predicted by bias in the TMs (Tables 8 and 9), but, again, those shifts were lower than the ones we found in Experiment 1.

Once again, comments focused largely on the negative TMs– 98 out of 313 people in the bias groups (31.3%) mentioned negative TMs, compared to only 1 person who mentioned positive TMs and only 6 people (1.9%) who commented on possible bias in the tweets.

The findings from Experiments 1, 2, and 3 suggest that Twitter displays its TMs the way it does–in a manner that makes them difficult to distinguish from organic user tweets–to maximize their impact on users.

Table 5. Experiment 2: Pre- and post-manipulation opinion ratings of candidates.

Compliance Level		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	**Post**	**Diff**	**Pre**	**Post**	**Diff**	$z^{\dagger}$
High	**Impression**	7.05 (1.76)	7.63 (1.86)	0.58	7.00 (1.74)	3.90 (1.97)	-3.10	-13.7^{***}
	Trust	6.04 (2.04)	6.94 (2.05)	0.90	5.94 (2.00)	3.57 (1.94)	-2.37	-13.5^{***}
	Likeability	6.95 (1.82)	7.39 (1.83)	0.44	6.91 (1.70)	4.02 (1.97)	-2.89	-13.6^{***}
Low	**Impression**	7.38 (1.54)	7.52 (1.89)	0.14	7.40 (1.75)	4.17 (2.22)	-3.23	-4.76^{***}
	Trust	6.40 (1.83)	6.73 (2.24)	0.33	6.54 (1.89)	3.88 (2.30)	-2.66	-4.54^{***}
	Likeability	7.13 (1.41)	7.23 (1.75)	0.10	7.17 (1.59)	4.38 (2.64)	-2.79	-4.35^{***}

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate.

$^{***}p < 0.001$.

https://doi.org/10.1371/journal.pone.0284495.t005

Table 6. Experiment 2: Pre- and post-manipulation mean voting preferences on 11-point scale (corrected so that positive values indicate preference for the favored candidate).

Compliance Level	Bias Groups *n*	Pre Voting Preference on 11-Point Scale (*SD*)	Post Voting Preference on 11-Point Scale (*SD*)	Mean Difference	*z*	*p*
High	322	0.20 (2.61)	2.64 (2.39)	2.44	-12.56	< 0.001
Low	48	0.12 (2.66)	2.06 (2.62)	1.94	-3.68	< 0.001

https://doi.org/10.1371/journal.pone.0284495.t006

This leaves us with (at least) two intriguing questions: To what extent can a *single* TM shift opinions and voting preferences, if at all? And how much more impactful might a single *negative* TM be than a single *positive* TM? We address these questions in Experiment 4.

5. Experiment 4: Impact of a single low-contrast, verified targeted message on opinions and voting preferences

5.1 Methods

5.1.1 Participants. After cleaning, our participant sample for this experiment consisted of a new group of 529 eligible US voters recruited through the MTurk subject pool and screened by Cloud Research. The cleaning procedure was identical to that of Experiment 1. Once again, the group was demographically diverse. See S1 Table for details about demographic characteristics.

5.1.2 Procedure. The procedure in Experiment 4 was identical to that of Experiment 1, except that only one TM appeared in the Twitter feed. It appeared in position 2 for each bias group, and the blue checkmark was *present* in the TM.

Given the obvious preoccupation that participants had with negative TM content in Experiments 1 through 3, in Experiment 4 we looked at how positive and negative TMs impacted participants in the bias groups. Because people saw only one TM in this experiment, it was well suited for comparing the impact of positive and negative TMs. In the Pro-Morrison group, the TM could either be a pro-Morrison tweet (content: "Scott Morrison awarded an honorary doctorate from the University of Melbourne, in recognition for his humanitarian efforts during the Australian wildfires") or an anti-Shorten tweet (content: "Bill Shorten charged with driving under the influence while vacationing in Adelaide"). In the Pro-Shorten group, the TM could either be a pro-Shorten tweet (content: "Bill Shorten has been nominated for The Innovation in Politics Award, which recognizes creative politicians who have the courage to break new ground to find innovative solutions for today's challenges") or an anti-Morrison tweet (content: "Scott Morrison caught spending taxpayer money on lush vacation in Mexico"); again, one or the other appeared at random.

5.2 Results

At first glance, the pattern of VMPs we found in Experiment 4 looks surprising (Table 10). In Experiments 1 to 3, the VMPs in the High Compliance groups were always substantially larger than the VMPs in the Low Compliance group. In Experiment 4 we found the opposite pattern,

Table 7. Experiment 3: VMPs by compliance level.

Compliance Level	Total *n*	Bias Groups *n*	Bias Groups Scroll % Mean (*SD*)	VMP (%)	McNemar's Test X^2	*p*
High	446	287	95.3 (11.0)	81.1	114.29	< 0.001
Low	55	44	34.3 (9.6)	40.7	8.06	< 0.01

https://doi.org/10.1371/journal.pone.0284495.t007

Table 8. Experiment 3: Pre- and post-manipulation opinion ratings of candidates.

Compliance Level		Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
		Pre	Post	Diff	Pre	Post	Diff	$z^{\dagger}$
High	**Impression**	6.92 (1.61)	5.61 (2.77)	-1.31	6.92 (1.71)	3.29 (2.11)	-3.63	-10.0***
	Trust	5.93 (1.84)	6.91 (2.19)	0.98	5.97 (1.92)	3.19 (1.91)	-2.78	-12.7***
	Likeability	6.79 (1.73)	7.43 (1.95)	0.64	6.74 (1.78)	3.65 (1.93)	-3.09	-13.3***
Low	**Impression**	7.75 (1.78)	5.48 (3.02)	-2.27	7.32 (1.88)	3.60 (2.53)	-3.72	-2.15***
	Trust	6.75 (1.92)	7.43 (2.22)	0.68	5.93 (2.27)	3.45 (2.28)	-2.48	-4.55***
	Likeability	7.55 (1.76)	7.84 (2.13)	0.29	7.39 (2.18)	4.09 (2.26)	-3.30	-4.82***

†z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non-favored candidate.
***$p < 0.001$.

https://doi.org/10.1371/journal.pone.0284495.t008

most likely because people in the Low Compliance group saw, on average, only 36.2% of the tweets following the TM in position 2, whereas people in the High Compliance group saw, on average, 97.4% of the tweets following that TM. Exposure to a large number of relatively bland tweets following a biased TM apparently dilutes the power of that TM. Finally, once again, very few people claimed that they saw any bias in the Twitter feed we showed them; only 3 out of the 399 people in the bias groups (0.75%) expressed concerns about possible bias in the content, and 76 of those people (19.0%) specifically mentioned the negative (but not the positive) things being said about the candidates in the TMs.

Breaking down the impact of positive TMs versus negative TMs on the VMPs in Experiment 4 confirms the enormous power that negative content has to alter people's thinking (Table 11). The positive TMs had virtually no impact on VMPs in either the Low Compliance or High Compliance groups. The negative TMs, on the other hand, had a relatively large impact on High Compliance participants (VMP = 51.2%) and shifted *all* of the 17 Low Compliance participants (VMP = 100.0%). Only one of the 399 people in the bias groups expressed any concerns about possible bias in the tweets (0.003%), whereas 76 of these individuals (19.0%) specifically singled out the negative content of the single TM (regarding the candidate's DUI conviction) as a reason for not supporting him. No participants mentioned the contents of the positive version of the TM (regarding the candidate receiving The Innovation in Politics Award) in their typed comments. The possibility of bias was mentioned somewhat more frequently in comments in Experiments 1 through 3, presumably because people in the bias groups in those experiments saw as many as five TMs that shared the same bias; in Experiment 4, people saw only one TM.

Most opinion shifts in the bias groups in Experiment 4 occurred in the predicted direction (Table 12), but they were smaller than the shifts found in the earlier experiments, presumably because participants had less information on which to base their opinions. Changes in voting

Table 9. Experiment 3: Pre- and post-manipulation mean voting preferences on 11-point scale (corrected so that positive values indicate preference for the favored candidate).

Compliance Level	Bias Groups *n*	Pre Mean Voting Preferences on 11-Point Scale (SD)	Post Mean Voting Preferences on 11-Point Scale (SD)	Mean Difference	*z*	*p*
High	287	0.03 (2.64)	2.95 (2.06)	2.92	-12.63	< 0.001
Low	44	0.95 (3.00)	2.77 (2.51)	1.82	-3.86	< 0.001

https://doi.org/10.1371/journal.pone.0284495.t009

Table 10. Experiment 4: VMPs by compliance level.

Compliance Level	Total n	Bias Groups n	Bias Groups Scroll % Mean (SD)	VMP (%)	McNemar's Test X^2	p
High	445	356	97.4 (8.5)	32.4	31.54	< 0.001
Low	50	43	36.2 (11.2)	40.0	6.40	< 0.05

https://doi.org/10.1371/journal.pone.0284495.t010

preferences as expressed on the 11-point scale also occurred in the predicted direction, but, again, they were smaller than in the previous experiments (Tables 13 and 14).

Experiment 4 suggests that a single biased TM in a Twitter feed can impact people's decision making, at least as it pertains to political candidates running for office.

6. Discussion

Recent news about the Twitter company is relevant to our research findings. According to an August 23rd, 2022, investigative story in the *Washington Post* [74], "an explosive whistleblower complaint" from Peter Zatko, former head of security at Twitter–an 84-page document filed simultaneously with the Securities and Exchange Commission, the Federal Trade Commission, and the Department of Justice [75]–Twitter had lax security that allowed false content to be posted easily by hackers, bots, foreign powers, and company employees. Regarding employees, the *Post* reported that "about half of Twitter's roughly 7,000 full-time employees had wide access to the company's internal software and that access was not closely monitored, giving them the ability to tap into sensitive data and alter how the service worked." According to Zatko, Twitter algorithms also determined what content gets suppressed or "amplified" [76].

Given Elon Musk's purchase of the company in October, 2022 [77] and his subsequent firing of most of Twitter's employees, Zatko's concerns about the security of the company's operations might understate the nature of the problems that might be emerging in a new and relatively unstable version of the company. Given the apparent power that tweets–especially tweets containing negative content–can have on opinions and voting preferences–we believe that Twitter's operations should be examined closely not only by Twitter's corporate leaders, but also by government officials and public policy makers in countries worldwide. Twitter currently has 480 million daily users, and it serves as an official platform for world leaders, government agencies, news services, and thousands of companies and organizations; even Pope Francis has a Twitter account. All those Twitter feeds are vulnerable to hacking and hijacking, according to Zatko's complaint, which contains examples of such interference.

Our experiments suggest that TME is a remarkably large effect, especially when Twitter itself sends people sensational tweets that have certain visual properties (Experiment 1): tweets with white backgrounds (matching the backgrounds of organic tweets), a brief headline (such as "Breaking News"), and Twitter's trademark blue checkmark. Experiment 1 yielded a VMP of 87%, with only 2.1% of the participants in the two bias groups expressing any concerns

Table 11. Experiment 4: VMPs by type of TM (positive or negative).

Type of TM	Compliance Level	Bias Groups n	VMP (%)	McNemar's Test X^2	p
Negative	**High**	178	51.2	31.72	< 0.001
	Low	17	100.0	6.00	< 0.05
Positive	**High**	178	15.6	4.08	< 0.05
	Low	26	14.3	1.00	0.32 NS

https://doi.org/10.1371/journal.pone.0284495.t011

Table 12. Experiment 4: Pre- and post-manipulation opinion ratings of candidates.

			Favored Candidate Mean (SD)			Non-Favored Candidate Mean (SD)			
	Compliance		Pre	Post	Diff	Pre	Post	Diff	$z^{\dagger}$
Negative TM									
	High	Impression	7.03 (1.92)	7.13 (1.92)	-0.10	7.11 (1.93)	5.39 (2.20)	1.72	-7.91***
		Trust	6.05 (1.80)	6.51 (2.01)	-0.46	6.19 (1.88)	5.03 (2.25)	1.16	-7.38***
		Likeability	6.72 (1.89)	6.94 (1.86)	-0.22	7.04 (1.86)	5.53 (2.13)	1.51	-7.99***
	Low	Impression	6.53 (1.62)	6.65 (1.93)	0.12	7.06 (1.85)	4.65 (1.90)	-2.41	-3.08**
		Trust	6.18 (2.27)	6.29 (1.40)	0.11	6.18 (2.19)	4.59 (2.29)	-1.59	-2.34*
		Likeability	6.88 (1.80)	6.47 (1.97)	-0.41	7.12 (1.62)	5.18 (2.53)	-1.94	-1.79 NS
Positive TM									
	High	Impression	7.31 (1.84)	7.26 (1.85)	0.05	7.17 (1.92)	7.05 (1.86)	0.12	-0.77 NS
		Trust	6.30 (1.95)	6.56 (1.97)	-0.26	6.22 (2.07)	6.39 (2.02)	-0.17	-0.85 NS
		Likeability	7.19 (1.87)	7.18 (1.93)	0.01	7.01 (1.91)	6.98 (1.88)	0.03	-0.59 NS
	Low	Impression	7.00 (2.21)	7.19 (2.32)	0.19	7.12 (1.93)	7.38 (1.92)	0.26	-0.29 NS
		Trust	6.04 (2.39)	6.23 (2.76)	0.19	6.12 (2.16)	6.50 (2.25)	0.38	-0.18 NS
		Likeability	6.85 (2.43)	6.81 (2.28)	-0.04	7.27 (1.93)	7.15 (1.87)	-0.12	-0.21 NS

† z values represent Wilcoxon signed ranks test comparing post-minus-pre ratings for the favored candidate to the post-minus-pre ratings for the non- favored candidate.

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$.

https://doi.org/10.1371/journal.pone.0284495.t012

about possible bias in the Twitter feed we showed them. That VMP shift means that in a group of 1,000 undecided voters–split, by definition, 500/500 before exposure to a biased Twitter feed–after viewing that feed, the split will now be 65/935, which means that interacting with the Twitter feed changed a win margin of 0% to a win margin of 87% among vulnerable voters. That shift could occur, in theory, with nearly 98% of the people in such a group having no idea they were manipulated.

On its face, a shift that big might seem impossible. In the real world, certainly, people are being influenced by many sources of information, not just by Twitter, and we currently have no reason to believe that Twitter's content is systematically biased to support just one candidate or political party. But our experiments show the *potential* that Twitter feeds have to shift opinions and votes. Twitter is a private company that is not accountable to the public, and no laws or regulations exist at this writing that would in any way restrict Twitter's ability to send highly biased content to users. Indeed, some people have claimed that Twitter's content already shows significant political bias at times [cf. 76]. Trump supporters cried foul, for example, when Twitter permanently shut down the President's Twitter account just after the January 6, 2021 insurrection in Washington, D.C. [78], and objections were raised when Twitter

Table 13. Experiment 4: Negative TM Pre- and post-manipulation mean voting preferences on 11-point scale (corrected so that positive values indicate preference for the favored candidate).

Compliance Level	*n*	Pre Voting Preference on 11-Point Scale (SD)	Post Voting Preference on 11-Point Scale (SD)	Mean Difference	*z*	*p*
High	178	-0.11 (2.85)	1.31 (2.88)	1.42	-6.42	< 0.001
Low	17	-0.47 (2.76)	1.41 (2.50)	0.64	-2.60	< 0.01

https://doi.org/10.1371/journal.pone.0284495.t013

Table 14. Experiment 4: Positive TM pre- and post-manipulation mean voting preferences on 11-point scale (corrected so that positive values indicate preference for the favored candidate).

Compliance Level	*n*	Pre Voting Preference on 11-Point Scale (SD)	Post Voting Preference on 11-Point Scale (SD)	Mean Difference	*z*	*p*
High	178	0.21 (2.77)	0.74 (2.78)	0.53	-2.58	= 0.01
Low	26	0.58 (2.86)	0.65 (3.03)	0.07	-0.20	= 0.84

https://doi.org/10.1371/journal.pone.0284495.t014

apparently suppressed news stories related to content found on Hunter Biden's laptop computer in October 2020 [79]. In this case, some of the facts about the laptop originally reported by the *New York Post* on October 14, 2020 were subsequently confirmed by both the *New York Times* and the *Washington Post* [80, cf. 81]. That Twitter content might show political bias should surprise no one given that, according to OpenSecrets.org, more than 96% of donations from Twitter and its employees in recent years have gone to one political party [43].

No experiments can show that a source of influence like TME is actually being used. Since 2016, however, our team has been building increasingly larger and more sophisticated systems that capture the ephemeral content being shown to users by Google, YouTube, Bing, Yahoo, and other companies [51,52,82]. In 2020, we preserved and analyzed more than 1.5 million online ephemeral experiences that would normally have been lost [52]. In 2022, we preserved more than 2.4 million online ephemeral experiences related to the US midterm elections, including, this time around, content from Twitter which we are currently analyzing.

We acknowledge that if, at some point, we detect political or other bias in Twitter feeds being displayed to certain groups, that will still tell us nothing about the origin of such bias. Bias in ephemeral content can be programmed deliberately [83], generated by unconscious bias on the part of programmers [84], or generated by user behavior [85]. No matter what the original of such bias, given the apparent power it has to shift opinions and voting preferences, we believe that if large-scale bias is ultimately found to exist in actual Twitter feeds, this is an issue that Twitter executives and government officials will need to examine. Otherwise, extreme bias–especially bias targeted toward certain groups–could easily undermine the integrity of the free-and-fair election. Moreover, if monitoring systems are not in place to preserve ephemeral content such as Twitter feeds, democracy might be undermined without the electorate knowing. Based on Mr. Zatko's recent revelations [74,75], along with documented cases in which Twitter content has been hacked by bad actors [86,87], it now appears that extreme bias in Twitter content can be introduced fairly easily by agents of foreign powers, by aggressive Twitter employees, or even by mischievous teenagers [74,75].

6.1 Limitations and future research

We have restricted ourselves in this report to TME as it might impact users of Twitter, but targeted messages can also be sent to users of Google and other search engines (S3 and S5 Figs), to users of Instagram and Facebook (S1 and S2 Figs), and even to users of personal assistants such as Siri and Alexa [4]. On platforms such as Google, the home page of which is viewed more than 500 million times a day in the US, we are especially concerned about targeted messages that remind people to vote or to register to vote in an election (S3 and S5 Figs). If such reminders were sent mainly or exclusively to members of one political party, they could presumably have a substantial partisan effect on voter turnout. We currently have research underway to help us understand and quantify the impact that TME might have on Google and other online platforms.

We are also concerned about the possibility that a number of major US tech companies all appear at the moment to share a similar political bias [88], and we are currently studying the impact of exposing people to similar or dissimilar bias experienced on more than one platform–research on what we call the Multiple Platform Effect (MPE). We have also expanded our research program to look at how new sources of influence made possible by the internet are affecting children.

Our findings in the present study should not be overinterpreted. We have shown, with a sample of 2,133 eligible US voters that biased, targeted tweets can shift opinions and voting preferences in predictable ways with only a small percentage of people showing any awareness that they have been manipulated. The effect proved to be especially large when the content of such tweets was derogatory (content that linguists might call "low-valence and high-arousal" [89]. But our participants were not real voters in the middle of real elections. Rather, they were US research subjects who had indicated that they were unfamiliar with two candidates who ran for Prime Minister of Australia in 2019. SEME has been shown to impact real voters in a real election [1], but TME has not yet been tested that way.

In a real election people are being subjected to dozens, if not hundreds of different sources of influence that might affect their voting decisions. Other sources of impact could presumably override the impact of biased tweets, and yet there are still, we believe, three reasons why we should be concerned about TME in general and corporation-generated biased tweets in particular. As we noted in our introduction, the bias in TMs is almost always invisible to people, which leads people, mistakenly, to believe that they have made up their own minds. Second, TMs are ephemeral, so unless permanent monitoring systems are in place, we will never know for sure how or even whether TMs are being used to affect people's opinions and decisions. And third, TMs generated by large online monopolies are inherently noncompetitive; when Twitter, Facebook, or Google deploys biased TMs favoring one candidate, the opposing candidate has no way to counteract them. In other words, online TMs are a uniquely powerful new form of influence.

Supporting information

S1 Fig. Facebook vote reminder, screenshotted in Georgia, January 5, 2021.
(DOCX)

S2 Fig. Instagram vote reminder, screenshotted in Georgia January 5, 2021.
(DOCX)

S3 Fig. Google home page with vote reminder, 2020 Presidential election, screenshotted October 27, 2020.
(DOCX)

S4 Fig. Google home page with no vote reminder, 2020 Presidential election, screenshotted October 27, 2020.
(DOCX)

S5 Fig. Google home page with vote reminfer, 2020 Presidential election, screenshotted on Election Day, November 3, 2020.
(DOCX)

S6 Fig. Twitter home page with vote reminder, 2022 Midterm election, screenshotted November 7, 2022.
(DOCX)

S7 Fig. Twitter home page with a "You might like" promoted tweet containing a vote reminder, screenshotted November 8, 2022.
(DOCX)

S8 Fig. Example of a control tweet presented to all participants in Experiments 1–4.
(DOCX)

S9 Fig. Example of a strongly negative targeted message about Morrison with a blue checkmark, presented to participants in the Pro-Shorten group in Experiment 1.
(DOCX)

S10 Fig. Example of a strongly negative targeted message about Morrison without a blue checkmark, presented to participants in the Pro-Shorten group in Experiment 2.
(DOCX)

S1 Table. Demographics characteristics across Experiments 1 to 4.
(DOCX)

S2 Table. Experiment 1: Demographic analysis by educational attainment.
(DOCX)

S3 Table. Experiment 1: Demographic analysis by gender.
(DOCX)

S4 Table. Experiment 1: Demographic analysis by age.
(DOCX)

S5 Table. Experiment 1: Demographic analysis by race/ethnicity.
(DOCX)

S6 Table. Experiment 2: Demographic analysis by educational attainment.
(DOCX)

S7 Table. Experiment 2: Demographic analysis by gender.
(DOCX)

S8 Table. Experiment 2: Demographic analysis by age.
(DOCX)

S9 Table. Experiment 2: Demographic analysis by race/ethnicity.
(DOCX)

S10 Table. Experiment 3: Demographic analysis by educational attainment.
(DOCX)

S11 Table. Experiment 3: Demographic analysis by gender.
(DOCX)

S12 Table. Experiment 3: Demographic analysis by age.
(DOCX)

S13 Table. Experiment 3: Demographic analysis by race/ethnicity.
(DOCX)

S14 Table. Experiment 4: Demographic analysis by educational attainment.
(DOCX)

S15 Table. Experiment 4: Demographic analysis by gender.
(DOCX)

S16 Table. Experiment 4: Demographic analysis by age.
(DOCX)

S17 Table. Experiment 4: Demographic analysis by race/ethnicity.
(DOCX)

S18 Table. Experiment 1: Pre-and post-manipulation opinions by group.
(DOCX)

S19 Table. Experiment 2: Pre-and post-manipulation opinions by group.
(DOCX)

S20 Table. Experiment 3: Pre-and post-manipulation opinions by group.
(DOCX)

S21 Table. Experiment 4: Pre-and post-manipulation opinions by group.
(DOCX)

S22 Table. Experiments 1–3: Pre- and post-manipulation votes on 11-point scale (-5 to +5).
(DOCX)

S23 Table. Experiment 4: Pre- and post-manipulation vote on 11-point scale (-5 to +5).
(DOCX)

S1 Text. Experiment 1: Candidate biographies.
(DOCX)

S2 Text. Experiments 1 to 4: Instructions immediately preceding Twitter simulation.
(DOCX)

S3 Text. Experiments 1 to 4: Textual content and positions of the five targeted messages.
(DOCX)

S4 Text. Vote Manipulation Power (VMP) calculation.
(DOCX)

Acknowledgments

We thank M. Voillot for help in preparing and running the experiments and V. Zankich and C. Sutlick for help in preparing the manuscript.

Author Contributions

Conceptualization: Robert Epstein.

Formal analysis: Robert Epstein, Christina Tyagi, Hongyu Wang.

Project administration: Robert Epstein.

Supervision: Robert Epstein.

Writing – original draft: Robert Epstein.

Writing – review & editing: Robert Epstein, Christina Tyagi.

References

1. Epstein R, Robertson RE. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proceedings of the National Academy of Sciences USA. 2015 Aug 4. https://doi.org/10.1073/pnas.1419828112 PMID: 26243876
2. Epstein R. Manipulating minds: The power of search engines to influence votes and opinions. In: Moore M, Tambini D, editors. Digital dominance: The power of Google, Amazon, Facebook, and Apple. Oxford: Oxford University Press; 2018. pp. 294–319.
3. Epstein R. Why Google poses a serious threat to democracy, and how to end that threat. Congressional Record of the United States. 2019 July 16 [Cited 2022 Mar 11]. Available from: https://www.judiciary.senate.gov/imo/media/doc/Epstein%20Testimony.pdf.
4. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating Google output. J Med Internet Res, 2014; 16(4):e100. https://doi.org/10.2196/jmir.2642 PMID: 24694866.
5. Haas A, Unkel J. Ranking versus reputation: perception and effects of search result credibility. Behav Inf Technol, 2017; 36(12):1285–1298. https://doi.org/10.1080/0144929X.2017.1381166.
6. Ludolph R, Allam A, Schulz PJ. Manipulating Google's knowledge box to counter biased information processing during an online search on vaccination: Application of a technological debiasing strategy. J Med Internet Res, 2016; 18(6):e137. https://doi.org/10.2196/jmir.5430 PMID: 27255736.
7. Eslami M, Vaccaro K, Karahalios K, Hamilton K. "Be Careful; Things Can Be Worse than They Appear": Understanding Biased Algorithms and Users' Behavior Around Them in Rating Platforms. Proceedings of the 11th International AAAI Conference on Web and Social Media [Internet]. 2017 May 3; 11(1):62–71. Available from: https://ojs.aaai.org/index.php/ICWSM/article/view/14898.
8. Pogacar FA, Ghenai A, Smucker MD, Clarke CLA. The positive and negative influence of search results on people's decisions about the efficacy of medical treatments. Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval [Internet]. 2017 Oct 1– 4;:209–216. Available from: https://dl.acm.org/doi/10.1145/3121050.3121074.
9. Trielli D, Diakopoulos N. Search as news curator: The role of Google in shaping attention to news information. CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems [Internet]. 2019 May 4–9; 453:1–15. Available from: https://dl.acm.org/doi/10.1145/3290605.3300683.
10. Epstein R, Mohr R Jr, Martinez J. The Search Suggestion Effect (SSE): How search suggestions can be used to shift opinions and voting preferences dramatically. Paper presented at: Western Psychological Association; 2018 Apr; Portland, OR. Available from: https://aibrt.org/downloads/EPSTEIN_MOHR_&_MARTINEZ_2018-WPA-The_Search_Suggestion_Effect-SSE-WP-17-03.pdf.
11. Epstein R, Lee V, Mohr R, Zankich VR. The answer bot effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. PLOS ONE. 2022; 17(6). https://doi.org/10.1371/journal.pone.0268081.
12. Epstein R. The Search Suggestion Effect (SSE): How autocomplete can be used to impact votes and opinions. Paper presented at the 2nd biennial meeting of the International Convention of Psychological Science; 2017 Mar; Vienna, Austria. Available from: https://aibrt.org/downloads/EPSTEIN_2017-The_Search_Suggestion_Effect-SSE-ICPS_Vienna-March_2017.pdf.
13. Epstein R, Huang Y. The Opinion Matching Effect (OME): A subtle but powerful new form of influence that is being widely used on the internet without user awareness. Poster presented at the 102nd annual meeting of the Western Psychological Association; 2022 Apr; Portland, OR. Available from: https://aibrt.org/downloads/EPSTEIN_&_HUANG_2022-WPA-Opinion_Matching_Effect-OME.pdf.
14. Epstein R, Voillot M. The YouTube Manipulation Effect (YME): The power that bias in YouTube's up-next algorithm has to shift votes and opinions, and preliminary evidence that such bias exists. Paper presented at the 102nd annual meeting of the Western Psychological Association; 2022 Apr; Portland, OR. Available from: https://aibrt.org/downloads/EPSTEIN_&_VOILLOT_2022-WPA-YouTube_Manipulation_Effect-YME.pdf.
15. Berelson B, Lazarsfeld P, MacPhee WN. Voting: A study of opinion formation in a presidential campaign. Chicago, Illinois: University of Chicago Press; 1986.
16. Lazarsfeld P, Berelson B, Gaudet H. The People's Choice: How the voter makes up his mind in a presidential campaign. Ann Arbor, Michigan: UMI; 1993.
17. Smith RB. A legacy of Lazarsfeld: Cumulative Social Research on voting. International Journal of Public Opinion Research. 2001; 13(3):280–98 https://doi.org/10.1093/ijpor/13.3.280.
18. Gerber AS, Karlan D, Bergan D. Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. American Economic Journal: Applied Economics. 2009; 1(2):35–52. https://doi.org/10.1257/app.1.2.35

19. Ceron A, d'Adda G. E-campaigning on Twitter: The effectiveness of distributive promises and negative campaign in the 2013 Italian election. New Media & Society. 2016; 18(9):1935–55. https://doi.org/10.1177/1461444815571915.

20. McInerney S. Political Advertising's role in elections: How messaging swings voters [Internet]. The Badger Herald. 2022 [Cited 2023 Mar 8]. Available from: https://badgerherald.com/news/2022/11/05/political-advertisings-role-in-elections-how-it-swings-voters/.

21. Lee E-J. Effectiveness of politicians' soft campaign on Twitter versus TV: Cognitive and experiential routes. Journal of Communication. 2013; 63(5):953–74. https://doi.org/10.1111/jcom.12049.

22. Coppock A, Guess A, Ternovski J. When treatments are tweets: A network mobilization experiment over Twitter. Political Behavior. 2015; 38(1):105–28. https://doi.org/10.1007/s11109-015-9308-6.

23. Kovic M, Rauchfleisch A, Metag J, Caspar C, Szenogrady J. Brute force effects of mass media presence and social media activity on electoral outcome. Journal of Information Technology & Politics. 2017; 14(4):348–71. https://doi.org/10.1080/19331681.2017.1374228.

24. Kim H. The mere exposure effect of tweets on vote choice. Journal of Information Technology & Politics. 2021; 18(4):455–65. https://doi.org/10.1080/19331681.2021.1900979.

25. Nevo A. Empirical models of consumer behavior. Annual Review of Economics. 2011; 3(1):51–75. https://doi.org/10.1146/annurev-economics-061109-080402.

26. Malter MS, Holbrook MB, Kahn BE, Parker JR, Lehmann DR. The past, present, and future of Consumer Research. Marketing Letters. 2020; 31(2–3):137–49. https://doi.org/10.1007/s11002-020-09526-8. PMID: 32836799

27. Chaudhary K, Alam M, Al-Rakhami MS, Gumaei A. Machine learning-based mathematical modelling for prediction of social media consumer behavior using Big Data Analytics. Journal of Big Data. 2021; 8(1). https://doi.org/10.1186/s40537-021-00466-2.

28. Guo L, Zhang B, Zhao X. A consumer behavior prediction model based on multivariate real-time sequence analysis. Mathematical Problems in Engineering. 2021; 2021:1–5. https://doi.org/10.1155/2021/6688750.

29. Cialdini RB, Goldstein NJ. Social Influence: Compliance and conformity. Annual Review of Psychology. 2004; 55(1):591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015 PMID: 14744228

30. Smith NC, Goldstein DG, Johnson EJ. Choice without awareness: Ethical and policy implications of defaults. Journal of Public Policy & Marketing. 2013; 32(2):159–72. https://doi.org/10.1509/jppm.10.114.

31. T Martin SJ, Goldstein NJ, Cialdini RB. The small big: Small changes that Spark Big Influence. London, United Kingdom: Profile Books; 2015.

32. Lehoucq F. Electoral Fraud: Causes, types, and consequences. Annual Review of Political Science. 2003; 6(1):233–56. https://doi.org/10.1146/annurev.polisci.6.121901.085655.

33. Menkhaus K. Elections in the hardest places: The case of Somalia. Journal of Democracy. 2017; 28(4):132–46. https://doi.org/10.1353/jod.2017.0073

34. Berners-Lee T, Fischetti M. Weaving the web: The original design and Ultimative destiny of the World Wide Web by its inventor. Manhattan, New York: Harper Business; 2011.

35. Hern A. Tim Berners-Lee on 30 years of the World Wide Web: 'We Can Get the Web we want' [Internet]. The Guardian. Guardian News and Media; 2019 [Cited 2023 Mar 8]. Available from: https://www.theguardian.com/technology/2019/mar/12/tim-berners-lee-on-30-years-of-the-web-if-we-dream-a-little-we-can-get-the-web-we-want.

36. Paterson N. Walled Gardens: the New Shape of the Public Internet. Proceedings of the 2012 iConference. 2012Feb7;:97–104. https://doi.org/10.1145/2132176.2132189.

37. Barnett E. Tim Berners-Lee criticises Facebook's 'Walled Garden' [Internet]. The Telegraph. Telegraph Media Group; 2010 [Cited 2023 Mar 8]. Available from: https://www.telegraph.co.uk/technology/facebook/8151101/tTim-Berners-Lee-criticises-Facebooks-walled-garden.html.

38. Bhandari A, Bimo S. Why's everyone on TikTok now? the algorithmized self and the future of self-making on social media. Social Media + Society. 2022; 8(1):205630512210862. https://doi.org/10.1177/20563051221086241.

39. Montag C, Yang H, Elhai JD. On the psychology of TikTok use: A first glimpse from empirical findings. Frontiers in Public Health. 2021; 9. https://doi.org/10.3389/fpubh.2021.641673. PMID: 33816425

40. Lorenz T. The political pundits of TikTok [Internet]. The New York Times. The New York Times; 2020 [Cited 2023 Mar 8]. Available from: https://www.nytimes.com/2020/02/27/style/tiktok-politics-bernie-trump.html.

41. Java A, Song X, Finin T, Tseng B. Why we twitter: An analysis of a microblogging community. Paper presented at the 1st International Workshop on Social Networks Analysis, SNA-KDD; 2007 Aug; San Jose, CA. Available from: https://doi.org/10.1007/978-3-642-00528-2_7.

42. Lozic J, Vojkovic G, Lozic I. Analysis of the financial results of the Twitter platform at the beginning of Covid 19. Paper presented at the 44th International Convention on Information, Communication and Electronic Technology (MIPRO); 2021 Sep-Oct; Opatija, Croatia. Available from: https://doi.org/10.23919/MIPRO52101.2021.9597016
43. OpenSecrets.org [Internet]. Twitter Totals. [Cited 2023 Mar 8]. Available from: https://www.opensecrets.org/orgs/twitter/totals?id=D000067113.
44. OpenSecrets.org [Internet]. Meta Totals. [Cited 2023 Mar 8]. Available from: https://www.opensecrets.org/orgs/meta/totals?id=D000033563.
45. OpenSecrets.org [Internet]. Alphabet Inc Totals. [Cited 2023 Mar 8]. Available from: https://www.opensecrets.org/orgs/alphabet-inc/totals?id=D000067823.
46. Zittrain J. Facebook could decide an election without anyone ever finding out: The scary future of digital gerrymandering—and how to prevent it. The New Republic. 2014 June 1 [Cited 2022 Mar 11]. Available from: https://newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering.
47. Bond RM, Fariss CJ, Jones JJ, Kramer ADI, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. Nature. 2012 Sept 12. https://doi.org/10.1038/nature11421 PMID: 22972300
48. Bump P. Donald Trump will be president thanks to 80,000 people in three states. The Washington Post. 2016 Dec 1 [Cited 2022 Mar 11]. Available from: https://www.washingtonpost.com/news/the-fix/wp/2016/12/01/donald-trump-will-be-president-thanks-to-80000-people-in-three-states/.
49. Epstein R. How Google shifts votes: A "go vote" reminder is not always what you think it is. The Epoch Times. 2019 Jan 2 [Cited 2022 Mar 11]. Available from: https://www.theepochtimes.com/another-way-google-manipulates-votes-without-us-knowing-a-go-vote-reminder-is-not-what-you-think-it-is_2754073.html.
50. Butera J. Partisan Gerrymandering and the Qualifications Clause. Boston University Law Review. 2015; 95(1): 303–334.
51. Epstein R. Taming Big Tech: The case for monitoring. Hacker Noon. 2018 May 13 [Cited 2022 Mar 11]. Available from: https://hackernoon.com/taming-big-tech-5fef0df0f00d.
52. Bock S, Epstein R, Peirson L, Wang H. Large-scale monitoring of Big Tech political manipulations in the 2020 Presidential election and 2021 Senate runoffs, and why monitoring is essential for democracy. Paper presented at; 2021 June 14; the 24th annual meeting of the American Association of Behavioral and Social Sciences (AABSS). Available from: https://aibrt.org/downloads/EPSTEIN_et_al_2021-Large-Scale_Monitoring_of_Big_Tech_Political_Manipulations-FINAL_w_AUDIO.mp4.
53. Aral S, Eckles D. Protecting elections from social media manipulation. Science. 2019 Aug 30. https://doi.org/10.1126/science.aaw8243 PMID: 31467206
54. Farkas J. Disguised Propaganda on Social Media: Addressing Democratic Dangers and Solutions. Brown Journal of World Affairs. 2018; 25(1): 1–16.
55. Ruck DJ, Rice NM, Borycz J, Bentley RA. Internet Research Agency Twitter activity predicted 2016 U.S. election polls. First Monday. 2019; 24(7): 1.
56. Hinds J, Williams EJ, Joinson AN. "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. International Journal of Human-Computer Studies [Internet]. 2020 Nov [Cited 2022 Apr 21];143:N.PAG. Available from: https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=145318574&site=ehost-live.
57. Rheault L, Musulan A. Efficient detection of online communities and social bot activity during electoral campaigns. Journal of Information Technology & Politics. 2021 Feb 2. https://doi.org/10.1080/19331681.2021.1879705
58. Berners-Lee T. The web can be weaponised–and we can't count on big tech to stop it. The Guardian. 2018 Mar 12 [Cited 2022 Mar 11]. Available from: https://www.theguardian.com/commentisfree/2018/mar/12/tim-berners-lee-web-weapon-regulation-open-letter.
59. Hardy Q. The web's creator looks to reinvent it. The New York Times. 2016 June 8 [Cited 2022 Mar 11]. Available from: https://www.nytimes.com/2016/06/08/technology/the-webs-creator-looks-to-reinvent-it.html.
60. Steer G. 'Go Vote.' Google's Redesigned Homepage Is Encouraging People to Head to the Polls. TIME USA. 2018 Nov 6 [Cited 2022 March 11]. Available from: https://time.com/5445991/google-go-vote/.
61. Altman M, McDonald M. The promise and perils of computers in redistricting. Duke Journal of Constitutional Law and Public Policy. 2010; 5(1). Available from: https://scholarship.law.duke.edu/djclpp/vol5/iss1/5.
62. Berghel H. Chasing Elbridge's ghost: The digital gerrymander. Computer. 2016 Nov. https://doi.org/10.1109/MC.2016.330

63. Carr DF. Wanted: Honest algorithms for voter redistricting. InformationWeek. 2014 Aug 7 [Cited 2022 Mar 11]. Available from: https://www.informationweek.com/open-government/wanted-honest-algorithms-for-voter-redistricting.

64. McLaughlin D. Amazon Ads Deceive Consumers, Unions Say in Complaint to U.S. Regulator. Bloomberg.com [Internet]. 2021 Dec 9 [cited 2022 Mar 18];N.PAG.

65. Trott V, Li N, Fordyce R, Andrejevic M. Shedding light on "dark" ads. Continuum: Journal of Media & Cultural Studies [Internet]. 2021 Oct [cited 2022 Mar 18]; 35(5):761–74. https://doi.org/10.1080/10304312.2021.1983258

66. Brennan E, Dunstone K, Vittiglia A, Mancuso S, Durkin S, Slater MD, et al. Testing the effectiveness of Alcohol Health Warning label formats: An online experimental study with Australian adult drinkers. PLOS ONE. 2022; 17(12). https://doi.org/10.1371/journal.pone.0276189. PMID: 36476743

67. Shadel WG, Martino SC, Setodji CM, Dunbar M, Scharf D, Creswell KG. Do graphic health warning labels on cigarette packages deter purchases at point-of-sale? an experiment with adult smokers. Health Education Research. 2019; 34(3):321–31. https://doi.org/10.1093/her/cyz011. PMID: 30932154

68. Puglisi R, Snyder JM. Empirical studies of media bias. Handbook of Media Economics. 2015; 1(1):647–67. https://doi.org/10.1016/B978-0-444-63685-0.00015-2.

69. DellaVigna S, Kaplan E. The Fox News Effect: Media Bias and voting. 2006; 122(3):1187–1234. https://doi.org/10.1162/qjec.122.3.1187.

70. Barberá P, Rivero G. Understanding the political representativeness of Twitter users. Social Science Computer Review. 2014; 33(6):712–29. https://doi.org/10.1177/0894439314558836

71. Hunt NC, Scheetz AM. Using MTurk to Distribute a Survey or Experiment: Methodological Considerations. Journal of Information Systems [Internet]. 2019 Spring [cited 2022 Apr 21];33(1):43–65. Available from: https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=137598052&site=ehost-live.

72. Loftus EF. Leading Questions and the Eyewitness Report. Cognitive Psychology [Internet]. 1975 Jan 1 [cited 2022 Apr 21]; 7(4):560. Available from: https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ133673&site=ehost-live.

73. Kerzel D, Schönhammer J. Salient stimuli capture attention and action. Attention, Perception & Psychophysics [Internet]. 2013 Nov [cited 2022 Apr 21];75(8):1633–43. Available from: https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=91944280&site=ehost-live. PMID: 23918550

74. Dwoskin E, Menn J, Zakrzewski C. Former security chief claims Twitter buried 'egregious deficiencies.' The Washington Post. 2022 August 23 [Cited 2022 August 30]. Available from: https://www.washingtonpost.com/technology/interactive/2022/twitter-whistleblower-sec-spam/?itid=hp-top-table-main.

75. Zatko P. Re: Protected Disclosures of Federal Trade Commission Act Violations, Material Misrepresentations and Omissions, and Fraud by Twitter, Inc. (NASDAQ: TWTR) and CEO Parag Agrawal, SEC TCR#. Documentcloud.org [Internet]. 2022 July 6 [Cited 2022 August 31]. Available from: https://www.documentcloud.org/documents/22186683-twitter-whistleblower-disclosure.

76. Belli L, Hardt M, Huszar F, Ktena SI, O'Brien CO, Schlaikjer A. Algorithmic amplification of politics on Twitter. Proceedings of the National Academy of Sciences. 2022 January. https://doi.org/10.1073/pnas.2025334119 PMID: 34934011

77. Conger K, Hirsch L. Elon Musk completes $44 billion deal to own Twitter [Internet]. The New York Times. The New York Times; 2022 [Cited 2023 March 8]. Available from: https://www.nytimes.com/2022/10/27/technology/elon-musk-twitter-deal-complete.html.

78. Rindler D, Shin Y, Tan S. How one of America's ugliest days unraveled inside and outside the Capitol. The Washington Post. 2021 January 9 [Cited 2022 September 1]. Available from: https://www.washingtonpost.com/nation/interactive/2021/capitol-insurrection-visual-timeline/.

79. Bond, S. Facebook And Twitter Limit Sharing 'New York Post' Story About Joe Biden. National Public Radio. 2020 October 14 [Cited 2022 September 1]. Available from: https://www.npr.org/2020/10/14/923766097/facebook-and-twitter-limit-sharing-new-york-post-story-about-joe-biden.

80. Tiffany K, Why Hunter Biden's Laptop Will Never Go Away. The Atlantic. 2022 April 28 [Cited 2022 September 1]. Available from: https://www.theatlantic.com/technology/archive/2022/04/tech-companies-suppressed-biden-laptop/629680/.

81. Benner K, Schmidt MS, Vogel K. Hunter Biden Paid Tax Bill, but Broad Federal Investigation Continues. The New York Times. 2022 March 16 [Cited 2022 September 1]. Available from: https://www.nytimes.com/2022/03/16/us/politics/hunter-biden-tax-bill-investigation.html.

82. Epstein R, Peirson L. How we preserved more than 2.4 million online ephemeral experiences in the 2022 midterm elections, and what this content revealed about online election bias. Paper to be presented at the 103rd annual meeting of the Western Psychological Association; 2023 Apr; Riverside, CA. Available from: https://aibrt.org/downloads/EPSTEIN-Submitted_2022-How_We_Preserved_More_Than_2.4_Million_Online_Ephemeral_Experiences_in_the_Midterm_Elections.pdf.

83. McKinnon JD, McMillan D. Google Workers Discussed Tweaking Search Function to Counter Travel Ban: Company says none of proposed changes to search results were ever implemented. The Wall Street Journal. 2018 September 20 [Cited 2022 September 1]. Available from: https://www.wsj.com/articles/google-workers-discussed-tweaking-search-function-to-counter-travel-ban-1537488472.

84. Obermeyer Z, Mullainathan S, Powers B, Vogeli C. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 October. https://doi.org/10.1126/science.aax2342 PMID: 31649194

85. Chen W, Menczer F, Pacheco D, Yang KC. Neutral bots probe political bias on social media. Nature Communications. 2021 September. https://doi.org/10.1038/s41467-021-25738-6 PMID: 34552073

86. Gonzales R. 2 Former Twitter Employees Charged With Spying For Saudi Arabia. National Public Radio. 2019 November 6. Available from: https://www.npr.org/2019/11/06/777098293/2-former-twitter-employees-charged-with-spying-for-saudi-arabia.

87. Kelly H. Twitter hacked; 250,000 accounts affected. CNN. 2013 February 1. Available from: https://edition.cnn.com/2013/02/01/tech/social-media/twitter-hacked/index.html.

88. Fearnow B. Big Tech employees donate overwhelmingly to 2020 democrats. Newsweek. 4 July 2020. Available from: https://www.newsweek.com/big-tech-employees-donate-overwhelmingly-2020-democrats-1515430.

89. Shuman V, Sander D, Scherer KR. Levels of valence. Front Psychology. 2013. https://doi.org/10.3389/fpsyg.2013.00261 PMID: 23717292