

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

Robert Epstein (re@aibrt.org) & Michelle Voillot
American Institute for Behavioral Research and Technology

Abstract

The internet has made possible a number of powerful new forms of influence, some of which are entirely invisible to users. Most of these techniques are controlled exclusively by four American companies and are now impacting more than 3 billion people worldwide. Currently, no laws or regulations exist to prevent companies from using these techniques. In the present study, we describe a relatively new form of influence we call the YouTube Manipulation Effect (YME): the power that YouTube videos have to change thinking and behavior. We quantify YME in a series of randomized, controlled, counterbalanced, double-blind experiments in which a total of 1,463 politically-diverse, eligible US voters were asked to weigh in on two candidates who ran for Prime Minister of Australia in 2019 (thus assuring that our participants were initially undecided). Participants were first given brief, innocuous descriptions of each candidate and then asked 8 questions about them. They were then given up to 15 minutes to watch videos about the candidates on our YouTube simulator (10 minutes minimum), after which they were asked those opinion questions again. Participants had been randomly assigned to one of 3 groups: pro-Candidate 1, pro-Candidate 2, or alternating (control). All videos and metadata were real, equalized by candidate only for average length. The only difference between the groups was the order of the videos on the right-hand panel, with the “up-next” video at the top of each list. In the bias groups, higher ranking videos favored one candidate. Before the videos, voting preferences were roughly equal; after the videos, preferences shifted by 52% toward the preferred candidate, with only 31% of participants showing awareness of bias in the videos. In a condition that masked the bias, we were able to produce a significantly larger shift in voting preferences (66%) while reducing bias awareness to 16%. Based on data from a large-scale election monitoring project we conducted in 2020, we also present evidence that in the days leading up to the Presidential election, 93 percent of the videos shown to voters across the political spectrum by YouTube's up-next algorithm had a strong liberal bias, thus potentially shifting a large number of votes in one direction. We believe that both public policy makers and the general public should be especially concerned about the power that YME has to shift opinions and voting preferences given information about YouTube that has leaked from Google in recent years. For example: (a) In a video that leaked from Google in 2019, Susan Wojcicki, the CEO of YouTube, tells her staff that YouTube's up-next algorithm was being modified so that it boosted content that Google staff believed was legitimate and demoted content that they believed was not. (b) Also in 2019, Google software engineer Guillaume Chaslot revealed that 70 percent of the videos that people watch on YouTube worldwide are suggested by YouTube's up-next algorithm. In our talk, we conclude that in a national election in the US in which 100 million people voted, bias in YouTube sequences has the power to shift between 5.2 and 6.6 million votes over time without people knowing and without leaving a paper trail for authorities to trace.