

Vandana Anand
Fareya Ikram
Daniel McDonough

Research Paper

Hypothesis

The information that we are exposed through search engines and websites is inherently bias. Taking three different search engines, we will perform searches on different political figures.

We will search the names of two conservative figures and two liberal figures. In the scope of this experiment liberal figures are those that currently belong to the democratic party and conservative figures that are currently members of the republican party. *We hypothesize that more than half of the first ten articles about the liberal figures will be positive articles, praising the figure and that more than half of the articles about the conservative figures will be negative, criticizing their action.* This will prove that search engine results are bias. The political figures that we will use are: Donald Trump (Republican), Marco Rubio (Republican), Hillary Clinton (Democratic), and Barack Obama(Democratic).

Definition: A bias search engine is a search engine that has prejudice in favor of or against one thing, person, or group. The search results of such a search engine will show this prejudice.

Research Methods:

- 1) Collect the first 10 articles/search results for each political figure in three separate search engines: Google, Yahoo, DuckDuckGo
 - a) Note: for this experiment we excluded social media and Wikipedia
- 2) For each article determine if the political figure is being presented in a positive, negative or neutral manner
 - a) Do this by scanning the websites for opinions and key words or phrases that show the point of view of the author
 - i) i.e Author chose to use “lack of success” vs failure
 - ii) <http://www.aim.org/on-target-blog/media-bias-in-strategic-word-choice/>
 - b) Check website against bias checking tool
 - i) <https://mediabiasfactcheck.com/>
- 3) Compare the results within each search engine and compare the results across the search engines and determine if there is a bias
- 4) Note: We will use the terms positive, negative, and neutral throughout our paper. Positive means that they are using words that make them look as if they are in favor of the politician. Negative means that they are using language that makes them appear to be critical, and against the political figure. Neutral means that we could not detect a positive or negative bias.

Data Collection:

Link to excel file:

<https://docs.google.com/spreadsheets/d/1N5pXGvKTBkP0dHHkuUkt1EJirIpRp4dMp3hy0VFhBag/edit?usp=sharing>

Note: This file contains the first 10 articles (search results) for each search engine
It also marks the articles as positive, negative, and neutral.

Hillary Clinton

| | | Google | Yahoo | Duck Duck GO |
|-----------------|--------------|--------|-------|--------------|
| Hillary Clinton | Positive | 5 | 4 | 6 |
| | Negative | 2 | 0 | 0 |
| | Neutral | 3 | 6 | 4 |
| | Passed Test? | NO | NO | YES |

This table shows # articles categorized as neutral, positive, or negative and shows whether it passed our hypothesis test, italicized in the hypothesis section

Barack Obama

| | | Google | Yahoo | Duck Duck GO |
|--------------|--------------|--------|-------|--------------|
| Barack Obama | Positive | 7 | 8 | 8 |
| | Negative | 0 | 0 | 0 |
| | Neutral | 3 | 2 | 2 |
| | Passed test? | YES | YES | YES |

This table shows # articles categorized as neutral, positive, or negative and shows whether it passed our hypothesis test, italicized in the hypothesis section

Donald Trump

| | | Google | Yahoo | Duck Duck GO |
|--|--|--------|-------|--------------|
|--|--|--------|-------|--------------|

| | | | | |
|--------------|--------------|----|----|----|
| Donald Trump | Positive | 3 | 5 | 5 |
| | Negative | 5 | 2 | 3 |
| | Neutral | 2 | 3 | 1 |
| | Passed Test? | NO | NO | NO |

This table shows # articles categorized as neutral, positive, or negative and shows whether it passed our hypothesis test, italicized in the hypothesis section

Marco Rubio

| | | | | |
|-------------|--------------|--------|-------|--------------|
| | | Google | Yahoo | Duck Duck GO |
| Marco Rubio | Positive | 1 | 3 | 3 |
| | Negative | 2 | 4 | 2 |
| | Neutral | 7 | 2 | 5 |
| | Passed Test? | NO | NO | NO |

This table shows # articles categorized as neutral, positive, or negative and shows whether it passed our hypothesis test, italicized in the hypothesis section

Results:

Altogether, our team collected 120 data points. We tested our hypothesis againsts 12 times over three different search engines. Out of those 12 times, our hypothesis passed only four times. Therefore, our hypothesis that search engines will show more than 5 positive results for liberal figures and more than 5 negative results for conservatives has failed.

Analysis:

In our experiment, we learned that there were more positive results than we expected. This was a result from the fact that many of these websites were self-promoting websites that were sponsored by the figure or the supporter of the figures. Another point that our hypothesis did not address is that there are a significant amount of neutral articles such as biographies.

However, although our hypothesis failed, the experiment did show that liberal figures had more positive articles than conservative figures. (Liberal: 38 Conservative: 20)

Possible Confounding Variables:

- Major events that take place pertaining a certain political figure may temporarily, significantly change search results
- Values of American political parties are continuously changing, and therefore the labels of conservative and liberal are not exact
- Timing of the searches and most recent events may have influenced current views of the chosen individuals

Conclusion:

In this experiment, we were not able to prove that search engines have a *statistically significant* liberal bias, as was suggested by our preliminary research. However, we were able to collect data that suggests that a bias does exist, as many of these figures have self-promoting websites as the top results. If we were to revise our experiment, we would focus on how difficult it is to get different views on the same figures, as many of the articles show the figures in a positive light. Self-promoting websites ranked highly across all search engines, and were often more than one for each figure.

We also learned that the search engine results differed from one another, but not significantly. This shows that they use similar algorithms. This can also be the basis of a possible new hypothesis that states that search engines do not show different perspectives on news stories and figures.