Reading Between the (Party) Lines

by

Sophie Beiying Chou

Submitted to the MIT Media Lab,
School of Architecture and Planning
in partial fulfillment of the requirements for the degree of

MS in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author .....
MIT Media Lab
May 5, 2016

Certified by .....
Deb Roy
Associate Professor
Thesis Supervisor

Accepted by .....
Pattie Maes
Academic Head, Program in Media Arts and Sciences

Reading Between the (Party) Lines

by

Sophie Beiying Chou

Submitted to the MIT Media Lab
on May 5, 2016, in partial fulfillment of the
requirements for the degree of
MS in Media Arts and Sciences

Abstract

TO-DO

Thesis Supervisor: Deb Roy
Title: Associate Professor

The following people served as readers for this thesis:

Sepandar Kamvar . . . . . .
Associate Professor of Media Arts and Sciences
MIT Media Lab

Iyad Rahwan . . . . . .
Associate Professor of Media Arts and Sciences
MIT Media Lab

Acknowledgments

[FILL IT WITH GRATITUDE]

Contents

Chapter 1

Introduction

Does anyone trust the news anymore? Not according to the latest Gallup Poll, which showed that only 4 in 10 Americans believe that mass media does a good job of reporting the news "fully, fairly, and accurately." It's a major decline since the poll was first taken in 1999, back when more than half (55%) of Americans believed the news was trustworthy [19].

And the trend has been steadily downward: in short, the majority of Americans have had little to no trust in mass media news coverage since 2007: a discouraging view for a tumultuous time in journalism.

But beyond frustrated readers and reporters, why does distrust in the news matter? For one, media bias—or at the very least, the *belief of* a biased media bias—may have a significant impact on the practice of democracy. A 2006 study from Georgetown University shows that those with more negative attitudes towards the news tend to be more highly influenced by their partisan prior beliefs and less by contemporary issues and messages when voting [13]. This implies that distrust of media plays a large role in the polarization of American politics.

In light of the upcoming 2016 elections, this thesis explores perceptions of media

trust and fairness in coverage of the presidential candidates. Claims of media bias and favoritism are especially high-stakes in election years, where trust has been shown to plummet [19]. And in this election cycle, cries of bias have been especially loud: Analysis at the New York Times showed that the news media gave Republican candidate Donald Trump a $1.9 billion advantage in free publicity, an amount 190 times as much as paid advertising [4].

In this thesis, we examine some of the factors that contribute to the perception of media bias. In particular, how does the *content* of a story (reading level and vocabulary) affect the reader versus the *context* (publication and author)?

We break down the larger question of media bias in two dimensions: trust and fairness in reporting, and examine the role of language in influencing the reader. Although studies have been conducted to both examine the psychological effect of wording on believability and the impact of media brands and bias, separating and comparing these two factors remains largely unexamined [23, 8]

Two studies are performed: one focusing on detecting reading level effects, and the second a follow-up on media brand effects, to collect reader's perceptions of news stories through crowdsourcing. We manipulate the source of the story to examine effects of media brands on the reader, and also compare trust and fairness rankings between high and low reading level stories.

Although the general consensus of mistrust is clear, perception of media bias is a complex phenomenon to dissect, as it combines social and psychological effects with the traits of the story itself. This thesis hopes to shed new light on understanding what motivates readers' trust and distrust of news media, and pave pathways for positive intervention.

Chapter 2

In Media We ... Trust?

Despite the news media ecosystem's rapid evolution in the past decade, the question of fairness in reporting remains a valued one. Although counterarguments for subjective reporting exist (Glenn Greenwald, most famous for his coverage of whistleblower Edward Snowden's leaks, said that "All journalism is a form of activism. Every journalistic choice necessarily embraces highly subjective assumptions–cultural, political or nationalistic–and serves the interests of one faction or another"), fair treatment of subjects and sources remain a central tenant to most publications [3].

But an attempt at fairness on the side the reporter is not always perceived in equal effect under the eyes of the reader. Presenting contradictory facts to a reader's beliefs can even sometimes *strengthen* their oppositions to it, a concept known as "motivated skepticism" [21].

In this section, we examine the impact of distrust in media, and explore the theories behind three main potential sources of media distrust: the characteristics of the reader, the source of the story and its use of language.

2.1 Why Does Media Trust Matter?

The idea that mass media has a large influence on the ramifications of democracy is nothing new. In 1922, political commentator, reporter, and writer Walter Lippman wrote about its central role in shaping public opinion:

> Each of us lives and works on a small part of the earth's surface, moves in a small circle, and of these acquaintances knows only a few intimately. Of any public event that has wide effects we see at best only a phase and an aspect. This is as true of the eminent insiders who draft treaties, make laws, and issue orders, as it is of those who have treaties framed for them, laws promulgated to them, orders given at them. Inevitably our opinions cover a bigger space, a longer reach of time, a greater number of things, than we can directly observe. They have, therefore, to be pieced together out of what others have reported and what we can imagine. Yet even the eyewitness does not bring back a naive picture of the scene. [15]

Many of the worries that Lippman had about the effects of poorly disseminated truth have been later confirmed in experimental studies. In short, when faced with a large and mistrusted news environment, we tend to rely on *confirmation bias* when searching for information. This term, first coined in 1988, describes the psychological phenomeon of seeking or analyzing new information in ways that align with one's existing beliefs, expectations or prior hypotheses [17].

Using a Bayesian voting model, a study from Georgetown University in 2005 was able to show that voters with low trust and a high dislike for the news media are significantly more influenced by their existing party identifications in casting ballots than current economic factors [13]. The study attributes increasing polarization in the American political sphere with increasing lack of trust in the news, a serious implication for the highly polarized 2016 presidential elections. Moreover, distrust of media implies a large information loss in the public, whose avoidance of diverse ideas

means a narrowing of information flow [12].

2.2 How is Media Trust Formed?

2.2.1 The Role of the Reader

The perception of media bias is a cornerstone component of distrust in the news. After all, most Americans claim that they want to read news that's unbiased. A survey from Pew Research in 2012 showed that more than two-thirds (68%) of readers want to read political articles with a neutral stance, compared to just a little less than a quarter (23%) of those who want to read those stories that share their point of view [5]. But what exactly does that entail?

It comes as no surprise that our own political stances have a significant effect in our perceptions of bias in the media. On whole, conservative readers tend to view media as more biased than both Democrats and Independents (49% to 32% and 35%, respectively)[5]. Partisans have also been shown to view the news as antagonistic to their beliefs, a phenomenon known as the "hostile media effect".

The effect, first studied in the 1980s, showed that when faced with the same piece of news media about the Sabra and Shatila massacre in Beirut, pro-Israeli and pro-Palestinian students both claimed the news clip was biased in favor of the other side [22]. It has since been repeated in a variety of contexts to the same effect.

What the story is reporting does not matter so much as the individual's attitude *towards* that issue. In 1988, Albert Gunther found a curvlinear effect between the viewer's polarization towards an issue and their trust in the media to fairly cover it [10]. In doing so, he suggests two models of persuasion to help understand media processing: first, the cognitive response theory, which predicts more portential for attitude change when the reader is highly involved in the content, as they are pro-

cessing information more deeply [9]. Social judgement theory, on the other hand, expects less change in attitude when the reader is highly involved or polarized about a subject, as they will simply reject the new information [20]. These two opposing theories help explain the presence of a curvlinear relationship to exposure to news media and resulting media trust.

2.2.2 The Role of Media Brands

The media, of course, is not just one unified mass, and in an increasingly fragmented ecosystem, the role of brands is a crucial factor in media trust. With the rise of the internet, the past decade has seen an explosion of new media platforms and publications, as well as significant transformations in style and audience in existing outlets.

Although the studies above present the media as one unified mass, there is a significant amplifying effect of hostility and bias perception depending on the reader's prior connotations of a news outlet. In 2008, researchers Matthew Baum and Phil Gussin showed significant differences in the evaluation of a piece of news content depending on whether it was labeled to be from CNN, Fox, or a fictional news outlet [8]. They concluded that media bias is very much "in the eye of the beholder," as viewers make information shortcuts dependent on media brand to jump to conclusions beyond their own partisanship and the content of the story.

2.2.3 The Role of Language

Chapter 3

Tools for Dissecting Trust

3.1 Computing Reading Level

3.1.1 Flesch-Kincaid Readability Tests

In this study, we focus primarily on the Flesch-Kincaid (F-K) tests for estimating text readability. Originally developed for the U.S. Navy in 1975 for assessing the difficulty of technical manuals, the F-K reading level corresponds roughly to U.S. grade level and the reading ease score is inversely proportional to the grade level on a scale from 0 to approximately 120 [11].

We chose the F-K tests over other comparable ones due to its popularity in educational assessment and other applications, including in legislation. For example, it is required by law in Florida that life insurance policies have a Flesch reading ease of 45 or greater (less than 12th grade in reading level) [14]. The F-K tests are also bundled in many common word processing services, including Microsoft Office Word. As a comparison, basic article analysis is also computed using the Gunning fog index (see Section 5.2.1).

The formula for Flesch reading ease is as follows:

$$206.835 - 1.015\left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6\left(\frac{\text{total syllables}}{\text{total words}}\right)$$

And for reading grade level:

$$0.39\left(\frac{\text{total words}}{\text{total sentences}}\right) + 11.8\left(\frac{\text{total syllables}}{\text{total words}}\right) - 15.59$$

The two formulas are not directly comparable due to the difference in weighting factors. For ease of metaphor, we use the grade level tests in our analysis. Syllable length is highly weighted in this formula, so it is possible to generate a story of very high reading level that consists of a single word in a single sentence (the longest English word, *pneumonoultramicroscopicsilicovolcanoconiosi*, a type of lung disease, has a reading grade level of 197.2), which is a limitation of the method, since texts with polysyllabic words are not always necessarily more difficult to read.

3.1.2 Comparison to Other Reading Tests

3.2 Crowdsourcing Science

Talk about platform, vs. turk, basic demographics (later show ones we found)

Chapter 4

Reading Level Effects

4.1 Motivation

What was the purpose of the study? What were the hypotheses?

4.2 Experimental Design

4.2.1 Dataset

Data Selection

For this study, we chose to analyze stories collected between January 1, 2016 (the start of the election year) and March 1, 2016 (Super Tuesday). Since a large number of states hold primary elections and caucuses on Super Tuesday, it is seen as an early indicator of candidate electability. All stories had been filtered through both the election (see section 3.3) and topic (see section 3.4) classifiers.

Based on the results of Super Tuesday, we selected four candidates for this study by

delegate count: Hillary Clinton (1,279), Bernie Sanders (1,027), Donald Trump (743), and Ted Cruz (517) [7].

News articles were then separated into single-candidate stories (i.e. articles featuring primarily one candidate in the headline) to be able to measure more clearly the perceived bias per candidate. This was done programatically using regular expressions to determine if a headline contained one candidate and one candidate only. A dictionary of related names was created to make sure that stories were correctly categorized (i.e. "Hillary", "Clinton", and "Hillary Clinton" were to be categorized as pertaining to "Hillary Clinton" but not if preceded by "Bill").

Publication Selection

For the purposes of this study, stories were examined from five outlets:

- CNN
- Fox News
- The New York Times
- The Wall Street Journal
- The Associated Press

The choices consist of two pairs of outlets in both print and television across the liberal-conservative divide, plus a wire service. Of the 14 outlets above, both Fox News and the Wall Street Journal have an audience that leans conservative compared to the overall population (27% mostly conservative viewers versus 17% in the overall population for Fox News and 22% mostly conservative viewers versus 17% in the overall population) measured by a 2014 Pew survey [16].

On the other hand, the New York Times and CNN both have audiences that lean mostly liberal (25% liberal versus 22% in all respondents for CNN and 25% for the New York Times). The Associated Press, which was not included in the survey, has members in outlets across the political divide and was chosen as an experimental

control.

[MIGHT INCLUDE THOSE DISTRIBUTIONS HERE]

Topic Selection

The top four topics by volume (Immigration, Abortion, Campaign Finance, Foreign Policy/National Security) were chosen for the survey to ensure a significant number of stories for each candidate for each topic. For overall topic distributions, see 9 in the Appendix.

Story Redaction

For those readers in Group B (no disclosure of source), stories were further redacted to remove *all mentions of news publications* in the body of the text and headline, to prevent readers from making assumptions about the source. The redacted words were replaced with the placeholder "xxx" and explained in the instructions for the task.

4.2.2 Survey

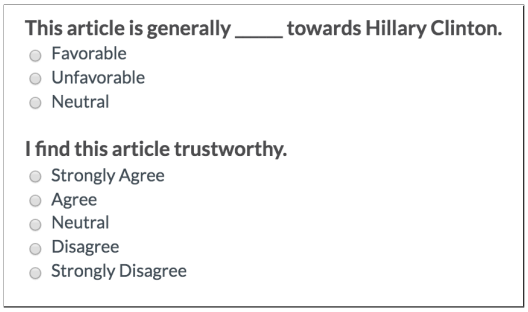

Figure 4-1: Survey Questions for Study 1

4.2.3 Quality Control

CrowdFlower has a built-in "Test Question" feature that allows for the rejection of a annotator whose answers to specific questions do not lie within a threshold (default 70%) of the "correct" answer or whose answers lay outside the standard variation compared to others.

However, since the questions we asked were by nature subjective and therefore outliers and disagreements in answers could imply signal rather than noise, we chose to monitor for quality using other metrics instead. CrowdFlower was not designed explicitly for survey-like tasks, and therefore there were no options for different screening methods or questions. Gold Questions on the platform are selected by the creator within the set of all questions being recorded.

Because of this, we monitored quality of results in two ways:

First, by setting a minimum of time of 180 seconds to complete the task of reading 5 stories for a task to be accepted,

And second, by selecting only Level 3 contributors on CrowdFlower as suggested on their website for handling survey-like tasks [2].

Level 3 contributors are described as those who "have completed over a hundred Test Questions across hundreds of different Job types, and have a near perfect overall Accuracy" [6]. This is the highest category of contributor.

Users were also only allowed to answer the set of questions once.

Average response time was 07:31 minutes. $0.50 was given per survey, as suggested by MIT Committee on the Use of Humans as Experimental Subjects [1].

4.3 Analysis

4.4 Conclusions

4.5 Limitations

From our exploratory study on reading level effects, we were able to obtain a significant but weak effect between disclosing the source and the levels of trust marked by readers towards an article.

We also observed trends that suggested an interaction between disclosing the source and the reading level of a story.

However, the study faced several limitations: first, we did not obtain enough samples to show a statistically significant result for interactions between source and reading level.

Furthermore, multiple levels of independent variables (ie: 5 levels for input source) made modeling complex and the results less clear.

The dataset was also unbalanced and sparse (ie, because of large numbers of input variables we did not have complete representation for each category, such as high, low, and mid-reading level stories for every outlet and topic). We tried to control for those factors by randomization, however it made more difficult to analyze specific correlations between source and trust.

To further explore the interaction between disclosing the source and the reading level of the story, we set up another crowdsourcing experiment on CrowdFlower, this time targeting this specific interaction, to see if there is a significant effect between the two, detailed in the following chapter.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

Media Brand Effects

5.1 Motivations

Following the limitations and patterns from our first study, we designed a follow-up study to verify interactions between news source and reading level.

This study focused on two main hypotheses:

- H1
- H2

5.2 Experimental Design

For the second study, our experiment was revised to have a 4 x 2 mixed-factorial design. In this study, reading level of articles and candidates featured in the articles were treated as within-subject variables, and the source of the story between-subjects.

	Source: None	**Source: AP**	**Source: Fox**	**Source: CNN**
High Reading Level	Clinton, Cruz, Sanders, Trump	Clinton, Cruz, Sanders, Trump	Clinton, Cruz, Sanders, Trump	Clinton, Cruz, Sanders, Trump
Low Reading Level	Clinton, Cruz, Sanders, Trump	Clinton, Cruz, Sanders, Trump	Clinton, Cruz, Sanders, Trump	Clinton, Cruz, Sanders, Trump

Table 5.1: Main Study Design

This time, we reduced the number of stories to N=8, and also changed reading level from a 3-level to 2-level variable (low, high) for clarity.

Most significantly, since we observed some significant effect from disclosing source to the reader in Study 1, we added a manipulation in this experiment to further study the effect of revealing the source:

Following Baum's research in showing the effects of media brands and reader bias by manipulating reported brands, all eight stories in Study 2 were in fact written by the Associated Press, however, we manipulated the source shown to the reader [8]. In group A, readers were shown the headline and text of the story with no other context. In group B, readers were additionally shown that the story was from the Associated Press (true label). In groups C and D, readers were shown that the story was from CNN and Fox News, respectively.

This setup was created to eliminate some of the confounding effects from using stories from different sources (writing style, focus of content, slant, etc.), while directly observing the effect of revealing a specific source to the reader. The Associated Press was chosen as the source of the stories as it is the highest circulation newswire service in the United States, and has 14,000 members that use its content [18]. Notably, both CNN and Fox News publish content in full or part from the Associated Press, although the specific stories chosen had not been published in full by either to avoid bias.

We removed the favorability question from Study 1 (as the 3-point scale did not yield significant results), instead asking the reader more directly about media bias

by ranking the fairness of the story on a 5-point Likert scale. The trustworthiness question from Study 1 was kept, also on the same 5-point Likert scale.

5.2.1 Dataset

Eight stories were chosen for this study: two (high and low reading level) per candidate. All eight stories were written by reporters from the Associated Press (although they may have been republished elsewhere).

Reading level cutoffs were made by taking the bottom and top 25% percentile of Flesch-Kincaid scores for each candidate. From stories written by the Associated Press that made the cutoff, we formed pairs of high and low reading level stories from each topic. The topic with the highest distance between reading level in the pair was chosen for each candidate.

5.2.2 Survey

Figure 5-1: Survey Questions for Study 2

Demographics What is your gender? Male Female Other What is your age? 18-29 30-49 50-64 64+ What is your political party affiliation? Democrat Republican Independent/Other No Affiliation If the election were tomorrow, I would vote forâĂę Hillary Clinton Bernie Sanders Donald Trump Ted Cruz John Kasich Other Might not vote Have you read any of the stories above before this study? Yes No Might have, but I am not sure

5.2.3 Quality Control

As in the first study, users were filtered by minimum time taken to complete the task. Again, only Level 3 workers were chosen from Crowdflower, and users from Study 1 were forbidden to complete tasks in Study 2.

Because Study 2 presented a longer task of reading 8 versus 5 news stories, we set a longer minimum time of 6 minutes to complete the task. A payment of $0.80 per survey was given based off guidelines by MIT [1].

The average response time for the task was 09:20 min.

5.3 Analysis

5.4 Conclusions

How do your trustworthiness findings line up with the findings from Pew surveys and prior work? What hypotheses did you verify from prior work?

5.5 Limitations

Just 8 stories I know Yes the diff candidates had diff topics i know we could have included all candidates

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 6

Future Work

So... what? We found these effects, so what?

–> interventions here is a possible design.

6.1 Designing Interventions

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 7

Conclusion

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 8

Tables

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 9

Figures

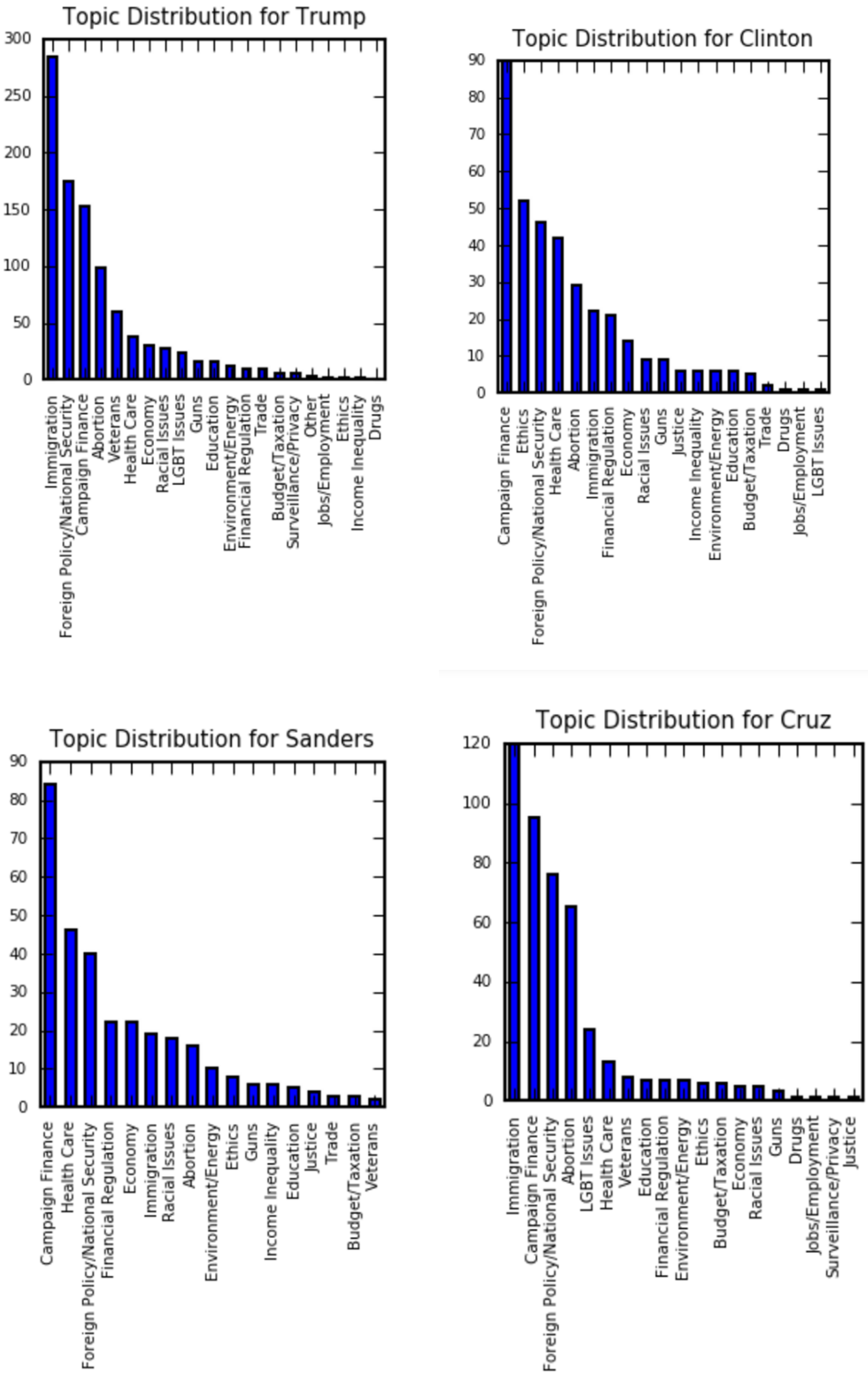

Figure 9-1: Topic Distributions for Candidates

Bibliography

[1] Couhes policy for using amazon's mechanical turk | welcome to couhes. https://couhes.mit.edu/guidelines/couhes-policy-using-amazons-mechanical-turk. (Accessed on 04/21/2016).

[2] Guide to: Running surveys âĂŞ- crowdflower success center. https://success.crowdflower.com/hc/en-us/articles/201855969-Survey-Guide-To-Running-Surveys. (Accessed on 04/21/2016).

[3] Is glenn greenwald the future of news? – the new york times. http://www.nytimes.com/2013/10/28/opinion/a-conversation-in-lieu-of-a-column.html?mtrref=www.google.com&gwh=450466A4B52C2124639CE220312812D5&gwt=pay&assetType=opinion. (Accessed on 04/20/2016).

[4] Measuring donald trump's mammoth advantage in free media - the new york times. http://www.nytimes.com/2016/03/16/upshot/measuring-donald-trumps-mammoth-advantage-in-free-media.html. (Accessed on 04/20/2016).

[5] Perceptions of bias, news knowledge. *Pew Research Center*, feb 2012. (Accessed on 04/06/2016).

[6] Introducing contributor performance levels! - crowdflower community. http://crowdflowercommunity.tumblr.com/post/80598014542/introducing-contributor-performance-levels, mar 2014. (Accessed on 04/21/2016).

[7] March 1 election results 2016 updates | the washington post. https://www.washingtonpost.com/2016-election-results/super-tuesday/, mar 2016. (Accessed on 04/06/2016).

[8] Matthew A Baum, Phil Gussin, et al. In the eye of the beholder: How information shortcuts shape individual perceptions of bias in the media. *Quarterly Journal of political science*, 3(1):1–31, 2008.

[9] John T Cacioppo and Richard E Petty. Effects of message repetition and position on cognitive response, recall, and persuasion. *Journal of personality and Social Psychology*, 37(1):97, 1979.

[10] Albert Gunther. Attitude extremity and trust in media. *Journalism & Mass Communication Quarterly*, 65(2):279–287, 1988.

[11] J Peter Kincaid, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Technical report, DTIC Document, 1975.

[12] Jonathan M Ladd. *Why Americans hate the media and how it matters*. Princeton University Press, 2011.

[13] Jonathan McDonald Ladd. Attitudes toward the news media and voting behavior. In *Annual Meeting of the Midwest Political Science Association*, 2005.

[14] The Florida Legislature. The 2015 florida statutes title xxxvii. `http://www.leg.state.fl.us/Statutes`. (Accessed on 04/05/2016).

[15] Walter Lippman. *Public Opinion*, chapter 6.1. Free Press, 1922.

[16] Amy Mitchell, Jeffrey Gottfried, Jocelyn Kiley, and Katerina Eva Matsa. Political polarization & media habits. *Pew Research Center*, Oct 2014. (Accessed on 04/02/2016).

[17] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175, 1998.

[18] Associated Press. Faqs. `http://www.ap.org/company/FAQs`, 2016. (Accessed on 04/21/2016).

[19] Rebecca Riffkin. Americans' trust in media remains at historical low. `http://www.gallup.com/poll/185927/americans-trust-media-remains-historical-low.aspx`. (Accessed on 04/20/2016).

[20] Muzafer Sherif and Carl I Hovland. Social judgment: Assimilation and contrast effects in communication and attitude change. 1961.

[21] Charles S Taber and Milton Lodge. Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3):755–769, 2006.

[22] Robert P Vallone, Lee Ross, and Mark R Lepper. The hostile media phenomenon: biased perception and perceptions of media bias in coverage of the beirut massacre. *Journal of personality and social psychology*, 49(3):577, 1985.

[23] Deena Skolnick Weisberg, Frank C Keil, Joshua Goodstein, Elizabeth Rawson, and Jeremy R Gray. The seductive allure of neuroscience explanations. *Journal of cognitive neuroscience*, 20(3):470–477, 2008.