Burst Your Bubble! An Intelligent System for Improving Awareness of Diverse Social Opinions

Mingkun Gao, Hyo Jin Do, Wai-Tat Fu
Department of Computer Science, University of Illinois at Urbana-Champaign
Urbana, USA
{mgao19, hjdo2, wfu}@illinois.edu

ABSTRACT
Social media users are overloaded with diverse opinions by people with opposing stances. Previous research shows that people often look for opinions that reinforce their pre-existing beliefs and stances, which may lead to social polarization. Traditional social media present opinions in a linear list format, which not only lacks structures for people to explore diverse viewpoints but also aggravates their selective exposure to agreeable opinions. To address this problem, we designed an intelligent system that improves awareness of diverse social opinions by providing visual hints and recommendations of opinions (e.g. news articles and comments) on different sides with different indicators. We evaluated our system with news articles about Obamacare repeal issue and their corresponding user comments from Facebook. Results demonstrate that our system could increase people's awareness of their stances and opinion selection preferences, which mitigates selective exposure and thereby leads to a more balanced perception of social opinions.

Author Keywords
Intelligent System; Social Opinion; Selective Exposure;

INTRODUCTION
Social media have been playing an increasingly important role in spreading and shaping social opinions. Pew research [28] showed that, in 2016, 62% of U.S. adults get news from social media. Contrary to the expectation that people are exposed to more diverse opinions, some research has shown that social media may have the opposite effect, i.e., they may play a role in encouraging people to be more selective in their information consumption – a phenomenon often referred to as *selective exposure* [21, 22, 23]. Selective exposure to information may not only prevent people from receiving diverse opinions but also cause polarization of social opinions. This effect may be exacerbated by implicit personalization algorithms that selectively guess what information users would like based on their history of information selection, a process that may keep users separated from information that is inconsistent with their beliefs or attitudes. Users are often not aware of this "filter bubble" [50], which may have a negative consequence on social opinions. For example, the lack of awareness has raised some concerns over how selective exposure to "fake news" on social media has influenced the US presidential election. The goal of the current paper is to investigate strategies for designing intelligent features that may mitigate this type of implicit behavioral tendencies.

When people are selectively exposed to information, they may not be fully aware that their attitudes are influencing their selection [30]. Some people may have traces of past experiences that mediate their favorable or unfavorable feelings about certain social issues, and these feelings may influence their information selection behavior that they may not be aware of. For example, when people are asked whether they support or oppose the repeal of Obamacare, they may think that they are neutral. However, when they encounter information related to various perspectives of the issue, their experiences or beliefs (e.g., their general beliefs about the role of the government) may lead to varying level of behavioral dispositions that mediate their behavior (e.g., more likely selecting information favoring the repeal) in ways that they are not aware of. Such behavioral dispositions may be magnified as people become selectively exposed, as their beliefs are reinforced by attitude-consistent information, leading to an *echo chamber* effect.

Online interface design often simply presents social opinions (e.g. news article comments) in a linear format. There is a general lack of structure that helps improve people's awareness of their behavioral dispositions and opinion selection preferences. In fact, research shows that the lack of sensitivity to people's attitudes towards different social issues in the design of social information interfaces may aggravate selective exposure, as people have a natural tendency to process information that they agree with.

To address this problem, we proposed a novel intelligent system consisting of design features that aim to improve people's awareness of their stances and selection preferences by encouraging them to attend to more diverse opinions. To be more specific, our system provides people with novel visual hints, including showing a trace of people's stances based on previously read news articles, as well as highlighting visu-

IUI'18, March 7–11, 2018, Tokyo, Japan

DOI: https://doi.org/10.1145/3172944.3172970

ally on the interface when people are selectively exposed to one side of opinions. In addition, our system recommends attitude-inconsistent and attitude-consistent comments according to people's stances with different priority and different recommendation indicators. When recommending attitude-inconsistent opinions to people at first, the system will label these opinions as "Recommended". Later, when recommending attitude-consistent opinions, these opinions will be labelled as "Not Recommended". We expect that the recommendation mechanism and the indicators could *nudge* people to attend to the connection between the indicator and the stance expressed in the recommended comments so that they could realize their stances and the existence of the "filter bubble". For example, people may be accustomed to being recommended attitude-consistent opinions by personalization algorithms on social network. We hope the mismatch of the indicator and stances expressed in recommended opinions in our system could motivate people to inspect their own stances, as well as exposing diverse opinions to people.

To summarize, we proposed the following research questions:

- RQ1: Can our system help people become more aware of their stances and social opinion selection preferences?
- RQ2: Can our system mitigate selective exposure when people use it to read social opinions?

RELATED WORK

Unconscious Bias and Awareness of Stance

Previous research [30, 41, 46, 47] shows that unconscious bias is common. It has been shown that sometimes people do not realize that they have implicit attitudes towards social issues, certain organizations, or groups of people. Unconscious bias could affect people's social behaviors. For example, [12, 41, 46] showed that unconscious bias influenced people's social behaviors in various scenarios, such as recruiting people from different races and evaluating public sector performance. In addition to unconscious bias, sometimes people are not aware of their own stances. The lack of awareness of implicit attitudes imply that people are often not aware of what and why certain information is recommended by a personalization algorithm such as the news aggregators from Google or Facebook. Given that how these algorithms work is often not transparent, people are also not aware how the information is selected. As a result, users may often form an inaccurate impression on other people's opinions. In order to mitigate implicit bias in information selection, we designed novel visual hints to improve people's awareness of their own stances.

Selective Exposure to Information

The current design of social media stacks opinions using syntactic features without considering their stances. Researchers [21, 22, 23] found that such interface design of social opinions may aggravate the selective exposure tendency of users who seek for attitude-consistent opinions while avoiding attitude-inconsistent ones. Zaller [57] also showed that people often show a biased assimilation [45] of information, which refers to a tendency of accepting attitude-consistent messages more than attitude-challenging messages. These effects often lead to other undesirable consequences, such as the echo chamber effect [43, 53], a creation of filter bubbles that arise from social homophily and personalization algorithms [11], and false consensus bias [16, 17, 40, 52], in which people estimate their values are more common than others who hold opposite values. These effects are considered undesirable because they may lead to polarization of opinions in the society.

Researchers have been trying to understand the selective exposure effect in various ways [31, 48, 49]. For instance, [31] argued that selective exposure behavior can be mitigated by exposing users to weak ties and visualizing the common ground facts. Many researchers attempted to mitigate the effect of selective exposure at the interface level. For instance, Liao, et al. [44] investigated the effect of aspect indicators on moderating selective exposure in the medical decision-making process. The results showed that aspect indicators facilitated users to seek other aspects, thus mitigating selective exposure. In [43], they studied whether the presence of source expertise indicators could influence participants' selective exposure. In addition, Garimella and his colleagues [27] attempted to bridge echo chambers using algorithmic techniques with endorsement graph and random-walk controversy score [26]. In our study, we designed a system with novel visual hints and recommendation mechanism to mitigate selective exposure.

Interface Design for Social Opinions

A torrent of research has presented novel interface designs to help users to get balanced and global insight from inundating social opinions online. Hoque and his colleagues [35, 36, 37] presented an interactive text analytic system to help people gain insight from long conversations in blogs by using topic segmentation [24] and topic labeling [59] on Fragment Quotation Graph [39]. Zhang, et al. [58] presented a system that helps readers or editors to effectively break a large-scale discussion down into subtopics using a summary tree. Faridani, et al. [20] proposed Opinion Space that automatically highlights insightful opinions from a range of perspectives.

Among these design features, sentiment is one of the most frequently utilized features in interfaces for social opinions. Many customer websites such as Amazon or IMDB leverage sentiments from feedback to gain better insights. Moreover, many social interfaces utilize sentiments or emotions to organize and visualize user-generated comments. For instance, Gao, et al.[25] presented an intelligent interface that organizes comments from Reddit based on their stances and emotions, and found that the interface helped people to get global insight about controversial topics. In our study, in addition to reminding people of their stances and opinions selection preference with visual features, we deploy sentiment analysis technique [54] to detect users' stances and propose a novel recommendation mechanism to mitigate selective exposure.

Sentiment Analysis

Despite the difficulty in sentiment classification for social media text, mining sentiment on such platforms can be practically useful to gain insights of a huge volume of opinionated information. Companies can organize consumer reviews of their products and services using sentiment analysis and therefore improve their marketing strategies [29]. Governments can observe public sentiments for rapid policy-making [14]. Do, et al. [19] used sentiment classification to investigate the trend of public opinions during the Middle East Respiratory Syndrome (MERS) outbreak in the Republic of Korea. Bollen, et al. [13] used sentiment to predict the stock market. Many recent works attempt to analyze sentiment in non-English texts, such as Japanese [55], Chinese [54] and Korean [18]. Similar to sentiment analysis, there have been a surge of studies in stance and reason classification [10, 32]. In our paper, we crawled Facebook comments and reactions to build a sentiment analysis model specifically attuned to social media domain. We leveraged this model to identify users' stances from their comments about the news article, thus, can recommend attitude-consistent or attitude-inconsistent information.

SYSTEM DESIGN AND IMPLEMENTATION

Data

Facebook allows users to select emoticons, among *Like*, *Love*, *Haha*, *Wow*, *Sad* and *Angry*, to express their emotions to the articles. Many Facebook users not only write their comments to articles but also select the emoticons to express their emotional reactions. This gives us cues to know users' emotions when they write their comments. We intended to categorize emotions into positive and negative sentiments. We found that the *Like* emoticon is so general that many Facebook users always selected the *Like* emoticon regardless of the emotions expressed in their comments. In addition, the *Haha* and *Wow* emoticons are too ambiguous. Some of the comments with the *Haha* emoticon expressed the emotion of jeer whereas others may express the emotion of happiness. The *Wow* emoticon is for the feeling of surprise. However, the sentiment of surprise could be either negative or positive. On the contrary, comments with emoticons of *Love*, *Sad* and *Angry* express relatively consistent emotions with their labels in most cases. Thus, in our study, the *Love* emoticon represents positive sentiment whereas the *Sad* or *Angry* emoticons denotes negative sentiment.

We selected the Obamacare repeal issue as a controversial topic in our study. In order to implement the system and mitigate the bias caused by different news sources, we collected CNN[1] and FoxNews[2] news articles with comments from Facebook[3] using Facebook API[4]. We crawled most recent 100 CNN news articles published from May 8th 2017 to August 10th 2017 and 100 FoxNews articles published from May 1st 2017 to August 16th 2017 from Facebook. Note that the time coverage is slightly different because we were not able to collect the same number of articles in the same period of time. But we tried to remove the potential bias stemmed from different collection times by crawling articles in overlapping periods that the median dates are similar. All these articles we collected have the keywords "health care" in their Facebook news messages. We analyzed people's sentiment (positive vs. negative) towards these articles by comparing how many people selected the *Love* emoticon and how many people selected the *Sad* or *Angry* emoticons. On average, the comparison between being positive and being negative is 40.03% vs. 59.97% for CNN news articles and 45.57% vs. 54.43% for FoxNews articles. This suggested that Obamacare repeal issue is a typical controversial topic on social network so that we select this as a representative in our study.

[1]http://www.cnn.com/

[2]http://www.foxnews.com/

[3]https://www.facebook.com/

[4]https://developers.facebook.com/

In order to train the sentiment classifier and build our system, for each article, we also collected the corresponding Facebook comments. Thus, unlike previous sentiment analysis studies where researchers collected data first and then recruited external annotators to add sentiment labels to their data [42, 55], we only collected Facebook comments provided by people who also selected one or more emoticons to indicate their emotions. For each comment we collected, there could be one or more emotion labels selected by comment providers. We labelled comments with the *Love* emoticon as positive sentiment whereas those with the *Sad* or *Angry* emoticons as negative sentiment. These positive and negative labels are treated as gold standard sentiment labels.

Data for Sentiment Analysis

We randomly selected 4,000 positive comments and 4,000 negative comments from our dataset to train and test the sentiment classifier. To further evaluate the classifier, we randomly selected extra 200 comments from each side. These extra 400 comments are used as test data for the comparison of performance between our classifier and human annotators (e.g. online crowd-workers).

Data for System Implementation

To implement the system prototype, we selected four CNN news articles [1, 2, 3, 5] that oppose to repeal Obamacare and four FoxNews articles [4, 6, 7, 8] that support to repeal Obamacare from our article pool. For each selected article, we picked eight comments that support this article and eight comments that oppose this article to obtain diverse opinions.

Sentiment and Stance Analysis

We trained a classifier to predict people's sentiment to an article from their comments. We worked on the corpus where there are 4,000 comments on each side. We preprocessed the comment texts by stemming and removing stop words, and we converted comments into unigram feature vectors. Since different terms have different levels of importance when used to decide the sentiment, we adopted χ^2 statistic (*CHI*) test [54, 56] on each unique term in the training set of the corpus. The *CHI* statistic measures the degree of association between the term and the sentiment category. The definition of *CHI* statistics between term t and sentiment category c_i is:

$$\chi^2(t,c_i) = \frac{N \times (AD - CB)^2}{(A+C)\times(B+D)\times(A+B)\times(C+D)} \qquad (1)$$

where A is the number of times t and c_i co-occur, B is the number of times t occurs without c_i, C is the number of times c_i occurs without t, D is the number of times neither t nor c_i occurs and N is the total number of comments. If t and c_i are independent, the value of the χ^2 statistic is zero. For each

	Positive	Negative
Precision	0.742	0.683
Recall	0.638	0.778
F_1 score	0.686	0.727

Table 1: 10-fold cross validation result of sentiment classifier.

unique term t, we calculated the χ^2 statistic for each sentiment category and finally we assigned the weight, which shows the importance of the term, to each unique term t with:

$$\chi^2_{max}(t) = \max_{i=1}^{m}\{\chi^2(t, c_i)\} \tag{2}$$

where m is the total number of categories. In our case, m equals 2 since we have two sentiment categories (positive vs. negative).

In the feature vector of each comment, the normalized weight $Weight(t,D)$ of a unique term t in comment D is calculated based on:

$$Weight(t,D) = \frac{tf(t,D) \times \chi^2_{max}(t)}{Total\_Weight(D)} \tag{3}$$

where $tf(t,D)$ is the term frequency of the unique term t in comment D and $Total\_Weight(D)$ is the total weight of all terms in comment D ($Total\_Weight(D) = \sum_{t \in D} tf(t,D) \times \chi^2_{max}(t)$).

We trained a Random Forest classifier [33, 34] on feature vectors using Python scikit-learn package[5]. We evaluated our sentiment classifier via 10-fold cross validation using the corpus. We calculated the precision, recall and F_1 score for both positive and negative sentiment categories and results are shown in Table 1.

Furthermore, we compared the performance between our classifier and human annotators on the extra test set which included 200 randomly selected comments in each sentiment category. To evaluate the performance of human annotators, we recruited Turkers on Amazon Mechanical Turk platform (AMT)[6] to annotate sentiment labels (positive vs. negative) to each comment. In each Human Intelligent Task (HIT), we presented 20 comments and we asked Turkers to identify the sentiment expressed in each comment. For each comment, we collected five sentiment labels and used majority vote to determine final sentiment. Ultimately, 100 Turkers participated into our study with a reimbursement of $1.00 per HIT on September 30th 2017. Among the 400 testing comments, 102 comments received either two positive votes and three negative votes or three positive votes and two negative votes from Turkers. This result suggested that among the 400 testing comments, about 25% of them are ambiguous. We compared model-predicted and human-annotated labels with the gold standard respectively and calculated the precision, recall and F_1 score for both methods. Results are shown in Table 2. We found that our classifier's performance is comparable to human annotators' performance on this difficult task.

[5]http://scikit-learn.org/

[6]https://www.mturk.com

		Positive	Negative
Random Forest	Precicion	0.734	0.678
	Recall	0.635	0.770
	F_1 score	0.681	0.721
Human Annotator	Precision	0.791	0.671
	Recall	0.585	0.845
	F_1 score	0.673	0.748

Table 2: Comparison of performance between our classifier and human annotators.

With the predicted sentiment label of the comment to an article, we could predict comment providers' stances as we know the stance of these articles in our study. Here, stance means the opinion (support vs. oppose) towards an issue. For example, if the sentiment classifier predicted a comment to be negative to an article which opposes to repeal Obamacare, we would predict the comment provider's stance as supporting Obamacare repeal.

System Design Principle

Our system aims to improve people's awareness of their own stances and social opinion selection preferences, and mitigate selective exposure. Based on previous studies about nudging the bias of selection using various visualization techniques [25, 51] and recommendation systems [9, 15], we propose the design principles as below:

1. **Highlight visual difference when opinion selection bias is detected:** The system should provide highlighted difference visually when social opinion selection bias is detected. We expect that visually salient hint is effective to remind people that they may be selectively exposed to a specific side of opinions.

2. **Show the trace of people's stances:** When the system detects stances of users, it should provide explicit feedback to improve users' awareness of their stances. We expect that this feature not only will help them realize their own stances, but also be more conscious of the stances of other people's opinions relative to their own stances.

3. **Recommend social opinions from an opposite stance with positive indicators:** People who are biased towards one side of the opinions may intentionally avoid opinions on the other side, or they may be unintentionally kept in a filter bubble by personalization algorithms. In our study, we investigate whether recommending opinions opposite to a user's stance would help increase the diversity of exposure to social opinions, thereby raising their awareness of their own stances. In the current design, the system will recommend attitude-inconsistent opinions with a positive indicator (i.e. labelling these opinions as "Recommended")

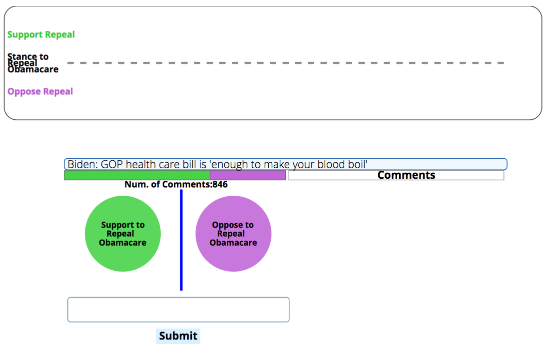


Figure 1: An overview of our new interface.

to encourage people to attend to these opinions – i.e., we expect that this will help burst the filter bubble.

4. **Recommend social opinions in the same stance with negative indicators:** If the system only recommends attitude-inconsistent opinions to people, people may leave the platform when they feel dissatisfied with the recommended information. In the current design, we investigate whether recommending attitude-consistent opinions will fulfill the implicit information needs of users. However, to avoid the effect of a filter bubble, we will provide a negative indicator to the opinions (i.e. labelling these opinions as "Not Recommended"). The negative indicator may also stimulate people's curiosity about why the system shows opinions which are not recommended. By nudging people to figure out the connection between the indicator and the stances expressed in the opinions, our system may help people realize their stances.

Visual Encoding and Implementation

To achieve our design goals, our system incorporates visual hints and opinion recommendations. There are four articles (two FoxNews articles and two CNN news articles) in our system. The system only shows people one article at a time and recommends the next article based on people's stances to the previous one. Figure 1 shows the overview of our new interface. On the top of the article title, there's a panel to trace people's stances on previous articles. The color bar under the article title indicates how controversial the article may be. The length of the green bar indicates the proportion of comments which support to repeal Obamacare and the length of the purple bar indicates the proportion of comments on the other side. There are two round clickable stance labels under the color bar. Green stance label represents comments which support to repeal Obamacare and purple stance label represents comments in the opposite stance. There's a text entry field under the stance labels where people could write their own comments. For each article, we selected eight comments on both sides respectively. We split comments on each side into two groups and each group has four comments. On each side, the first group of the comments will be presented when people click the corresponding stance label to read comments. The second group of comments will be used as recommended

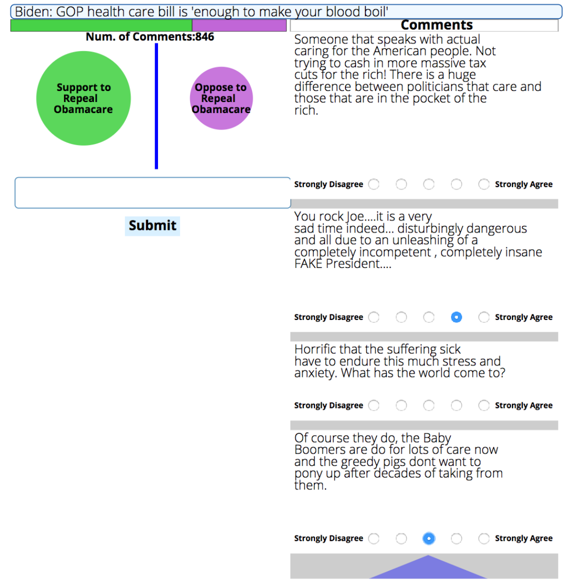


Figure 2: An example of size changeable stance labels.

comments when the system makes recommendations based on people's stances.

Visual Hints (DP-1, DP-2)

Our system provides various visual hints to help people realize their opinion selection bias and stances.

The interface highlights the visual difference in terms of the stance label size (*DP-1*). Figure 2 shows an example that after a user clicked a stance label to read comments on that side, the stance label became smaller. This could improve the user's awareness of their social opinion selection preference. Under each comment, the interface allows the user to rate the degree of (dis)agreement on the comment in a 5 Point Likert scale ranging from strongly disagree to strongly agree. If necessary, these ratings could be used to evaluate the user's stance after submitted by the user.

Figure 3 shows that after the user wrote and submitted a comment, our system predicted the user's stance on the Obamacare repeal issue. In case of inaccurate classification, our system gives the user a chance to correct the predicted stance label and confirm. Previous research [38] shows that most people are willing to correct the prediction mistakes caused by auto-systems.

Figure 4 shows an example of the function of stance record panel. After submitting the ratings to comments which were selected to read and confirming the stance for the first article, the user moved on to next article. The stance record panel showed the trace of the user's stances on previous article(s) to improve the self-awareness of his/her own stance (*DP-2*). The dashed line in the middle represents the neutral stance. The small round purple label below the dashed line is a user stance indicator showing the user's confirmed stance (oppose the repeal). If the user confirmed the stance as "support to repeal

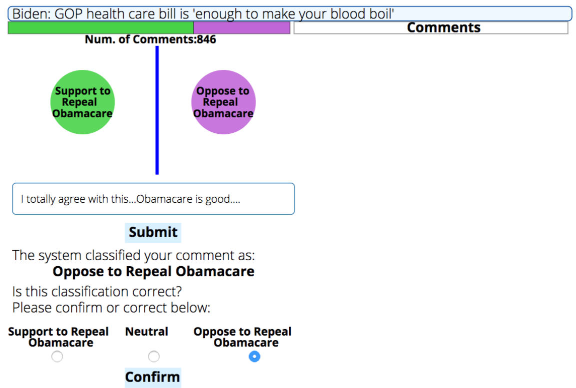


Figure 3: An example of our new interface after people write and submit a comment.

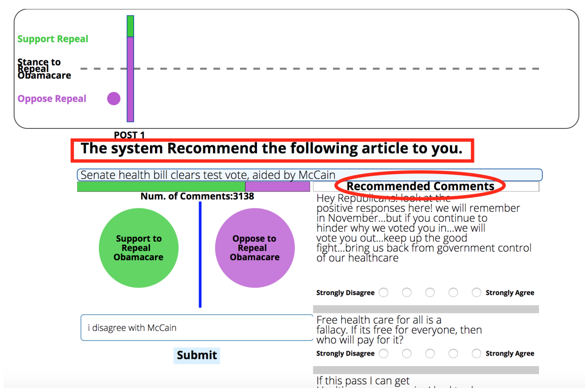


Figure 4: An example of our new interface after people move on to the next article and the system conducts *Type 1* article recommendation and *Type 1* comment recommendation.

Obamacare", there would be a green user stance indicator above the dashed line. For neutral stance confirmation, a grey user stance indicator would appear on the dashed line. There is a bar colored with green and purple on the right of the user stance indicator. The length of the green part indicates the mean of ratings given to comments which support to repeal Obamacare while the length the purple part indicates the mean of ratings given to comments on the other side. The ratio between the lengths of these two parts indicates the ratio between the average ratings for comments on each side. In the example shown in Figure 4, the purple part exceeds the dashed line, which means the user rated comments which oppose to repeal Obamacare higher than comments on the other side.

Social Opinion Recommendation (DP-3,DP-4)

Our system recommends diverse social opinions to mitigate selective exposure. Before recommending articles or comments, the system needs to detect people's stances. If a user confirmed a non-neutral stance expressed in his/her own comment, the user's stance would be the same as the confirmed one. If the user didn't confirm or confirmed a neutral stance

Type	Recommendation Mechanism
Type 1	Recommend attitude-inconsistent opinions with a positive indicator (*DP-3*)
Type 2	Recommend attitude-consistent opinions with a negative indicator (*DP-4*)

Table 3: Two types of recommendations.

and the mean of his/her ratings given to each side of comments are different, the user's stance would follow the stance of the side with higher average rating. Furthermore, the user's stance would be neutral if he/she confirmed neutral stance and the average rating for comments on each side is the same. If the user confirmed a neutral stance and only rated comments on one side, the user's stance label will be chosen based on the mean of ratings given to the comments on that side. If the mean value was higher than three, the user's stance would be the same as the stance of this side. Otherwise, if the mean value was lower than three, the user's stance would be opposite to the stance of this side. If it's three, the user's stance would be considered as neutral.

In order to recommend diverse opinions to users, our system provides two types of recommendations (*Type 1* vs. *Type 2*; See Table 3) at both the article and comment level. For both article and comment level recommendations, *Type 1* recommendation has higher priority than *Type 2* recommendation since attitude-inconsistent opinions could make people be conscious of the existence of opinions with different stances, which could mitigate selective exposure. In our system, as long as there are attitude-inconsistent opinions available, our system will conduct *Type 1* recommendation. *Type 2* recommendation will be conducted after all attitude-inconsistent opinions (e.g. articles and comments) have been recommended.

Figure 4 shows an example of *Type 1* article recommendation. Based on the user's ratings to comments and stance confirmation, our system determined his/her stance as opposing to repeal Obamacare. Since there were articles with an opposite stance available in our article database, the system conducted *Type 1* recommendation by recommending an article expressing an opposite stance with a positive indicator "The system Recommend the following article to you" (See the red rectangle in Figure 4). However, if at this moment, there were only articles with the same stance available in the system, the system would conduct *Type 2* recommendation by recommending the same-stance article with a negative indicator "The system Doesn't Recommend the following article to you".

Comment recommendation is conducted in the same way as the article recommendation. Figure 4 also shows an example of *Type 1* comment recommendation. The user confirmed the stance as opposing to repeal Obamacare. The system recommended four comments which support to repeal Obamacare with a positive indicator "Recommended Comments" (See the red circle in Figure 4). After the user clicked the "readmore" button (the upside-down blue triangle at the bottom of the comment list), the system conducted *Type 2* recommendation by presenting four agreeable comments with a negative indicator "Not Recommended Comments" to the user. Last but not least, if the system determines the user's stance as neutral, there's no recommendation for articles and comments. The system will randomly select articles and comments for the user to read.

USER STUDY

We conducted a user study to evaluate how our system could improve users' awareness of their stances and selection preferences of social opinions. In addition, we examined how our interface could mitigate selective exposure effect and encourage users to read more diverse opinions.

Participant

In this study, 12 subjects (age range from 18 to 64, 6 females) from the Midwest of the U.S. were recruited via email. They all reported that they have considerable knowledge about the Obamacare issues. When they were asked about their stances on the Obamacare repeal issue prior to the user study, four of them reported neutral and eight of them reported opposing to repeal Obamacare.

Experimental Design and Task

As shown in Figure 5, the control interface shows 16 comments in a linear list format for each article. These comments are shown in a randomized order. To compare our new interface (system) with the control interface, we designed a within-subject study where the interface is the within-subject factor. To avoid potential carry-over effect caused by the order of exposure to specific interface and group of articles, we first split eight selected news articles into two groups where each group had two CNN news articles and two FoxNews articles. Then, we changed the order of the interfaces and the article groups.

In the beginning of the user study, participants took a background survey about their demographic information and their experience in reading controversial topics. Then, participants did the following tasks for each interface: 1) Participants watched a tutorial video about how to use the interface; 2) Participants were asked to read each article, rate comments they were interested in and to write a comment to the article; 3) After finishing the task, participants worked on a two-part in-study survey. The first part asked about how the interface could help them realize their stances and comment selection preferences, and whether the interface could help them discover diverse opinions easily. The second part asked about usability of the interface, which focused on the following measures: 1) *Usefulness*: "I found this interface to be useful for reading controversial issues"; 2) *Ease of use*: "This interface is easy to use when reading controversial topics"; 3) *Enjoyable*:"I found this interface enjoyable to use"; 4) *Effective*: "The interface is effective in helping me complete the tasks"; 5) *Overall Satisfaction*: "Overall, I am satisfied with this system." All the questions in the in-study survey were measured on a 5 Point Likert scale. Then, after finishing the tasks and in-study surveys for both interfaces, participants were asked to finish a post-study questionnaire about the usefulness of each design feature in our system and their perception of how the system recommends articles to them. Finally, we conducted a short interview. The user study lasted about 1 hour and 20 minutes with a payment of $12.

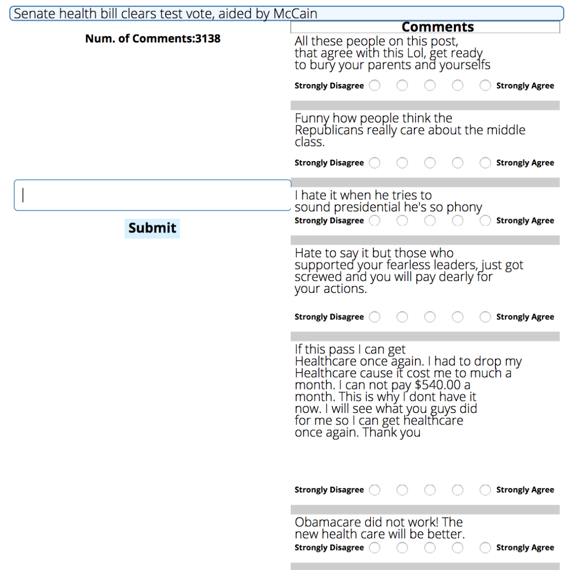


Figure 5: An overview of the control interface.

RESULTS

Improving Awareness of Diverse Opinions (RQ1,RQ2)

Figure 6 shows the comparison between two interfaces for the first part of the in-study survey. We performed Wilcoxon Signed Ranks Test to each measure and found that our new interface is significantly more helpful than the control interface in terms of helping people realize their stances ($Z = -2.859$, $p < 0.01$) and their selection preferences in comments ($Z = -1.982$, $p < 0.05$). Furthermore, results indicated that our new interface is significantly more helpful for users to discover diverse opinions than the control interface ($Z = -2.214$, $p < 0.05$).

Since participants' interface usage behaviors (e.g. which comments they rated, what rating they gave to each comment, participants' stances) were recorded, we conducted analysis

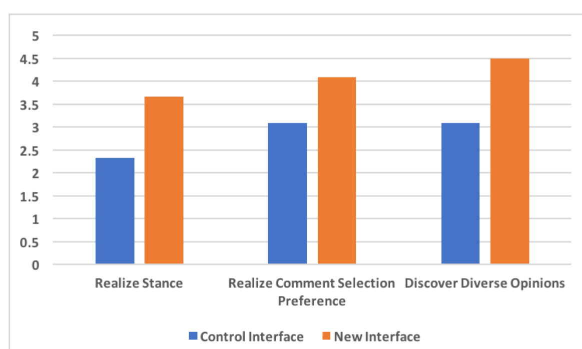


Figure 6: Comparison of average ratings between two interfaces in terms of whether the interface could help users realize their stance, comment selection preference and discover diverse opinions.

	New Interface	Control Interface
Number of comments rated (n = 48)	12.33 ($\pm$ 3.49)	6.50 ($\pm$ 2.61)
SEI (n = 48)	0.49 ($\pm$ 0.08)	0.57 ($\pm$ 0.21)

Table 4: For both interfaces, we calculated the mean and the standard deviation of the numbers of comments rated and SEIs.

on their social opinion browsing behaviors to evaluate whether our system could mitigate selective exposure. During the user study, each participant was asked to read four different articles on each interface. The system recorded which comment(s) they selected to rate with corresponding ratings and their stances to Obamacare repeal issue after reading each article. Since we asked participants to only select and rate comments they were interested in, we could evaluate how they were selectively exposed to attitude-consistent opinions by comparing the number of comments they selected on different sides. To measure selective exposure effect, we defined Selective Exposure Index (SEI) for an article as:

$$SEI = \frac{N_{consistent}}{N_{consistent} + N_{inconsistent}} \tag{4}$$

where $N_{consistent}/N_{inconsistent}$ is the number of selected comments which are consistent/inconsistent with participants' stances respectively. An SEI higher than 0.50 means the participant tends to be selectively exposed to agreeable social opinions.

We counted the total number of comments participants selected for each article and calculated the SEI for each article in both interfaces. Since every participant read four articles using each interface, in total, for each interface, we have 48 data entries collected from 12 participants for each measure. The results of comparison between these two interfaces for each measure are shown in Table 4. A two-tailed paired-sample t-test was conducted on each of these measures. The result showed that participants read more comments on our new interface than the control interface and the difference is significant ($t(47) = 12.717$, $p < 0.001$). In addition, the SEI of our new interface is significantly lower than that of the control interface ($t(47) = -2.280$, $p < 0.05$), which suggests that our new interface motivates participants to read more attitude-inconsistent comments and helps to mitigate selective exposure.

To further investigate the efficacy of our interface in mitigating selective exposure on people with different reading behaviors, for each participant, we calculated the average SEI of all four articles he/she read on the control interface and separated participants into High SEI group (with 9 participants) and Low SEI group (with 3 participants). High SEI group participants have an average SEI higher than 0.50 and Low SEI group participants have an average SEI lower than 0.50. We evaluated the effect of our new interface on different groups of participants and Table 5 shows the mean and standard deviation of SEIs on different groups and interfaces. A two-tailed paired-sample t-test showed that there is a significant difference between two interfaces ($t(35) = -2.621$, $p < 0.05$) for High

	New Interface	Control Interface
High SEI group (n = 36)	0.50 ($\pm$ 0.08)	0.61 ($\pm$ 0.23)
Low SEI group (n = 12)	0.47 ($\pm$ 0.05)	0.45 ($\pm$ 0.12)

Table 5: For both High and Low SEI groups, we calculated the mean and the standard deviation of SEIs for both interfaces.

SEI group. Meanwhile, no significant difference is found for Low SEI group ($t(11) = 0.457$, $p = 0.657$). This result indicated that High SEI group participants have more balanced comment selection behaviors when using our interface. However, there's no significant difference in comment selection behaviors for Low SEI group participants when using different interfaces.

These results show that our new interface (system) improves people's awareness of their stances and social opinion selection preferences (RQ-1), and mitigates selective exposure (RQ-2).

Usability Improvement

Figure 7 shows the comparison between two interfaces for the second part of the in-study survey. We conducted Wilcoxon Signed Ranks Test on each measure of usability (Table 6) and found that for the measures of *EaseofUse*, *Enjoyable*, *Effective* and *Overall Satisfaction*, our new interface are rated better than the control interface with a significant difference. However, we only found marginal significant difference in the *Usefulness* measure when comparing our new interface with the control interface.

Evaluation of System Features

In the post-study questionnaire, we asked participants how system features (e.g. size-changeable stance labels, stance recorder panel, *Type 1*/*Type 2* recommendations of articles/comments) help them realize their stances and comment selection preferences.

Figures 8a and 8b show participants' evaluation about these system features. Based on the result, we found that the stance

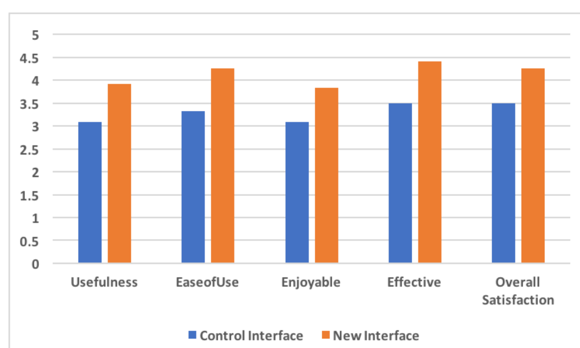


Figure 7: Comparison of average ratings between two interfaces in terms of usability measures: Usefulness, EaseofUse, Enjoyable, Effective and Overall Satisfaction. Higher rating means better interface efficacy.

Measures	New Interface vs. Control Interface
Usefulness	Z = -1.831; p = 0.067
EaseofUse	Z = -2.598; p <0.01
Enjoyable	Z = -2.460; p <0.05
Effective	Z = -2.428; p <0.05
Overall Satisfaction	Z = -2.165; p <0.05

Table 6: Wilcoxon Signed Ranks Test on Usefulness, EaseofUse, Enjoyable, Effective and Overall Satisfaction.

	SEI (Mean $\pm$ SD)
Control Interface	0.57 ($\pm$ 0.21)
Stance Label Comments	0.50 ($\pm$ 0.14)
Recommended Comments	0.48 ($\pm$ 0.11)

Table 7: We calculated the mean and the standard deviation of SEIs in each comment browsing scenario.

recorder panel is the most important feature to help participants realize both of their stances and opinion selection preferences with an agreement rate of 9/12 and 11/12 respectively. In addition, all other features have an agreement rate equal to or higher than 50%, which means participants found the proposed features useful in general.

Unlike the control interface where all comments are shown together at once, there are two types of comments in our interface: stance label comments (comments which are presented after people click the corresponding stance label) and recommended comments (comments which are recommended by both *Type 1* and *Type 2* recommendations). We compared participants' comment selection behaviors in the measure of SEI among different comment browsing scenarios: 1) browsing comments on control interface; 2) browsing stance label comments on our interface and 3) browsing recommended comments on our interface. Table 7 shows the mean and standard deviation of SEIs in each scenario. A repeated measures ANOVA (with Greenhouse-Geisser correction for sphericity) found that the SEI differs significantly between scenarios ($F(1.708, 80.280) = 4.176$, $p < 0.05$). Post-hoc tests using the Bonferroni correction indicated that there's a significant difference between scenario 1) and 3) ($p < 0.05$) but no significant difference between 2) vs. 3) ($p = 0.968$) and 1) vs. 2) ($p = 0.242$). This result indicates that recommendation feature mitigates selective exposure more effectively than the size-changeable stance label feature when comparing with the control interface.

Qualitative Analysis

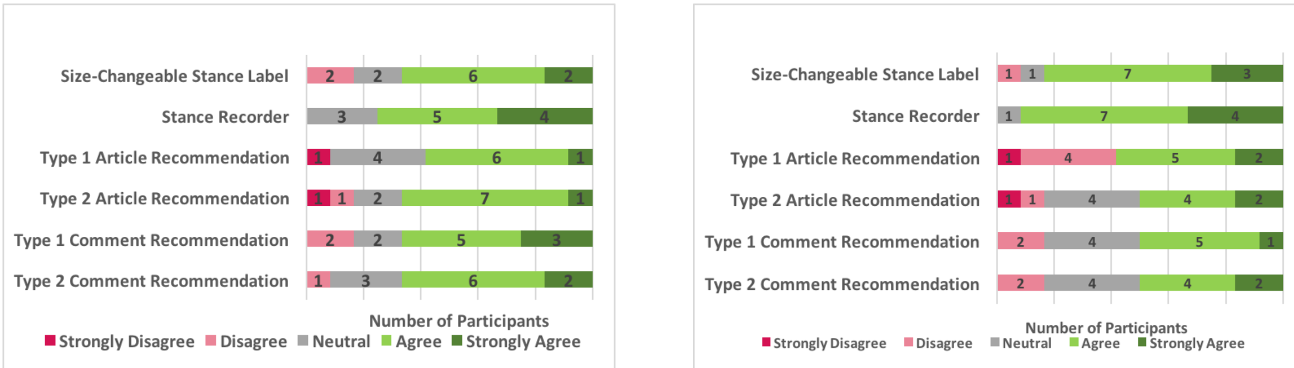
We asked participants about their perception of how the system recommends articles to them. Specifically, we asked 1) whether they could figure out why sometimes the system showed a "Recommended" article but sometimes showed a "Not Recommended" article, 2) whether they thought the recommendation was related to their stances, and 3) how the system recommended articles to them. As a result, 10 out of 12 participants agreed to the first question, saying that they could figure out how the system made recommendations. For the second question, 8 out of 12 participants agreed, which illustrates that most of the users thought the recommendations were made based on their stances. For the third question, 6 out of 12 participants correctly described how the system made recommendations. For example, participant *U7* reported "I think that the recommended articles are trying to get you to see different perspectives on the issue" and participant *U12* reported "The counter-recommendation mechanism is there to provide an opportunity to 'burst your bubble' so the reader is exposed to opposing points of view...This whole mechanism that was implemented is interesting considering the buzz around the whole notion of being trapped in a bubble of one's own opinions: as in everyone they talk to or read from agrees with them and only reinforces their beliefs." These results show that our recommendation mechanism nudged participants to figure out the connection between the indicator and stance expressed in recommended opinions, which improved self-awareness of their own stances. However, the other 6 participants couldn't accurately describe the recommendation mechanism, which suggests that people's perception of how the system made recommendations may vary. In addition, all participants showed an overall preference to our system and 11 out of 12 participants agreed our system helped them to get global insight across different opinions during our short interview.

DISCUSSION

Past research has pointed out the need to design informational structures or personalization algorithms for the vast amount of user-generated contents on online social media platforms. These informational structures or algorithms often focus on helping users find and select information they desire. There is, however, a general lack of attention to how these structures may lead to undesirable effects, such as selective exposure to information and social polarization. The design goal of our system is different from these approaches: Instead of *only* helping users find information they desire, our system *also* aims to help user correctly perceive and understand different perspectives expressed by other users. In this paper, we showed how the general techniques of sentiment analysis and opinion recommendation can be repurposed to satisfy this design goal.

Analyzing People's Sentiments for Controversial Topics

Sentiment analysis technique is typically used to find emotional reactions to a topic. However, we utilized the technique to infer people's stances, whether they support or oppose a controversial topic. Specifically, in our study, we wanted to infer people's stances on Obamacare issue using a classifier that predicts sentiment labels of people's comments about Obamacare issue. We conducted feature engineering with the χ^2 statistic test on each unique term. We found some terms, such as "president", "Trump" and "McCain" were assigned a very high weight. These terms are highly related to the topic of the issue and people's stereotypes of political figures related to the issue. However, they are not the common terms people use to express sentiments. This indicates that when analyzing people's sentiments on controversial issues, we should consider not only their emotional expressions but also their stereotypes of the topic and people related to the issue.



(a) (b)

Figure 8: (a) Evaluation for specific system features in helping participants realize their stances. (b) Evaluation for specific system features in helping participants realize their comment selection preferences.

Social Opinion Recommendation

Traditional personalized recommendation systems select and recommend agreeable information or opinions to people based on their records of information selection. This mechanism aims to increase people's satisfaction with the system but it often separates people from attitude-inconsistent information and may further aggravate the selective exposure effect. We designed our recommendation mechanism differently. Although our system recommends opinions on both sides, the system recommends attitude-consistent opinions (*Type 2* recommendation) after the attitude-inconsistent opinions (*Type 1* recommendation). We found that our new recommendation mechanism could mitigate selective exposure. We should point out that, in more realistic situations where there are a large amount of opinions on each side, the system may need to use an appropriate ratio between *Type 1* recommendation and *Type 2* recommendation. We believe that the question of how to set this ratio will be an interesting research topic in the future.

Limitations

In our study, we tested user behavior using only four articles and a limited number of comments on Obamacare topic. In the future, we plan to investigate people's social opinion browsing behavior with more articles and comments on other controversial topics over a longer period of time in a larger-scale user study. This would generate a richer set of data to achieve more comprehensive evaluation and inform better design of such interfaces.

In addition, we only considered positive and negative sentiments when designing and implementing the system. However, people may have other types of emotions when browsing social opinions. Also, people may get a better understanding of comments by checking scores or ratings about sentiments. We plan to improve the sentiment classifier to support multi-category emotion classification as well as generating numeric sentiment scores in the future.

Lastly, although users could be motivated to inspect both stances at the same time, we decided to present opposite viewpoints asynchronously for the sake of testing our design ideas more effectively and conveniently. However, we will consider this alternative design in our future studies.

CONCLUSION

In this paper, we presented an interactive intelligent system that organizes social opinions using a combination of novel visual cues and recommendation mechanisms to increase self-awareness of users' stances and to mitigate effects of selective exposure. Compared to a control interface in which opinions are organized in a linear format, the new system was found to help users become more aware of their own stances and their selection preferences of opinions. We also found that users attended to more diverse opinions and showed less selective exposure to attitude-consistent information in the new interface. We concluded that the proposed new system is promising in raising users' awareness of their own stances and preferences, as well as mitigating selective exposure to information.

ACKNOWLEDGMENTS

This research was supported by the NSF SES grant #1419297 and the NSF IIS grant #1441563. We would like to thank Yang Du for her help to this project and all the reviewers for their valuable feedback.

REFERENCES

1. 2017a. Bernie Sanders: Every state should have a public option. CNN Facebook. (30 July 2017).
2. 2017b. Bernie Sanders to headline 'Don't Take Our Health Care' tour. CNN Facebook. (24 June 2017).
3. 2017c. Biden: GOP health care bill is 'enough to make your blood boil'. CNN Facebook. (18 July 2017).
4. 2017a. He Must Be 'Clairvoyant': Huckabee Blasts Schumer for Immediate Rebuttal of Health Care Bill. FoxNews Facebook. (25 June 2017).
5. 2017d. Obama on Senate bill: It's 'not a health care bill'. CNN Facebook. (23 June 2017).

6. 2017b. ObamaCare replacement bill approved in House. FoxNews Facebook. (4 May 2017).
7. 2017c. Pro-Trump group launches $1M ad campaign against Sen. Heller over ObamaCare stance. FoxNews Facebook. (23 June 2017).
8. 2017d. Senate health bill clears test vote, aided by McCain. FoxNews Facebook. (25 July 2017).
9. Gediminas Adomavicius and Alexander Tuzhilin. 2005. 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* 17, 6 (2005), 734–749.
10. Isabelle Augenstein, Tim Rocktäschel, Andreas Vlachos, and Kalina Bontcheva. 2016. Stance Detection with Bidirectional Conditional Encoding. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Austin, Texas, 876–885. https://aclweb.org/anthology/D16-1084
11. Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.
12. Irene V Blair, John F Steiner, Diane L Fairclough, Rebecca Hanratty, David W Price, Holen K Hirsh, Leslie A Wright, Michael Bronsert, Elhum Karimkhani, David J Magid, and others. 2013. Clinicians' implicit ethnic/racial bias and perceptions of care among black and Latino patients. *The Annals of Family Medicine* 11, 1 (2013), 43–52.
13. Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of computational science* 2, 1 (2011), 1–8.
14. Andrea Ceron, Luigi Curini, Stefano M Iacus, and Giuseppe Porro. 2014. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society* 16, 2 (2014), 340–358.
15. Dan Cosley, Shyong K Lam, Istvan Albert, Joseph A Konstan, and John Riedl. 2003. Is seeing believing?: how recommender system interfaces affect users' opinions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 585–592.
16. Robyn M Dawes. 1989. Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology* 25, 1 (1989), 1–17.
17. Robyn M Dawes and Matthew Mulford. 1996. The false consensus effect and overconfidence: Flaws in judgment or flaws in how we study judgment? *Organizational Behavior and Human Decision Processes* 65, 3 (1996), 201–211.
18. Hyo Jin Do and Ho-Jin Choi. 2015. Korean Twitter Emotion Classification Using Automatically Built Emotion Lexicons and Fine-Grained Features.. In *PACLIC*.
19. Hyo Jin Do, Chae-Gyun Lim, You Jin Kim, and Ho-Jin Choi. 2016. Analyzing emotions in twitter during a crisis: A case study of the 2015 middle east respiratory syndrome outbreak in Korea. In *Big Data and Smart Computing (BigComp), 2016 International Conference on*. IEEE, 415–418.
20. Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1175–1184.
21. Leon Festinger. 1962. *A theory of cognitive dissonance*. Vol. 2. Stanford university press.
22. Peter Fischer, Andreas Kastenmüller, Tobias Greitemeyer, Julia Fischer, Dieter Frey, and David Crelley. 2011. Threat and selective exposure: the moderating role of threat and decision context on confirmatory information search after decisions. *Journal of Experimental Psychology: General* 140, 1 (2011), 51.
23. Dieter Frey. 1986. Recent research on selective exposure to information. *Advances in experimental social psychology* 19 (1986), 41–80.
24. Michel Galley, Kathleen McKeown, Eric Fosler-Lussier, and Hongyan Jing. 2003. Discourse segmentation of multi-party conversation. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*. Association for Computational Linguistics, 562–569.
25. Mingkun Gao, Hyo Jin Do, and Wai-Tat Fu. 2017. An Intelligent Interface for Organizing Online Opinions on Controversial Topics. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. ACM, 119–123.
26. Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2016. Quantifying Controversy in Social Media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (WSDM '16)*. ACM, New York, NY, USA, 33–42. DOI: http://dx.doi.org/10.1145/2835776.2835792
27. Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2017. Reducing Controversy by Connecting Opposing Views. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (WSDM '17)*. ACM, New York, NY, USA, 81–90. DOI: http://dx.doi.org/10.1145/3018661.3018703
28. Jeffrey Gottfried and Elisa Shearer. 2016. News Use Across Social Media Platforms 2016. (26 May 2016).
29. Dietmar Gräbner, Markus Zanker, Günther Fliedl, and Matthias Fuchs. 2012. *Classification of Customer Reviews based on Sentiment Analysis*. Springer Vienna, Vienna, 460–470. DOI: http://dx.doi.org/10.1007/978-3-7091-1142-0_40

30. Anthony G Greenwald and Mahzarin R Banaji. 1995. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review* 102, 1 (1995), 4.
31. Catherine Grevet, Loren G Terveen, and Eric Gilbert. 2014. Managing political differences in social media. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1400–1408.
32. Kazi Saidul Hasan and Vincent Ng. 2014. Why are You Taking this Stance? Identifying and Classifying Reasons in Ideological Debates.. In *EMNLP*, Vol. 14. 751–762.
33. Tin Kam Ho. 1995. Random decision forests. In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, Vol. 1. IEEE, 278–282.
34. Tin Kam Ho. 1998. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence* 20, 8 (1998), 832–844.
35. Enamul Hoque and Giuseppe Carenini. 2014. Convis: A visual text analytic system for exploring blog conversations. In *Computer Graphics Forum*, Vol. 33. Wiley Online Library, 221–230.
36. Enamul Hoque and Giuseppe Carenini. 2015. Convisit: Interactive topic modeling for exploring asynchronous online conversations. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, 169–180.
37. Enamul Hoque and Giuseppe Carenini. 2016. Multiconvis: A visual text analytics system for exploring a collection of online conversations. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, 96–107.
38. Shih-Wen Huang, Pei-Fen Tu, Wai-Tat Fu, and Mohammad Amanzadeh. 2013. Leveraging the crowd to improve feature-sentiment analysis of user reviews. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 3–14.
39. Shafiq Joty, Giuseppe Carenini, and Raymond T Ng. 2013. Topic segmentation and labeling in asynchronous conversations. *Journal of Artificial Intelligence Research* 47 (2013), 521–573.
40. Joachim Krueger and Russell W Clement. 1994. The truly false consensus effect: an ineradicable and egocentric bias in social perception. *Journal of personality and social psychology* 67, 4 (1994), 596.
41. Audrey J Lee. 2005. Unconscious bias theory in employment discrimination litigation. *Harv. CR-CLL Rev.* 40 (2005), 481.
42. Weiyuan Li and Hua Xu. 2014. Text-based emotion classification using emotion cause extraction. *Expert Systems with Applications* 41, 4 (2014), 1742–1749.
43. Q Vera Liao and Wai-Tat Fu. 2014. Expert voices in echo chambers: effects of source expertise indicators on exposure to diverse opinions. In *Proceedings of the 32nd annual ACM conference on human factors in computing systems*. ACM, 2745–2754.
44. Q Vera Liao, Wai-Tat Fu, and Sri Shilpa Mamidi. 2015. It is all about perspective: An exploration of mitigating selective exposure with aspect indicators. In *Proceedings of the 33rd annual ACM conference on Human factors in computing systems*. ACM, 1439–1448.
45. Charles G Lord, Lee Ross, and Mark R Lepper. 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology* 37, 11 (1979), 2098.
46. John D Marvel. 2015. Unconscious bias in citizens' evaluations of public sector performance. *Journal of Public Administration Research and Theory* 26, 1 (2015), 143–158.
47. Horace McCormick. 2015. The real effects of unconscious bias in the workplace. *UNC Executive Development, Kenan-Flagler Business School. DIRECCIÓN* (2015).
48. Sean A Munson, Stephanie Y Lee, and Paul Resnick. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget.. In *ICWSM*.
49. Sean A Munson and Paul Resnick. 2010. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1457–1466.
50. Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.
51. Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 443–452.
52. Lee Ross, David Greene, and Pamela House. 1977. The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of experimental social psychology* 13, 3 (1977), 279–301.
53. Cass R Sunstein. 2009. *Republic. com 2.0*. Princeton University Press.
54. Songbo Tan and Jin Zhang. 2008. An empirical study of sentiment analysis for chinese documents. *Expert Systems with applications* 34, 4 (2008), 2622–2629.
55. Ryoko Tokuhisa, Kentaro Inui, and Yuji Matsumoto. 2008. Emotion classification using massive examples extracted from the web. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*. Association for Computational Linguistics, 881–888.
56. Yiming Yang and Jan O Pedersen. 1997. A comparative study on feature selection in text categorization. In *Icml*, Vol. 97. 412–420.

57. John Zaller. 1992. *The nature and origins of mass opinion*. Cambridge university press.

58. Amy X Zhang, Lea Verou, and David R Karger. 2017. Wikum: Bridging Discussion Forums and Wikis Using Recursive Summarization.. In *CSCW*. 2082–2096.

59. Ding Zhou, Sergey A Orshanskiy, Hongyuan Zha, and C Lee Giles. 2007. Co-ranking authors and documents in a heterogeneous network. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*. IEEE, 739–744.