

To Label or Not to Label: The Effect of Stance and Credibility Labels on Readers' Selection and Perception of News Articles

MINGKUN GAO*, ZIANG XIAO*, KARRIE KARAHALIOS, and WAI-TAT FU,
University of Illinois at Urbana-Champaign, USA

Social media sites use different labels to help users find and select news feeds. For example, Blue Feed, Red Feed, a news feed created by the Wall Street Journal, use stance labels to separate news articles with opposing political ideologies to help people explore diverse opinions. To combat the spread of fake news, Facebook has experimented with putting credibility labels on news articles to help readers decide whether the content is trustworthy. To systematically understand the effects of stance and credibility labels on online news selection and consumption, we conducted a controlled experiment to study how these labels influence the selection, perceived extremeness, and level of agreement of news articles. Results show that stance labels may intensify selective exposure - a tendency for people to look for agreeable opinions - and make people more vulnerable to polarized opinions and fake news. We found, however, that the effect of credibility labels on reducing selective exposure and recognizing fake news is limited. Although originally designed to encourage exposure to opposite viewpoints, stance labels can make fake news articles look more trustworthy, and they may lower people's perception of the extremeness of fake news articles. Our results have important implications on the subtle effects of stance and credibility labels on online news consumption.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**;

Additional Key Words and Phrases: Selective Exposure; Social Opinion Perception; Stance Label; Credibility Label

ACM Reference Format:

Mingkun Gao*, Ziang Xiao*, Karrie Karahalios, and Wai-Tat Fu. 2018. To Label or Not to Label: The Effect of Stance and Credibility Labels on Readers' Selection and Perception of News Articles. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW, Article 55 (November 2018), 16 pages. <https://doi.org/10.1145/3274324>

1 INTRODUCTION

As an example of helping people exploring diverse social opinions, the Wall Street Journal proposed a novel feed design "Blue Feed, Red Feed"¹ by presenting liberal and conservative labelled news articles in side-by-side columns. The purpose of such design is to promote awareness of news articles with different standpoints to readers. Although recent research has shown that stance labels will likely promote awareness of opposing arguments and thereby reduce selective exposure to information [17], it is not clear to what extent they can truly reduce attitude polarization and lead to productive cross-ideological dialogue. For example, theories of motivated reasoning [27]

¹<http://graphics.wsj.com/blue-feed-red-feed/>

*The two first authors contributed equally to this work.

Authors' address: Mingkun Gao*, mgao19@illinois.edu; Ziang Xiao*, zxiao5@illinois.edu; Karrie Karahalios, kkarahal@illinois.edu; Wai-Tat Fu, wfu@illinois.edu,

University of Illinois at Urbana-Champaign, 201 N. Goodwin Avenue, Urbana, IL, 61801, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-0142/2018/11-ART55 \$15.00

<https://doi.org/10.1145/3274324>

suggest that when presented with attitude-inconsistent arguments or opinions, people may engage in counter-arguing and will likely dismiss or downplay them as a way to reduce the psychological stress caused by cognitive dissonance [15]. As a result, people may become even more entrenched in their original beliefs and polarized than before [10]. This kind of "anticonformity boomerang effect" [22] may affect people's perception of these labelled news articles and lead to undesirable news browsing behavior. For example, when a person who is self-identified as supporting gun control encounters an anti-gun control article on his/her news feed, he/she will expect certain degree of cognitive conflict if he/she reads the news article. The presence of the "conservative" label on the article may trigger this person's previous stressful memory of reading other conservative news articles, which in turn aggravates his/her expectation of the level of cognitive conflict. To avoid such unpleasant experience, this person may simply discontinue reading, which incubates potential selective bias.

For the sake of preventing the dissemination of fake news, Facebook² proposed a user-self-report mechanism that allows users to label news articles with a disputed flag to indicate potential misleading information in a news article. However, after a year of testing, a Medium article [46] found that the dispute flag label didn't work effectively to prevent fake news from spreading. One reason is that the strong wording or visuals of the label may backfire and consequently reinforce users' belief in fake information [28]. People who wanted to believe these false stories were found to share them even more when they were labelled, according to a TechCrunch article [9]. What's more, when only some news articles are flagged, those that are not flagged (but contain false stories) are often presumed to be trustworthy. These unintended consequences of labels prompted the use and testing of new strategies to combat the spread of misinformation.

Given the aforementioned efforts by Wall Street Journal and Facebook in helping people explore diverse opinions and combating the spread of fake news, understanding the effect of these stance and credibility labels is crucial to learning how designs of news feed platforms could be further improved. However, to the best of our knowledge, questions about the effect of stance labels and credibility labels have not been well examined in the social opinion, selective exposure and fake news literature. Since people's social opinion selection and perception (i.e. perceived level of agreement and extremeness on social opinions) may be relevant to the extent of their opinion polarization [38, 48] and their judgment on public opinions [8], we are interested in how stance labels and credibility labels impact people's social opinion selection and perception. Thus, we ask the following research questions:

- RQ-1: How stance labels and credibility labels impact people's news article selection when browsing news for controversial topics?
- RQ-2: How stance labels and credibility labels impact people's perception of news articles in terms of perceived extremeness and level of agreement?

Even though labels are commonly used in social media design to categorize contents and facilitate people finding information, little has been known about the effect of labels beyond facilitating information seeking. Previous research [27] suggests people are driven by two motives when seeking information: the defense motive for directional goals and the accuracy motive for accuracy goals. People tend to seek more supportive information for a desired and directional goal while they would expend more cognitive effort and seek for objective information for an accurate conclusion. In terms of the effect of labels, stance labels may trigger people's defense motive and push people away from opinions in the opposing side. Meanwhile, credibility labels may ignite people's desire for the accurate and objective truth and pull people to more opposing views. We are motivated

²<https://www.facebook.com/>

by the dual motives theory in information seeking and conducted an experiment to figure out the effect of these labels on social media.

Our analysis suggests that stance labels may exacerbate selective exposure while increase people's level of agreement and decrease their perceived extremeness on fake news articles from different sides. Thus, stance labels may have some negative effects on people's social opinion consumption and mislead people to give incorrect judgment on the credibility of fake news. Unlike stance labels, credibility labels may have more positive effects in our experiment even though those effects might be limited. We found that credibility labels could mitigate selective exposure, especially for people with liberal stances, and marginally decrease people's level of agreement to news articles on their own side, which, as a result, could have the potential to weaken social polarization. Our results modestly suggest that labels may have complex effects on people's news article browsing behavior and their perception of social opinions, and social media designers should be cautious about these complex effects when designing novel news feed.

2 RELATED WORK

2.1 Cognitive Dissonance

Social media users are exposed to social opinions (e.g. news articles and comments) from different stances. Some opinions may be aligned with users' own stances but others may be not. Cognitive dissonance theory [15] suggested that people may experience discomfort when they encounter conflicting opinions, which could break their internal psychological consistency. For example, [11] showed that cognitive dissonance made patients feel negative toward beneficial but unpleasant medical screenings. This may negatively influence patients' medical-decision making. [49] indicated that while deciding a travel destination between a green and nongreen resort, people with strong pro-environmental attitudes would feel dissonance when making a nongreen choice. They may actively avoid those conflicting beliefs and attitudes [15] to reduce the discomfort or dissonance. Consequently, they may mentally or even physically stay in their comfort zone, losing opportunities to experience and understand the outside world. According to these cognitive dissonance theories, moderating the cognitive dissonance with the help of external factors could get people out of their comfort zone. In our study, we want to see if hiding news articles' stance labels will affect people's perception of those articles and their decisions on whether to avoid stance-inconsistent news articles.

2.2 Selective Exposure

In terms of information selection, cognitive dissonance could cause selective exposure - a phenomenon that people tend to seek for stance-consistent opinions [16]. People's self-defense mechanism to reduce the discomfort of being exposed to stance-inconsistent social opinions motivated them to select more agreeable opinions. Furthermore, Zaller [50] suggested that biased assimilation [32] commonly existed among people for the tendency that stance-consistent opinions were more acceptable than stance-inconsistent opinions. Such bias (i.e. selection bias and assimilation bias) may cause the echo chamber effect [4, 30] and create the filter bubble [40], thus pushing people's pre-existing beliefs or attitudes to a polarizing level. As a consequence, a polarized society seems inevitable [38] and there will be less room to exchange opinions from people with different stances rationally.

Researchers [14, 19, 41] have been working on mitigating selective exposure through exploring novel moderating factors and designing new interfaces. Liao and Fu [29] explored how position indicator, including pro/con indicators and the corresponding magnitude could affect opinion selection for people with different levels of learning motivations. They found that for people with

higher learning motivation, these indicators could nudge them to explore more stance-inconsistent social opinions. In another work [31], Liao and her colleagues stated that aspect indicators could alleviate selective exposure when people read drug-related comments on side effects. In line with Liao et al. [31], Gao et al. [18] and Munson et al. [37] respectively illustrated that how a novel recommendation mechanism or browser widget could encourage people to read more stance-inconsistent view points, thereby relieving selective exposure.

Blue Feed, Red Feed helped people explore diverse opinions by presenting liberal and conservative news articles in two separate columns with corresponding stance labels. We investigated the effect of stance labels in the two-column layout on people's news article selection, which may further help to expose diverse opinions to people.

2.3 Fake News

Two BuzzFeed articles [44, 45] mentioned that social media platforms are vulnerable to fake news for the lack of effective methods to distinguish between fake and true stories. One article [44] claimed that during the 2016 presidential election campaign, among most popular fake stories from hoax website and true stories from mainstream media, fake stories were shared more than true stories on Facebook. The other [45] reported that many people believed those fake news were true, which may impact people's voting, according to a Guardian report [42].

In [1], Allcott and Gentzkow intended to figure out how much impact the fake news had on the 2016 presidential election theoretically and empirically. Their research showed that fake news were widely shared through Facebook and all U.S. adults may have been exposed to one or several pieces of fake news before the election. However, people were less likely to trust information from social media in contrast to information from national or local news organizations. Thus, they cannot provide a concrete assessment about whether fake news was the deciding factor in the 2016 presidential election.

To help users recognize fake news, originally, Facebook allowed users to label disputed flags to fake stories. However, in an announcement [46] from Facebook, they admitted that the disputed flag was not so effective in preventing people from sharing fake news, possibly for the reason that the intense disputed flag with strong wordings and visualization may strengthen one's belief [28]. In our study, we considered both disputed labels and trustworthy labels, and intended to see whether and how these credibility labels would have impact on people's selection and perception of news articles.

2.4 Effect of Labeling

Researchers have been investigating the effect of labeling in people's perception of different products [5, 24]. In [24], Jeddi and Zaiem studied how the quality labels of food products could affect consumers' purchase intention. In the field of food products, people have raising uncertainties and concerns about the quality. They found that labeling the quality of food products was an effective marketing tool and had positive impacts on consumers' purchase intention. Furthermore, they pointed out as the perceived risk of the food quality went higher, the impact of quality labels went stronger for consumers' purchase intention.

Beltramini [5] studied how warning labels of varying presenting forms on cigarette package influenced young adults' perceived belief that smoking was harm to health. He found that warning labels indicating specific risky consequences of smoking and specific remedial actions may be more believable than other warning labels. In addition, he showed that people's "mushiness" levels, namely how firm a person sticks to his/her positions, also had impact on their perceived believability. In another study about warning labels on cigarette package, Bansal-Travers et al. [3] suggested

that larger, pictorial and loss-framed warning labels could raise people’s awareness of the risks of smoking on health.

Kriplean et al. [26] proposed ConsiderIt where diverse social opinions were separated into pro and con lists with corresponding labels of "Pros" and "Cons" based on users’ stances and the stance of opinions in each column. They showed that such a design could encourage people to adopt opinions contributed by others and therefore enhance the public deliberation.

Epstein and Robertson [12] showed that the Search Engine Manipulation Effect (SEME) was powerful and biased search rankings could affect undecided voters’ preference. Meanwhile, this ranking manipulation may not be realized by people if the manipulation was somewhat masked. Epstein et al. [13] conducted experiments and argued that the voting shift towards the favored candidate caused by biased search rankings could be effectively mitigated by labeling and alerting the biased rankings and news in the search result.

According to a NPR (National Public Radio) article [20], political labels may trigger people to identify their own stances and consolidate their pre-existing beliefs. To reduce the possibility that people’s self-identified beliefs or stances to be even more biased by external political labels, politicians and social activists initiated a movement called "No Labels" to create a political force between the left-wing and the right-wing, establishing a "common ground" to solve critical issues. Political labels, such as Democratic and Republican, may encourage people to identify their core values and make people less tolerant to opposing opinions, which could polarize the society.

In our study, we examined whether and how stance labels of news articles could affect social media users’ dissonance when exposed to diverse social opinions. Moreover, we also studied how stance labels and credibility labels influence people’s news article selection and their perception of those news articles in terms of extremeness and level of agreement.

3 METHOD

3.1 Interface Design

We used 2×2 full factorial experiment design to see how stance labels and credibility labels affect participants’ preferential selection and perception of news articles. Our design elicited four interfaces, see Table 1 and Figure 1³.

	No Stance Label	Stance Label
No Credibility Label	Interface A	Interface B
Credibility Label	Interface C	Interface D

Table 1. Experiment Settings

Based on established research on social media design [17, 26] and the Blue Feed, Red Feed website, our basic interface template for all four interfaces adapted a two-column layout where each column only listed news articles with similar political leanings. To control for the potential bias from people’s reading habits (e.g. people may tend to start from left), the left-and-right position of those two columns was randomly assigned to each participant. The order of news articles within each column was randomly assigned initially but kept the same for all participants. To highlight the two-column design and meanwhile to control for people’s color preference, the background colors of each column were grays with two different shades.

³On the screenshot, the top two news articles on the feed are disputed. Their credibility are not labelled due to the experiment design.

Since our study was focused on labels' effects, all interfaces were based on the same two-column template and had exact same news articles. A total of 14 news articles⁴ with clear political leanings were selected one day prior to the experiment. The number of news articles with conservative or liberal leanings was balanced, i.e. 7 news articles with conservative leanings in one column and 7 news articles with liberal leanings in the other. Since news articles were collected from different news sources, for the sake of design consistency we created our own page for each article. We also removed the source media to control the effect of users' existing knowledge of media's political leanings. More details about news article selection process will be described in the later section.

3.2 Experiment Procedure

The study procedures were approved by IRB and included three major steps,

- 1) Pre-study Survey. Upon participants consented to join our study, they first took a pre-study survey of demographics including age, gender, political leaning and social media usage.
- 2) News Browsing Task. Participants were randomly assigned to one of our four interfaces. To ensure participants actually read the news, before the news reading starts, participants were told that there would be a follow-up task about the news articles they read. Next, participants were asked to select and read at least five different news articles to their most interest. Once participants finished news reading, they were asked to write a short paragraph to express their opinions of the news articles they just read. Then, participants were asked to rate those news articles accordingly in terms of level of agreement and level of extremeness.
- 3) Post-study Survey. The post-study survey contains a questionnaire measuring participants' level of state anxiety.

After all steps, participants were thoroughly debriefed with the purpose of the study and a list of fake news articles they might read during the study. All data we collected were anonymized and stored locally. The study was published on Amazon Mechanical Turk⁵ and framed as a task to gather information from news articles. Each participant was paid \$6/hr as compensation.

3.3 Selected News Articles

Due to the scope of the study, articles on our news feed interface covered two major trending and controversial topics, President Trump and Gun Control. The number of articles about those topics appeared in our interfaces were balanced.

To mimic the mundane world where we are often exposed to both real and fake news articles, our news feed contained both real and fake news articles as well. In our study, "fake news" denotes news articles which contain false or misleading information with political implications [1, 23]. 6 out of 14 news articles on our news feed were fake news articles directly collected from Snopes.com's⁶ and Politifact.com's⁷ fake news archive or modified from other news articles. We manually changed people's name and location and exaggerated numbers appeared in the original article to ensure those articles contain false information. All the fake news articles have a clear political leaning and related to either President Trump or Gun Control. The stance labels of those news articles were assigned independently and verified by authors.

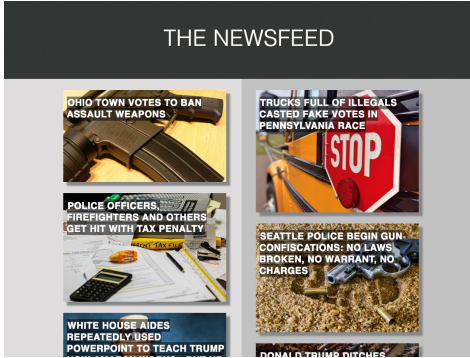
The rest eight news articles were collected from Wall Street Journal's Blue Feed, Red Feed. The stance label of each news article follows Blue Feed, Red Feed's category (liberal or conservative),

⁴News articles were retrieved from dailyworldupdate.com, libertyheadlines.com, Breitbart.com, cleveland19.com, forbes.com, bluebirddaily.com, rawstory.com, immediatesafety.org, dailykos.com, dailycaller.com, thinkprogress.org, conservativetri-bune.com and thehill.com.

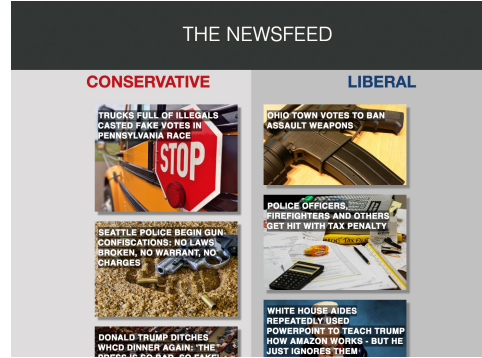
⁵<https://www.mturk.com/>

⁶<https://www.snopes.com/>

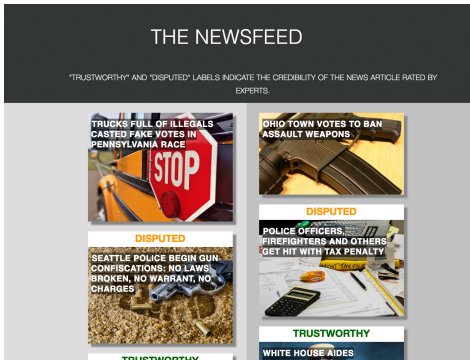
⁷<http://www.politifact.com/>



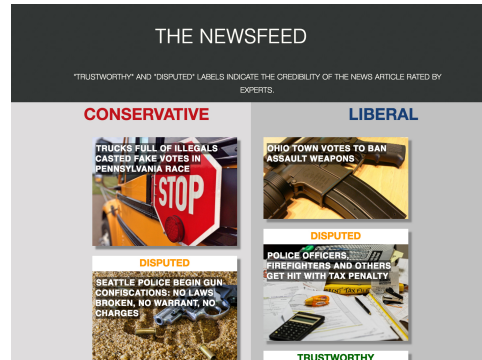
(a) Interface A, Without Stance Labels and Credibility Labels



(b) Interface B, With Stance Labels but no Credibility Labels



(c) Interface C, With Credibility Labels but no Stance Labels



(d) Interface D, With both Stance Labels and Credibility Labels

Fig. 1. Screenshots from the 4 interfaces in our 2x2 experimental design.

which is based on self-described political leanings of article-sharers on Facebook [2], and was verified independently by authors.

3.4 Credibility Label

There were two kinds of credibility labels on Interface C and Interface D, e.g. "Trustworthy" and "Disputed". To best mimic real world scenarios where even on a credibility-labelled news feed, such as Facebook, not all news articles were clearly labelled, therefore, on Interface C and Interface D, only 4 of the 8 real news articles were labelled as "Trustworthy", and 4 of the 6 fake news articles were assigned with label "Disputed". All those credibility labelled news articles were chosen randomly. The rest 6 news articles had no credibility label. The credibility label assignment was also balanced regarding political leanings and topics.

3.5 Measures

For the purpose of fair comparison among conditions, we defined the following measures for later analyses.

3.5.1 Independent Variables.

- **Stance Label.** A binary categorical variable, 1) with stance label 2) without stance label, indicating if participants were assigned to an interface with stance labels.
- **Credibility Label.** A binary categorical variable, 1) with credibility label 2) without credibility label, indicating if participants were assigned to an interface with credibility labels.
- **Self-Reported Political Stance.** Participants' pre-existing political stances were collected from a five-point Likert scale from "Very Liberal" to "Very Conservative" with "Neutral" in the middle. Based on the participant's response, he/she was coded as a member of the liberal group if he/she selected "Very Liberal" or "Liberal", as a member of the conservative group if he/she selected "Very Conservative" or "Conservative", or as a member of the neutral group if "Neutral" was selected.

3.5.2 Dependent Variables.

- **Selective Exposure Index.** According to Gao et al. [18], the selective exposure index (SEI) is defined as:

$$SEI = \frac{N_{consistent}}{N_{consistent} + N_{inconsistent}} \quad (1)$$

where $N_{consistent}$ is the number of stance-consistent news articles that the participant selected to read and $N_{inconsistent}$ denotes the number of stance-inconsistent news articles which were selected and read.

- **Perceived Level of Agreement** After finishing news browsing on the assigned interface, participants were asked to rate their perceived level of agreement of all news articles they just read. The measure of perceived level of agreement is a five-point Likert scale from "Strongly Disagree" to "Strongly Agree" on "How much you agree with the opinion expressed in the news article". The rating was then coded as 1 (Strongly Disagree) to 5 (Strongly Agree).
- **Perceived Level of Extremeness** Similar to Perceived Level of Agreement, after finishing news browsing on the assigned interface, participants were asked to rate their perceived extremeness of each individual news articles they just read. The measure of perceived extremeness is a five-point Likert scale from "Least Extreme" to "Most Extreme" on "How extreme you perceive the news article is". The rating was then coded as 1 (Least Extreme) to 5 (Most Extreme).
- **The Level of Cognitive Dissonance.** The level of cognitive dissonance measured by a Six-Item State Anxiety Scale derived from the State-Trait Anxiety Inventory [33, 47]. The State-Trait Anxiety Inventory has been examined and widely used in cognitive psychology and marketing research as a measure for accessing the level of cognitive dissonance [36, 39]. While the original State-Trait Anxiety scale contains 20 items, to prevent the effects of survey fatigue, we decided to use the shorter version which has six items. The six-item scale has been examined as effective as the original scale in terms of reliability and validity[33].

4 RESULT

To answer our research questions, we ran the 2×2 full factorial experiment on Amazon Mechanical Turk. The qualification for participants to join our study was a 90% approval rate. In a period of 3 days, a total of 132 participants joined our study. 62 of them self-identified as members of the liberal group and 60 of them self-identified as members of the conservative group. Since the study focused on people who have an explicit political leaning, 10 participants' data were removed due to their self-identified neutral leanings or incomplete data. The number of participants in the liberal and the conservative group is roughly balanced for all interface conditions (see Table 2). Of those

	Interface A	Interface B	Interface C	Interface D	All
Liberal	16	18	16	12	62
Conservative	18	13	15	14	60
All	34	31	31	26	122

Table 2. Number of Participants under Each Interface Condition

Source	df	MS	F	p	Cohen's F
Stance Label(A)	1	0.57	17.14	0.00**	0.38
Credibility Label(B)	1	0.19	5.64	0.02*	0.22
A x B	1	0.00	0.06	0.81	0.02
Residuals	118	0.03			

Table 3. Two-Way ANOVA Results on Selective Exposure Index

122 participants in the analysis, 40% were male and 60% were female. 79.51% of participants were between 25 and 54 years old, 4.10% of participants were between 18 and 25 years old, and 16.39% of participants were above 54 years old. All participants used Facebook and 98.36% of them reported Facebook as an frequent source for reading news and opinions. On average, participants spent 11.25 minutes on our interface. Participants followed the instructions, clicked and read at least five news articles ($M = 5.5$, $SD = 1.29$). Across all four interfaces, there was no significant difference in the number of clicked news articles ($F = 1.25$, $p = 0.29$). Additionally, a logistic regression on news article clicks showed participants had no significant preference between fake news articles and real news articles ($\beta = 0.10$, $z = 1.01$, $p = 0.31$).

In the following subsections, we will present the effects of stance labels and credibility labels on participants' preferential selection in 4.1 and on their perception of news articles in 4.2.

4.1 The Effect of Stance Labels and Credibility Labels on News Article Selection (RQ-1)

To measure participants' selective exposure, we calculated the selective exposure index for each participant, which is defined as the percentage of news articles the participant read that align with his/her pre-existing political stance [18].

The two-way ANOVA showed significant main effects of both stance labels and credibility labels on participants' selective exposure but no interaction effect, see Table 3. Therefore, in the following analyses, we will discuss the effect of stance labels and the effect of credibility labels on selective exposure separately.

4.1.1 The Effect of Stance Labels on Selective Exposure. The ANOVA results (Table 3) suggested the presence of stance labels significantly affected participants' opinion selection preference (Cohen's $F = 0.38$, $p < 0.01^{**}$). Participants ($N = 57$) who saw the news article with a stance label exhibited a significantly higher selective exposure bias ($M = 0.65$, $SD = 0.19$) than participants ($N = 65$) who saw the news article without a stance label ($M = 0.51$, $SD = 0.18$). In other words, participants were more likely to click and read news articles from their own side if political leanings of news articles on their assigned interfaces were labelled. The effect held for both participants in the liberal group and participants in the conservative group, see Figure 2.

These results support our hypothesis that even though news articles from both sides have equal opportunity to appear on a user's screen, the user exhibits a higher-level of selective exposure bias

if news articles' political leanings are labelled. In stance label conditions (Interface B and D), on average, 65% of the news articles that one participant read were stance-consistent.

To further understand the result, we looked into users' level of state anxiety [33] after using our interfaces. The level of state anxiety indicates the level of cognitive dissonance users experienced after they read news articles through our interfaces. We found users who used the interface with stance labels ($M = 1.18$, $SD = 0.28$) had a higher level of state anxiety than those who used the interface without stance labels ($M = 1.08$, $SD = 0.27$). A two-sample t-test showed the difference was statistically significant ($t = 2.10$, $p = 0.03^*$). The difference between interfaces with and without stance labels suggested that the presence of the stance label might trigger a higher level of cognitive dissonance that may affect participants' preferential selection processes.

4.1.2 The Effect of Credibility Labels on Selective Exposure. From Table 3, we found that showing the credibility label of news articles significantly reduced participants' selective exposure (Cohen's $F = 0.22$, $p = 0.02^*$). Participants ($N = 57$) who used interfaces (Interface C and D) with credibility labels showed a significantly lower level of selective exposure bias ($M = 0.53$, $SD = 0.18$) than users ($N = 65$) who used interfaces (Interface A and B) without credibility labels ($M = 0.61$, $SD = 0.20$). When considering participants' political stances, we found a marginal interaction effect between showing credibility labels and users' political stances (Cohen's $F = 0.18$, $p = 0.06$). Through a post-hoc test, we found that showing credibility labels to participants in the liberal group significantly reduced selective exposure bias ($M_{w/\text{credibility label}} = 0.47$, $SD = 0.19$; $M_{w/o\text{ credibility label}} = 0.61$, $SD = 0.20$; $p = 0.01^*$) while the effect was not significant for participants in the conservative group ($M_{w/\text{credibility label}} = 0.59$, $SD = 0.14$; $M_{w/o\text{ credibility label}} = 0.61$, $SD = 0.21$; $p = 0.98$), see Figure 3.

To further understand the effect of credibility labels on news article selection, we divided our news articles into three categories based on their credibility labels. According to our interface design, we had three categories of news articles, 1) Trustworthy news articles, a subset of real news articles that are labelled as "Trustworthy" on interfaces with credibility labels (Interface C and D). 2) Disputed news articles, a subset of fake news articles that are labelled as "Disputed" on interfaces with credibility labels (Interface C and D). 3) Unknown news articles, a mix of real and fake news articles that are not labelled on all interfaces.

Since selective exposure theory suggests users would click and read stance-consistent and stance-inconsistent news articles differently, we further divided each of the three news article categories into two subcategories – stance-consistent and stance-inconsistent. Therefore, in the following analysis, we will look at the composition of what kinds news articles each participant clicked and read in terms of those six categories (Table 4).

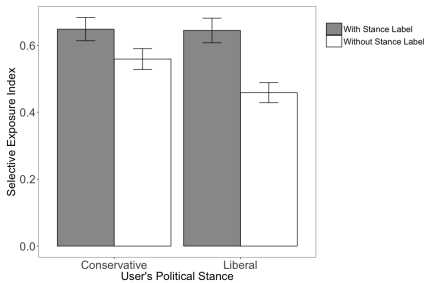


Fig. 2. The Effect of Showing Stance Labels on Users' Selective Exposure Index

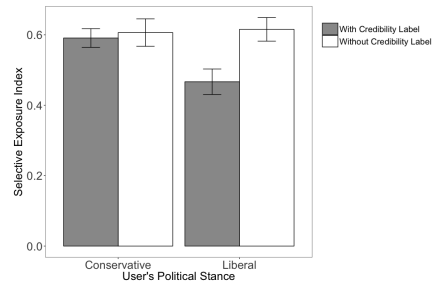


Fig. 3. The Effect of Showing Credibility Labels on Users' Selective Exposure Index

	Stance Consistent	Stance Inconsistent
Trustworthy	Stance Consistent Trustworthy	Stance Inconsistent Trustworthy
Disputed	Stance Consistent Disputed	Stance Inconsistent Disputed
No Label	Stance Consistent Unknown	Stance Inconsistent Unknown

Table 4. 6 types of News Articles Based on Credibility and Stance-consistency

Since the previous result showed that credibility labels affect participants in the liberal group and participants in the conservative group differently, we placed participants' political stances as an interactive term in the two-way ANOVA analysis for each categories in Table 4.

The main effect of showing the credibility label was found in stance-inconsistent trustworthy news articles, where users clicked and read a larger proportion of stance-inconsistent trustworthy news articles on the interfaces with credibility labels (Interface C and D) ($M_{w/ \text{credibility label}} = 0.17$, $SD = 0.14$; $M_{w/o \text{credibility label}} = 0.12$, $SD = 0.11$; $p = 0.03^*$). There was no interaction effect from participants' political leanings.

For stance-inconsistent unknown news articles and stance-consistent disputed news articles, the ANOVA analysis revealed a significant interaction effect between credibility labels and participants' political leanings but no main effect of credibility labels. The post-hoc test showed participants in the liberal group read significantly more stance-inconsistent unknown news articles ($M_{w/ \text{credibility label}} = 0.27$, $SD = 0.13$; $M_{w/o \text{credibility label}} = 0.18$, $SD = 0.14$; $p = 0.02^*$) and a significantly lower proportion of stance-consistent disputed labelled news articles ($M_{w/ \text{credibility label}} = 0.08$, $SD = 0.12$; $M_{w/o \text{credibility label}} = 0.16$, $SD = 0.12$; $p = 0.03^*$) if they used the interfaces with credibility labels (Interface C and D).

No significant effect was found in stance-consistent trustworthy news articles, stance-consistent unknown news articles or stance-inconsistent disputed news articles.

The results from above analyses provide insights on how credibility labels mitigate selective exposure bias, especially for users with liberal stances. Showing credibility labels nudged users to read more stance-inconsistent trustworthy news articles. For users with liberal stances, showing credibility labels helped them avoid disputed news articles from their own ideology and nudged them to explore stance-inconsistent news articles without any credibility labels. The result is consistent with previous studies on the individual difference between people with conservative and liberal stances that found people with liberal stances in general might be more curious [7, 10, 25]. Users with liberal stances may select news articles driven by curiosity. They may be less familiar with conservative news articles and fake news articles in general, but showing credibility labels may have made them more curious about the conservative "no label" articles and want to understand these news articles' credibility by reading them. However, they may be less curious about the disputed news articles on their own side because of the potential cognitive conflicts of reading fake news articles that supported their beliefs. The change in the proportion of different types of news articles (see Table 4) between conditions with and without credibility labels indicates how credibility labels of news articles influenced people's selections.

4.2 The Effect of Stance Labels and Credibility Labels on Perceived Extremeness and Level of Agreement (RQ-2)

We tested the effects of stance labels and credibility labels on people's perception of extremeness and their level of agreement on all news articles with a mixed effect model, in which we used Stance Label, Credibility Label, and whether news articles were stance-consistent as independent variables (details of these variables were described in the Method section). In general, people expressed

higher level of agreement on stance-consistent news articles ($M_{\text{stance-consistent}} = 3.57, SE = 0.08; M_{\text{stance-inconsistent}} = 2.58, SE = 0.09; p < 0.01^*$) and perceived stance-consistent news articles as less extreme ($M_{\text{stance-consistent}} = 3.10, SE = 0.07; M_{\text{stance-inconsistent}} = 3.51, SE = 0.07; p < 0.01^*$). This shows the general trend for existing attitudes to influence their perception of news articles. The more interesting result was the effect of labels on their perception – we found that stance labels had a significant effect on increasing people’s level of agreement on news articles ($M_{\text{w/ stance label}} = 3.50, SE = 0.20; M_{\text{w/o stance label}} = 3.12, SE = 0.13; p = 0.02^*$), suggesting that the mere presence of the stance labels could increase their level of agreement of news articles, regardless of whether they were stance-consistent or not. We also found that credibility labels had a marginal effect on decreasing people’s level of agreement, but the effect was only found in stance-consistent news articles ($M_{\text{w/ credibility label}} = 2.10, SE = 0.52; M_{\text{w/o credibility label}} = 3.08, SE = 0.25; p = 0.09$). This could be because people had lower level of agreement with stance-inconsistent articles in general, and therefore credibility labels did not lead to significantly lower level of agreement (i.e., there could be a floor effect). Interestingly, we did not find stance labels and credibility labels had any effect on participants’ perceived extremeness of news articles, regardless of their stances. There was also no significant interaction effect among those three independent variables on the level of agreement and perceived extremeness of articles.

Motivated by the idea that large amounts of fake news can influence people’s general perception of public opinion, and worse still, encourage the formation of extreme and polarized viewpoints [34], we further investigated how the labels affected people’s perception of fake news articles using the same mixed effect model with the same independent and dependent variables only on fake news articles. In general, we found that people agreed more on stance-consistent fake news articles ($M_{\text{stance-consistent}} = 3.25, SE = 0.12; M_{\text{stance-inconsistent}} = 2.15, SE = 0.12; p < 0.01^*$) and a marginally significant effect of people perceiving stance-inconsistent fake news as more extreme ($M_{\text{stance-consistent}} = 3.47, SE = 0.11; M_{\text{stance-inconsistent}} = 3.82, SE = 0.10; p = 0.08$). Our results showed that the presence of stance labels led to a significantly higher level of agreement on fake news articles ($M_{\text{w/ stance label}} = 3.44, SE = 0.27; M_{\text{w/o stance label}} = 2.72, SE = 0.18; p < 0.01^{**}$) and led to a significantly lower level of perceived extremeness of fake news articles ($M_{\text{w/ stance label}} = 2.13, SE = 0.21; M_{\text{w/o stance label}} = 3.66, SE = 0.14; p < 0.01^{**}$) regardless of news articles’ political stances. We did not find that credibility labels had any effect on people’s perception of fake news, suggesting that, surprisingly, stance labels might have a stronger impact on people’s perception of fake news than credibility labels. There was no interaction effect among those three independent variables on people’s level of agreement and perceived extremeness on fake news articles.

Similar analysis was conducted on real news articles as well. People agreed more on stance-consistent real news articles ($M_{\text{stance-consistent}} = 3.81, SE = 0.09; M_{\text{stance-inconsistent}} = 2.84, SE = 0.11; p < 0.01^*$) and perceived stance-consistent real news articles as less extreme ($M_{\text{stance-consistent}} = 2.82, SE = 0.09; M_{\text{stance-inconsistent}} = 3.24, SE = 0.10; p < 0.01^*$). Neither stance labels nor credibility labels had significant effects on participants’ level of agreement and perceived extremeness of real news articles. No interaction effect was found.

Our results indicate that stance labels have an effect on people’s perception of fake news articles. In general, stance labels increase people’s level of agreement on fake news, regardless of whether those news articles are consistent with their political stances. This indicates that stance labels may have the undesirable effect of increasing the perceived trustworthiness of fake news. In addition, stance labels lower people’s perceived extremeness of fake news on both sides. These effects of stance labels may be potentially dangerous since people might be misled by extreme and polarized opinions in fake news when they are categorized under different stance labels. Meanwhile, credibility labels could marginally decrease people’s level of agreement on news articles on their

own side, which may lead to a more moderate opinion space. However, consistent with previous findings, the effect of credibility labels was in general limited.

5 CONCLUSION AND GENERAL DISCUSSION

Overall, our results suggest that stance and credibility labels can affect people's news article selection and perception in ways that deviate from their original intended effects. As a result, such effects are often neglected in online news platform designs. Thus, social media designers should be cautious about using labels. Our results show that, consistent with previous studies, assigning credibility labels to articles (such as "Disputed") is not effective for combating fake news. Interestingly, stance labels, which are often designed to help people find information, may have undesirable effects on facilitating the spread of fake news. Specifically, we found that stance labels can make fake news articles look more trustworthy, and they seem to lower people's perception of the extremeness of fake news articles. Without systematic evaluation, those unintended, yet undesirable, effects of stance labels could not have been detected. In fact, our results suggest that stance labels may reinforce users' existing beliefs and indirectly exacerbate polarization of social opinions.

At a higher level, our results demonstrate the importance of studying the role of readers' pre-existing stances on online news platform designs, as seemingly benign design choices, such as stance labels, could lead to complex interactions among people's stances, their perception of the news articles, as well as how much they agree with and consequently influenced by the news articles, including fake news articles. Social and behavioral theories that study these complex interactions are therefore important to guide the designs of online social information platforms. In general, one important message from decades of social behavioral research is that humans should not be treated as simple information processors that search for and consume information using indices like machines do. Rather, humans have complex internal traits and states, such as their existing beliefs, attitudes, and opinions, that influence how they look for relevant information react to information that deviate from their existing beliefs.

Our results show that news articles' stance labels may exacerbate selective exposure. There are multiple possible causes of the effect. First of all, showing stance labels helps people quickly target stance-consistent news articles in their areas of interest, thus paving the way for them to only read articles to their interest. Another possible explanation could be that stance labels elevate people's cognitive dissonance and trigger people's defense motives for desired and directional goals, which further makes them look for more stance-consistent news articles. Even though no causal relationship between showing the stance label and the level of cognitive dissonance can be derived from the current experiment design, our analyses showed that the state anxiety level of people who used the stance-labelled interfaces is significantly higher than those who used interfaces without stance labels. The results are consistent with the idea that stance labels may induce uncomfortable feeling, which may prompt people to adapt reading strategies to resolve them.

On the other hand, our results also suggest that credibility labels might mitigate people's selective exposure, especially for people with liberal stances; furthermore, our findings indicate that people tend to read more stance-inconsistent news articles that are labelled as trustworthy. It is possible that people read more trustworthy news articles from the opposite side since credibility labels could trigger people's accuracy motives for objective and accurate information. To further understand the underlying mechanism for the effect of labels, more sophisticated study and evaluation is needed. To figure out whether different types of labels could ignite different types of motives and thus alter people's news article reading behavior, the future studies could measure the level of accuracy and defense motives in different labelling conditions and then analyze the relationship between motives and social opinion selection and/or perception.

From the perspective of information seeking strategies, Munson and Resnick [38] proposed that there exist both challenge-averse and diversity-seeking people. Challenge-averse people tend to avoid disagreeable social opinions while diversity-seeking people are open-minded for opinions with different stances. It would be fruitful to address and explore how stance and credibility labels affect social opinion selection and perception for challenge-averse and diversity-seeking people, respectively. Labels may have different effects on people with different social opinion reading behavior. For example, showing stance labels may facilitate challenge-averse people avoiding opposite opinions but help diversity-seeking people to find different opinions more easily. Further studies need to be conducted to figure out the exact effect of labels on people with different social opinion reading behaviors, which could help to design better news feed to mitigate selective exposure across different groups of people in terms of social opinion consumption trait.

6 LIMITATION AND FUTURE WORK

We want to address several limitations of the study. First, news articles in our study cover only two topics, President Trump and Gun Control. Even though during the time of the study, those two topics were trending and controversial, users' news feed often bombards with a massive amount of stories. Therefore, including more topics can further benefit our study's ecological validity. Second, our study only demonstrates how the presence of labels affects readers' news article preferential selection and perception in the short term. However, people's opinions can be affected by information in the long run. For example, if people repeatedly encounter misinformation in a long term, misinformation may retain even though people have been informed with the correction [35]. Therefore, a future long-term field study on social media websites can best help us understand how labels may affect people's viewpoints in a long run.

There are many possible directions for future work. For example, although the design of our news feed is directly adapted from Blue Feed, Red Feed, the popular news feed, more research needs to be done in other contexts. For example, friends' activities may also affect users' behavior on news feed [21]. Although a news article is labelled with opposite political leanings, users may still be willing to read the news article because their friends commented on it. Another important research direction is to study the effects of labels as their positions, colors, fonts or background colors are changed [6, 43]. These design choices may also influence readers' behavior in ways that nudge them to read opposite viewpoints and eventually become more open-minded and less polarized.

ACKNOWLEDGMENTS

This research was supported by the NSF awards SES-1419297 and IIS-1441563. Part of the work was also supported by the University of Illinois Academic for Excellence In Engineering Education program. We would like to thank all the reviewers for their valuable feedback.

REFERENCES

- [1] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 2 (2017), 211–36.
- [2] Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.
- [3] Maansi Bansal-Travers, David Hammond, Philip Smith, and K Michael Cummings. 2011. The impact of cigarette pack design, descriptors, and warning labels on risk perception in the US. *American journal of preventive medicine* 40, 6 (2011), 674–682.
- [4] Pablo Barberá, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. 2015. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science* 26, 10 (2015), 1531–1542.

- [5] Richard F Beltrami. 1988. Perceived believability of warning label information presented in cigarette advertising. *Journal of Advertising* 17, 2 (1988), 26–32.
- [6] Manuel Cabrera, Leandro Machín, Alejandra Arrúa, Lucía Antúnez, María Rosa Curutchet, Ana Giménez, and Gastón Ares. 2017. Nutrition warnings as front-of-pack labels: influence of design features on healthfulness perception and attentional capture. *Public health nutrition* 20, 18 (2017), 3360–3371.
- [7] Dana R Carney, John T Jost, Samuel D Gosling, and Jeff Potter. 2008. The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology* 29, 6 (2008), 807–840.
- [8] Alexander Chernev. 2004. Extremeness aversion and attribute-balance effects in choice. *Journal of consumer research* 31, 2 (2004), 249–263.
- [9] Josh Constine. 2018. Facebook shrinks fake news after warnings backfire. *techcrunch.com* (April 2018).
- [10] Sanorita Dey, Karrie Karahalios, and Wai-Tat Fu. 2018. Effects of Socially Stigmatized Crowdfunding Campaigns in Shaping Opinions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 242.
- [11] Michael R Ent and Mary A Gerend. 2016. Cognitive dissonance and attitudes toward unpleasant medical screenings. *Journal of health psychology* 21, 9 (2016), 2075–2084.
- [12] Robert Epstein and Ronald E Robertson. 2015. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences* 112, 33 (2015), E4512–E4521.
- [13] Robert Epstein, Ronald E Robertson, David Lazer, and Christo Wilson. 2017. Suppressing the search engine manipulation effect (SEME). *Proceedings of the ACM: Human-Computer Interaction* 1 (2017), 42.
- [14] 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.
- [15] Leon Festinger. 1962. *A theory of cognitive dissonance*. Vol. 2. Stanford university press.
- [16] Dieter Frey. 1986. Recent research on selective exposure to information. *Advances in experimental social psychology* 19 (1986), 41–80.
- [17] 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.
- [18] Mingkun Gao, Hyo Jin Do, and Wai-Tat Fu. 2018. Burst Your Bubble! An Intelligent System for Improving Awareness of Diverse Social Opinions. In *23rd International Conference on Intelligent User Interfaces*. ACM, 371–383.
- [19] 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.
- [20] Liz Halloran. 2010. New 'No Labels' Movement Seeks Bipartisanship. National Public Radio (Dec. 2010).
- [21] Aniko Hannak, Drew Margolin, Brian Keegan, and Ingmar Weber. 2014. Get Back! You Don't Know Me Like That: The Social Mediation of Fact Checking Interventions in Twitter Conversations.. In *ICWSM*.
- [22] Carl I Hovland, Irving L Janis, and Harold H Kelley. 1953. Communication and persuasion; psychological studies of opinion change. (1953).
- [23] Caroline Jack. 2017. Lexicon of Lies: Terms for Problematic Information. *Data & Society* 3 (2017).
- [24] Nabil Jeddi and Imed Zaiem. 2010. The Impact of Label Perception on the Consumers' Purchase Intention: An application on food products. *IBIMA Business Review* (2010).
- [25] John T Jost, Jack Glaser, Arie W Kruglanski, and Frank J Sulloway. 2003. Political conservatism as motivated social cognition. *Psychological bulletin* 129, 3 (2003), 339.
- [26] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 265–274.
- [27] Ziva Kunda. 1990. The case for motivated reasoning. *Psychological bulletin* 108, 3 (1990), 480.
- [28] Stephan Lewandowsky, Ullrich KH Ecker, Colleen M Seifert, Norbert Schwarz, and John Cook. 2012. Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest* 13, 3 (2012), 106–131.
- [29] Q Vera Liao and Wai-Tat Fu. 2014. Can you hear me now?: mitigating the echo chamber effect by source position indicators. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 184–196.
- [30] 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 SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2745–2754.
- [31] 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.
- [32] 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.

- [33] Theresa M Marteau and Hilary Bekker. 1992. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology* 31, 3 (1992), 301–306.
- [34] Alice Marwick and Rebecca Lewis. 2017. Media manipulation and disinformation online. *New York: Data & Society Research Institute* (2017).
- [35] Nicco Mele, David Lazer, Matthew Baum, Nir Grinberg, Lisa Friedland, Kenneth Joseph, Will Hobbs, and Carolina Mattsson. 2017. Combating fake news: An agenda for research and action.
- [36] Michael B Menasco and Del. I. Hawkins. 1978. A field test of the relationship between cognitive dissonance and state anxiety. *Journal of Marketing Research* (1978), 650–655.
- [37] Sean A Munson, Stephanie Y Lee, and Paul Resnick. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget.. In *ICWSM*.
- [38] 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.
- [39] Sadaomi Oshikawa. 1972. The measurement of cognitive dissonance: Some experimental findings. *The Journal of Marketing* (1972), 64–67.
- [40] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.
- [41] 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.
- [42] Hannah Jane Parkinson. 2016. Click and elect: how fake news helped Donald Trump win a real election. *The Guardian* 14 (2016).
- [43] Jonathon P Schuldt. 2013. Does green mean healthy? Nutrition label color affects perceptions of healthfulness. *Health Communication* 28, 8 (2013), 814–821.
- [44] Craig Silverman. 2016. This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook. *Buzzfeed* (Nov 2016).
- [45] Craig Silverman and Jeremy Singer-Vine. 2016. Most Americans who see fake news believe it, new survey says. *BuzzFeed News* (2016).
- [46] Jeff Smith, Grace Jackson, and Seetha Raj. 2017. Designing Against Misinformation. *Medium.com* (Dec 2017).
- [47] Charles D Spielberger, RL Gorsuch, and RE Lushene. 1970. STAI. *Manual for the State-Trait Anxiety Inventory (Self Evaluation Questionnaire)*. Palo Alto California: Consulting Psychologist 22 (1970), 1–24.
- [48] Cass R Sunstein. 2002. The law of group polarization. *Journal of political philosophy* 10, 2 (2002), 175–195.
- [49] Sarah Tanford and Rhonda Montgomery. 2015. The effects of social influence and cognitive dissonance on travel purchase decisions. *Journal of Travel Research* 54, 5 (2015), 596–610.
- [50] John Zaller. 1992. *The nature and origins of mass opinion*. Cambridge university press.

Received April 2018; revised July 2018; accepted September 2018