How Partisan Crowds Affect News Evaluation

Maurice Jakesch
Cornell University
107 Hoy Rd
Ithaca, NY 14850
mpj32@cornell.edu

Moran Koren
Stanford University
450 Serra Mall
Stanford CA 94305
korenm@stanford.edu

Anna Evtushenko
Cornell University
107 Hoy Rd
Ithaca, NY 14850
ae392@cornell.edu

Mor Naaman
Cornell Tech
2 West Loop Road
New York 10044
mn469@cornell.edu

Abstract

Social influence is ubiquitous in politics and online social media. Here we explore how social signals from partisan crowds influence people's evaluations of political news. For example, are liberals easily persuaded by a liberal crowd, while resisting the influence of conservative crowds? We designed a large-scale online experiment (N=1,000) to test how politically-annotated social signals affect participants' opinions. In times rife with misinformation and polarization, our findings are optimistic: the mechanism of social influence works across political lines, that is, liberals are reliably influenced by majority-Republican crowds and vice versa. At the same time, we replicate findings showing that people are inclined to discard news claims that are inconsistent with their political views. Considering that people show negative reactions to politically dissonant news but not to social signals that oppose their views, we point to the possibility of depolarizing social rating systems.

1 Introduction

Political polarization – the vast and growing disagreement on political values – has become a significant concern in American politics and globally (Baldassarri and Gelman, 2008; Sides and Hopkins, 2015). In their 2017 report, Pew Research finds record levels of disagreement on questions of welfare, race, immigration and foreign policy in the U.S. (Center, 2017). Whether social media sites support political polarization has been intensely debated by scholars and journalists alike. Early studies claim that social clustering (McPherson et al., 2001; Conover et al., 2011; Vosoughi et al., 2018) paired with content recommendation systems lead to online "filter bubbles" in which individuals are exposed mainly to conforming views (Pariser, 2011). More recent research has challenged this notion, showing that social media users see a more diverse set of information than previously assumed (Bakshy et al., 2015; Beam et al., 2018; Barberá et al., 2015).

This paper explores an alternative mechanism that might cause polarization on social media sites: *social rating systems may be polarizing in and of themselves when they aggregate opinions of partisan crowds.* We hypothesize that exposure to a rating from a partisan crowd may have different effects on liberal and conservative viewers. Given the pervasiveness of rating systems on social media sites (Dellarocas, 2003), it is critical to understand how they interact with political biases.

Research has shown that the aggregated crowd opinions displayed on social media platforms exert both normative and informative influence over individuals (Asch, 1955; Deutsch and Gerard, 1955; Cosley et al., 2003). While the political makeup of the crowd is typically not made explicit, viewers may infer crowd partisanship based on platform and content cues in many cases. For example, a YouTube video on the Fox News channel is likely rated by a majority-conservative crowd, whereas comments under a New York Times article come from a more liberal crowd. Being aware of the ideological slant of a crowd may help to assess the rating's informativeness, but may also lead to dismissal of or identification with the crowd opinion. We propose three hypotheses that extend basic social influence to political situations:

(a) **Out-group resistance:** Users may only be influenced by social signals that come from a group they identify with. A conservative may be moved by a conservative crowd, but indifferent towards signals from a liberal crowd. Research showing that the effectiveness of social influence depends on identification (Kelman, 1961) and perceived similarity (Berscheid, 1966; Simons et al., 1970) sup-

Proceedings of the 2020 Truth and Trust Online (TTO 2020), pages 1–11,
Virtual, October 16-17, 2020.

ports this hypothesis.

(b) **Selective attention:** Research on cultural cognition (Kahan et al., 2011) and motivated reasoning (Kunda, 1990; Taber and Lodge, 2006) has shown that individuals find ways to discard information that is inconsistent with their cultural identities. When encountering crowd opinions, people may selectively attend to those social signals that affirm their preexisting political views. A liberal may consider crowd support for Democrat-consistent claims, but ignore support for Republican-consistent claims[1].

(c) **Signal informativeness:** Opinions may be affected most by informative or surprising signals. For example, seeing a majority-Democrat crowd reject a Democrat-consistent claim is highly informative. Seeing a majority-Republican crowd reject a Democrat-consistent claim is less informative. Research showing that people attend more to unusual information (Hope and Wright, 2007) motivates this hypothesis that corresponds to rational learning from biased information (Calvert, 1985).

We designed a large-scale online experiment (N=1,000) to test how crowd partisanship affects news evaluations. Specifically, we evaluated how politically-annotated social signals affected the likelihood that participants thought the news they saw was true. We found that people – as expected – tended to discard claims that did not align with their political views. However, they were reliably influenced by social signals independent of crowd partisanship. We find no evidence of selective attention, and partial support influence based on signal informativeness among liberals only, who were more likely to change their evaluation when they saw an informative signal from a majority-Democrat crowd.

Generally speaking, our results imply that the mechanism underlying social influence is non-political. We discuss how the finding that people reject politically dissonant information but are equally influenced by group opinions regardless of politics may allow for configurations of social rating system that mitigate platform polarization.

[1]The paper is written in the context of the U.S political system, using left/liberal and right/conservative labels.

2 Background

Our work connects three areas of prior research that we briefly touch on: social influence, political bias, and rating systems.

Asch's conformity experiments (Asch, 1955; Bond and Smith, 1996) in the 1950s initiated an extensive field of research on social influence and group conformity (Mason et al., 2007). People's opinions are affected by those of their peers, as groups exert normative pressure and are a source of information (Deutsch and Gerard, 1955). Most related to our study, researchers have found that similarity increases the strength of social influence (Kelman, 1961; Berscheid, 1966; Simons et al., 1970) and that partisan groups may sustain themselves through social influence (Flache and Macy, 2011). Pronin et al. have shown that in political debates, peer influence may affect opinions more than policies' actual content (Pronin et al., 2007). In cases where opinions are not well-formed, observing the choice of others creates *social defaults* (Huh et al., 2014). However, it is hard to separate social influence from homophily in real-world settings (Shalizi and Thomas, 2011).

Researchers have extensively studied biases in political thinking. We know that biased cognitive processes affect the way people process information through *motivated reasoning* (Kunda, 1990): individuals have various strategies to be skeptical of information that may falsify their firmly held beliefs. Motivated reasoning is associated with attitude polarization (Taber and Lodge, 2006; Slothuus and De Vreese, 2010), especially for those with high levels of political sophistication (Flynn et al., 2017). For example, both liberals and conservatives tend to believe the economy is improving when their party is in power and vice versa (Bartels, 2002). Individuals may avoid incongruent information due to the discomfort of *cognitive dissonance* (Festinger, 1957) it evokes. While some studies have found that exposing individuals to politically misaligned views may even increase political polarization (Bail et al., 2018; Nyhan and Reifler, 2010), such *backfire effects* have not been reliably reproduced in follow-up studies (Wood and Porter, 2019; Weeks and Garrett, 2014).

Researchers have explored the role of social rating systems in opinion formation. The *digitization of word of mouth* is a pervasive feature of websites (Dellarocas, 2003). It affects users' opinions (Cosley et al., 2003) and shapes the diffusion

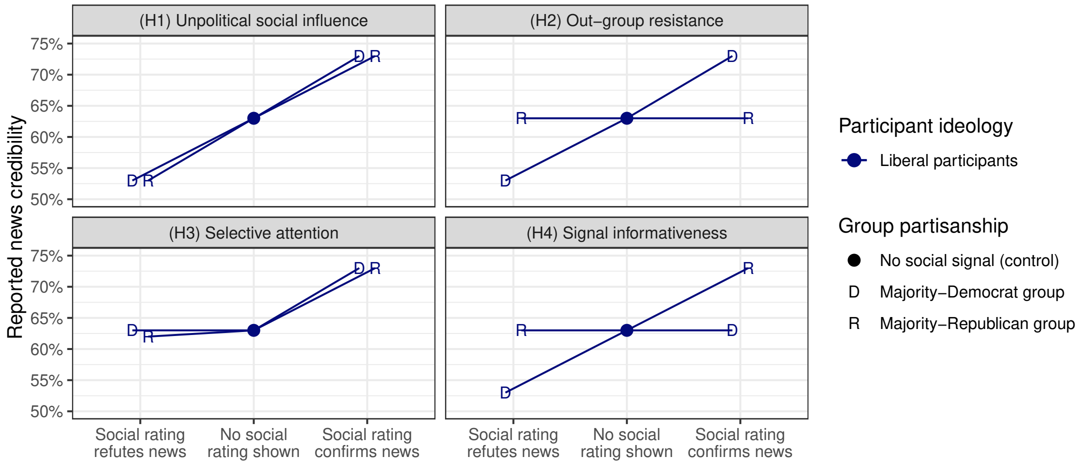

Figure 1: Hypotheses on how a partisan crowd could affect a liberal's evaluation of *Democrat-consistent news*.

of user-generated content (Bakshy et al., 2009) in complex ways (Romero et al., 2011). The feedback of earlier users affects the behavior of those who follow (Muchnik et al., 2013), creating system dynamics with unpredictable and unequal outcomes (Salganik et al., 2006). For example, individuals may take arbitrary political sides on an issue based on the votes and political association of the first few users who rate it (Macy et al., 2019). Most related to our experiment, Messing and Westwood (Messing and Westwood, 2014) have studied the effect of social indicators on the selection of news. They found that social endorsements are more predictive of news selection than source indicators and that their effect is particularly strong for partisans selecting news from ideologically misaligned sources. We add to this line of work by investigating how social signals affect the *evaluation* of news claims depending on different political constellations.

3 Methods

We designed and executed an online experiment to quantify the effect of a social signal from a politically-annotated crowd on the evaluations of news. We randomly assigned respondents to either a control group where they saw no social signal or one of two treatment groups where they saw social signals from majority-Democrat or majority-Republican crowds. We showed participants political news claims and asked them whether they thought they were true or false.

3.1 Hypotheses and preregistration

We formalize our research question into a set of empirically testable hypotheses, drawing on the types of influence mentioned in the introduction:

(H1) Unpolitical social influence: participants conform to a group opinion independently of the politics of the group and the news claim.

(H2) Out-group resistance: Participants follow the crowd, but only if they politically identify with the crowd.

(H3) Influence through selective attention: Participants follow the crowd only if the crowd opinion supports their pre-existing views.

(H4) Influence through signal informativeness: Participants are more influenced by informative signals where a crowd votes against its political agenda.

The different hypotheses are summarized in Figure 1, showing the expected opinion shift (Y-axis) under each hypothesis for a liberal participant evaluating a Democrat-consistent claim. For example, the first graph in the second row shows that, under the selective attention hypothesis (H3), a liberal would attend to any group that supports a Democrat-consistent claim, but ignore social signals that refute the same. We test these hypotheses separately for liberal and conservative participants. We preregistered the study, the analysis and hypotheses (except for H3), available at `http://aspredicted.org/blind.php?x=hx82az`.

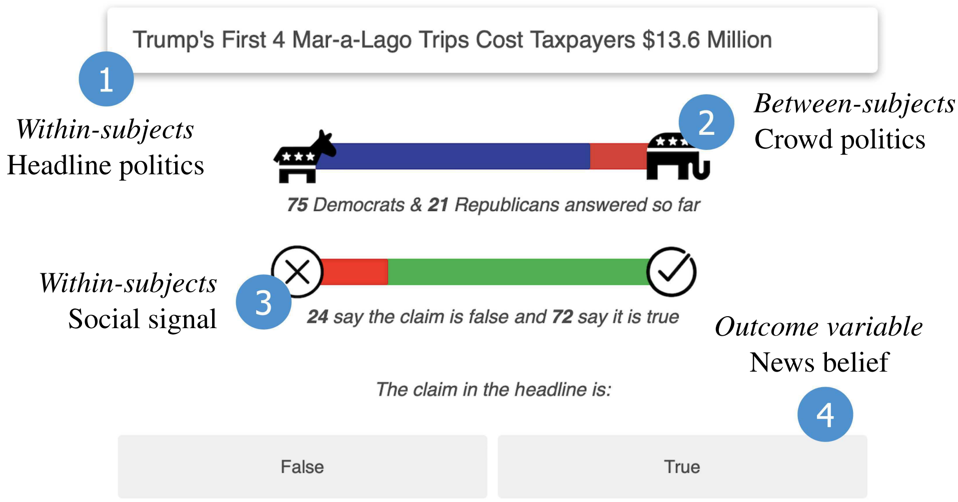

Figure 2: Screenshot of the experiment with annotated variables.

3.2 Experimental design

Our experiment imitated three common elements of people's online experience: a news claim rendered as a news headline without the actual article (element 1 in Figure 2); the aggregated opinion of prior users who have rated the claim (element 3); and buttons to rate the claim as either *true* or *false*, similar to rating features on social media platforms (4). Besides, we displayed the political makeup of users who supposedly rated the headline before the participant (2). We varied both the politics of the headline (1) and the social signal (3) as part of our experimental manipulation. The composition of the crowd (2) was explained before the study and did not change within-subject. The different elements successively appeared so that participants spent some time focusing on each. We used a 2x3x2 mixed factorial design:

(1) *Headline politics (within-subjects)*: Each participant rated 16 headlines. Four of these headlines were Democrat-consistent, and four were Republican-consistent, randomly selected from a pool of true-but-hard-to-evaluate political headlines (detailed below). The remaining eight headlines were non-political decoys designed to disguise the purpose of the study.

(2) *Crowd politics treatment (between-subjects)*: Participants were randomly assigned to one of three conditions: the control, the majority-Democrat, or the majority-Republican group. While participants in the control group saw no social signal, participants in the treatment groups were lead to believe that 96 people had rated the headline before them and 75 of them–chosen to reflect a clear majority–were either Democrats or Republicans. We chose a fixed population of 96 as it is large enough for a meaningful vote, allows for easy proportion calculations due to its proximity to 100, and may look less artificial than 100.

(3) *Social signal treatment (within-subjects)*: Participants in the treatment groups saw a social signal of those who supposedly rated the headline before them. We manipulated the signal to overwhelmingly (by a large majority, e.g., 72 versus 24) affirm two and reject two Democrat- and Republican-consistent claims each. The social signal for decoy headlines was more balanced to give a realistic impression of crowd deliberation. We chose to show an aggregate social signal, not separating the opinions of Democrats and Republicans, in order to reflect the setup of most current popular social media sites where only see aggregated social ratings.

We iterated on the design based on multiple usability tests with local volunteers and one pilot study (N=100) using workers from Amazon Mechanical Turk. The tests were performed to confirm that people understood the manipulation. As a result of the tests, we animated the display so that elements (1) to (4) were added to the screen in succession, calling attention to each element. We also added additional textual explanations to each indicator. The $N = 100$ pilot provided an initial indication that the manipulation had an effect on the dependent variable and helped us plan the size of the experiment.

We allowed 20 seconds to rate each claim and disabled the copying of headlines to prevent participants from looking them up. Also, we conducted an attentiveness check at the beginning of the study and asked participants whether they had searched for headlines online after completion. After the main rating task, participants provided demographic information and answered questions about their political affiliation.

3.3 News claim selection

The experimental design required a set of factual news claims that can be evaluated as either true or false, that support either U.S. Democrat or Republican views, and are not widely known or easily evaluated. To increase the external validity of the study, we also wanted the headlines to be representative of people's ordinary news consumption.

To collect a set of such headlines, we identified the top four liberal and top four conservative online news websites using Alexa traffic rank data[2]. From this initial pool of 10,660 headlines, we computationally extracted the ones that were likely to contain a claim based on ClaimBuster scores (Hassan et al., 2017), leaving 899 headlines. We then randomly sampled 40 headlines from each news organization, manually excluding headlines that did not contain a claim or were not related to U.S. politics. As we expect the effect of social influence to be stronger and more robust for claims of which participants had little prior knowledge, we evaluated the headlines in a preliminary survey with N=85 workers on Amazon Mechanical Turk. We asked the workers to label which claims they thought were true and removed headlines that were marked as true in more than 75% of cases. We asked a second set of 85 workers to label whether the claims aligned with U.S. Democrat or Republican views. Based on these labels, we selected the 15 most pro-Democrat and the 15 most pro-Republican headlines for our study. Besides, we collected a set of 12 politically neutral decoy headlines, six of which were randomly shown to participants in the experiment but were not part of the analysis. We added two easily discernible and politically neutral fake headlines from a prior study (Pennycook and Rand, 2018) to assure participants that some of the headlines they were reviewing were indeed false. We randomized headline assignments and order for each participant. We include the full set of headlines in the project's Open Science Repository (https://https://osf.io/stjer/).

[2] https://www.alexa.com/topsites/countries/US

3.4 Participants

We recruited 1,000 participants through Amazon Mechanical Turk (AMT) (Buhrmester et al., 2011). While not nationally representative, samples from AMT have been shown to reproduce treatment effects in political research (Coppock, 2018; Clifford et al., 2015) and behavioral research reliably (Mason and Suri, 2012). Recruitment was limited to U.S. participants of age 18+ with an approval rate of $\geq 98\%$. To counterbalance the over-representation of liberal workers on AMT, we posted part of our recruitment ($n = 250$) as a separate task available for workers that identified as politically conservative only. Participants received compensation of \$0.80 based on an estimated participation time of 4-5 minutes for a projected \$10-12 hourly wage. Participants were debriefed upon study completion, explained the purpose of the study and the deception involved, and given the option to withdraw. The study protocols were approved by the Institutional Review Board at the primary investigator's institution. We excluded participants who had failed an attentiveness check, participants who indicated they had searched for headlines on the web, and participants who rated all headlines as either true or false. We also deleted data of two participants who withdrew after the debrief, leaving us with a final sample of N=971 participants. As a result of the recruitment strategy detailed above, our participant sample was politically balanced, with 47.8% of our participants identifying as politically right or right-center. Participants on average, were 39.8 years old, 53.7% identified as female.

3.5 Open Science Repository

The experimental data, analysis code and experiment preregistration are available at https://osf.io/stjer/.

4 Results

The results show that social influence is effective and primarily non-political (H1). The results also provide partial evidence for social influence based on informativeness (H4). Figure 3 reports the ag-

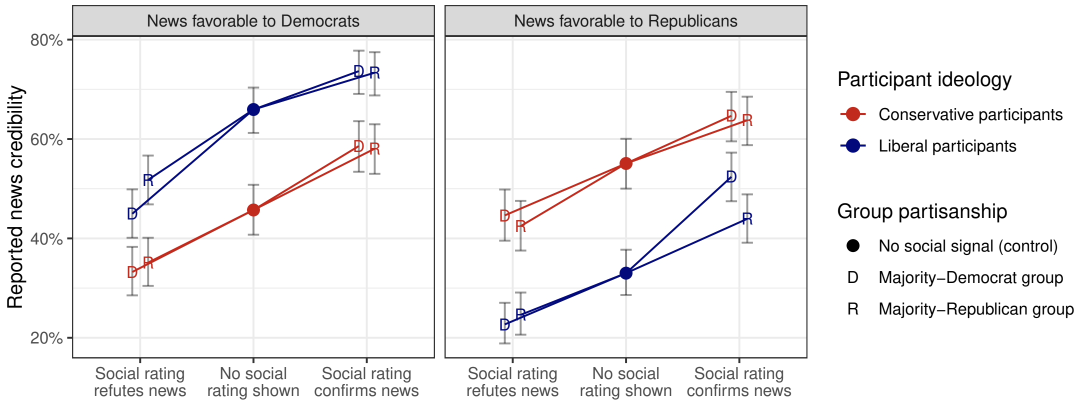

Figure 3: **Liberals are reliably influenced by majority-Republican crowds and vice versa.** N = 1000; error bars represent 95% CIs. The Y-axis measures the percent of participants who evaluated a headline as "true", broken down by the participant's ideology and the social signal shown. The line slope indicates the strength of the social influence effect. The effect is similar for majority-Republican and majority-Conservative crowds, except when liberal participants saw a majority-Democrat crowd confirm news favoring Republican views (bottom right).

gregated evaluations of news headlines. The left panel shows how participants evaluated headlines that supported Democrat views, the right panel show evaluations of news in favor of Republicans. The middle column on each panel shows the responses in the control group. Participants here did not see any social signals, their evaluations serve as a baseline. The Y-axis shows how often they said the news they saw was true. For example, liberal participants, shown in blue, said that the Democrat-consistent claims they saw were true in 66% of cases when they saw no social signal. Republicans found Democrat-consistent headlines significantly less credible. To the left and right of the control group we see how participants evaluated news when the group of prior raters refuted or rejected a claim. Ratings are split depending on the ideology of the crowd. For example, liberal participants thought Democrat-consistent headlines were true 45% of the time when a majority-Democrat (D) crowd rejected them. If the same headlines were rejected by a majority-Republican crowd (R) instead, liberals thought they were true 51.8% of the time.

We first test whether the social signal had a significant effect on evaluations in the treatment groups using logistic regression (full model specifications in the preregistration). We find that across all 16 combinations of treatments and politics, participants significantly changed their evaluations compared to the control group. We see the smallest change for liberals evaluating a Democrat-consistent claim rejected by a majority-Republican crowd (65.9% to 73.4%; +7.5%, OR=1.42, $p<0.05$) and the most substantial change for liberals evaluating a Democrat-consistent claim rejected by a majority-Democrat crowd (65.9% to 45%; -20.9%, OR=0.42, $p<0.001$). We can confidently reject the null hypothesis that the social signal had no effect.

We now turn to the differences due to the partisanship of the crowd. In almost all cases, the political makeup of the crowd did not have a significant effect on participants' evaluations. Only when a majority-Democrat crowd affirmed a Republican-consistent claim (bottom right; 52.1% compared to 43.9%; -8.18%, OR=0.72, $p<0.01$), the indicator of the crowd's political makeup had a significant effect.

In the final part of the analysis, we perform a regression to estimate how much the different hypothesized types of social influence contributed to the result. We create three composite variables that correspond to our hypotheses:

- *Group identification (H2):* Does the participant politically identify with the crowd?
- *Selective attention (H3):* Does the social signal support the participant's political views?
- *Signal informativeness (H4):* Does the crowd vote against its political agenda?

We model hypotheses (H2) to (H4) separately for Democrats and Republicans through interactions of these terms and the social signal. Note that since we did not include this model in the preregistration,

Table 1: Model coefficients predicting participants' evaluations

	OR
(Intercept)	2.07^{***}
1: Conservative participant	1.097
2: Republican-consistent claim	0.46^{***}
3: Politically misaligned claim	0.39^{***}
4: Liberal : Social signal (H1)	1.76^{***}
5: Liberal : Signal : Identification (H2)	1.104^{*}
6: Liberal : Signal : Select. attention (H3)	0.861
7: Liberal : Signal : Informativeness (H4)	1.068
8: Conservative: Social signal (H1)	1.62^{***}
9: Cons. : Signal : Identification (H2)	0.984
10: Cons. : Signal : Select. attention (H3)	0.956
11: Cons. : Signal : Informativeness (H4)	1.013
σ	2.1^{***}
N = 7,643, $P < .00001$, $\chi^2 = 472.8$, LL $= -4786.7$	

Significance codes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

the result should be considered exploratory. The predicted outcome (y_j) is the claim evaluation:

$$
\begin{aligned}
\mathrm{y}_j = \alpha &+ \beta_1 \cdot \text{ConservativeParticipant}_j \\
&+ \beta_2 \cdot \text{RepublicanConsistentClaim}_j \\
&+ \beta_3 \cdot \text{PoliticallyMisalignedClaim}_j \\
&+ \beta_4 \cdot \text{SubjectPolitics}_j : \text{SocialSignal}_j \\
&+ \beta_5 \cdot \text{SPol}_j : \text{Signal}_j : \text{Identification}_j \\
&+ \beta_6 \cdot \text{SPol}_j : \text{Signal}_j : \text{SelectiveAttention}_j \\
&+ \beta_7 \cdot \text{SPol}_j : \text{Signal}_j : \text{Informativeness}_j
\end{aligned}
$$

Table 1 shows the regression results. The baseline (see intercept) is a liberal participant evaluating a Democrat-consistent claim. We see that conservative participants did not evaluate headlines differently from liberal participants on average (line 1 in Table 1). The largest change in evaluations is predicted by political misalignment with the claim: Participants were 21.5% less likely to rate a claim as true when it did not align with their political views (line 3).

For both liberals and conservatives, we see a significant social influence effect: the social signal raised or lowered their evaluations by about 12% compared to the control group, independently of crowd politics (line 4 and 8 in Table 1). In Figure 3, this corresponds to an average drop of ratings by 12% to the left and an average raise of 12% to the right of the control group. The model does not predict any interactions between social influence and crowd partisanship for conservatives. For liberals, the model finds a significant reaction to the politics of the crowd, which it captures as combination influence based on identification (H2, line 5 in Table 1) and informativeness (H4, line 7), corresponding to the observation that the social influence was about 8% stronger when liberals saw an informative signal from a majority-Democrat crowd. We have repeated variations of the analysis while excluding respondents who identified as independents or only mildly liberal or conservative, with similar results.

5 Discussion

Our results confirm that political news claims polarize people. When participants saw a claim that did not align with their political views, they were 21% less likely to evaluate it as true, making political misalignment with a claim the principal predictor of disbelief. Previous studies have observed such strong adverse reactions to politically misaligned content (Hart et al., 2009; Stroud, 2010) and recent work has shown that they affect evaluations more than the news source (Jakesch et al., 2019).

Our findings also show that individuals reliably respond to social influence even in polarized political settings. Both liberals and conservatives were highly influenced by the opinion of the crowd, regardless of its political partisanship. For all types of crowds respondents were about 12% more likely to say a claim was true if the crowd supported it, and 12% less if the crowd rejected it.

While evaluations were significantly affected by claim politics and the social signal, the strength of social influence was mostly independent of the political makeup of the crowd. Conservatives, in particular, did not change how they evaluated news claims based on crowd partisanship. Majority-Democrat and Republican crowds equally influenced conservative views on all issues. This finding supports the idea that social influence is a non-political mechanism (H1), which is surprising given research showing that people are more influenced by politically like-minded others (e.g., (Marks et al., 2019)).

Participants identifying as liberals changed their opinions more when they saw an *informative* (that is, unusual) signal from a majority-Democrat crowd. Specifically, a majority-Democrat crowd that affirmed a Republican-consistent claim had an unusually large influence on liberals. This result partially supports the idea that influence depends

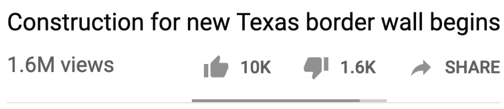

Figure 4: A slanted social signal from a partisan crowd on YouTube may reinforce opinions

on informativeness (H4), but only for liberals in a majority-Democrat crowd (H2). We found no support for selective influence (H3) and saw no evidence of backfire effects.

The results raise the question of whether the experiment was adequate to produce the hypothesized political effects. The design has 90% statistical power for small effect sizes (Cohen's $f^2 = 0.25$). We also found that our manipulations worked as intended: Participants believed a claim was true in only 49% of cases, showing that people had little prior knowledge of the claims. Claims we selected as Democrat-consistent were more likely to be believed by Democrats and vice versa. The social influence manipulation had a significant effect across all groups, and the crowd partisanship indicator changed the evaluation of liberals, showing that all manipulations worked as intended.

5.1 Implications

Our findings show that the mechanism of social influence is largely unpolitical. However, even a non-political social influence mechanism can increase polarization due to selective exposure.

Polarizing content that appeals to a specific political audience is likely to be viewed by a politically slanted crowd (McPherson et al., 2001; Conover et al., 2011; Vosoughi et al., 2018). Take as an example the ratings of a video on the Mexico-U.S. border barrier posted by CBS on April 2018 (see Figure 4). Representative polls show that Americans are divided on their support for a border barrier, with recent polls finding a majority of 60% opposing it (Gallup, 2019). The rating system, however, shows that 86% of those who rated the border barrier video liked it. This social signal not only creates a false impression of unanimity, but it influences subsequent (and mostly conservative) viewers to evaluate the video more favorably. Based on our results, even if viewers are aware that the crowd is partisan and the signal is uninformative, they are significantly influenced by the social signal. The rating system may thus reinforce the audience's prior views by exposing them to their own, non-representative opinion.

One way to counteract this would be showing content to a more diverse audience. Past attempts at reducing polarization by exposing people to a more balanced set of views have had mixed success due to the adverse reactions evoked by politically misaligned content (Bail et al., 2018; Nyhan and Reifler, 2010; Wood and Porter, 2019).

As an alternative, we propose to statistically debias ratings. Platforms do not need to not show raw social signals if they know the audience is non-representative. Instead, they could use the data they collect about users to estimate what a rating would look like if it came from a more representative population, using techniques such as inverse probability weighting or methods developed to train recommender systems on biased data (Schnabel et al., 2016). Social signals with population corrections would remain informative even in politically homogeneous settings and would avoid creating false impressions of unanimity. They would tend to be more moderate, e.g. showing only 40% support in the case of the border barrier video (Figure 4). Our results demonstrate that partisans are reliably influenced by social signals even if the rating does not align with their prior opinion. More representative and moderate social ratings would influence users on both sides of the political aisle towards a common ground of more moderate views.

5.2 Limitations

Our study has several important limitations. First, expressive responding (Jakesch et al., 2019) or demand characteristics (Nichols and Maner, 2008) could be affecting our results, with participants providing answers that are not their true beliefs. However, we have analyzed the results separately for the subset of more politically extreme participants, which would be expected to be more prone to these biases, and found no shift in our results. Second, our results are limited to the context of rating systems on websites. Our design models the experience of users exposed to political claims on social media sites and our findings can not be extended to crowd situations in the offline world. Finally, we have exposed people to claims and signals once and asked for their immediate evaluations. How our findings apply to cases of complex exposures (Centola and Macy, 2007) and the self-reinforcing political processes in social networks (Muchnik et al., 2013) is to be explored in future studies.

6 Conclusion

We performed a novel online experiment to evaluate how social signals from a partisan crowd affect people's evaluations of political news claims. We find that while individuals tend to discard news claims that do not align with their political views, they are influenced by social signals independently of the crowd's political makeup. Based on the observation that people have negative reactions to politically opposed content but not to politically opposed social signals, we have proposed a design of depolarizing social rating systems.

References

Solomon E Asch. 1955. Opinions and social pressure. *Scientific American*, 193(5):31–35.

Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221.

Eytan Bakshy, Brian Karrer, and Lada A Adamic. 2009. Social influence and the diffusion of user-created content. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 325–334. ACM.

Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132.

Delia Baldassarri and Andrew Gelman. 2008. Partisans without constraint: Political polarization and trends in american public opinion. *American Journal of Sociology*, 114(2):408–446.

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):1531–1542.

Larry M Bartels. 2002. Beyond the running tally: Partisan bias in political perceptions. *Political behavior*, 24(2):117–150.

Michael A Beam, Myiah J Hutchens, and Jay D Hmielowski. 2018. Facebook news and (de) polarization: reinforcing spirals in the 2016 us election. *Information, Communication & Society*, 21(7):940–958.

Ellen Berscheid. 1966. Opinion change and communicator-communicatee similarity and dissimilarity. *Journal of Personality and Social Psychology*, 4(6):670.

Rod Bond and Peter B Smith. 1996. Culture and conformity: A meta-analysis of studies using asch's (1952b, 1956) line judgment task. *Psychological bulletin*, 119(1):111.

Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. 2011. Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological science*, 6(1):3–5.

Randall L Calvert. 1985. The value of biased information: A rational choice model of political advice. *The Journal of Politics*, 47(2):530–555.

Pew Research Center. 2017. The partisan divide on political values grows even wider.

Damon Centola and Michael Macy. 2007. Complex contagions and the weakness of long ties. *American journal of Sociology*, 113(3):702–734.

Scott Clifford, Ryan M Jewell, and Philip D Waggoner. 2015. Are samples drawn from mechanical turk valid for research on political ideology? *Research & Politics*, 2(4):2053168015622072.

Michael D Conover, Jacob Ratkiewicz, Matthew Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. 2011. Political polarization on twitter. In *Fifth international AAAI conference on weblogs and social media*.

Alexander Coppock. 2018. Generalizing from survey experiments conducted on mechanical turk: A replication approach. *Political Science Research and Methods*, pages 1–16.

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*, pages 585–592. ACM.

Chrysanthos Dellarocas. 2003. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science*, 49(10):1407–1424.

Morton Deutsch and Harold B Gerard. 1955. A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology*, 51(3):629.

Leon Festinger. 1957. *A theory of cognitive dissonance*, volume 2. Stanford university press.

Andreas Flache and Michael W Macy. 2011. Local convergence and global diversity: From interpersonal to social influence. *Journal of Conflict Resolution*, 55(6):970–995.

DJ Flynn, Brendan Nyhan, and Jason Reifler. 2017. The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*, 38:127–150.

Gallup. 2019. Americans' views on immigration issues.

William Hart, Dolores Albarracín, Alice H Eagly, Inge Brechan, Matthew J Lindberg, and Lisa Merrill. 2009. Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, 135(4):555.

Naeemul Hassan, Fatma Arslan, Chengkai Li, and Mark Tremayne. 2017. Toward automated fact-checking: Detecting check-worthy factual claims by claimbuster. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1803–1812. ACM.

Lorraine Hope and Daniel Wright. 2007. Beyond unusual? examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 21(7):951–961.

Young Eun Huh, Joachim Vosgerau, and Carey K Morewedge. 2014. Social defaults: Observed choices become choice defaults. *Journal of Consumer Research*, 41(3):746–760.

Maurice Jakesch, Moran Koren, Anna Evtushenko, and Mor Naaman. 2019. The role of source and expressive responding in political news evaluation. *Computation and Journalism Symposium 2019*.

Dan M Kahan, Hank Jenkins-Smith, and Donald Braman. 2011. Cultural cognition of scientific consensus. *Journal of risk research*, 14(2):147–174.

H.C. Kelman. 1961. Processes of opinion change. *Public Opinion Quarterly*, 25:57–78.

Ziva Kunda. 1990. The case for motivated reasoning. *Psychological bulletin*, 108(3):480.

Michael Macy, Sebastian Deri, Alexander Ruch, and Natalie Tong. 2019. Opinion cascades and the unpredictability of partisan polarization. *Science Advances*, 5(8):eaax0754.

Joseph Marks, Eloise Copland, Eleanor Loh, Cass R Sunstein, and Tali Sharot. 2019. Epistemic spillovers: Learning othersâĂŹ political views reduces the ability to assess and use their expertise in nonpolitical domains. *Cognition*, 188:74–84.

Winter Mason and Siddharth Suri. 2012. Conducting behavioral research on amazon's mechanical turk. *Behavior research methods*, 44(1):1–23.

Winter A Mason, Frederica R Conrey, and Eliot R Smith. 2007. Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and social psychology review*, 11(3):279–300.

Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.

Solomon Messing and Sean J Westwood. 2014. Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication research*, 41(8):1042–1063.

Lev Muchnik, Sinan Aral, and Sean J Taylor. 2013. Social influence bias: A randomized experiment. *Science*, 341(6146):647–651.

Austin Lee Nichols and Jon K Maner. 2008. The good-subject effect: Investigating participant demand characteristics. *The Journal of general psychology*, 135(2):151–166.

Brendan Nyhan and Jason Reifler. 2010. When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2):303–330.

Eli Pariser. 2011. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Gordon Pennycook and David G Rand. 2018. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*.

Emily Pronin, Jonah Berger, and Sarah Molouki. 2007. Alone in a crowd of sheep: asymmetric perceptions of conformity and their roots in an introspection illusion. *Journal of personality and social psychology*, 92(4):585.

Daniel M Romero, Brendan Meeder, and Jon Kleinberg. 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th international conference on World wide web*, pages 695–704. ACM.

Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *science*, 311(5762):854–856.

Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as treatments: Debiasing learning and evaluation. *arXiv preprint arXiv:1602.05352*.

Cosma Rohilla Shalizi and Andrew C Thomas. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research*, 40(2):211–239.

John Sides and Daniel J Hopkins. 2015. *Political polarization in American politics*. Bloomsbury Publishing USA.

Herbert W Simons, Nancy N Berkowitz, and R John Moyer. 1970. Similarity, credibility, and attitude change: A review and a theory. *Psychological Bulletin*, 73(1):1.

Rune Slothuus and Claes H De Vreese. 2010. Political parties, motivated reasoning, and issue framing effects. *The Journal of Politics*, 72(3):630–645.

Natalie Jomini Stroud. 2010. Polarization and partisan selective exposure. *Journal of communication*, 60(3):556–576.

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

Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science*, 359(6380):1146–1151.

Brian E Weeks and R Kelly Garrett. 2014. Electoral consequences of political rumors: Motivated reasoning, candidate rumors, and vote choice during the 2008 us presidential election. *International Journal of Public Opinion Research*, 26(4):401–422.

Thomas Wood and Ethan Porter. 2019. The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, 41(1):135–163.