Partisan politics, polarization and information processing:
An experimental study
Proposal

Abhinash Borah and Reem Qamar

September 26, 2021

The experiment aims to study the processing and diffusion of information in settings marked by polarization and partisan divides. In recent years with deeper penetration of the internet and rise in social media popularity, two trends are discernible: one, a growing tendency towards heightened social and political polarization; and two, societies being increasingly overwhelmed by viral disinformation (Dryzek et al., 2019). From misleading information regarding vaccine usage during the current pandemic to the narrative of the "stolen" US presidential election of 2020, recent events have reiterated the importance of understanding how individuals consume, process and diffuse information when such information speaks to existing patterns of polarization in society. This concern motivates the experiment and it seeks to deepen our understanding of the polarization–disinformation nexus.

The key questions in this context that the experiment will attempt to investigate are the following:

1. What are the psychological and cognitive factors responsible for biased information processing and individuals succumbing to and participating in disinformation cascades? In this regard, we are particularly interested in understanding the role that identity-based motivated reasoning plays and whether such reasoning is primarily driven by biased beliefs [System 1 channel] or by the desire to enjoy expressive/affective utility from divisive partisan communication [System 2 channel].[1]

2. What kind of social interactions produce more biased information processing and greater distortions of the truth? Are such distortions greater in environments where individuals interact only with other like-minded individuals with similar world-views (aka echo chambers) as has often been claimed (Sunstein, 1999); or are such distortions greater when individuals with opposing world-views interact and communicate?[2] Fur-

---

[1]See Pennycook and Rand (2019) and Kahan (2013), respectively, for discussions of the System 1 and System 2 channels of motivated information processing.

[2]Although communication across groups has been considered as a source of reducing conflict and bias (Gaertner et al., 1994), more recent evidence shows that interaction between diverse groups on social media can result in even more polarized viewpoints (see Bail et al. (2018); An et al. (2019); Matthes et al. (2019); Zhang et al. (2019)).

1

ther, does moderates' presence help to reduce such distortions, or do moderates end up holding more extreme biased beliefs when put in deeply partisan settings?

3. Do incentives provided for making correct inferences help in reducing biased information processing? Or, is the sway of identity-based motivated reasoning unaffected by such incentives?

4. What are the significant correlates of biased information processing? We will check for variables like response time, cognitive reflection, over-confidence, and self-control.

Our experimental set-up will use pre-existing partisan divides in society to classify participants into groups—to fix ideas think of a stylized Leftist-Centrist-Rightist classification on the ideological/partisan political spectrum. Along with this, as a theoretical benchmark, we will draw on the canonical structure of models of herding and information cascades in economics, e.g., Bikhchandani et al. (1992). In the experiment, subjects will be asked to assess the accuracy of news items in an incentivized manner. Our choice of these news items will be such that they are either Left-favoring or Right-favoring. If motivated reasoning is a salient consideration, Leftist subjects will be biased towards assigning higher probabilities to Left-favoring news being true and Right-favoring ones being false; and the opposite will be the case with Rightist subjects. These a priori beliefs about the accuracy of the news item act as their private information. Each subject will also be shown the response of other subjects who have assessed the accuracy of the same news item before them, and this serves as public information. Under the experimental design, subjects will do the assessment sequentially. This will allow us to observe the entire sequence of responses and see where this sequence converges. E.g., is there convergence to the truth (as there is under Bayesian information processing when priors have unbounded support)? If not, what can be learned about how these assessments diverge from the truth and how do these distortions relate to the structure of interactions and information? Experimental treatments will manipulate whether subjects interact in an echo chamber and observe only the responses of others with similar political leanings, or experience mixed interactions and observe responses of those with different ideological/partisan viewpoints as well. Other treatments will give subjects the choice of which information set-up they want to be in—echo chamber or mixed interactions; check how the presence of political moderates impacts outcomes; see whether partisan divergence is more significant in non-incentivized settings where the item to be assessed is a subjective opinion rather than one with an objective truth value. A complete list of all treatments and their details is provided in Appendix A.1.

A Appendix

A.1 Treatments

Political group identities. In all of the treatments, the participants are classified and divided into 3 groups – Left (L), Right (R) and Center (C) – based on their natural political

identities elicited through a survey.[3] Their group identities are made salient before the information processing task.

Task. In each experimental session, participants will be presented with a political news item. They will be asked to assign a probability $\alpha \in [0, 1]$ to the news item being true. Their monetary payoff from any such task is ₹$[\alpha X]$, if the news is true; and ₹$[(1 - \alpha)X]$, if it is false. Participants answer sequentially. Before a participant answers, she is shown the responses of participants who have answered before her. What varies across the treatments is who these other participants are, specifically, what their political identities are. In each of the control and treatments below, data on four types of news items will be elicited: $\{\text{Left-favoring } (Lf), \text{Right-favoring } (Rf)\} \times \{\text{true } (t), \text{false } (f)\}$, i.e., (Lf, t), (Lf, f), (Rf, t), (Rf, f).

In terms of measurement, what particularly interests us is the "limiting" behavior of any sequence of responses. Let $\overline{\alpha}_0$, $\overline{\alpha}_L$, and $\overline{\alpha}_R$ denote the average α-s of, respectively, the last k respondents, the last k Leftist respondents and the the last k Rightist respondents along any sequence of responses. We define information processing *distortion* associated with a news item and reference group $g \in \{0, L, R\}$ by $|\alpha^* - \overline{\alpha}_g|$, where $\alpha^* = 1$, if the news is true, and 0, if it is false. We define Left-Right *divergence* associated with a news item by $|\overline{\alpha}_L - \overline{\alpha}_R|$

C0. **Control Group**

  – Participants are not informed about the presence of groups and no group identity is made salient before the experiment.

  – Baseline results

  – The data from the control will help to check if there is convergence to the truth when group identities are not made salient, i.e., distortion is not significantly different from 0.

  – In the Control (as opposed to the treatments), the L-R-C classification of subjects will be made after they perform the inference task. We can check for distortions for Leftist, Centrist and Rightists

T1. **Echo Chamber** (Participants from Groups L and R)

  – Participants are restricted to an "echo chamber" where they only see the responses of others from their group.

  – Results from this treatment will show the impact on information processing of participants who have been exposed to the answers of a sample pool with similar priors as them.

---

[3]This classification in terms of Left, Right and Center is for illustrative purposes only. The groups have to be formed drawing on existing partisan divides of the place where the experiment will be conducted. In India, such a natural partisan divide exists around how favorably people look at Prime Minister Narendra Modi. So the three groups can be Modi disapproving, Modi approving and Modi neutral.

– This will allow us to check for motivated reasoning: For Leftists, is there zero distortion for (Lf, t) [resp., (Rf, f)] news but positive distortion for (Rf, t) [resp., (Lf, f)] news. Similar checks for Rightists. Specifically, this would be indicative of System 1 motivated reasoning.

T2. **Mixed Interactions** (Participants from Groups L and R)

– Participants can see the responses, along with the group identities, of participants from their own group as well as those of participants from the other group.

– The results, compared to T1, will show the impact on information processing of participants who have been exposed to the answers of a sample pool with priors at the opposite ends of the partisan spectrum.

– We can check whether distortions and divergence for each type of news is greater or smaller than in T1. If they are indeed greater, then we will have evidence that more than echo chambers, it is cross cutting political interactions that is more responsible for disinformation. Such a finding will also indicate that a System 2 channel of motivated reasoning may be in play. This can be also checked by comparing response times.

Treatments below not included

T3. **Information Choice** (Participants from Groups L and R)

– Participants can choose to either see responses only from their group or from both their group and the other group.

– This choice will reflect their preference to either be in an echo chamber or expose themselves to the other side as well.

– In mixed interactions, if distortions remains the same or increases (compared to T2), it will show that even those participants who want to expose themselves to a larger argument pool do not end up improving the accuracy of their inferences. If so, it raises a question about why, in the first place, did individuals want to be exposed to the answers of both sides. Is it because they derive a direct expressive utility from being in oppositional environments and opposing the other group, independent of a quest for the truth (i.e, the System 2 channel at work)?

T4. **Mixed Interactions + Moderates** (Participants from Groups L, R and C)

– Same as T2 with the addition of the centrists.

– A change in distortion and disagreement, compared to T2, will show the impact of the presence of a neutral and moderate group in a polarized setting in reducing information processing biases.

– On the other hand, if we see greater distortion amongst the moderates in T4 compared to control, it will suggest that moderates turn more partisan in polarized settings.