Demand for Online News, Inertia, and Misperceptions

Luca Braghieri, Ro'ee Levy, and Hannah Trachtman

29 January 2025

Pre-Analysis Plan

Note 1: The key difference between Version 1.1 of the pre-analysis plan and Version 1.0 is that, in September 2024, we initially attempted to run the experiment but quickly realized that recruitment costs were too high. At those costs, we would not have been able to gather a sufficiently large sample to conduct the study. As a result, we halted the September attempt before reaching the randomization stage. Since then, we have improved certain technical aspects of our survey to enhance participant retention and adjusted our approach to assigning participants to treatment arms to reduce costs.

Note 2: the key difference between Version 1.2 of the pre-analysis plan and Version 1.1. is that we edited footnote 2 to include a caveat about how to relax our recruitment targeting criteria in case recruitment costs become prohibitively high.

1 Introduction

Over the last two decades, social media has become one of the primary ways in which people consume information and access the news (Newman et al., 2023). In 2023, approximately half of U.S. adults indicated that they get their news from social media "often" or "sometimes", as opposed to "rarely" or "never" (Pew Research Center, 2023). Moreover, social media is the preferred medium of news consumption for individuals aged 19 to 29 (Pew Research Center, 2023).

The increased reliance on social media for news consumption has generated widespread concern, especially in the wake of momentous political events such as the 2016 Brexit Referendum and U.S. Presidential election (Allcott and Gentzkow, 2017; Sunstein, 2017). There are two primary reasons why the social media environment might not be conducive to healthy news consumption: first, the structure of the social network and the algorithms governing users' newsfeeds might primarily expose people to content that matches their ideology, thus generating so-called "echo-chambers" and "filter-bubbles" that promote the consumption of highly partisan news (Sunstein, 2017; Pariser, 2011). Second, the lack of editorial oversight on social media can lead to the proliferation of low-quality and fake news (Allcott and Gentzkow, 2017; Grinberg et al., 2019; Vosoughi et al., 2018).

Many of the worries about the diffusion of low-quality and highly partisan news on social media find support in the academic literature. Specifically, evidence shows that social media users are indeed more likely to be exposed to news that is in line with their political ideology (Bakshy et al., 2015; Conover et al., 2011; Halberstam and Knight, 2016), and that some social media algorithms, like the one employed by Facebook, downgrade counter-attitudinal news (Levy, 2021). Similarly, a large body of evidence documents the proliferation of misinformation and fake news on social media, especially around elections (Allcott and Gentzkow, 2017; Grinberg et al., 2019; Mocanu et al., 2015; Vosoughi et al., 2018). These worrisome findings highlight the pressing need to sharpen our understanding of the drivers of people's news choices on social media and to develop cost-effective remedies to low-quality and highly partisan news consumption on social media platforms.

In this project, we plan to run a large-scale field experiment aimed at answering two questions. First, to what extent is the consumption of low-quality and partisan news on Facebook driven by behavioral factors such as inertia in decision-making and misperceptions about the slant and quality of news outlets? Second, can interventions that target those behavioral factors lead to an increase in the quality and a decrease in the partisanship of the news users consume on Facebook? The interventions we propose build on key insights from the behavioral economics literature and involve: i) countering inertia in decision-making by requiring Facebook users to make an active choice about the news outlets that they follow on the platform; ii) countering misperceptions by providing information that might help users make such active choice more deliberately.

2 Experimental Design

The proposed field experiment will take place in the United States. The experiment will last around ten weeks and will be conducted in three overlapping phases, each of which last five weeks. For the first 10-15 days of each phase, we will deploy Facebook ads to recruit participants. The ads will not contain details about the nature of the experiment to avoid recruiting a sample with particular opinions about news on social media. We will stratify recruitment based on whether a person is browsing from a tablet/smartphone or a computer.

Upon clicking one of the ads, users will be shown a consent form that explicitly informs them that, in order to be part of the experiment, they need to log into the study via Facebook and grant the research team permission to collect the list of Facebook pages that they follow on the platform.[1] As in Levy (2021), a subset of participants will also be asked to allow the researchers to track their browsing behavior on news websites over the subsequent six months by means of a browser extension. We aim to recruit 4,000 participants, 1,100 of which provide browser tracking permissions.

Subjects who consent to participate in the experiment and who log into the study through Facebook will begin a baseline survey and will be paid for survey completion.[2] The baseline survey will contain demographic questions, questions about participants' news consumption habits, and pre-treatment observations of survey-based outcome variables. Furthermore, the baseline survey will include a set of incentivized questions eliciting participants' perceptions of the slant and quality of major news outlets, as well as their perceptions of the slant and quality of the news pages that they currently follow on Facebook. Our primary measure of slant is the one from Bakshy et al. (2015), who assign a measure of slant to hundreds of English-language outlets based on the fraction of Democrats vs. Republicans who shared articles from those outlets on social media. Following Aslett et al. (2022), our primary measure of quality relies on the ratings assigned by NewsGuard, an independent organization that rates the reliability of news and information websites along nine criteria related to a source's journalistic practices. We will also ask participants to report their bliss points in terms of the slant and quality of the news they would want to appear on their Facebook feeds.

One week after closing the baseline survey for participants in a certain recruitment wave, those participants will be invited to a midline survey. Most participants—specifically, those from whom we do not collect browsing data—will be block-randomized into a control group and five treatment groups.[3] We will stratify randomization on average baseline slant and quality of the initial portfolio of news pages followed by participants. We plan for $\sim 3,100$ individuals to complete the midline survey and be randomized.

The social media behavior of participants in the control group will not be interfered with. Members of the

[1] Whenever we use the phrase "following a news page" we refer to two distinct but very similar ways of interacting with pages on Facebook: "following" and "liking." Liking and following pages on Facebook both allow users to receive updates from those pages in their News Feed, with the key difference being that following is a newer feature that offers more control over notification settings and privacy.

[2] Early in the baseline survey, we will screen out individuals who meet any of the following criteria: i) do not reside in the United States, ii) use a VPN, iii) do not follow any news pages on Facebook, iv) have a professional user profile on Facebook, v) do not provide permissions to observe the pages they follow, iv) fail twice the understanding question about our definition of slant or our definition of reliability, vi) attempt to participate in the study more than once. If recruitment costs become prohibitively high, we will eliminate requirement iii). For the impact evaluation analysis assessing the average slant and quality of news pages followed by participants on Facebook, we will impute data for users who do not follow any news pages at baseline. Specifically, we will assign them the average baseline slant and quality of the news pages followed by participants in the same recruitment wave who follow at least one news page.

[3] One of these treatment groups, *experimenter-demand*, is deployed for robustness purposes and described in Section 5.2.

passive-choice-quality-info treatment group will only be given information about the quality of the various outlets that they follow on Facebook, as measured by the NewsGuard reliability ratings. The remaining three treatment groups will be asked to make an active choice about the news pages they follow on Facebook using a simple and user-friendly interface built into the survey. Specifically, participants in those treatment groups will be shown a set of news outlets, some of which they currently follow on Facebook and some of which they currently do not follow on Facebook, and will be asked which of them they would like to follow going forward. The outlets offered will span the spectrum of slant and quality.[4] In the *active-choice-no-info* treatment, the re-optimization exercise will not be accompanied by any information; in the *active-choice-slant-info* treatment, participants will be given information about the slant of all the news outlets in the re-optimization task, along with slant information about the outlets they currently follow, using the measure from Bakshy et al. (2015); in the *active-choice-quality-info* treatment, participants will be given information about quality according to the NewsGuard ratings rather than slant. Our survey will interface automatically with Facebook so that, if a participant follows (unfollows) an outlet within the context of our survey, the participant will automatically follow (unfollow) that outlet also on Facebook.

Those who provided permissions for us to track their browsing behavior will be randomized only to the control group and to the *active-choice-no-info*, *active-choice-slant-info*, and *active-choice-quality-info* treatment groups.[5]

Approximately four weeks after the midline survey, participants will be invited to an endline survey and receive a payment for survey completion. In the endline survey, participants will be asked a set of questions measuring exposure to news, satisfaction with one's news diet and trust in news. Furthermore, to check whether participants are satisfied with the portfolio of news pages that they chose in the midline survey, we will offer them another opportunity to re-optimize their portfolio using an interface identical to the one in the midline survey.

3 Main Outcome Variables

We plan to collect the following outcome variables:

1. The participants' incentivized perceptions of the slant and quality of various news outlets, including those they follow on Facebook. When analyzing participants' misperceptions in the context of the regression equation described in Section 5.1, we will calculate, for each participant and news outlet,

[4] We exclude outlets with very low Newsguard reliability ratings from our offers in order to avoid promoting misinformation.

[5] Since recruiting these participants is costly, we aim to maximize the study's power by selecting the treatments that are most likely to produce significant changes in the set of liked pages. We will ultimately pool the treatment groups for the analysis of browsing data, as described below.

the absolute distance between a participant's perceptions of the slant and quality of the news outlet and the news outlet's actual slant and quality. We will consider both cardinal and ordinal rankings of outlets.

2. The set of news pages that participants follow on Facebook at various points in the experiment. Specifically, after receiving a subject's consent, we will be able to collect the set of news pages that the subject follows on Facebook directly from the Facebook API. We will use that list of pages to determine the average slant and quality of the portfolio of news outlets that each subject follows on Facebook both before and at various points in time after the intervention. When analyzing slant in the context of the regression equation described in Section 5.1, we will also consider the absolute value of the average slant of the portfolio of news outlets that each subject follows on Facebook.

3. Each participant's bliss point in a two dimensional space where one dimension captures the average slant of the participant's news consumption portfolio on Facebook, measured using the scale from Bakshy et al. (2015), and the other dimension captures the average quality of the portfolio, measured using the NewsGuard reliability ratings.

4. When analyzing the degree of optimality of participants' news portfolios on Facebook in the context of the regression equation described in Section 5.1, we will also calculate, for each participant and for both slant and quality, the distance between the participant's bliss point and the actual average slant or quality of the portfolio of news pages that the participant follows on Facebook. We will also calculate a single measure of distance by first standardizing both the slant and quality measures and then calculating the Euclidean distance. We will measure distances both before and at various points in time after our intervention.

5. News consumption data from the subset of participants who agreed to let the research team track their browsing behaviors on news websites. Here we will measure both the number of visits to articles from various outlets as well as the overall slant and quality of the outlets visited. We will consider separately referrals to news websites coming from Facebook and referrals to news websites coming from any domain.

6. News exposure using self reports. Specifically, we will ask participants about the average quality and slant of the news they saw on their Facebook feeds over the last four weeks.

4 Secondary Outcome Variables

1. Survey measures about news consumption, trust in news, and satisfaction with one's news diet on Facebook. Whenever we elicit multiple survey measures related to the same outcome (e.g., satisfaction with news), we will construct an equally-weighted or inverse-covariance weighted index and analyze it using the regression equation described in Section 5.1.

5 Proposed Analysis

We will begin the analysis by documenting the role of behavioral factors such as inertia and misperceptions in online news consumption. In order to study inertia, we will consider the behavior of the random subset of participants who are asked to make an active choice about the news pages that they follow on Facebook, but who are not given any information about slant or quality (the *active-choice-no-info* treatment group). Observing a substantial degree of re-optimization among those participants can be interpreted as evidence of inertia in news consumption on social media, because the transaction cost of following a news outlet in our experiment is not meaningfully different from the transaction cost of following a news outlet in the normal Facebook environment.[6] In order to study the participants' misperceptions about the slant and quality of news outlets, including the ones that they currently follow on Facebook, we will compare the participants' incentivized perceptions of slant and quality to the true measures of slant and quality using simple t-tests.

The next piece of the analysis will focus on the portfolios of news pages that participants follow on Facebook to study the effects of our treatments on the slant and quality of such portfolios. Specifically, after the midline survey, we will compare the portfolios of news pages that participants in the control and treatment groups follow on Facebook along the dimensions of slant and quality. We will also compare, across the control and treatment groups, the distance between a participant's bliss point and the actual average slant or quality of the portfolio of news pages that the participant follows on Facebook.[7] Lastly, we will compare the portfolio of news pages that a participant follows before our intervention to the one she follows after the intervention along the dimensions of slant and quality.[8] We will perform most comparisons using the

[6] One might be concerned that the re-optimization is driven by experimentation motives. Such concern can be assuaged by noticing that: a) unfollowing an outlet at midline that one is already following on Facebook cannot easily be explained by experimentation motives; b) experimentation motives would not easily explain the behavior of participants who start following an outlet they claim to be already familiar with; c) experimentation motives would likely lead to additional re-optimization in our endline survey. Specifically, under the assumption that a month of experimentation is sufficient to understand whether one appreciates the content of an outlet, we would expect that, at endline, a meaningful fraction of participants would unfollow outlets that they started following at midline.

[7] We will briefly present some descriptive statistics about participants' bliss points as well.

[8] For ethical reasons, we do not offer low-quality outlets to participants in the active-choice treatment groups. As a result, any increase in the quality of the portfolio of news pages that a participant follows on Facebook will partly be mechanical. In order to account for such mechanical effect, we plan to perform a robustness check where we only consider medium and high quality outlets,

regression equation described in the next section.

Leveraging data collected using our browser extension, we will be able to also study: i) whether following the page of an outlet on Facebook increases consumption of articles from that outlet, and ii) whether treatment assignment affects the slant and quality of the news articles that people consume. In order to gain power, we will pool all our treatments and deploy an IV strategy (discussed in detail below) when measuring whether following the page of an outlet on Facebook increases consumption of articles from that outlet.[9]

Lastly, we will focus on the survey questions eliciting participants' satisfaction with their news diets on Facebook and overall trust in the news they see on Facebook to investigate some of the downstream implications of re-optimizing the portfolio of news pages that one follows on social media. Once again, in order to increase power, we will pool the *active-choice-no-info* treatment, the *active-choice-slant-info* treatment, and the *active-choice-quality-info* treatments. We will pool these treatments since we expect that they will substantially change the pages people follow, in contrast to the *passive-choice-quality-info* treatment, which we expect will have rather small effects. We will perform most comparisons using the regression equation described in the next section. If we find that these pooled treatments affect our downstream outcomes, we will also run regressions for each treatment separately to attempt to assess which treatment is driving the effect.

5.1 Regression Equation

For our impact evaluation, we estimate the effects of treatment assignment on our outcome variables.

Let Y_i denote one of our outcome variables, and $\mathbf{Y_i^b}$ denote the baseline value of the outcome and the baseline value of the index that includes the outcome.[10] Let $T_i^j \in \{0, 1\}$ be a set of indicators for membership in treatment group $j \in \{1, 2, 3, 4\}$ and μ_s a vector of recruitment and randomization strata dummies.[11] Let $\mathbf{X_i}$ denote a vector of controls that includes: baseline political ideology, baseline degree of reliance on Facebook for news, baseline interest in politics, wave fixed effects, and device fixed effects (desktop or mobile). We can estimate the average treatment effect of our interventions using the following regression equation:

and where we study whether re-optimization increases average quality among those outlets.

[9] We will most likely exclude the *passive-choice-quality-info* treatment from this analysis, unless we find that its effect on pages followed is statistically indistinguishable from zero. In that case, we will likely pool it with the control group.

[10] $\mathbf{Y_i^b}$ does not include the baseline value of the outcome for outcomes that do not have a baseline value. Similarly, it does not include the baseline value of the index for outcomes that are not part of an index. When Y_i is an index, $\mathbf{Y_i^b}$ is simply the baseline value of the index.

[11] As described in Section 5, we will sometimes pool treatments to increase power. The regression equation will be of course modified accordingly.

$$Y_i = \alpha + \sum_{j=1}^{4} \beta_j T_i^j + \rho \cdot \mathbf{Y_i^b} + \eta \cdot \mathbf{X_i} + \mu_s + \varepsilon_i$$

where ε_i is an idiosyncratic error term. We will estimate the equation above using OLS and cluster standard errors at the participant level. In some specifications, we will restrict the analysis only to the active choice groups, so that we can compare pages followed (or unfollowed) in the *active-choice-slant-info* and the *active-choice-quality-info* groups to those followed in the *active-choice-no-info* group. In specifications of outcomes on the *passive-choice-quality-info* and/or *experimenter-demand* treatment groups, we will restrict our focus to those for whom we do not track browsing behavior in order to keep the samples comparable.

In order to study whether following the page of an outlet on Facebook increases consumption of articles from that outlet, we will deploy the following IV strategy:

$$V_{i,k} = \alpha + \widehat{F_{i,k}} + \rho \cdot \mathbf{Y_i^b} + \eta \cdot \mathbf{X_i} + \mu_s + \varepsilon_{i,k}$$

$$F_{i,k} = \alpha + G_i + \nu_{i,k}$$

where $V_{i,k}$ captures the number of visits of individual i to outlet $k \in K$, K denotes the baseline set of outlets that we offer to all treated participants who do not already follow them, $F_{i,k}$ is an indicator for whether individual i follows outlet k, and $G_i \in \{0, 1\}$ is an indicator for whether participant i belongs to any of our primary treatment groups other than the *passive-choice-quality-info* treatment.

5.2 Robustness

Since the intervention is quite transparent and participants are not blind to treatment status, it is natural to worry about experimenter demand effects: the possibility that subjects might alter their behaviors in order to conform to what they perceive the hypothesis of the experiment to be. We plan to measure experimenter demand effects by means of an additional treatment that, but for one important aspect, is identical to the *active-choice-quality-info* treatment. Specifically, in our experimenter-demand treatment we will inform participants that we will not collect their likes in the midline survey. Furthermore, our permissions to collect likes will be removed and participants will have the option of verify that the permissions are indeed removed (and to remove the permissions themselves otherwise). Thus, when making their following decisions in the midline survey, participants should not be concerned about experimenter demand effects.

In the endline survey, we will ask subjects whether we can collect the set of pages that they follow on

Facebook. This request will come as a surprise to subjects, and we expect that most of them will accept it. Since the Facebook API reveals not only which pages a user follows, but also the last date in which the user starter following a page, we will be able to retroactively collect information about the news pages that participants followed in the midline survey.[12]

5.3 Heterogeneous Treatment Effects

We will study whether treatment effects are heterogeneous along the following margins: political ideology, the average slant and quality of a participant's initial collection of news pages on Facebook, the degree to which a participant relies on Facebook for news, political interest, gender, age, and initial misperceptions about slant and quality.

5.4 Quality flags

We will flag the following participants as low quality and, depending on their prevalence, exclude them from the analysis: participants who are in the bottom 5% – or 10% depending on data quality – of survey duration, participants who fail an attention check that we included in the survey, participants who reported consuming outlets that do not exist.

In recent recruitment efforts on Facebook, the research team found a higher proportion of suspicious responses than in the past. We will implement scammer detection algorithms and additional flagging mechanisms to identify and mitigate fraudulent activity. Most fraudolent activity should be weeded out before the midline survey in which randomization takes place.

[12] We will not be able to observe pages which were followed in the midline survey and unliked before the endline survey, but in any case we should be able to compare the set of pages followed in endline across treatments

References

Allcott, H. and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36.

Aslett, K., Guess, A. M., Bonneau, R., Nagler, J., and Tucker, J. A. (2022). News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions. *Science advances*, 8(18).

Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to Ideologically Diverse News and Opinion on Facebook. *Science*, 348(6239):1130–1132.

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Flammini, A., and Menczer, F. (2011). Political polarization on twitter. In *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. (2019). Fake news on twitter during the 2016 u.s. presidential election. *Science*, 363(6425):374–378.

Halberstam, Y. and Knight, B. (2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from twitter. *Journal of Public Economics*, 143:73–88.

Levy, R. (2021). Social Media, News Consumption, and Polarization: Evidence from a Field Experiment. *American Economic Review*, 111(3):831–70.

Mocanu, D., Rossi, L., Zhang, Q., Karsai, M., and Quattrociocchi, W. (2015). Collective attention in the age of (mis)information. *Computers in Human Behavior*, 51:1198–1204.

Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., and Kleis Nielsen, R. (2023). Reuters institute digital news report 2023.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Pew Research Center (2023). American trends panel, wave 135 september 2023.

Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

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