sociological science

Supplement to:

Barrie, Christopher, Aybuke Atalay, and Alia ElKattan. 2026. “Information Diets Are More Diverse in Attention Than in Engagement” Sociological Science 13: 864-883.

Online Supplement

Information diets are more diverse in attention than in engagement

Appendix

A1 Ethics and Consent

This study was reviewed and approved by two institutional ethics committees: University of Edinburgh (Ethics submission ID 287283) and New York University (IRB-FY2025-9473). Participants were recruited via Prolific and RIWI and provided informed consent prior to participation. No personally identifiable information was collected, and all responses were anonymized before analysis. The study adhered to all relevant institutional and national ethical guidelines.

A2 Prompt design

1. Model version: `gpt-4o-mini-2024-07-18`.

Listing 1: Prompt Design for Tweet Classification

```
from utils import classify_content, process_dataset
from config import pd, openai, datetime, os, json

## Classify static tweets dataset

function_name_tweets = "classify_tweet" # Note function name cannot contain spaces
function_description_tweets = "This task involves reading some tweets by
    ↪ individual UK MPs or users who are or will later become candidates in the
    ↪ General Election in the UK in 2024. On the basis of its content, you are
    ↪ asked to score the political stance of the text on a scale from Very Left to
    ↪  Very Right."
output_variables_tweets = {
    "type": "object",
    "properties": {
```

```
        "tweet": {
            "type": "string",
            "description": "The content of the tweet to be classified."
        },
        "code": {
            "type": "string",
            "enum": ["Very left", "Somewhat left", "Neither left nor right", "
                ↪ Somewhat right", "Very right"],
            "description": "You are asked to give your best judgment of how left or
                ↪  right wing the text is. If you think it expresses a very right-
                ↪ wing stance, label as \"Very right\" for political leaning (and
                ↪ vice versa). If you think it has a stance that is left-wing, but
                ↪ not very left-wing, label as \"Somewhat left (and vice versa).\"
                ↪ If you believe the text expresses a centrist stance or does not
                ↪ express any clear position, you should label as \"Neither left
                ↪ nor right\"."
        }
    },
    "required": ["tweet", "code"]
}
```

The US classification script used the same schema and response categories, but adapted the country/source description as follows:

Listing 2: US Prompt Adaptation for Tweet Classification

```
function_description_tweets = "This task involves reading some tweets by
    ↪ individual US legislators. On the basis of its content, you are asked to
    ↪ score the political stance of the text on a scale from Very Left to Very
    ↪ Right."
```

A3 Clone environment

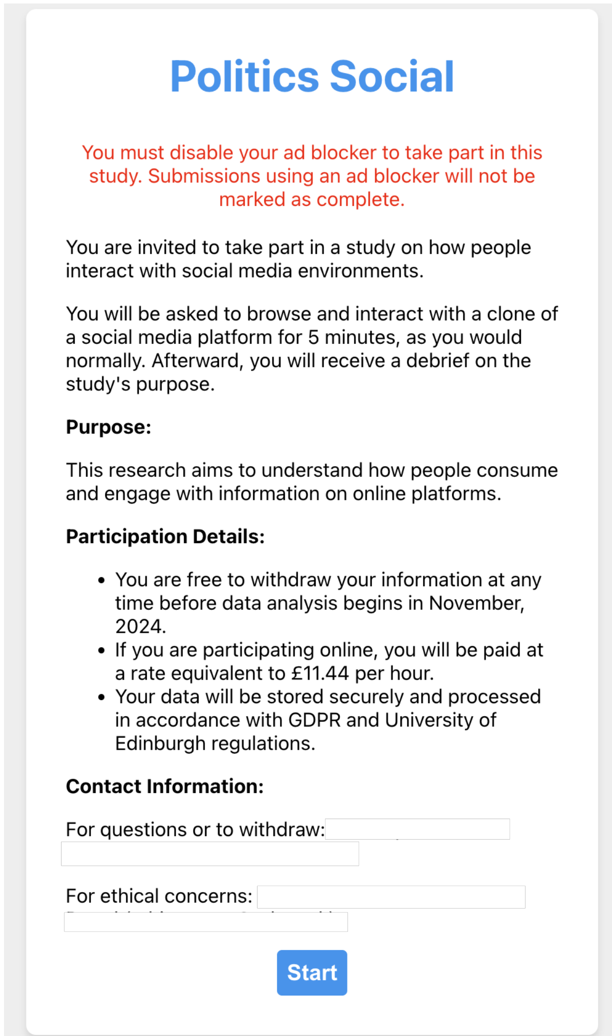

Figure A3.1: Clone environment welcome screen information and study description.

A4 Study Questionnaire

The following questions were asked to participants after they interacted with the clone social media environment.

We selected the questions based on observed relationships in the literature. First, we know that individuals with stronger ideological beliefs are more likely to select information they agree with (pro-attitudinal information) (Iyengar and Hahn, 2009; Kim and Lu, 2020; Knobloch-Westerwick and Meng, 2009). As such, we asked respondents to place themselves on a scale that runs from 0 (Very left) to 10 (Very right). Second, scholars have also found that strong partisans, i.e., those who identify strongly with a given political party, are more likely to select into pro-attitudinal content (Iyengar and Hahn, 2009; Garrett and Stroud, 2014; Rodriguez et al., 2017). We therefore asked respondents the party they most identify with, and how strongly they identify with them. Third, we know that higher political interest is generally associated with a greater likelihood of engaging with counter-attitudinal content, as demonstrated by Dahlgren (2019) and Skovsgaard et al. (2016)[1]. We measured political interest with a 5-point Likert scale ranging from 'Not at all interested' to 'Very interested.'

We also asked participants a series of questions designed to act as controls. We asked participants to report their age, gender, education, income, and frequency of social media use and include an attention check. Note that we exclude income for the US as this question gave us minimal additional insight for the UK. We also did not ask a question on diversity of news consumption (which we had originally asked for the UK study) as the question is cognitively taxing and resulted in higher survey costs for minimal additional insight.

A4.1 Political/Ideological Orientation

Question: In political matters, people talk of "the left" and "the right." How would you place your views on this scale, generally speaking?
Scale: 0 (Left) to 10 (Right)
Source: Fieldhouse et al. (2024)

A4.2 Party Identification (Party ID)

Question: Generally speaking, do you think of yourself as Labour, Conservative, or some other party?
Options (UK): Conservative, Labour, Other (with a box for entry)
Options (US): Republican, Democrat, Others (with a box for entry)
Source: Fieldhouse et al. (2024)

[1]Note though this effect might be diminished in contexts with strong public service broadcasting traditions such as the UK (Castro-Herrero et al., 2018).

A4.3 Strength of Party ID

Question: Would you call yourself a very strong supporter of this party, fairly strong, or not very strong?
Response format: 3-point scale (Very strong, Fairly strong, Not very strong)
Source: Fieldhouse et al. (2024)

A4.4 Political Interest

Question: How interested would you say you are in politics?
Response format: 5-point Likert scale ranging from 'Not at all interested' to 'Very interested'
Source: Fieldhouse et al. (2024)

A4.5 Social Media Use

Question: On which social media sites do you currently have an account or profile?
Options:

- Facebook
- TikTok
- Instagram
- LinkedIn
- Pinterest
- Snapchat
- Twitter
- YouTube
- Reddit
- Other

Response format: Yes/No/Don't know
Source: Dubois and Blank (2018)

A4.6 Age

Question: What is your age? (Age in years)

A4.7 Gender

Question: What is your gender (Male; Female; Other; Prefer not to say)

A4.8 Educational Attainment

Question: At what age did you finish full-time education?
Options:

- 15 or under
- 16
- 17-18
- 19
- 20+
- Still at school/Full time student
- Can't remember
- Not applicable

Source: Fieldhouse et al. (2024)

A4.9 Income

Question: What is your total household income before taxes? (not included for US)
Options:

- Less than £5,000
- £5,000–£9,999
- £10,000–£14,999
- £15,000–£19,999
- £20,000–£24,999
- £25,000–£29,999
- £30,000–£34,999
- £35,000–£39,999

- £40,000–£44,999
- £45,000–£49,999
- £50,000–£59,999
- £60,000–£69,999
- £70,000–£99,999
- £100,000–£149,999
- £150,000 or more
- Prefer not to say

Source: Fieldhouse et al. (2024)

A4.10 News Diet Diversity (not included for US)

Question: When looking for information about political news, issues, or elected officials, how often do you go to the following sources (either online or offline)?
Options:

- BBC News
- ITV News
- Sky News
- Channel 4 News
- Daily Mail/Mail on Sunday
- Commercial radio news
- Regional or local newspaper
- GB News
- Guardian/Observer
- Metro
- Sun/Sun on Sunday
- The Times/Sunday Times
- Al Jazeera

- Daily Telegraph/Sunday Telegraph
- Channel 5 News
- CNN

Response format: 5-point Likert scale ranging from 'Never' to 'Very often'
Operationalization: Sum of the product of Likert scale scores and political leaning score of each outlet, which went from -1 (Left) to 0 (Non-biased) to 1 (Right).
Source: Newman et al. (2024)

A4.11 Attention Check

Question: Did you see a tweet by the following individuals on the simulated social media feed you just used? (Yes/No)
Individuals (UK):

- Boris Johnson
- Theresa May
- Nicola Sturgeon

Individuals (US):

- George W. Bush
- Kamala Harris
- Barack Obama

Note: Should all be "No." Score out of 3.

A5 Post-level metrics

Each post in our clone environment displayed a "Share" button, a "Like" button, and a "Read more" clickable text, as illustrated in the main text. To operationalize the (randomized) endorsement in the regression setup, we simply added likes and retweets.

A6 User-level comparisons

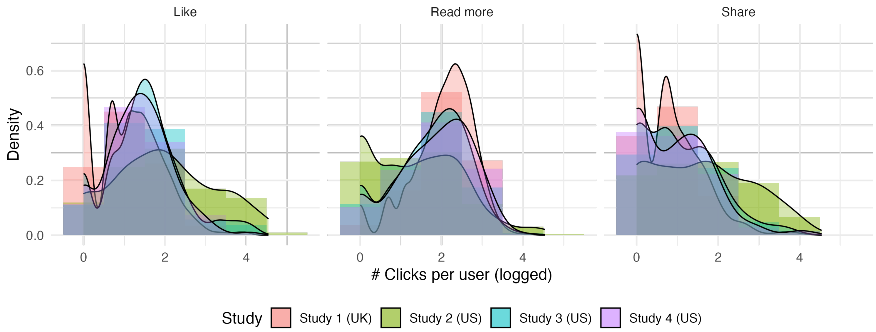

Figure A6.1: Density plots of US and UK experiment participant behaviors by click type.

A7 Robustness

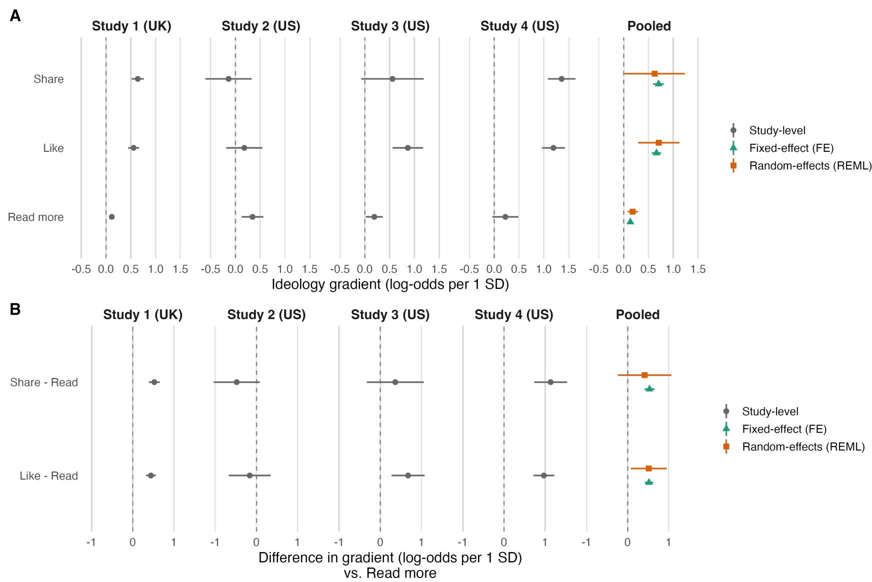

Figure A7.1: Joint fixed-effects estimates of ideology gradients across four studies and pooled meta-analytic models for attention-passing respondents. The top row shows, for each study and for the pooled estimates, the ideology–congruence gradient for each interaction type (“Read more”, “Like”, “Share”): the change in log-odds of acting on a right- versus left-aligned post associated with a one–standard deviation shift to the right in respondent ideology. The bottom row shows the corresponding *differences* between engagement and attention gradients (Like − Read more, Share − Read more), capturing how much more strongly ideology predicts partisan-congruent likes and shares than partisan-congruent read-more clicks. Gray points represent study-specific estimates; green and orange points show pooled common-effect (fixed-effect) and random-effects (REML) meta-analytic estimates, respectively. Error bars denote 95% confidence intervals.

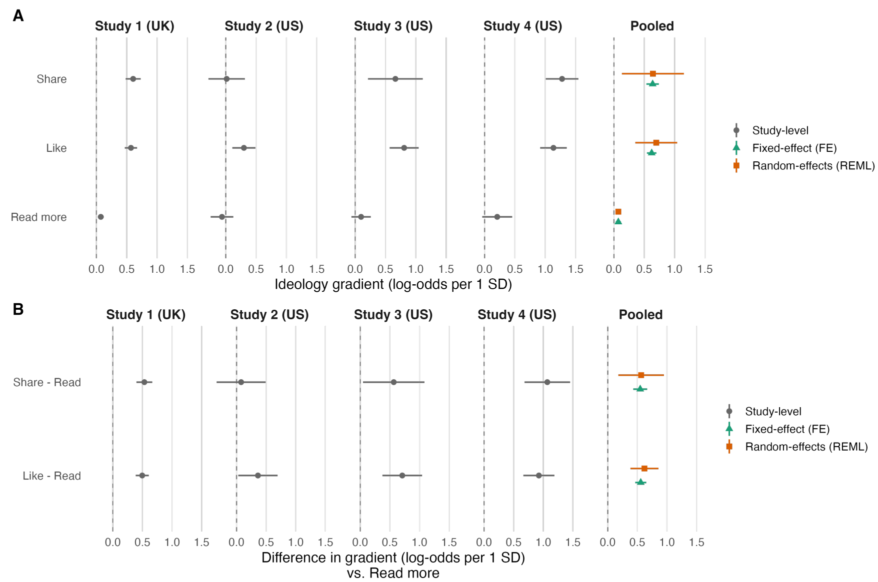

Figure A7.2: Joint fixed-effects estimates of ideology gradients across four studies and pooled meta-analytic models, expanding the panel to "non-clicker" respondents. The top row shows, for each study and for the pooled estimates, the ideology–congruence gradient for each interaction type ("Read more", "Like", "Share"): the change in log-odds of acting on a right- versus left-aligned post associated with a one–standard deviation shift to the right in respondent ideology. The bottom row shows the corresponding *differences* between engagement and attention gradients (Like − Read more, Share − Read more), capturing how much more strongly ideology predicts partisan-congruent likes and shares than partisan-congruent read-more clicks. Gray points represent study-specific estimates; green and orange points show pooled common-effect (fixed-effect) and random-effects (REML) meta-analytic estimates, respectively. Error bars denote 95% confidence intervals.

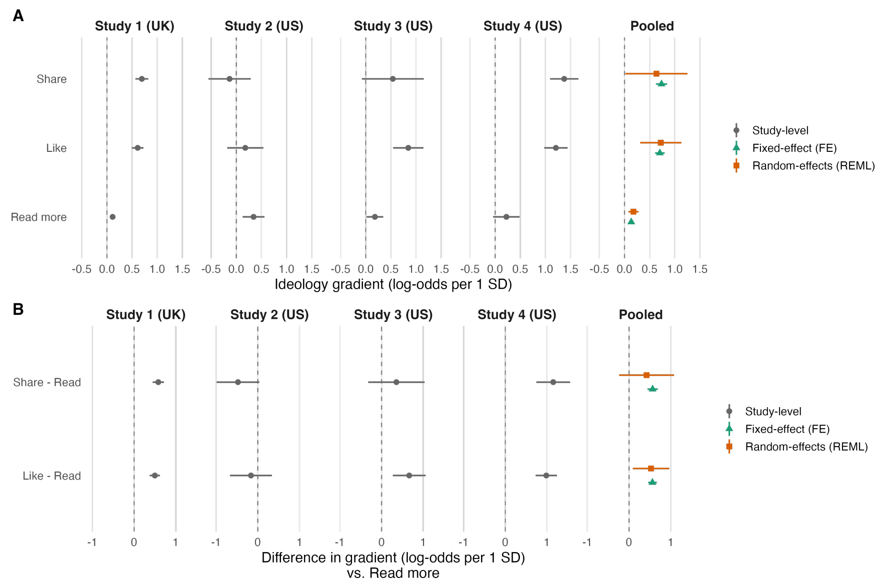

Figure A7.3: Joint fixed-effects estimates of ideology gradients across four studies and pooled meta-analytic models for attention-passing respondents, expanding the panel to "non-clicker" respondents. The top row shows, for each study and for the pooled estimates, the ideology–congruence gradient for each interaction type ("Read more", "Like", "Share"): the change in log-odds of acting on a right- versus left-aligned post associated with a one–standard deviation shift to the right in respondent ideology. The bottom row shows the corresponding *differences* between engagement and attention gradients (Like − Read more, Share − Read more), capturing how much more strongly ideology predicts partisan-congruent likes and shares than partisan-congruent read-more clicks. Gray points represent study-specific estimates; green and orange points show pooled common-effect (fixed-effect) and random-effects (REML) meta-analytic estimates, respectively. Error bars denote 95% confidence intervals.

A8 Regression tables

Table A8.1: Joint FE Logit: Baseline Clicker-Restricted Assigned-Post Panel

	Study 1 (UK) (1)	Study 2 (US) (2)	Study 3 (US) (3)	Study 4 (US) (4)
Ideology × Right (Read more)	0.076†	−0.063	0.097	0.204
	(0.024)	(0.097)	(0.085)	(0.125)
Ideology × Right × Like (Delta vs. Read)	0.487†	0.374*	0.760†	0.937†
	(0.057)	(0.177)	(0.183)	(0.134)
Ideology × Right × Share (Delta vs. Read)	0.525†	0.069	0.596*	1.078†
	(0.069)	(0.213)	(0.285)	(0.198)
Observations	677 100	151 560	137 520	84 960
Respondent FE	X	X	X	X
Tweet FE	X	X	X	X

* $p < 0.05$, †$p < 0.01$

Table A8.2: Joint FE Logit: Attention-Check Passers

	Study 1 (UK) (1)	Study 2 (US) (2)	Study 3 (US) (3)	Study 4 (US) (4)
Ideology × Right (Read more)	0.115†	0.343†	0.191*	0.225
	(0.026)	(0.112)	(0.088)	(0.134)
Ideology × Right × Like (Delta vs. Read)	0.441†	−0.165	0.676†	0.970†
	(0.060)	(0.260)	(0.205)	(0.129)
Ideology × Right × Share (Delta vs. Read)	0.525†	−0.483	0.366	1.135†
	(0.068)	(0.287)	(0.353)	(0.204)
Observations	431 400	88 560	117 360	77 400
Respondent FE	X	X	X	X
Tweet FE	X	X	X	X

* $p < 0.05$, †$p < 0.01$

Table A8.3: Joint FE Logit: Expanded Panel Including Non-Clickers

	Study 1 (UK) (1)	Study 2 (US) (2)	Study 3 (US) (3)	Study 4 (US) (4)
Ideology × Right (Read more)	0.073†	−0.063	0.098	0.208
	(0.024)	(0.095)	(0.081)	(0.126)
Ideology × Right × Like (Delta vs. Read)	0.497†	0.361*	0.709†	0.926†
	(0.056)	(0.169)	(0.171)	(0.134)
Ideology × Right × Share (Delta vs. Read)	0.533†	0.077	0.565*	1.068†
	(0.068)	(0.211)	(0.265)	(0.196)
Observations	677 100	151 560	137 520	84 960
Respondent FE	X	X	X	X
Tweet FE	X	X	X	X

* $p < 0.05$, †$p < 0.01$

Table A8.4: Joint FE Logit: Attention-Check Passers, Expanded Panel Including Non-Clickers

	Study 1 (UK) (1)	Study 2 (US) (2)	Study 3 (US) (3)	Study 4 (US) (4)
Ideology × Right (Read more)	0.113†	0.343†	0.183*	0.224
	(0.027)	(0.111)	(0.085)	(0.135)
Ideology × Right × Like (Delta vs. Read)	0.498†	−0.165	0.664†	0.984†
	(0.062)	(0.257)	(0.201)	(0.130)
Ideology × Right × Share (Delta vs. Read)	0.581†	−0.477	0.356	1.150†
	(0.070)	(0.262)	(0.345)	(0.207)
Observations	431 400	88 560	117 360	77 400
Respondent FE	X	X	X	X
Tweet FE	X	X	X	X

* $p < 0.05$, †$p < 0.01$

A9 Sensitivity analyses

A9.1 Random misclassification

We conduct two complementary sensitivity analyses to assess how fragile our main engagement–attention differences are to misclassification of tweet ideology. First, we simulate *random misclassification* of tweet labels. For a grid of misclassification probabilities p_{flip} (0–80%), we repeatedly (i) randomly flip the left/right label for that share of tweets in each study, (ii) rebuild the exposure panels, (iii) re-estimate the joint fixed-effects models, and (iv) recompute the random-effects meta-analytic estimates of $\Delta^{(\text{Like})}$ and $\Delta^{(\text{Share})}$. For each misclassification probability p_{flip}, we re-estimate, in each simulation replicate, the pooled random-effects (REML) meta-analytic engagement–attention differences $\Delta^{(\text{Like})}$ and $\Delta^{(\text{Share})}$ derived from the joint fixed-effects logit models. The sensitivity curves plot, for each p_{flip}, the average of these pooled REML estimates across simulation runs, together with average model-based 95% confidence limits obtained by averaging the lower and upper REML confidence limits across simulations. This analysis asks: if tweet partisanship were measured with substantial, but nonsystematic, error, how quickly would our congruence–engagement gradients attenuate toward zero and lose statistical significance? We find that even in the presence of a large majority of mislabelled posts, the main meta-analytic result would still not disappear.

A9.2 Worst-case misclassification.

Second, we implement a *worst-case misclassification* strategy that deliberately targets the part of the data most supportive of our conclusions. Using the baseline exposure panels, we quantify, for each tweet, a "congruence score" based on the tweet-level slope of likes and shares on respondent ideology, oriented so that positive values indicate that the tweet behaves as ideologically congruent content. We then adversarially mislabel the top fraction q of tweets on this score (e.g., 10–60%), flipping their left-/right assignments in order to maximally erode the estimated ideology-by-orientation gradients. For each q, we re-estimate the study-specific joint models, pool the resulting $\Delta^{(\text{Like})}$ and $\Delta^{(\text{Share})}$ via random-effects meta-analysis, and plot the pooled estimates and corresponding z-statistics and p-values as a function of q. This analysis asks a stricter question: even if the tweets that most strongly drive congruent engagement were systematically mislabeled in the least favorable way for our hypothesis, how much of that adversarial misclassification would be required to eliminate the pooled engagement–attention gap? We find that only at around 40% of tweets mislabelled do the results vanish.

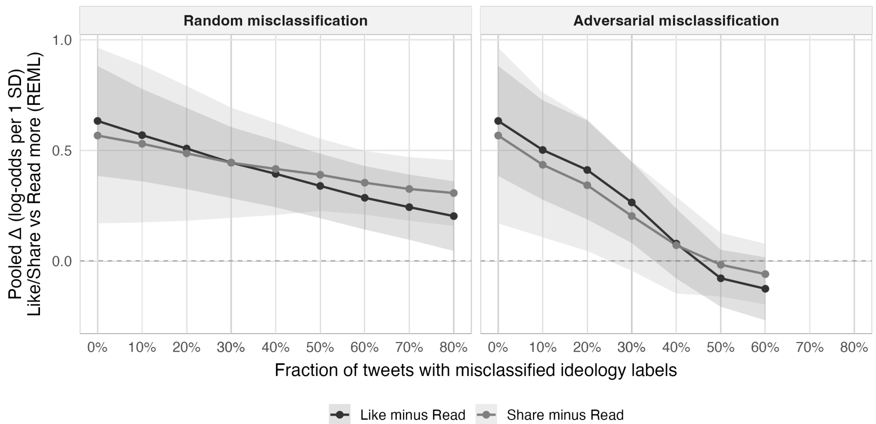

Figure A9.1: Sensitivity of pooled engagement–attention differences to misclassification of tweet ideology. Each panel plots the pooled random-effects meta-analytic gradient differences, $\Delta^{(\text{Like})}$ and $\Delta^{(\text{Share})}$ (log-odds per 1 SD, likes/shares vs. read-more), as a function of the fraction of tweets whose left/right labels are perturbed. In the left panel ("Random misclassification"), labels are flipped uniformly at random across tweets; for each misclassification rate p_{flip}, we re-estimate the model over multiple independent flip patterns and plot the mean pooled REML estimate with average model-based 95% confidence limits across simulations, obtained by averaging the lower and upper REML confidence limits. In the right panel ("Adversarial misclassification"), we instead flip the labels of the tweets that most strongly exhibit congruent engagement (based on tweet-level ideology slopes for likes and shares), and for each fraction q of such tweets misclassified we plot the pooled REML estimate together with its conventional 95% confidence interval.

References

Castro-Herrero, Laia, Lilach Nir, and Morten Skovsgaard. 2018. “Bridging gaps in cross-cutting media exposure: The role of public service broadcasting.” *Political Communication* 35:542–565.

Dahlgren, Peter M. 2019. “Selective exposure to public service news over thirty years: The role of ideological leaning, party support, and political interest.” *The International Journal of Press/Politics* 24:293–314.

Dubois, Elizabeth and Grant Blank. 2018. “The echo chamber is overstated: the moderating effect of political interest and diverse media.” *Information, Communication & Society* 21:729–745.

Fieldhouse, Edward, Jane Green, Geoffrey Evans, Jonathan Mellon, Chris Prosser, Joe Bailey, Justin Griffiths, and Sarah Perrett. 2024. “British Election Study Internet Panel Waves 1-26.” doi:10.5255/UKDA-SN-8202-2.

Garrett, R Kelly and Natalie Jomini Stroud. 2014. “Partisan paths to exposure diversity: Differences in pro-and counterattitudinal news consumption.” *Journal of Communication* 64:680–701.

Iyengar, Shanto and Kyu S Hahn. 2009. “Red media, blue media: Evidence of ideological selectivity in media use.” *Journal of Communication* 59:19–39.

Kim, Minchul and Yanqin Lu. 2020. “Testing partisan selective exposure in a multidimensional choice context: Evidence from a conjoint experiment.” *Mass Communication and Society* 23:107–127.

Knobloch-Westerwick, Silvia and Jingbo Meng. 2009. “Looking the other way: Selective exposure to attitude-consistent and counterattitudinal political information.” *Communication Research* 36:426–448.

Newman, Nic, Richard Fletcher, Craig T. Robertson, Ana R. Arguedas, and Rasmus Kleis Nielsen. 2024. *Reuters Institute Digital News Report 2024*. Oxford: Reuters Institute for the Study of Journalism.

Rodriguez, Cristian G, Jake P Moskowitz, Rammy M Salem, and Peter H Ditto. 2017. “Partisan selective exposure: The role of party, ideology and ideological extremity over time.” *Translational Issues in Psychological Science* 3:254.

Skovsgaard, Morten, Adam Shehata, and Jesper Strömbäck. 2016. “Opportunity structures for selective exposure: Investigating selective exposure and learning in Swedish election campaigns using panel survey data.” *The International Journal of Press/Politics* 21:527–546.