A Comparative Study of News Exposure and Consumption On and Off Facebook

NARDJES AMIEUR, CNRS, INRIA, Institut Polytechnique de Paris, France
SALIM CHOUAKI, CNRS, INRIA, Institut Polytechnique de Paris, France
OANA GOGA, INRIA, Institut Polytechnique de Paris, France
BEATRICE ROUSSILLON, Univ. Grenoble Alpes, GAEL, France

Social media giants like Meta, Google, and X leverage powerful algorithms to personalize user feeds, a practice now under intense public scrutiny. These algorithms can inadvertently skew the information users consume, potentially influencing political opinions and voting decisions. This raises critical questions: Do social media platforms foster misinformation and contribute to echo chambers?

To address this ongoing debate, our study directly compares news exposure on Facebook (where algorithmic influence is strong) with news consumption off-platform (where user behavior plays a larger role). Specifically, we investigate: (1) *Are users exposed to more/less misinformation on Facebook compared with their off-platform misinformation consumption?* (2) *Is news exposure on Facebook more/less diverse than off-platform news consumption?* (3) *To what extent do socio-demographic and psychological factors influence misinformation exposure on Facebook and consumption off Facebook?* (4) *Is there a relationship between socio-demographic and psychological factors and news diversity on and off Facebook?* and (5) *Is users' exposure to misinformation on Facebook correlated to off-platform news consumption?*

The longstanding biggest barrier to answering these questions has been the lack of access to data on what information users see and consume while browsing the Internet. In this paper, we use a measurement approach that asks a panel of users to donate data about the content they see online. For this, we designed a tool to collect traces of all news articles that individuals encounter on their desktop Facebook timeline and while they browse the Internet (off Facebook), along with signals about how users interact with them (e.g., clicks, time spent reading). Our tool observes content and interactions on and off Facebook on 4,149 news media domains sourced from Media Bias Fact Check and NewsGuard. Alongside the news post and article collection, we conduct surveys to gather socio-demographic and psychological data from our participants.

Our study of 123,995 news-related posts on Facebook and 70,587 news articles visits off Facebook, collected from 642 users during 12 weeks, reveals the following central findings: (1) Only a small fraction 4% of users' news consumption off Facebook is driven by news exposure on Facebook, and only 5.7% of misinformation consumption off Facebook is driven by news exposure on Facebook. (2) There is a higher prevalence of misinformation in user-received content on Facebook compared to deliberately consumed content off-platform. On Facebook, 5.9% of our users' news exposure comes from sources known for spreading misinformation, while off-platform, only 2.6% of our users' news consumption is from misinformation sources. Conversely, Facebook presents more diverse content – 22% of users received content from only one political leaning on Facebook, compared to 36% of users who consumed content from only one political leaning off-platform. (3) Several socio-demographic and psychological factors showed a statistically significant correlation with misinformation exposure on Facebook but not misinformation consumption off Facebook. (4) The proportion of misinformation consumed off Facebook emerged as a statistically significant predictor of users' exposure to misinformation on Facebook, independent of news consumption on Facebook.

Authors' Contact Information: Nardjes Amieur, nardjes.amieur@inria.fr, CNRS, INRIA, Institut Polytechnique de Paris, Palaiseau, France; Salim Chouaki, salim.chouaki@inria.fr, CNRS, INRIA, Institut Polytechnique de Paris, Palaiseau, France; Oana Goga, oana.goga@cnrs.fr, INRIA, Institut Polytechnique de Paris, Palaiseau, France; Beatrice Roussillon, beatrice.Roussillon@univ-grenoble-alpes.fr, Univ. Grenoble Alpes, GAEL, Grenoble, France.

ACM 2573-0142/2025/11-ARTCSCW359
https://doi.org/10.1145/3757540

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: Misinformation exposure, Misinformation consumption, Socio-demographic factors, Psychological factors, Content diversity.

ACM Reference Format:
Nardjes Amieur, Salim Chouaki, Oana Goga, and Beatrice Roussillon. 2025. A Comparative Study of News Exposure and Consumption On and Off Facebook. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW359 (November 2025), 34 pages. https://doi.org/10.1145/3757540

1 Introduction

Social media platforms have become a primary news source for many individuals, with nearly half of U.S. adults relying on these platforms, particularly Facebook, as their main news source [125]. This shift has raised concerns about the influence of social media on public debate and opinion [17, 73, 96, 102, 116]. Unlike traditional news sources, news on social media appears passively in feeds through friends' posts, targeted ads, and algorithmic recommendations, often without users intentionally seeking it [71, 128]. This passive exposure means social media platforms can potentially have a strong role in shaping public opinion on a large scale. High-profile incidents underscore this potential. For instance, the Cambridge Analytica scandal collected personal data from millions of Facebook users to create psychological profiles for highly targeted political advertising in the 2016 U.S. presidential election. These ads were designed to sway voter behavior, potentially impacting the outcomes of these elections [69]. Another example is the Internet Research Agency (IRA) interference, a Russian organization that used fake accounts to target U.S. voters to influence public opinion during the 2016 election [99]. More recently, during the COVID-19 pandemic, misinformation about vaccines, treatments, and the virus itself spread widely on social media, leading to public confusion and increased vaccine hesitancy [107].

Social media algorithms play a crucial role in shaping news exposure by determining what content appears in users' feeds. These algorithms, while designed to maximize engagement by showing users content that aligns with their interests, can inadvertently create echo chambers. This occurs when users are primarily exposed to information that reinforces their existing beliefs, limiting their access to diverse perspectives. Furthermore, these algorithms may promote misleading and false content because such content may generate more engagement. As a result, users encounter misinformation more frequently, contributing to its spread across social networks. Given these concerns, previous research has examined the biases introduced by social media algorithms in shaping information exposure. Studies have attempted to understand the role of these algorithms in fostering echo chambers. However, the findings have been varied, leading to an ongoing debate. One perspective challenges the notion that echo chambers are primarily created by social media algorithms [14, 28, 39, 61, 104]. Conversely, other studies have provided evidence that social media algorithms can indeed increase political polarization by limiting exposure to counter-attitudinal news [76]. In addition to examining polarization, previous research has also investigated exposure to misinformation [13, 20, 32, 65, 67, 77, 126]. Moreover, previous research, particularly in psychology, have investigated the characteristics of users who are more susceptible to misinformation [3, 10, 19, 74, 83, 90, 92, 97, 106, 113, 117, 120].

To contribute to the debate, this study employs a unique dataset we built by asking a panel of users to share data on the news content they encounter online. We developed a tool to collect the news articles users see on their desktop Facebook timelines and while browsing the Internet (off Facebook), including information about their interactions (e.g., clicks, time spent reading). Our tool tracks news content and interactions across 4,149 news media domains, sourced from Media Bias

Fact Check and NewsGuard. Additionally, we conducted surveys to gather socio-demographic and psychological data from participants. Overall, our dataset includes 123,995 news-related posts on Facebook and 70,587 news articles visited off Facebook, collected from 642 U.S. users over 12 weeks between November 2020 and February 2021 [1].

The dataset contrasts data on news exposure on Facebook, where algorithms likely play a stronger role, with data on news consumption off Facebook, where user behavior may have a greater influence. This provides a unique opportunity to determine whether social media indeed leads to greater exposure to misinformation and echo chambers compared to news consumption on media websites. Additionally, the dataset allows us to examine whether different user attributes, such as socio-demographics and psychological factors, shape news exposure on Facebook and news consumption off Facebook.

We start in Section 3 by addressing two core questions: (1) *Are users exposed to more or less misinformation on Facebook compared to their off-platform news consumption?* and (2) *Is news exposure on Facebook more or less diverse than off-platform consumption?* Our findings indicate that a majority of users 72% of users encounter equal or more misinformation on Facebook than off it, and a substantial 22% of users experience twice the fraction of misinformation on Facebook than off it. Users are also more likely to find politically balanced content on Facebook compared to their off-platform news consumption. We find that 22% of users received content from only one political leaning on Facebook, compared to 36% of users who consumed content from only one political leaning off-platform. We also observe that only 4% of users' news consumption off Facebook is driven by news exposure on Facebook, and only 5.7% of misinformation consumption off Facebook is driven by news exposure on Facebook.

We then explore in Section 4 the role of user attributes in shaping news exposure on Facebook and news consumption off Facebook. Specifically, we investigate (3) *To what extent do socio-demographic characteristics (ethnicity, religion, partisanship, community, education, age, gender) and psychological factors (authoritarian attitudes, stress and anxiety, threat management system and need for closure) influence misinformation exposure on Facebook and consumption off Facebook?* and (4) *Is there a relationship between these factors and the political balance of news on and off Facebook?* While previous works have addressed parts of these questions, they often relied on proxy data that could not capture precise exposure, they were unable to contrast factors influencing exposure with factors influencing consumption, and only focused on a subset of the factors we consider [90].

We find no strong correlation between these user attributes and misinformation off Facebook. However, we find that several factors such as "Community: Suburban and rural", "Education: High school", "Age: [25-34]", and "Partisanship: Republican" significantly correlated with increased exposure to misinformation on Facebook; raising questions about how algorithmic amplification may interact with specific user attributes to selectively surface particular types of content—most notably, misinformation. While Republican partisanship demonstrates a correlative relationship with both political balance in news exposure and consumption, alternative psychological variables exhibit a disjunctive correlation, influencing either news exposure or news consumption, but not both simultaneously.

Finally, Section 5 investigates (5) *Whether off-platform news consumption correlates with users' exposure to misinformation on Facebook*. Specifically, we examine whether users who consume misinformation off Facebook during time slice 1 tend to be exposed to more of it on the platform in time slice 2. External website tracking by Facebook, notably the Pixel [2] on news sites, could be the cause. Prior research indicates that news websites are more likely than others to include the

[1]https://github.com/Nardjes-Am/A-Comparative-Study-of-News-Exposure-and-Consumption-On-and-Off-Facebook.git
[2]A tracking tool that advertisers can install to track site visitors and facilitate ad retargeting.

Facebook Pixel [16, 22]. Our findings reveal that the proportion of misinformation consumed off Facebook emerged as a statistically significant predictor of users' exposure to misinformation on Facebook. Crucially, this association persists independently of its correlation with the fraction of misinformation consumed on Facebook. This suggests that users' news browsing habits beyond Facebook *may* influence the type of news they encounter on the platform.

Overall, our paper offers a unique perspective by contrasting news exposure on Facebook with news consumption off the platform. Our findings highlight several socio-demographic and psychological attributes that are correlated with the prevalence of misinformation and political polarization on Facebook, but not off Facebook. These findings raise important concerns and questions about algorithmic biases and the potential role of platforms in amplifying content that resonates with certain users. Addressing these questions requires more comprehensive data from online platforms and access to a broader and more representative user base. The recent adoption of the Digital Services Act by the European Union introduces a legal framework that mandates online platforms to provide data access to researchers and regulators for the purpose of assessing systemic risks [25]. We are hopeful that this data access will enable future research to answer the critical questions triggered by our findings.

2 Measurement methodology

The measurement methodology employed to produce the dataset we analyze in this study was developed for a broader project investigating risks associated with online platforms, such as targeted advertising and information exposure. Full details of the methodology are presented in our prior works [23, 24].

This section presents an overview of key components of this methodology, with a level of detail we think is necessary to understand and put into context the results presented in the paper.

At a high level, the methodology consists of a mixed-method approach to carefully and reliably [3] collect data on: (1) users' precise *exposure* to and *interaction* with news-related posts on Facebook, (2) users' news *consumption* behaviors on media websites (i.e., off Facebook), and (3) survey-based information about the socio-demographic and psychological profiles of users. This data collection is facilitated by a desktop browser extension that users have to install on their computers. Our previous study has only analyzed on Facebook data and has not analyzed consumption behavior off Facebook and survey-based data.

Our data collection procedures have undergone review and approval by our Institutional Review Board (IRB) and comply with GDPR, ensuring the protection of participants' rights, privacy, and anonymity.

2.1 Respondents

Our analysis is based on data collected from 642 users recruited on Prolific [98] between November 2020 and February 2021 during and immediately after the U.S. Presidential Election. The user sample is not representative; however, to reduce gender, political, and racial biases, users were recruited across 48 different states, across both urban and rural areas, and from various ethnic groups and genders. Among them, 65% identify as female and 35% as male, which deviates from the gender distribution observed among U.S. Facebook users (45% male and 55% female) [111] and the overall U.S. population (49% male and 51% female) [112]. In terms of racial background, 79% of users identify as white and 21% as non-white, compared to the respective percentages of 76% and 24% in the U.S. population [110]. For more socio-demographic statistics, please refer to Table 6 in

[3] Because platforms are adversarial to data collection, multiple validation steps were necessary to ensure the reliability of the data obtained.

Appendix A. Users were asked to install our browser extension and keep it active for 6 weeks, as well as to fill out surveys during the study period. Users received payment once at the beginning, upon installation of our browser extension, and once at the end of the study. Some users did not fill out all the surveys and did not keep the browser extension active, resulting in reduced sample sizes. Table 1 presents the filtered number of users for each data type.

Table 1. Number of users for which we have collected various data types.

Survey/Data type	Number of users
Data off Facebook	467
Data on Facebook	416
Socio-demographic survey	426
Authoritarian attitudes survey	441
Stress and anxiety survey	398
Threat management systems survey	394
Need for closure survey	305
Users completing all surveys	304

2.2 News exposure and consumption data

Our analysis draws on data about users' interaction with news content on Facebook and various news media websites off Facebook.

List of news outlets. The identification of news content—both on and off Facebook—relies on a predefined list of news sources. This list was compiled using data from Media Bias Fact Check [82] and NewsGuard [88], two organizations that list and evaluate news media outlets. Media Bias Fact Check lists 2,062 news sites, and NewsGuard includes 2,939 news sites, resulting in a combined total of 4,149 news outlets. Both Media Bias Fact Check and NewsGuard are widely used in previous research [100, 103, 129], and both follow publicly available methodologies with well-defined criteria in evaluating news sources [21, 89].

Media Bias Fact Check and NewsGuard evaluate the political bias and factual accuracy of the news sources they list. We leverage their data to assign each source in our list a quality label (misinformation or not) and a political bias label (left, center, or right). Overall, 64% of the outlets are classified as Center, 20% as Left (7% Far-Left and 13% as Slightly Left), and 17% as Right (10% Far-Right and 7% Slightly Right). In addition, based on ratings from Media Bias Fact Check and NewsGuard, we have 456 news websites with a misinformation quality label and 2,309 with a factual quality label. Further details on our labeling procedures are provided in Appendices B.1 and B.2.

Note that these labels are applied at the news source level. Hence, we assign the same value to all news items originating from a given source. We acknowledge this as a limitation of our work, as well as a limitation of similar research [6, 46, 52, 62, 103, 129]. Currently, there is no widely reliable methodology for labeling individual news items for misinformation and political bias, and existing approaches tend to suffer from excessively high false positive rates.

News consumption off Facebook. Our dataset includes users' web browsing histories on news media websites. It consists of the URLs of all news articles users visited through their desktop browsers and the time spent on each article. Among the 642 users, we identified 13 exhibiting abnormal behavior, defined as visiting more than 100 articles per day. These users were excluded from the analysis. In total, the dataset contains 70,587 news article URLs visited across 1,882 news domains. In median, each user accessed 38 news articles ($\overline{x} = 112, SD = 251$) and spent 40 seconds on each news article ($\overline{x} = 53, SD = 47$). Table 8 in Appendix B.4 presents the top 20 news domains that were most visited by our users.

For the regression analysis, we applied an additional activity threshold to ensure meaningful data. Specifically, we included only users who visited at least 10 URLs. This filter yielded a final sample of 69,862 URLs received by 467 users with sufficient off Facebook browsing activity.

News exposure on Facebook. Our dataset contains all the news-related posts that users received on Facebook through their desktop browsers. A Facebook post is considered news-related if it meets one of the following criteria: (a) it was published by a news media Facebook page, or (b) the landing URL of the post directs to a news media website. To identify posts published by news media Facebook pages, we relied on a list of Facebook pages associated with the news domains in our dataset. We built this list in a prior study [24], by querying Facebook for pages that had verified external websites and matched one of the news domains in our list. Overall, we compiled 4,323 Facebook pages associated with the 4,149 news domains.

In addition to the news-related posts, our dataset includes their visibility duration on users' screens, providing insights into the exposure time each post received. Appendix B.3 describes how we compute posts' visibility times.

In total, our dataset comprises 123,995 news-related posts received by 467 users. Each user received a median of 50 news-related posts ($\overline{x} = 266, SD = 707$). The median visibility time of a news-related post on a user's screen is 6 seconds ($\overline{x} = 7, SD = 5$). Table 7 in Appendix B.4 presents the top 20 Facebook news pages with the most received news posts in our dataset.

For the regression analyses, we restricted the sample to users who received at least 10 news-related posts on Facebook. This final subset included 123,708 posts received by 416 users.

2.3 Survey data

We conducted five surveys to collect self-reported socio-demographic characteristics and psychological attributes from respondents who have been shown to correlate with susceptibility to misinformation [83, 106, 117].

Socio-demographic survey. This survey gathered respondents' background information, including their age, gender, education level, income, community, ethnicity, religion, and political partisanship. Additionally, we obtained respondents' ethnicities and political orientations from Prolific. We found that 96% of their self-reported data matched Prolific's data, providing confidence in the accuracy of our survey responses.

Stress and anxiety survey. This survey evaluates an individual's stress and anxiety levels [78]. A higher score on this survey indicates greater levels of stress and anxiety. Respondents answered a series of statements regarding their feelings and in different situations, using a Likert scale ranging from "not at all" to "very much". We included this psychological test since previous research has shown a correlation between stress and anxiety and misinformation as well as political balance [41, 118, 122].

Authoritarian attitudes survey. This survey consists of questions designed to measure a person's agreement or disagreement with authoritarian principles and values [84]. A higher score is a stronger endorsement of authoritarian principles. The items in the survey assess various dimensions of authoritarianism, including conformity, obedience to authority, and skepticism towards out-groups. Respondents were presented with a series of statements related to these dimensions and were asked to indicate their level of agreement on a Likert scale ranging from "strongly oppose" to "strongly favor". We included this psychological test since previous research has shown a correlation between authoritarian attitudes and susceptibility to misinformation [7].

Threat management systems survey. This survey aims to assess users' understanding of potential threats, knowledge of threat management strategies, and engagement with relevant systems and protocols [87]. Respondents evaluated their feelings and beliefs regarding susceptibility to illness and societal threats by responding to a series of statements on a Likert scale ranging from "strongly disagree" to "strongly agree." A higher score reflects greater concern or belief in the threats presented.

Need for closure survey. This survey aims to assess an individual's preference or motivation for seeking closure, which reflects their preference for certainty and aversion to ambiguity [101]. A higher score indicates a stronger need for closure. Respondents evaluated their feelings about uncertainty and decision-making by responding to a series of statements on a Likert scale ranging from "strongly disagree" to "strongly agree". We included this psychological test since previous research has shown a correlation between need for closure and misinformation [127].

The full list of survey questions can be found in the GitHub repository [4]. We distributed the surveys to users in four waves, with a two-week interval between each wave. The first wave included the socio-demographic and stress and anxiety surveys. The second wave consisted of the authoritarian attitudes survey. The third wave included the threat management systems survey, and in the fourth wave, we sent the need for closure survey. The distribution of psychological scores among participants is presented in Figure 7 in Appendix A.

Table 1 presents the number of respondents who completed each survey. The table shows that each survey was completed by a subset of participants. Hence, for the analyses involving specific psychological traits, we included only those users who completed the corresponding survey. In total, 304 users completed all the surveys.

2.4 Measures

We focus on analyzing two variables related to news exposure and two variables related to news consumption:

A. Exposure to misinformation: This variable captures the fraction of news content from misinformation sources that each user was exposed to on Facebook. It is computed as the ratio of misinformation posts to the total number of news posts encountered by the user.

B. Consumption of misinformation: This variable measures the fraction of news content from misinformation sources that each user consumes on news media websites. It is computed as the ratio of misinformation articles to the total number of news articles visited by the user.

C. News exposure's political balance: This variable measures the ratio of left-leaning versus right-leaning sources a user is exposed to on Facebook. It is computed as the fraction of posts from left-leaning to right-leaning sources (or vice versa, depending on which is higher). It ranges from 0 to 1, where 0 indicates that the user was exposed exclusively to sources from one political leaning, 0.5 means the user was exposed to twice as many posts from one leaning compared to the other, and 1 represents equal exposure to posts from both left- and right-leaning sources. A higher value indicates a more balanced news diet.

D. News consumption's political balance: This variable is similar to the previous one but focuses on the political balance of news that a user consumes on news media websites. It also ranges from 0 to 1, reflecting the proportion of left-versus right-leaning sources users consume. Again, a higher value signifies a more balanced news consumption.

[4] https://github.com/Nardjes-Am/A-Comparative-Study-of-News-Exposure-and-Consumption-On-and-Off-Facebook.git

3 Characterization of news exposure and consumption

Previous studies have explored whether social media platforms promote a balanced or polarized news diet and whether they contribute to misinformation [3, 14]. Due to limited access to precise data on news exposure on social media platforms, previous research relied on proxy data, such as news articles shared on Twitter or identifying news articles accessed via social media in browsing histories, leading to mixed findings. For instance, some studies suggest that social networks enhance political diversity [14, 39, 104], while others indicate they may contribute to political polarization [28, 76].

Our dataset offers a unique opportunity to examine Facebook's role in shaping users' exposure to misinformation and the political balance or polarization of the content they encounter. The dataset includes actual news posts received by users on Facebook and enables a direct comparison of news encountered both on and off the platform, which, to our knowledge, has not been explored.

3.1 Quantity of news on and off Facebook

This section explores the quantity of news received by users on Facebook and the quantity of news accessed off Facebook. We look at both the frequency of news items and the time spent on them. Figure 1 presents the empirical cumulative distribution function (ECDF) of the median daily number of news posts users receive on Facebook compared to the median daily number of news articles they visit off Facebook per active day. The figure shows that, on the median, users are exposed to 3 news posts per active day on Facebook ($\overline{x} = 5, SD = 9$), while they consume a median of 2 news articles off Facebook per active day ($\overline{x} = 2, SD = 3$).

Next, we examine the extent to which news consumption off Facebook is driven by news exposure on the platform. To conduct this analysis, we systematically review all URLs of news articles consumed by users and identify those that were accessed from Facebook. This identification is made possible by specific query parameters added by Facebook to the URLs, indicating that they were accessed through the platform. Precisely, we search for the parameters "utm_source=facebook" [34] or "fbclid" [35] within the URLs. Our findings reveal that only 4% of users' news consumption off Facebook is driven by news exposure on Facebook. This figure is considerably lower than the approximately 14% of news website visits estimated by a 2024 Ofcom study to be driven by Facebook [131]. Complementing these results, a separate study [104] found that the probability of visiting a news site after Facebook was approximately 2.7% in 2018. This suggests that while Facebook plays a role in guiding users to news content, its influence on overall news consumption is relatively limited.

Figure 2 presents the median time users spend on looking at a news-related post on Facebook and looking at a news article off Facebook. The figure shows that users spend a median of 6 seconds on a news-related post on Facebook ($\overline{x} = 7, SD = 5$), with 95% of users spending between 1 and 15 seconds per news post. This behavior is well-documented in prior research and is often referred to as rapid scrolling, where users quickly scan headlines and visuals, spending only a few seconds on most posts [9, 37, 57]. While 6 seconds is not a lot, our previous research [24] showed that users spend more time on news-related posts compared to non-news-related posts on Facebook, suggesting that these news items effectively capture users' attention. In contrast, users spend a median of 40 seconds reading a news article off Facebook ($\overline{x} = 53, SD = 47$), with 95% spending between 0 and 143 seconds per news article.

Takeaways: Our findings reveal a distinct pattern of news engagement: users encounter a higher volume of news content on Facebook, but commit significantly less time per piece compared to off-platform news websites. Intriguingly, only 4% of news articles consumed directly on news media websites were accessed via Facebook. This challenges the perception of Facebook as a primary news source [125], demonstrating it accounts for merely a small fraction of users' actual news article consumption. These divergent exposure patterns—brief, high-volume encounters on Facebook versus extended engagement with a curated selection of articles off-platform—raise critical questions regarding their differential impact on users' beliefs and knowledge. Specifically, does this brief exposure on Facebook have the same impact on users' beliefs and knowledge as longer reading on news websites? How long does a user need to engage with a Facebook post to recall specific details of the content?

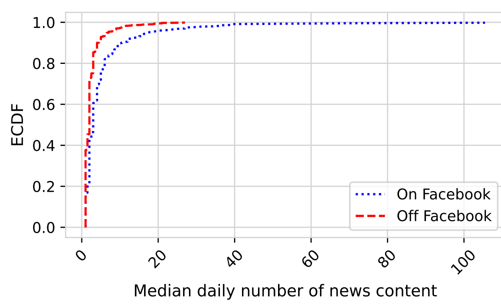

Fig. 1. ECDFs of the median daily number of news posts received on Facebook and news articles consulted off Facebook.

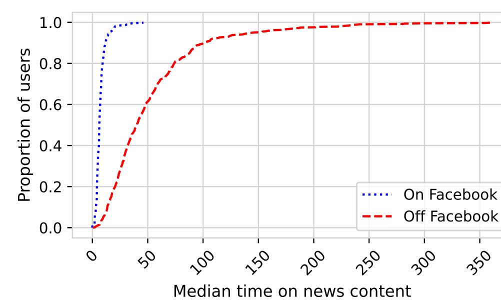

Fig. 2. ECDFs of the median time spent looking at news posts received on Facebook and news articles consulted off Facebook.

3.2 Prevalence of misinformation on and off Facebook

In this section, we examine the extent of *exposure to misinformation on Facebook and consumption of misinformation off Facebook*. We find that 5.9% of news exposure posts and 2.6% of news consumption articles in our dataset originate from misinformation news sources.

Figure 3 presents the ECDF of the fraction of misinformation users are exposed to on Facebook and the fraction of misinformation they consume off Facebook across all users. For each user, the misinformation fraction is computed as the ratio of news posts (or news articles) coming from sources known to spread misinformation to the total number of news posts encountered by the user (or news articles consumed by the user). The figure shows that 63% of users were exposed to misinformation at least once on Facebook, while 46% of users consumed misinformation at least once off Facebook. Furthermore, for the median user, the fraction of misinformation in their news exposure is 2% ($\overline{x} = 6\%, SD = 12\%$), while the fraction of misinformation in their news consumption is 0% ($\overline{x} = 3\%, SD = 8\%$). Finally, while misinformation posts represent a small fraction (less than 20%) of news exposure on Facebook for the majority of users 92%, a small subset 1% encounters a much higher proportion 60% of misinformation. Moreover, 53% of users who never consumed misinformation off Facebook were exposed at least once to it on the platform. 72% of users encounter equal or more misinformation on Facebook than off it, and a substantial 22% of users experience twice the amount of misinformation on Facebook than off it.

Figure 4 presents, for each user, the fraction of misinformation in their news exposure on Facebook compared to their consumption off Facebook. The plot's x-axis represents the average misinformation fraction between on and off Facebook, and the y-axis represents the difference between the fraction of misinformation on Facebook minus the fraction of misinformation off Facebook. The figure shows a subtle but consistent trend towards higher misinformation fractions on Facebook compared to off Facebook, as indicated by a mean positive difference of 0.03.

Takeaways: These findings provide evidence that users are exposed to more misinformation on Facebook than off Facebook. While the median user encounters only 2% of news-related posts from misinformation sources, suggesting a low individual rate, a substantial 22% of users experience twice the amount of misinformation on Facebook than off it. This disparity may be attributed to platform design features such as algorithmic amplification [36, 81], social sharing mechanisms [42, 119, 121], and content virality [59]. These features often prioritize content that drives high engagement, inadvertently favoring misinformation and hate speech [58].

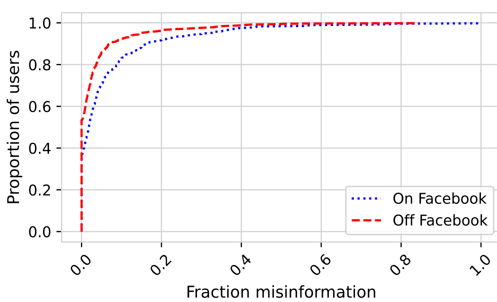

Fig. 3. ECDFs of the fraction of misinformation content on and off Facebook.

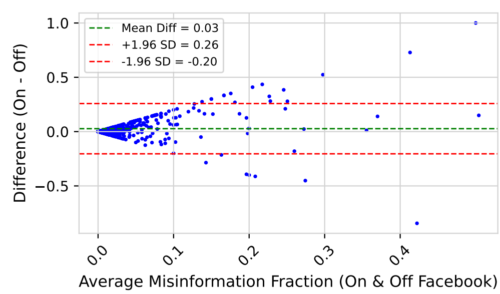

Fig. 4. Bland-Altman plot comparing the fraction of misinformation on vs. off Facebook across users.

3.3 Political balance on and off Facebook

This section analyzes the political balance of users' news exposure on Facebook and their news consumption off Facebook, as defined respectively by variables C and D in Section 2.4. Note that we do not filter news items based on their content (e.g., whether they are explicitly political), but rather based on the political bias of their source outlets. Specifically, we include only news content originating from sources that have been classified as either left-leaning or right-leaning by Media Bias Fact Check and NewsGuard. We find that 43% (53,564) of news exposure posts and 36% (25,400) of news consumption URLs in our dataset come from left-leaning sources. In contrast, only 11% (13,979) of news posts and 6% (4,262) of news articles are from right-leaning outlets.

Figure 5 presents the cumulative distribution of the political balance variables for both on and off Facebook. The figure shows that the median political balance value for news exposure on Facebook is 0.13 ($\overline{x} = 0.22, SD = 0.26$), while the median political balance for off Facebook news consumption is 0.08 ($\overline{x} = 0.16, SD = 0.22$). Recall that a political balance value of 0 indicates that a user encountered news from only one political leaning (either only from left-leaning sources or only from right-leaning sources), and 0.5 means the user was exposed to twice as many posts from one leaning compared to the other. The figure reveals that 22% of users were exposed to content from a single political leaning on Facebook, whereas 36% consumed content from a single political leaning off Facebook. Finally, 4% of users have almost perfectly balanced news consumption (political balance value >= 0.7).

Additionally, Figure 6 shows, for each user, a comparison between their political balance in news exposure on Facebook and off Facebook. The figure shows that, for most users 52%, the political balance on Facebook is higher than the political balance off Facebook. More intriguing, the figure shows that some users are exposed to politically balanced content on one side (either on or off Facebook) but not on the other.

Takeaways: Our data indicates that users encounter a greater degree of political balance in news content on Facebook compared to their off-platform news consumption. This supports prior studies [39, 104] and suggests that Facebook can play a role in promoting exposure to a wider range of political perspectives.

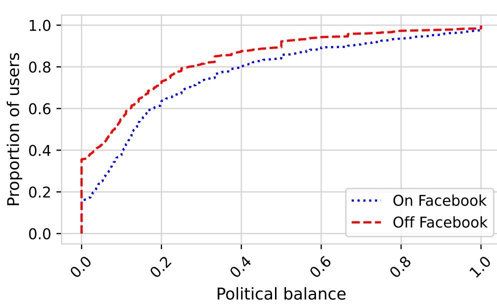

Fig. 5. ECDFs of political balance on and off Facebook.

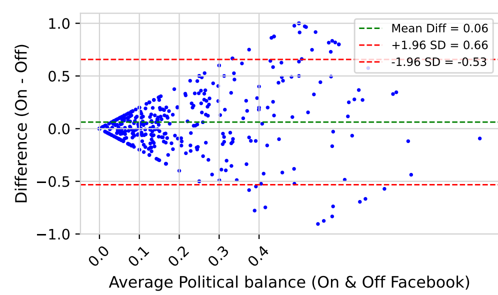

Fig. 6. Scatter plot of the political balance on and off Facebook across users.

This section showed differences in misinformation prevalence and political balance across users, both on and off Facebook. The next section investigates extent to which socio-demographic factors and psychological factors drive these disparities.

4 Predictors of news exposure and consumption

This section examines how users' socio-demographic characteristics and psychological factors affect the four dependent variables: exposure to misinformation, consumption of misinformation, news exposure's political balance, and news consumption's political balance (see Section 2.4). To do this, we employ Generalized Linear Models (GLMs) on a dataset of 95,001 news-related Facebook posts received by 264 users (for the exposure analysis), and 47,368 news articles visited by 244 users (for the consumption analysis), for which we have all socio-demographic and psychological data as well as enough activity both on and off the platform (see Section 2.3).

Table 2 presents the estimated coefficients and their associated robust standard errors (in parentheses) for the correlates included in the best-fitting Generalized Linear Model for each of the four dependent variables. We conducted a series of steps to ensure that the GLMs we employed fit the data distribution and provide reliable and interpretable results. First, we adopted a Gaussian distribution since our four dependent variables are continuous. Additionally, we tested beta regression [26], beta-binomial [55], and zero-inflated beta models [93], as our outcome variables represent fractions that can be exact zeros in some cases. To compare model performance and assess the statistical significance of the differences, we used the likelihood ratio test (LR test) [124] and ANOVA [109]. Second, we explored various link functions—log [1], logit [1], probit [1], and output transformation functions such as square root, and cube—to normalize the response variable and tested multiple predictor combinations for robust convergence. Third, we assessed the model fit using comprehensive diagnostic tools, including the DHARMa package [56], residual analysis plots, Q-Q plots, and information criteria (AIC, BIC) [123]. Fourth, we performed statistical tests (F-test [38], Kolmogorov-Smirnov test [8], outlier analysis, dispersion tests [85], and Breusch-Pagan test [18]) to validate model assumptions. We performed multicollinearity test [27] as well to get stable coefficient estimates and determine the true effect of each predictor. Fifth, we utilized weighted least squares regression [2] and calculated robust standard errors [60] to address potential heteroscedasticity. Finally, we computed the coefficient of determination (R-squared) [95] to quantify the model's explanatory power. Please refer to Appendix C for the results of the various tests. Further details concerning the selected models and their performance metrics are available in the project's GitHub repository [5].

4.1 Predictors of misinformation

4.1.1 Exposure to misinformation on Facebook. The first column of Table 2 examines user attributes influencing the prevalence of misinformation in users' news exposure on Facebook. The table presents in bold and marks with an $*$ correlates that achieved statistically significant values. The table shows that several socio-demographic categories correlate positively with exposure to misinformation on Facebook: "Community: Suburban and rural" (0.34,0.37), "Education: High school and less" (0.91), "Age: [25-34]" (0.51), and "Partisanship: Republican" (with a coefficient of 0.40), indicating that Republicans are more likely to encounter misinformation on Facebook compared to other groups. This finding aligns with previous research on patterns of misinformation exposure across social media platforms [48, 54, 115].

Regarding psychological factors, the table reveals a significant negative correlation (-0.32) between authoritarian attitudes and misinformation exposure on Facebook. This indicates that users with higher authoritarian attitude scores tend to encounter less misinformation content on the platform. This finding contrasts with previous psychological research by Altemeyer [7], which reported that individuals with high authoritarian attitudes are generally more susceptible to misinformation (though not specifically in the context of social media). This difference highlights the unique dynamics of news exposure on social media platforms. One possible explanation is that algorithmic personalization and platform-specific biases shape which users are exposed to misinformation, potentially reaching groups differently than traditional media. Furthermore, the way content is presented on Facebook may resonate more with users who are not typically seen as vulnerable to misinformation.

In contrast to previous research by [54, 132], which found that lower trust in mainstream media correlated with a higher likelihood of engaging with fake news content on Facebook, we did not capture any significant correlation in our analysis.

Not all forms of news exposure are the same. Some are shaped by algorithms—for example, targeted exposure, which occurs when third parties pay ad platforms to deliver specific news content to particular audiences, or algorithmic exposure, where content is shown based on the platform's prediction of user interest. Other types are driven by user behavior, such as incidental exposure—news posts appearing in users' feeds because they were shared by friends or communities—and selective exposure, where users intentionally follow news outlets on platforms like Facebook [24]. We investigate which factors most strongly influence each type of exposure.

Table 3 examines user attributes influencing the prevalence of misinformation in users' news exposure on Facebook across the different types of exposure. The table shows that individuals living in rural areas are significantly more likely to encounter misinformation through targeted and algorithmic means. This group is often perceived by others as the most vulnerable to misinformation because of assumed lower digital literacy skills [47]. In contrast, suburban residents are more likely to experience selective exposure, indicating that information ecosystems exhibit meaningful differences across rural, suburban, and urban areas.

In addition to geographic trends, educational background emerges as another key factor shaping exposure types. Users with higher education levels are less likely to encounter misinformation

[5]https://github.com/Nardjes-Am/A-Comparative-Study-of-News-Exposure-and-Consumption-On-and-Off-Facebook.git

Table 2. Coefficients for correlates of the prevalence of content from sources known for spreading misinformation and the political balance on and off Facebook.

		On Facebook: misinformation	**Off Facebook: misinformation**	**On Facebook: political balance**	**Off Facebook: political balance**
Ethnicity					
	White	-0.00(0.16)	**0.04* (0.02)**	0.37(0.60)	0.53(0.81)
Religion					
	Non-religious	0.17(0.16)	-0.03 (0.02)	-0.16(0.57)	-0.4(0.66)
	Other religion	-0.07(0.20)	-0.01(0.02)	-0.18(0.63)	-0.54(0.83)
Partisanship					
	Republican	**0.40*(0.20)**	**0.07*(0.03)**	**1.88**(0.60)**	**1.60*(0.71)**
	Other	0.04(0.26)	0.00(0.03)	0.36(0.86)	**2.26** (0.81)**
Community					
	Suburban	**0.34*(0.14)**	-0.01(0.02)	0.32(0.47)	0.30(0.59)
	Rural	**0.37†(0.2)**	-0.02(0.02)	0.06(0.67)	-0.70(0.94)
Education					
	Undergraduate	0.32(0.25)	0.01(0.02)	0.32(0.52)	-0.50(0.57)
	High school	**0.91**(0.30)**	0.02(0.02)	0.90(0.63)	-0.62(0.86)
	Community college	-0.12(0.33)	0.07(0.04)	-0.18(0.84)	-1.47(1.06)
Age					
	[25-34]	**0.51†(0.29)**	-0.04(0.02)	0.52(0.66)	0.65(0.88)
	[35-44]	-0.20(0.31)	0.00(0.03)	1.02(0.71)	1.03(0.97)
	[45-54]	0.26(0.35)	0.00(0.03)	0.10(0.82)	0.17(1.21)
	More than 55	-0.02(0.39)	-0.04(0.03)	1.34(0.86)	0.67(1.12)
Gender					
	Male	-0.00(0.21)	0.03(0.02)	0.35(0.42)	0.04(0.52)
Psychologica					
	Authoritarian	**-0.32**(0.12)**	-0.01(0.01)	-0.40(0.36)	0.60(0.48)
	Need for closure	0.01(0.12)	0.01(0.02)	-0.33(0.41)	**-1.32*(0.52)**
	Stress	-0.48(0.43)	-0.02(0.03)	**3.17***(0.95)**	-0.75(1.23)
	Threat manag	0.17(0.14)	0.00(0.02)	**-1.03**(0.46)**	0.33(0.57)
R2 score		21%	23%	20%	19%

*** p<0.001; ** p<0.01; * p<0.05; † p<0.10

through selective exposure since educated users are more discerning of false news; this finding is supported by previous research [4]. While individuals with lower educational backgrounds are more susceptible to incidental, targeted, and algorithmic exposure since they are more likely to accept misinformation [63].

Psychological traits also contribute to variations in how users encounter misinformation. Those with high levels of authoritarianism are significantly less exposed to misinformation incidentally. In contrast, individuals with a high need for closure are more prone to incidental exposure. Individuals with a high need for closure tend to seize on information that resolves uncertainty. Psychologically, this could make them more vulnerable to misinformation encountered incidentally as they might readily accept the first explanation they come across (even if false) rather than seeking out waiting in ambiguity. This aligns with previous research that found that individuals with a high need for cognitive closure are more likely to accept conspiracy narratives [66].

Table 3. Coefficients for correlates of the prevalence of content from sources known for spreading misinformation across exposure types.

		Selective exposure	Incidental exposure	Targeted exposure	Algorithmic exposure
Ethnicity					
	White	**-0.35(0.20)†**	0.09(0.22)	**-0.47(0.24)***	**0.39(0.20)†**
Religion					
	Non-religious	-0.16(0.19)	0.14(0.22)	-0.40(0.28)	0.22(0.18)
	Other religion	-0.08(0.26)	-0.24(0.31)	-0.12(0.56)	-0.07(0.27)
Partisanship					
	Republican	0.07(0.25)	0.40(0.25)	0.07(0.28)	0.16(0.27)
	Other	0.21(0.29)	0.00(0.31)	**-0.61(0.30)***	**0.45(0.25)†**
Community					
	Suburban	**0.37(0.17)***	0.09(0.19)	0.07(0.27)	0.03(0.16)
	Rural	0.38(0.25)	0.01(0.25)	**0.85(0.36)***	**0.77(0.21)*****
Education					
	Undergraduate	**-0.34(0.18)†**	0.32(0.21)	0.21(0.25)	-0.05(0.19)
	High school	-0.11(0.25)	**0.79(0.26)****	0.39(0.33)	**0.62(0.23)****
	Community college	-0.20(0.31)	**0.87(0.27)****	**0.81(0.48)†**	**-0.71(0.38)†**
Age					
	[25-34]	-0.01(0.23)	0.15(0.26)	-0.31(0.31)	-0.02(0.20)
	[35-44]	-0.27(0.28)	**-0.62(0.30)***	0.18(0.33)	**-0.47(0.23)***
	[45-54]	-0.10(0.30)	-0.07(0.32)	-0.22(0.35)	0.16(0.27)
	More than 55	-0.35(0.30)	-0.39(0.35)	0.04(0.44)	-0.53(0.35)
Gender					
	Male	-0.25(0.17)	-0.19(0.17)	0.06(0.24)	-0.20(0.17)
Psychologica					
	Authoritarian	-0.15(0.14)	**-4.76(1.32)*****	-0.09(0.23)	0.17(0.14)
	Need for closure	**-0.23(0.14)†**	**2.89(1.28)***	0.12(0.22)	-0.03(0.15)
	Stress	-0.16(0.39)	-1.94(1.57)	0.67(0.52)	-0.56(0.37)
	Threat manag	-0.01(0.16)	0.92(1.33)	0.05(0.22)	0.13(0.14)

*** $p<0.001$; ** $p<0.01$; * $p<0.05$; † $p<0.10$

4.1.2 Consumption of misinformation off Facebook. The second column of Table 2 examines user attributes influencing the prevalence of misinformation in users' news consumption off Facebook.

The table shows that consumption of misinformation off Facebook is positively correlated with only "Ethnicity: White" (0.04) and "Partisanship: Republican" (0.07). However, the coefficients are very low. Hence, even if they affect in a statistically significant manner the consumption of misinformation, they do it to a very small degree. Despite previous research indicating a correlation between stress and anxiety and misinformation consumption [118], our findings did not reveal a significant association.

Takeaways: We found no significant correlations between user attributes and misinformation consumption off Facebook. However, on Facebook, several factors—including suburban and rural community types, high school education levels, the 25-34 age group, and Republican partisanship—demonstrated statistically significant correlations with increased misinformation exposure on Facebook. This pattern suggests that social media algorithms may amplify misinformation for specific demographic or psychological profiles. This underscores the need for greater transparency and algorithmic auditing to understand why these socio-demographic traits correlate with on-platform misinformation exposure, a phenomenon absent from off-platform consumption.

Nevertheless, these psychosociodemographic factors only explain 21% of the variability in exposure to misinformation. This indicates that more granular online activity metrics would likely better explain exposure to misinformation. It is, however, noteworthy that these weak indicators alone account for a substantial portion of the variance in on-platform misinformation exposure.

4.2 Predictors of political balance

4.2.1 Political balance on Facebook. The third column of Table 2 investigates user attributes influencing the political balance in users' news exposure on Facebook. The table shows that political balance in Facebook news exposure is positively correlated with "Partisanship: Republican" (1.88), suggesting that Republicans tend to get exposed to a mix of both left-leaning and right-leaning news sources. This finding aligns with prior research, such as Bakshy et al. [15], which observed that conservative Facebook users were exposed to more politically balanced content compared to liberal users.

Regarding psychological factors, the political balance in Facebook news exposure correlates positively with stress and anxiety scores (3.17). This result aligns with previous research in two key ways. First, it supports findings that individuals with higher anxiety levels tend to be more open to diverse viewpoints, making them less likely to be trapped in echo chambers [108] and more inclined to seek out information that challenges their existing beliefs [130]. Second, it is consistent with studies showing that exposure to politically diverse content can itself increase anxiety in some users [122], suggesting a reciprocal relationship between anxiety and balanced news exposure.

The table shows a negative correlation between political balance and the threat management system score (-1.03), suggesting that users with greater concerns or beliefs in surrounding threats are exposed to less politically balanced content.

4.2.2 Political balance off Facebook. The fourth column of Table 2 explores user attributes affecting the political balance of users' news consumption off Facebook. The results show a positive correlation between political balance in off Facebook news consumption and "Partisanship: Republican" (1.6), indicating that Republican users are more likely to consume news from a balanced range of political perspectives. Previous research supports this finding: Garrett et al. [44] have shown that while Republicans tend to prefer stories aligning with their views, they also demonstrate a willingness to consider more balanced perspectives.

Regarding psychological factors, the table shows that the political balance of news consumption off Facebook negatively correlates with the need for closure scores (-1.32). This finding aligns with prior research, such as Webster et al. [127], which suggests that individuals with a high need for closure may exhibit polarization and decreased openness to diverse viewpoints.

Takeaways: Beyond Republican partisanship, which displays consistent correlations with political balance across platforms, psychological factors reveal context-specific associations. This disparity underscores the distinct mechanisms of information encounter: user-directed choice off-platform versus the combined influence of user behavior and algorithmic filtering on Facebook.

4.3 Attention to misinformation

While users might encounter posts and articles from sources known to spread misinformation, they can choose to pay more or less attention to them. This section analyzes the visibility duration of each misinformation post on Facebook as well as the time users spend on each misinformation article off Facebook.

4.3.1 Attention to misinformation on Facebook. The first column of Table 4 explores user attributes influencing attention to misinformation on Facebook. The table reveals that individuals aged "[25-34]", "[35-44]", and "[45-54]" have significant positive coefficients (0.29, 0.31, and 0.42, respectively), indicating that these age groups tend to spend more time looking at misinformation posts compared to the age group "More than 55." Additionally, the table shows a negative correlation (-0.38) between the time spent on misinformation posts on Facebook and users' stress and anxiety scores. This finding may be explained by previous research, which indicates that individuals experiencing higher levels of stress or anxiety tend to engage in avoidance coping strategies [105].

4.3.2 Attention to misinformation off Facebook. The second column of Table 4 explores user attributes influencing attention to misinformation off Facebook. The table shows that the categories of "Ethnicity: White", "Partisanship: Republican", and "Education level: Community college" have significant positive coefficients (0.04, 0.08, and 0.09, respectively). This suggests that individuals in these groups tend to spend more time interacting with misinformation content off Facebook compared to other groups. The finding regarding partisanship aligns with previous research [50], which shows that Republicans engage more with misinformation content and are more likely to share content from fake news sources.

Takeaways: On Facebook, the socio-demographic factors influencing attention to misinformation differ from those influencing exposure. This indicates that high exposure does not guarantee high attention, and vice versa, underscoring the need to separately investigate passive exposure and active engagement, along with their respective demographic and psychological correlates.

5 Correlation between misinformation consumption and misinformation exposure

Social media platforms collect extensive data to infer user attributes and build detailed profiles for personalized content delivery. These profiles are shaped by user activity both on Facebook and across other websites, which may include news sites that spread misinformation. This is particularly relevant given that prior research has shown news websites adopt the Facebook Pixel at disproportionately high rates (18.6%) compared to other types of websites [16, 22]. As a result, misinformation consumed off Facebook could influence the user profile Facebook builds, which in turn affects the type of content the user is shown on the platform.

In this section, we look into the correlation between misinformation consumption off Facebook and misinformation exposure on Facebook.

5.1 Splitting user data into consumption and exposure periods

To examine the relationship between misinformation consumption off Facebook and misinformation exposure on Facebook, we divided each user's data into two subsets based on the observed date. The first subset represents *the news consumption period*, which includes news and misinformation users consume both on and off Facebook. The second subset represents the *news exposure period*, which includes news and misinformation that users encounter on Facebook. All news consumption included in the first subset occurred prior to all news exposure on Facebook in the second subset. Our goal is to analyze the correlation between the prevalence of misinformation consumed during the first period and the prevalence of misinformation the user is exposed to in the second period.

Users exhibited varying levels of activity throughout the study period. They have different start dates, end dates, and active days of data collection. Therefore, we compute a split date for each user to divide their data separately. This approach ensures that there is sufficient data for analysis in both subsets. Specifically, we implemented two constraints: (1) a minimum of 10 news articles consumed during the first period, and (2) a minimum of 10 news-related Facebook posts encountered in the second period. When multiple split dates satisfy both conditions, we select the latest date. This

Table 4. Coefficients for correlates of time spent on content from sources known for spreading misinformation on and off Facebook.

		On Facebook: Median time spent	Off Facebook: Median time spent
Ethnicity			
	White	0.12(0.10)	**0.04*(0.02)**
Religion			
	Non religious	0.08(0.09)	-0.02(0.02)
	Other religion	-0.20(0.12)	-0.01(0.02)
Partisanship			
	Republican	0.05(0.12)	**0.08**(0.03)**
	Other	0.12(0.15)	0.01(0.03)
Community			
	Suburban	0.05(0.08)	-0.01(0.03)
	Rural	0.09(0.12)	-0.03(0.02)
Education			
	Undergraduate	0.11(0.09)	0.02(0.02)
	High school	0.13(0.12)	0.03(0.03)
	Community college	-0.18(0.14)	**0.09*(0.04)**
Age			
	[25-34]	**0.29**(0.11)**	-0.02(0.02)
	[35-44]	**0.31**(0.13)**	0.03(0.03)
	[45-54]	**0.42**(0.15)**	-0.00(0.03)
	More than 55	0.24(0.16)	-0.01(0.04)
Gender			
	Male	-0.10(0.08)	0.02(0.02)
Psychologica			
	Authoritarian	-0.08(0.07)	-0.01 (0.02)
	Need for closure	-0.01(0.07)	0.02(0.02)
	Stress and anxiety	**-0.38*(0.17)**	-0.04(0.03)
	Threat manag	0.05(0.08)	-0.00(0.02)
R2 score		17%	24%

*** p<0.001; ** p<0.01; * p<0.05

approach maximizes the number of articles in the first subset, reducing the potential influence of articles consumed before the start of our data collection (which are not included in the dataset) on the news exposure data we analyze.

Table 5. Coefficients for correlates of misinformation consumption off Facebook and misinformation exposure on Facebook. Robust standard errors are in parentheses.

	Misinformation fraction on Facebook	
Predictor	**Estimate**	**P value**
Fraction of misinformation consumed on Facebook	**0.99**	**2e-16**
Fraction of misinformation consumed off Facebook	**0.57**	**0.008**
Time spent on misinformation posts on Facebook	0.00	0.12
Time spent on misinformation articles off Facebook	0.00	0.13
R2 score	61%	

5.2 Analyzing predictors of misinformation exposure

In the first subset, representing the news consumption period, we compute four metrics for each user: (1) the fraction of misinformation consumed on Facebook, (2) the fraction of misinformation consumed off Facebook, (3) the median time spent on misinformation posts on Facebook, (4) the median time spent on misinformation articles off Facebook. We then use these four metrics as input variables in a Generalized Linear Model to predict the fraction of misinformation in users' exposure derived from the second subset representing the news exposure period.

Table 5 presents the coefficients for the correlates in the best-fitting Generalized Linear Model. The table shows that the fraction of misinformation in news exposure is positively correlated with the fraction of misinformation consumed off Facebook (estimate = 0.57). This suggests that users who consume misinformation off Facebook are more likely to be exposed to misinformation on Facebook compared to those who do not. Similarly, a positive correlation is observed between misinformation exposure and misinformation consumption on Facebook (estimate = 0.99). In contrast, the amount of time spent on misinformation—whether on or off Facebook—does not emerge as a statistically significant predictor.

The model's coefficient of determination ($R^2 = 0.61$) indicates that approximately 61% of the variance in misinformation exposure on Facebook is explained by the predictors, suggesting that the model provides a reasonably strong fit to the data and that the predictors collectively account for a meaningful degree of variability in misinformation exposure.

One potential concern is that the fraction of misinformation off Facebook may be influenced by prior exposure to misinformation on Facebook. If these two variables are highly correlated, any observed association between the fraction of misinformation consumed off Facebook and the outcome could be driven by its overlap with the fraction of misinformation consumed on Facebook, rather than representing an independent effect. To address this, we tested for multicollinearity using the Variance Inflation Factor (VIF) [114]. We found that the fraction of misinformation consumed off Facebook had a VIF of 2.12 and the fraction of misinformation consumed on Facebook had a VIF of 1.47, both below the commonly used threshold of 5. These results indicate low multicollinearity, supporting the interpretation that the two variables are measuring distinct information and that each contributes independently to the model. This suggests that the association observed for the fraction of misinformation consumed off Facebook is not merely an artifact of its correlation with the fraction of misinformation consumed on Facebook.

Takeaways: We observed a significant positive correlation between users' consumption of misinformation off Facebook and their subsequent exposure to similar content on the platform. This correlation indicates a potential link between off-platform news browsing habits and on-platform news encounters. While this association could be influenced by various factors, the widespread adoption of the Facebook Pixel on news media websites [16, 22] presents a plausible mechanism. Further research is necessary to determine the extent to which off-platform misinformation consumption, potentially tracked by the Pixel, influences on-platform exposure.

6 Discussion

This work presents several findings and opens up a number of discussion points.

Facebook's low contribution to overall news consumption. Our data reveals a striking contrast: while a small fraction of news articles consumed on news websites 4% originated from Facebook, and only a small percentage of news-related posts on Facebook lead to article clicks 2.6%, a majority of U.S. adults report Facebook as their primary news source [125]. This discrepancy suggests a hypothesis: users who identify Facebook as their main news source may primarily rely on social media headlines (and not the actual articles) to be informed about current events. Future research should investigate the depth of understanding, information retention, belief formation, and vulnerability to misinformation among individuals who primarily rely on news headlines on social media, given their limited exposure to the full context of news events.

Disparate exposure to misinformation on and off Facebook. The unique nature of our dataset allows us to show that users experience a greater prevalence of misinformation on Facebook than in their typical news consumption off the platform. We believe that this comparison, contrasting platform exposure with users' selective engagement on news websites, provides a more powerful argument than studies confined to social media analysis [5, 86, 94].

Correlation between psychosociodemographic attributes and exposure, but not consumption, of misinformation. Our analysis reveals correlations between specific user demographics and psychological attributes and their exposure to misinformation on Facebook. Notably, these same attributes do not correlate with consumption off Facebook. This discrepancy suggests that Facebook's algorithms might inadvertently learn and amplify misinformation towards users exhibiting certain traits. To further investigate this, controlled experiments are crucial to determine the extent to which Facebook's algorithms infer socio-demographic traits (e.g., partisanship) and psychological characteristics (e.g., stress, authoritarianism) and use these inferences for content personalization.

In light of the EU's Digital Services Act mandating systemic risk assessments, we encourage platforms to proactively investigate whether their optimization algorithms inadvertently generate biased correlations with psychosocial or demographic groups, drawing parallels to known biases in algorithmic hiring related to gender and ethnicity [33].

Impact of off-platform news consumption on on-platform news exposure. While previous research has examined patterns of misinformation exposure across social media platforms [51, 54], to the best of our knowledge, this is the first study to show that the proportion of misinformation consumed off Facebook is a statistically significant predictor of users' exposure to misinformation on Facebook. Although it is well known that social media platforms track user activity both on and off their services to build detailed user profiles, further analysis is needed to assess the extent to which Facebook can track users on misinformation websites—for example, by identifying whether these sites include Facebook's tracking scripts.

It is also important to investigate whether, and to what extent, this tracking data is used to build user profiles and personalize content. While we know from Facebook's privacy policies that such data is used to deliver targeted advertisements to the most relevant users [75], we believe it is important to explore whether the same data is also used to personalize all news feed posts.

Finally, from a policy perspective, we believe it is worth examining whether restricting Facebook's ability to track users off-platform—through regulation or other mechanisms—could help mitigate the issue of misinformation amplification.

Regulatory opportunities. We believe this study highlights several potential unwanted biases in users' exposure to misinformation on Facebook. Our findings point to the need for deeper audits—both of algorithmic biases in the ranking of posts in users' timelines, and of how tracking technologies may influence misinformation exposure. Unfortunately, for more robust conclusions, we need better access to data than we currently have.

The European Union's recently adopted Digital Services Act introduces a legal framework that requires online platforms to share data with researchers and regulators to assess systemic risks. We intend to take this study as a "worth investigating" base and submit requests through the Delegated Regulation on data access [31]. We hope this can facilitate more robust algorithmic audits by enabling independent researchers to analyze platform data and uncover potential biases.

7 Related work

This section reviews key studies on the prevalence of misinformation in news consumption and the characteristics of users who are more exposed to it. We also examine research on online news' political diversity, focusing on whether online platforms exacerbate or mitigate political polarization.

7.1 Prevalence of misinformation in online news

Misinformation has been an important focus in CSCW, with prior work exploring a wide range of aspects related to its spread, impact on users, and potential solutions. This includes (1) research on the prevalence of misinformation on social media [45, 79], (2) investigations into how platform design and algorithms shape users' exposure to false content [13, 126], (3) the role of collaborative sensemaking in evaluating and correcting misinformation [12, 72, 80], (4) how interface features, recommendation systems, and engagement-driven algorithms influence the visibility of misinformation [65, 67], and (5) the effectiveness of platform-driven solutions, such as content moderation, flagged content, alongside community-led approaches like fact-checking in countering misinformation [11, 43, 68, 70, 80].

Beyond the CSCW community, research from different disciplines has also extensively examined user exposure to and consumption of online misinformation and the individual-level characteristics associated with these behaviors [5, 29, 48, 51, 94, 115]. For instance, Guess et al. [51, 53] reported that, on average, articles from misleading news sites made up 5.9% of all news content consumed by users, and observed a significant age effect in misinformation sharing, with users aged 65 and older sharing nearly seven times more fake news articles than younger users. Further research by Mosleh et al. [86] examined the role of ideological partisanship in misinformation consumption, finding that ideological extremity correlated with higher misinformation consumption, particularly among conservatives. In terms of the effects of misinformation, Ognyanova et al. [91] revealed that exposure to online misinformation is associated with a decrease in trust in mainstream media across party lines.

Our study builds on this existing research and its findings. By analyzing data collected from real Facebook users—data that was not available in previous studies—we confirm several previously reported findings, such as the pervasive nature of misinformation exposure [30, 80], and the ways in which people respond to encountering misinformation [80]. Additionally, our unique dataset, which includes both users' news exposure and consumption habits, along with demographic and psychological profiles, allows us to address new questions and present previously unknown insights. Specifically, we investigate how users' demographics and psychological traits correlate with exposure to misinformation on Facebook and their consumption of misinformation on external news media websites. Furthermore, we explore whether engagement with misinformation on news media websites leads to increased exposure to false content on Facebook, shedding light on cross-platform dynamics.

7.2 Political balance in online news

Several studies have investigated political polarization and balance in online news consumption [40, 61, 64], with mixed findings on whether social media platforms contribute to polarization.

Two metrics have been employed to measure diversity in online news. First, studies such as those by Fletcher et al. [40] and Scharkow et al. [104], measure diversity by counting the number of distinct news sources a user encounters. Under this approach, exposure to a broader range of sources is interpreted as a more diverse news diet, whereas exposure to a limited number of sources is associated with political polarization. However, this metric may not fully capture political diversity, as similar-leaning sources can still result in a polarized news diet. Other studies [14, 28, 39, 61, 76] have employed more nuanced methods that focus on the political leanings of the news sources, similarly to our approach.

Some research, such as Flexed et al. [39] and Scharkow et al. [104], suggests that social networks and search engines increase exposure to diverse political perspectives, promoting political diversity. Similarly, Bakshy et al. [14] found that personal choices play a greater role than algorithms in limiting exposure to diverse content. On the other hand, Levy et al. [76] reported that Facebook's algorithm restricts exposure to opposing viewpoints, even when users follow diverse sources, potentially causing polarization. Supporting this, González-Bailón et al. [49] found high ideological segregation on Facebook, with conservatives primarily engaging with conservative content and misinformation, while liberals had less access to similar levels of partisan news. Additionally, Eady et al. [28] found an asymmetry in cross-cutting behavior, with conservatives more likely to engage with left-leaning media than liberals with right-leaning media, highlighting differences in openness to opposing views. Regarding the impact of political news consumption, Levy et al. [76] found that while encountering news from opposing perspectives can reduce negative attitudes toward rival political parties, it has minimal impact on changing overall political views.

8 Limitations

Despite our best efforts, our study has a few limitations:

Non-representative user sample. Our user sample is not representative. Recruiting participants was challenging, as it required individuals willing to install a browser extension and share their browsing data. The sample was further narrowed to those who maintained consistent activity both on and off Facebook and completed the required socio-demographic and psychological surveys. Additionally, our recruitment focused solely on U.S.-based users, preventing cross-country comparisons that might reveal broader or differing patterns.

Focus on desktop. Our dataset does not capture news exposure or consumption on mobile devices. Developing a comparable data collection tool for mobile platforms is technically challenging, if not impossible. Nevertheless, existing research offers no strong indication of major behavioral distinctions between desktop and mobile news consumption patterns. Importantly, our comparison of on and off Facebook activity relies on data collected from web browsers in both contexts, mitigating potential biases and supporting the validity of our findings.

Misinformation labels. Our evaluation of misinformation and political bias is done at the news source level rather than the individual article level. For instance, we report that 5.9% of news content on Facebook comes from sources labeled as repeatedly sharing misinformation by Media Bias Fact Check and NewsGuard. However, not all content from these sources is necessarily false. Thus, this 5.9% should be considered an upper bound, with the actual share of misinformation likely being lower. This limitation is common in similar research [6, 46, 52, 62, 103, 129], as no widely accepted methodology exists for reliably assessing misinformation and political bias at the individual article level.

To test the robustness of our regression findings, we conducted an additional analysis using NewsGuard's website scores as a proxy for misinformation likelihood, modeling exposure as a continuous variable. The results were consistent: most significant predictors aligned with those from the original model (see Appendix D for details).

While acknowledging these limitations, we believe our unique dataset and analysis offer interesting insights into understanding misinformation exposure and political content balance.

9 Conclusion

In this work, we aimed to address longstanding questions related to political polarization and the spread of misinformation in online news. Specifically, we examined the role that social media platform algorithms may play in shaping these phenomena and whether user demographics and psychological factors influence them.

While several previous studies have attempted to explore these questions, most were hindered by limited access to comprehensive data. The few that have accessed comparable datasets have not investigated our specific research question. Our study, therefore, fills a critical gap by being the first to use a dataset that integrates on-platform news exposure, off-platform news consumption, and detailed psychological and demographic attributes to address this novel inquiry. This approach not only yields unique and insightful findings but also opens new avenues for future research.

Our analysis allows us to confirm some previous findings in improved experimental conditions (the ability to observe users' precise exposure and consumption) and provide more nuanced views. Notably, only a small percentage of users' off Facebook news consumption 4% and misinformation consumption 5.7% originated from exposure on Facebook. Furthermore, most users were more frequently exposed to misinformation on Facebook than off it. However, political content was generally more balanced on Facebook; only 22% of users received news from a single political leaning on the platform, compared to 36% off Facebook. Our analysis also reveals new insights. Notably, our analysis of user demographics and psychological factors associated with exposure to misinformation on Facebook and off Facebook showed that certain attributes, including "Community: Suburban and rural", "Education: High school", "Age: [25-34]" and "Partisanship: Republican", were significantly associated with higher exposure to misinformation on Facebook, but not off Facebook. This raises important questions about how platform algorithms interact with user demographics to amplify specific types of content. Furthermore, we find that users who engage with misinformation off Facebook are likely to encounter more of it on the platform. This might be driven by tracking mechanisms, such as the Facebook Pixel, which tracks user behavior on news media websites. These results highlight the need for further exploration into the effects of tracking on news diversity and misinformation spread across social media platforms.

Overall, our findings highlight the crucial role that social media platforms play in amplifying specific types of content, particularly misinformation. With the new Delegated Regulation on data access under the Digital Services Act, researchers will have greater access to data, allowing deeper studies on how social media algorithms and user attributes influence news consumption and exposure. We hope our study can be used as an argument for data access requests and help develop strategies aimed at reducing misinformation and promoting content diversity across digital platforms.

10 Ethics

Our monitoring tool gathers sensitive and personal information from participants involved in the study. This data collection strictly complies with the guidelines outlined in the EU General Data Protection Regulation (GDPR), ensuring lawful, fair, and transparent data processing. Furthermore, we have obtained the necessary approvals from the Data Protection Officers and the Ethical Review Board of our institution confirm compliance with the required regulations. In addition, to maintain transparency, every user installing our tool is presented with a detailed page explaining the data collection process and its intended use. We request explicit consent from participants to contribute to their data and engage in the research study. Finally, participants were fully informed of their right to withdraw from the study at any time and request the removal of their data.

References

[1] Alan Agresti. 2015. *Foundations of linear and generalized linear models.* John Wiley & Sons.

[2] Alexander C Aitken. 1936. IV.—On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh* 55 (1936), 42–48.

[3] Hunt Allcott and Matthew Gentzkow. 2017. Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives* 31 (05 2017), 211–236. doi:10.1257/jep.31.2.211

[4] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *The journal of economic perspectives: a journal of the American Economic Association* 31, 2 (2017), 211–236. doi:10.1257/jep.31.2.211

[5] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. 2019. Trends in the diffusion of misinformation on social media. *Research & Politics* 6, 2 (2019). https://doi.org/10.1177/2053168019848554

[6] Jennifer Allen, Baird Howland, Markus Mobius, David M. Rothschild, and Duncan J. Watts. 2019. Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances* 6 (2019). https://api.semanticscholar.org/CorpusID:215745943

[7] Bob Altemeyer. 2006. *The Authoritarians.* University of Manitoba, Winnipeg.

[8] Kolmogorov An. 1933. Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att* 4 (1933), 89–91.

[9] Ian A Anderson. 2024. Beyond Active and Passive Social Media Use: Habit Mechanisms Are Behind Frequent Posting and Scrolling on Twitter/X. doi:10.31234/osf.io/824nb

[10] Jonathan Nagler Andrew M. Guess and Joshua Tucker. 2020. Who Falls for Fake News? The Roles of Bullshit Receptivity, Overclaiming, Familiarity, and Analytic Thinking. *Journal of Personality* 88, 2 (2020), 185–200. doi:10.1111/jopy.12476

[11] Scott Appling, Amy Bruckman, and Munmun De Choudhury. 2022. Reactions to Fact Checking. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 403 (Nov. 2022), 17 pages. doi:10.1145/3555128

[12] Ahmer Arif, John J. Robinson, Stephanie A. Stanek, Elodie S. Fichet, Paul Townsend, Zena Worku, and Kate Starbird. 2017. A Closer Look at the Self-Correcting Crowd: Examining Corrections in Online Rumors. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 155–168. doi:10.1145/2998181.2998294

[13] Ahmer Arif, Leo Graiden Stewart, and Kate Starbird. 2018. Acting the Part: Examining Information Operations Within BlackLivesMatter Discourse. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 20 (Nov. 2018), 27 pages. doi:10.1145/3274289

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

[15] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132. arXiv:https://www.science.org/doi/pdf/10.1126/science.aaa1160 doi:10.1126/science.aaa1160

[16] Paschalis Bekos, Panagiotis Papadopoulos, Evangelos P. Markatos, and Nicolas Kourtellis. 2023. The Hitchhiker's Guide to Facebook Web Tracking with Invisible Pixels and Click IDs. In *Proceedings of the ACM Web Conference 2023* (Austin, TX, USA) *(WWW '23)*. Association for Computing Machinery, New York, NY, USA, 2132–2143. doi:10.1145/3543507.3583311

[17] Priyanjana Bengani. 2019. Hundreds of 'pink slime' local news outlets are distributing algorithmic stories and conservative talking points. *Columbia Journalism Review* (2019). https://www.cjr.org/tow_center_reports/hundreds-of-pink-slime-local-news-outlets-are-distributing-algorithmic-stories-conservative-talking-points.php

[18] Trevor S Breusch and Adrian R Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the econometric society* (1979), 1287–1294.

[19] Dustin P. Calvillo, Alex León, and Abraham M. Rutchick. 2024. Personality and misinformation. *Current Opinion in Psychology* 55 (2024), 101752. doi:10.1016/j.copsyc.2023.101752

[20] John M. Carey, Victoria Chi, D. J. Flynn, Brendan Nyhan, and Thomas Zeitzoff. 2020. The effects of corrective information about disease epidemics and outbreaks: Evidence from Zika and yellow fever in Brazil. *Science Advances* 6, 5 (2020), eaaw7449. arXiv:https://www.science.org/doi/pdf/10.1126/sciadv.aaw7449 doi:10.1126/sciadv.aaw7449

[21] Media Bias/Fact Check. 2025. Media Bias/Fact Check methodology. https://mediabiasfactcheck.com/methodology/

[22] Quan Chen, Panagiotis Ilia, Michalis Polychronakis, and Alexandros Kapravelos. 2021. Cookie Swap Party: Abusing First-Party Cookies for Web Tracking. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*. Association for Computing Machinery, New York, NY, USA, 2117–2129. doi:10.1145/3442381.3449837

[23] Salim Chouaki, Islem Bouzenia, Oana Goga, and Beatrice Roussillon. 2022. Exploring the Online Micro-targeting Practices of Small, Medium, and Large Businesses. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 378 (Nov. 2022), 23 pages. doi:10.1145/3555103

[24] Salim Chouaki, Abhijnan Chakraborty, Oana Goga, and Savvas Zannettou. 2024. What News Do People Get on Social Media? Analyzing Exposure and Consumption of News through Data Donations. In *Proceedings of the ACM Web*

Conference 2024 (Singapore, Singapore) *(WWW '24)*. Association for Computing Machinery, New York, NY, USA, 2371–2382. doi:10.1145/3589334.3645399

[25] European Commission. 2025. Digital Services Act. https://ec.europa.eu/digital-single-market/en/digital-services-act-package

[26] Francisco Cribari-Neto and Achim Zeileis. 2010. Beta regression in R. *Journal of statistical software* 34 (2010), 1–24.

[27] Jamal I Daoud. 2017. Multicollinearity and regression analysis. In *journal of physics: Conference series*, Vol. 949. IOP Publishing, 012009.

[28] Gregory Eady, Jonathan Nagler, Andy Guess, Jan Zilinsky, and Joshua A Tucker. 2019. How many people live in political bubbles on social media? Evidence from linked survey and Twitter data. *Sage Open* 9, 1 (2019), 2158244019832705.

[29] Alexandros Efstratiou and Emiliano De Cristofaro. 2022. Adherence to Misinformation on Social Media Through Socio-Cognitive and Group-Based Processes. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–35.

[30] Alexandros Efstratiou and Emiliano De Cristofaro. 2022. Adherence to Misinformation on Social Media Through Socio-Cognitive and Group-Based Processes. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 488 (Nov. 2022), 35 pages. doi:10.1145/3555589

[31] European Commission. 2024. Commission Delegated Regulation (EU) 2024/1773 of 13 March 2024 supplementing Regulation (EU) 2022/2554 of the European Parliament and of the Council with regard to regulatory technical standards specifying the detailed content of the policy regarding contractual arrangements on the use of ICT services supporting critical or important functions provided by ICT third-party service providers. https://eur-lex.europa.eu/eli/reg_del/2024/1773/oj/eng. https://eur-lex.europa.eu/eli/reg_del/2024/1773/oj/eng Official Journal of the European Union, L, 25 June 2024.

[32] Solomon Messing Eytan Bakshy and Lada A. Adamic. 2015. Exposure to Ideologically Diverse News and Opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132. doi:10.1126/science.aaa1160

[33] Alessandro Fabris, Nina Baranowska, Matthew J. Dennis, David Graus, Philipp Hacker, Jorge Saldivar, Frederik Zuiderveen Borgesius, and Asia J. Biega. 2025. Fairness and Bias in Algorithmic Hiring: A Multidisciplinary Survey. *ACM Trans. Intell. Syst. Technol.* 16, 1, Article 16 (Jan. 2025), 54 pages. doi:10.1145/3696457

[34] Facebook Business Help Center. 2024. How to add URL parameters to Meta ads. https://www.facebook.com/business/help/1016122818401732

[35] Facebook Developers. 2024. Conversions API Parameters: fbp and fbc. https://developers.facebook.com/docs/marketing-api/conversions-api/parameters/fbp-and-fbc/?locale=fr_FR

[36] Miriam Fernandez, Alejandro Bellogín, and Iván Cantador. 2024. Analysing the Effect of Recommendation Algorithms on the Spread of Misinformation. In *Proceedings of the 16th ACM Web Science Conference* (Stuttgart, Germany) *(WEBSCI '24)*. Association for Computing Machinery, New York, NY, USA, 159–169. doi:10.1145/3614419.3644003

[37] Joseph Firth, John Torous, Brendon Stubbs, Josh Firth, Genevieve Steiner-Lim, Lee Smith, Mario Alvarez-Jimenez, John Gleeson, Davy Vancampfort, Christopher Armitage, and Jerome Sarris. 2019. The "online brain": how the Internet may be changing our cognition. *World psychiatry: official journal of the World Psychiatric Association (WPA)* 18 (05 2019), 119–129. doi:10.1002/wps.20617

[38] Ronald A Fisher. 1922. On the interpretation of χ 2 from contingency tables, and the calculation of P. *Journal of the royal statistical society* 85, 1 (1922), 87–94.

[39] Seth Flaxman, Sharad Goel, and Justin M. Rao. 2016. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly* (2016). https://doi.org/10.1093/poq/nfw006

[40] Richard Fletcher and Rasmus Kleis Nielsen. 2018. Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society* (2018). https://doi.org/10.1177/1461444817724170

[41] Isabelle Freiling, Nicole M. Krause, Dietram A. Scheufele, and Dominique Brossard. 2021. Believing and sharing misinformation, fact-checks, and accurate information on social media: The role of anxiety during COVID-19. *New Media Society* (2021), 14614448211011451–.

[42] Adrien Friggeri, Lada Adamic, Dean Eckles, and Justin Cheng. 2014. Rumor Cascades. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (May 2014), 101–110. doi:10.1609/icwsm.v8i1.14559

[43] Mingkun Gao, Ziang Xiao, Karrie Karahalios, and Wai-Tat Fu. 2018. To Label or Not to Label: The Effect of Stance and Credibility Labels on Readers' Selection and Perception of News Articles. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 55 (Nov. 2018), 16 pages. doi:10.1145/3274324

[44] R. Kelly Garrett and Natalie Jomini Stroud. 2014. Partisan Paths to Exposure Diversity: Differences in Pro- and Counterattitudinal News Consumption. *Journal of Communication* 64, 4 (2014), 680–701. doi:10.1111/JCOM.12105

[45] Amira Ghenai and Yelena Mejova. 2018. Fake Cures: User-centric Modeling of Health Misinformation in Social Media. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 58 (Nov. 2018), 20 pages. doi:10.1145/3274327

[46] Maria Glenski, Tim Weninger, and Svitlana Volkova. 2018. Propagation From Deceptive News Sources Who Shares, How Much, How Evenly, and How Quickly? *IEEE Transactions on Computational Social Systems* 5 (2018), 1071–1082.

https://api.semanticscholar.org/CorpusID:54462486

[47] Gregory Gondwe, Dani Madrid-Morales, Melissa Tully, and Herman Wasserman and. 2025. Misinformation and Digital Inequalities: Comparing How Different Demographic Groups Get Exposed to and Engage with False Information. *Mass Communication and Society* 0, 0 (2025), 1–15. arXiv:https://doi.org/10.1080/15205436.2025.2474139 doi:10.1080/15205436.2025.2474139

[48] Sandra González-Bailón, David Lazer, Pablo Barberá, William Godel, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, and Matthew Gentzkow. 2024. The Diffusion and Reach of (Mis)Information on Facebook During the U.S. 2020 Election. *Sociological Science* 11, 41 (2024), 1124–1146. doi:10.15195/v11.a41

[49] Sandra González-Bailón, David Lazer, Pablo Barberá, Meiqing Zhang, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, Matthew Gentzkow, Andrew M. Guess, Shanto Iyengar, Young Mie Kim, Neil Malhotra, Devra Moehler, Brendan Nyhan, Jennifer Pan, Carlos Velasco Rivera, Jaime Settle, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. 2023. Asymmetric ideological segregation in exposure to political news on Facebook. *Science* 381, 6656 (2023), 392–398. arXiv:https://www.science.org/doi/pdf/10.1126/science.ade7138 doi:10.1126/science.ade7138

[50] Andrew Guess, Jonathan Nagler, and Joshua Tucker. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances* 5, 1 (2019), eaau4586.

[51] Andrew Guess, Jonathan Nagler, and Joshua Tucker. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances* (2019). https://www.science.org/doi/abs/10.1126/sciadv.aau4586

[52] Andrew Markus Guess, Jonathan Nagler, and Joshua A. Tucker. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances* 5 (2019). https://api.semanticscholar.org/CorpusID:58025666

[53] Andrew M Guess, Brendan Nyhan, and Jason Reifler. 2020. Exposure to untrustworthy websites in the 2016 US election. *Nature human behaviour* 4, 5 (2020), 472–480.

[54] Katherine Haenschen. 2023. Curated Misinformation: Liking Facebook Pages for Fake News Sites. *American Behavioral Scientist* (2023). doi:10.1177/00027642231175638

[55] James W Hardin and Joseph M Hilbe. 2014. Estimation and testing of binomial and beta-binomial regression models with and without zero inflation. *The Stata Journal* 14, 2 (2014), 292–303.

[56] Florian Hartig. 2018. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R Packag version 020* (2018).

[57] Ethan Hilman. 2024. This is Your Brain on Social Media. https://research-archive.org/index.php/rars/preprint/view/1860/version/2002

[58] Matthew Hindman, Nathaniel Lubin, and Trevor Davis. 2022. Facebook Has a Superuser-Supremacy Problem. *The Atlantic* (Feb 2022). https://www.theatlantic.com/technology/archive/2022/02/facebook-hate-speech-misinformation-superusers/621617/

[59] Tuan-Anh Hoang and Ee-Peng Lim. 2021. Virality and Susceptibility in Information Diffusions. *Proceedings of the International AAAI Conference on Web and Social Media* 6, 1 (Aug. 2021), 146–153. doi:10.1609/icwsm.v6i1.14245

[60] Daniel Hoechle. 2007. Robust standard errors for panel regressions with cross-sectional dependence. *The stata journal* 7, 3 (2007), 281–312.

[61] Homa Hosseinmardi and Amir Ghasemian and Aaron Clauset and Markus Mobius and David M. Rothschild and Duncan J. Watts. 2021. Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences* (2021). https://www.pnas.org/doi/abs/10.1073/pnas.2101967118

[62] Benjamin D. Horne, Jeppe Nørregaard, and Sibel Adali. 2019. Different Spirals of Sameness: A Study of Content Sharing in Mainstream and Alternative Media. *ArXiv* abs/1904.01534 (2019). https://api.semanticscholar.org/CorpusID:91184294

[63] Yoori Hwang and Se-Hoon Jeong. 2023. Education-Based Gap in Misinformation Acceptance: Does the Gap Increase as Misinformation Exposure Increases? *Communication Research* 50, 2 (2023), 157–178. arXiv:https://doi.org/10.1177/00936502221121509 doi:10.1177/00936502221121509

[64] Hazem Ibrahim, Nouar AlDahoul, Sangjin Lee, Talal Rahwan, and Yasir Zaki. 2023. YouTube's recommendation algorithm is left-leaning in the United States. *PNAS Nexus* 2, 8 (08 2023), pgad264. arXiv:https://academic.oup.com/pnasnexus/article-pdf/2/8/pgad264/51841800/pgad264.pdf doi:10.1093/pnasnexus/pgad264

[65] Farnaz Jahanbakhsh, Amy X. Zhang, Adam J. Berinsky, Gordon Pennycook, David G. Rand, and David R. Karger. 2021. Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 18 (April 2021), 42 pages. doi:10.1145/3449092

[66] Alexander Jedinger and Lena Masch. 2025. Need for cognitive closure, political trust, and belief in conspiracy theories during the COVID-19 pandemic. *Frontiers in Social Psychology* 2 (2025), 1447313. doi:10.3389/frsps.2024.1447313

[67] Chenyan Jia, Alexander Boltz, Angie Zhang, Anqing Chen, and Min Kyung Lee. 2022. Understanding Effects of Algorithmic vs. Community Label on Perceived Accuracy of Hyper-partisan Misinformation. *Proc. ACM Hum.-Comput.*

Interact. 6, CSCW2, Article 371 (Nov. 2022), 27 pages. doi:10.1145/3555096

[68] Shan Jiang and Christo Wilson. 2018. Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 82 (Nov. 2018), 23 pages. doi:10.1145/3274351

[69] The Guardian Julia Carrie Wong. 2019. The Cambridge Analytica scandal changed the world–but it didn't change Facebook. Available at: https://www.theguardian.com/technology/2019/mar/17/the-cambridge-analytica-scandal-changed-the-world-but-it-didnt-change-facebook..

[70] Jan Kirchner and Christian Reuter. 2020. Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 140 (Oct. 2020), 27 pages. doi:10.1145/3415211

[71] Neta Kligler-Vilenchik, Alfred Hermida, Sebastián Valenzuela, and Mikko Villi. 2020. Studying incidental news: Antecedents, dynamics and implications. *Journalism* 21, 8 (2020), 1025–1030. doi:10.1177/1464884920915372

[72] Yubo Kou, Xinning Gui, Yunan Chen, and Kathleen Pine. 2017. Conspiracy Talk on Social Media: Collective Sensemaking during a Public Health Crisis. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 61 (Dec. 2017), 21 pages. doi:10.1145/3134696

[73] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096.

[74] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096. arXiv:https://www.science.org/doi/pdf/10.1126/science.aao2998 doi:10.1126/science.aao2998

[75] Facebook legal terms. 2020. Facebook Business Tools Terms. https://en-gb.facebook.com/legal/technology_terms

[76] Ro'ee Levy. 2021. Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review* 111, 3 (2021), 831–870.

[77] Stephan Lewandowsky, Ullrich KH Ecker, and John Cook. 2017. Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of applied research in memory and cognition* 6, 4 (2017), 353–369.

[78] P. F. Lovibond and S. H. Lovibond. 1995. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour research and therapy* 33, 3 (1995), 335–343. doi:10.1016/0005-7967(94)00075-u

[79] Jim Maddock, Kate Starbird, Haneen J. Al-Hassani, Daniel E. Sandoval, Mania Orand, and Robert M. Mason. 2015. Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 228–241. doi:10.1145/2675133.2675280

[80] Lisa Mekioussa Malki, Dilisha Patel, and Aneesha Singh. 2024. "The Headline Was So Wild That I Had To Check": An Exploration of Women's Encounters With Health Misinformation on Social Media. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 128 (April 2024), 26 pages. doi:10.1145/3637405

[81] Masoud Mansoury, Himan Abdollahpouri, Mykola Pechenizkiy, Bamshad Mobasher, and Robin Burke. 2020. Feedback Loop and Bias Amplification in Recommender Systems. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (Virtual Event, Ireland) *(CIKM '20)*. Association for Computing Machinery, New York, NY, USA, 2145–2148. doi:10.1145/3340531.3412152

[82] Media Bias Fact Check. 2014. https://mediabiasfactcheck.com/

[83] Harald Merckelbach, Madeleine Dalsklev, Daniël van Helvoort, Irena Boskovic, and Henry Otgaar. 2018. Symptom self-reports are susceptible to misinformation. *Psychology of Consciousness: Theory, Research, and Practice* 5, 4 (2018), 384.

[84] Samuel Messick and Douglas N. Jackson. 1958. The Measurement of Authoritarian Attitudes. *Educational and Psychological Measurement* 18, 2 (1958), 241–253. arXiv:https://doi.org/10.1177/001316445801800202 doi:10.1177/001316445801800202

[85] Lincoln E Moses. 1963. Rank tests of dispersion. *The annals of mathematical statistics* (1963), 973–983.

[86] Mohsen Mosleh and David G Rand. 2022. Measuring exposure to misinformation from political elites on Twitter. *Nature Communications* 13, 1 (2022), 7144.

[87] Steven L. Neuberg, Douglas T. Kenrick, and Mark Schaller. 2011. Human threat management systems: self-protection and disease avoidance. *Neuroscience and biobehavioral reviews* 35, 4 (2011), 1042–1051. doi:10.1016/j.neubiorev.2010.08.011

[88] News Guard. 2023. https://www.newsguardtech.com/

[89] NewsGuard. 2018. NewsGuard Rating Process and Criteria. https://www.newsguardtech.com/ratings/rating-process-criteria/

[90] Lisa Friedland Briony Swire-Thompson Nir Grinberg, Kenneth Joseph and David Lazer. 2019. Fake News on Twitter During the 2016 U.S. Presidential Election. *Science* 363, 6425 (2019), 374–378. doi:10.1126/science.aau2706

[91] Katherine Ognyanova, David Lazer, Ronald E. Robertson, and Christo Wilson. 2020. Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harvard Kennedy School Misinformation Review* 1 (06 2020). doi:10.37016/mr-2020-024

[92] Danny Osborne, Thomas H Costello, John Duckitt, and Chris G Sibley. 2023. The psychological causes and societal consequences of authoritarianism. *Nature reviews psychology* 2, 4 (2023), 220–232.

[93] Raydonal Ospina and Silvia LP Ferrari. 2012. A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis* 56, 6 (2012), 1609–1623.

[94] Kostantinos Papadamou, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Sirivianos. 2022. "It is just a flu": assessing the effect of watch history on YouTube's pseudoscientific video recommendations. In *Proceedings of the international AAAI conference on web and social media*, Vol. 16. 723–734.

[95] Karl Pearson. 1896. VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character* 187 (1896), 253–318.

[96] Gordon Pennycook and David G Rand. 2019. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences* 116, 7 (2019), 2521–2526.

[97] Gordon Pennycook and David G Rand. 2021. The psychology of fake news. *Trends in cognitive sciences* 25, 5 (2021), 388–402.

[98] Prolific. 2025. https://www.prolific.co

[99] Filipe N. Ribeiro, Koustuv Saha, Mahmoudreza Babaei, Lucas Henrique, Johnnatan Messias, Fabricio Benevenuto, Oana Goga, Krishna P. Gummadi, and Elissa M. Redmiles. 2019. On Microtargeting Socially Divisive Ads. *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Jan 2019). doi:10.1145/3287560.3287580

[100] Ronald Robertson, Jon Green, Damian Ruck, Katherine Ognyanova, Christo Wilson, and David Lazer. 2023. Users choose to engage with more partisan news than they are exposed to on Google Search. *Nature* 618 (05 2023), 1–7. doi:10.1038/s41586-023-06078-5

[101] Arne Roets and Alain Van Hiel. 2011. Item selection and validation of a brief, 15-item version of the Need for Closure Scale. *Personality and individual differences* 50, 1 (2011), 90–94. doi:10.1016/j.paid.2010.09.004

[102] Lauren L Saling, Devi Mallal, Falk Scholer, Russell Skelton, and Damiano Spina. 2021. No one is immune to misinformation: An investigation of misinformation sharing by subscribers to a fact-checking newsletter. *PloS one* 16, 8 (2021), e0255702.

[103] Mattia Samory, Vartan Kesiz Abnousi, and Tanushree Mitra. 2020. Characterizing the Social Media News Sphere through User Co-Sharing Practices. In *International Conference on Web and Social Media.* https://api.semanticscholar.org/CorpusID:215413736

[104] Michael Scharkow, Frank Mangold, Sebastian Stier, and Johannes Breuer. 2020. How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences* (2020). https://www.pnas.org/doi/abs/10.1073/pnas.1918279117

[105] Pia-Isabell Schmidt, Kristin Rosga, Celina Schatto, Anja Breidenstein, and Lars Schwabe. 2013. Stress reduces the incorporation of misinformation into an established memory. *Learning memory (Cold Spring Harbor, N.Y.)* 21, 1 (2013), 5–8. doi:10.1101/lm.033043.113

[106] Julia Schulte-Cloos and Veronica Anghel. 2024. Right-wing authoritarian attitudes, fast-paced decision-making, and the spread of misinformation about COVID-19 vaccines. *Political Communication* 41, 4 (2024), 608–626.

[107] Ingjerd Skafle, Anders Nordahl-Hansen, Daniel S Quintana, Rolf Wynn, and Elia Gabarron. 2022. Misinformation about COVID-19 vaccines on social media: rapid review. *Journal of medical Internet research* 24, 8 (2022), e37367.

[108] Xiaolei Song, Siliang Guo, and Yichang Gao. 2024. Personality traits and their influence on Echo chamber formation in social media: a comparative study of Twitter and Weibo. *Frontiers in Psychology* 15 (02 2024), 1323117. doi:10.3389/fpsyg.2024.1323117

[109] Lars St, Svante Wold, et al. 1989. Analysis of variance (ANOVA). *Chemometrics and intelligent laboratory systems* 6, 4 (1989), 259–272.

[110] Statista. 2023. Resident population of the United States by race from 2000 to 2020. Available at: https://www.statista.com/statistics/183489/population-of-the-us-by-ethnicity-since-2000/..

[111] Statista. 2024. Distribution of Facebook users in the United States as of July 2021, by gender. Available at: https://www.statista.com/statistics/266879/facebook-users-in-the-us-by-gender/..

[112] Statista. 2024. Total population in the United States by gender from 2010 to 2025. Available at: https://www.statista.com/statistics/737923/us-population-by-gender/..

[113] Zea Szebeni, Jan-Erik Lönnqvist, and Inga Jasinskaja-Lahti. 2021. Social psychological predictors of belief in fake news in the run-up to the 2019 Hungarian elections: the importance of conspiracy mentality supports the notion of

ideological symmetry in fake news belief. *Frontiers in psychology* 12 (2021), 790848.
[114] Christopher Glen Thompson, Rae Seon Kim, Ariel M Aloe, and Betsy Jane Becker. 2017. Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and applied social psychology* 39, 2 (2017), 81–90.
[115] Christopher K Tokita, Kevin Aslett, William P Godel, Zeve Sanderson, Joshua A Tucker, Jonathan Nagler, Nathaniel Persily, and Richard Bonneau. 2024. Measuring receptivity to misinformation at scale on a social media platform. *PNAS Nexus* 3, 10 (09 2024), pgae396. arXiv:https://academic.oup.com/pnasnexus/article-pdf/3/10/pgae396/59639674/pgae396_supplementary_data.pdf doi:10.1093/pnasnexus/pgae396
[116] Joshua A Tucker, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. 2018. Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)* (2018).
[117] Zahir Vally. 2021. Compliance with health-protective behaviors in relation to COVID-19: The roles of health-related misinformation, perceived vulnerability, and personality traits. *Mental health effects of COVID-19* (2021), 263–281.
[118] Natasha van Antwerpen, Deborah Turnbull, and Rachel A. Searston. 2021. The role of anxiety in mediating the relationship between information consumption and COVID-19 protective behaviours. *Psychology Health Medicine* (2021), 1–14.
[119] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113, 3 (2016), 554–559. arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.1517441113 doi:10.1073/pnas.1517441113
[120] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151. arXiv:https://www.science.org/doi/pdf/10.1126/science.aap9559 doi:10.1126/science.aap9559
[121] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359 (03 2018), 1146–1151. doi:10.1126/science.aap9559
[122] Emily K. Vraga, Kjerstin Thorson, Neta Kligler-Vilenchik, and Emily Gee. 2015. How individual sensitivities to disagreement shape youth political expression on Facebook. *Computers in Human Behavior* 45 (2015), 281–289. doi:10.1016/j.chb.2014.12.025
[123] Scott I Vrieze. 2012. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods* 17, 2 (2012), 228.
[124] Quang H Vuong. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: journal of the Econometric Society* (1989), 307–333.
[125] Mason Walker and Katerina Eva Matsa. 2022. News Consumption Across Social Media in 2021. https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/
[126] Yuping Wang, Chen Ling, and Gianluca Stringhini. 2023. Understanding the Use of Images to Spread COVID-19 Misinformation on Twitter. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 108 (April 2023), 32 pages. doi:10.1145/3579542
[127] Donna M Webster and Arie W Kruglanski. 1994. Individual differences in need for cognitive closure. *Journal of personality and social psychology* 67, 6 (1994), 1049.
[128] Brian E. Weeks, Daniel S. Lane, Dam Hee Kim, Slgi S. Lee, and Nojin Kwak. 2017. Incidental Exposure, Selective Exposure, and Political Information Sharing: Integrating Online Exposure Patterns and Expression on Social Media. *Journal of Computer-Mediated Communication* (2017). https://doi.org/10.1111/jcc4.12199
[129] Galen Weld, Maria Glenski, and Tim Althoff. 2021. Political Bias and Factualness in News Sharing across more than 100,000 Online Communities. *Proceedings of the International AAAI Conference on Web and Social Media* 15, 1 (May 2021), 796–807. doi:10.1609/icwsm.v15i1.18104
[130] Dag Wollebæk, Rune Karlsen, Kari Steen-Johnsen, and Bernard Enjolras. 2019. Anger, Fear, and Echo Chambers: The Emotional Basis for Online Behavior. *Social Media + Society* 5, 2 (2019), 2056305119829859. arXiv:https://doi.org/10.1177/2056305119829859 doi:10.1177/2056305119829859
[131] www.ofcom.org.uk. 2024. Understanding the impact of social media on online news. https://www.ofcom.org.uk/media-use-and-attitudes/media-plurality/social-media-online-news/
[132] Juan Xie and Yanqing Sun. 2024. Who Shares Misinformation on Social Media? A Meta-analysis of Individual Traits Related to Misinformation Sharing. *Computers in Human Behavior* 158 (Sep 2024). doi:10.1016/j.chb.2024.108271

Appendices

A Demographic and psychological sample of study

Table 6. Sample demographics

		Percentage
Ethnicity		
	White	79%
	Non-white	21%
Religion		
	Non-religious	47%
	Christian	37%
	Other religion	16%
Partisanship		
	Democrat	77%
	Republican	16%
	Other	7%
Community		
	Urban	39%
	Suburban	47%
	Rural	14%
Education		
	Undergraduate	49%
	Graduate	20%
	High school	18%
	Community college	13%
Age		
	[18-24]	21%
	[25-34]	34%
	[35-44]	21%
	[45-54]	13%
	More than 55	11%
Gender		
	Female	65%
	Male	35%

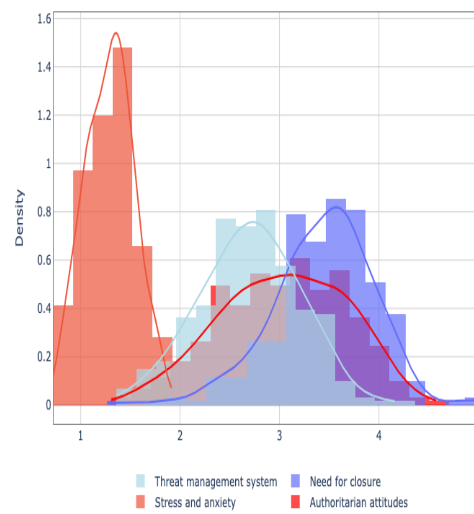

Fig. 7. Psychological scores across users.

B Data processing details

B.1 Determining the quality of news sources

To label each news source as factual or as a source of misinformation, we used data from NewsGuard and Media Bias Fact Check. NewsGuard indicates whether a news source has a history of sharing misinformation in the "Topics" column of their dataset, while Media Bias Fact Check provides these assessments under the "Detailed" section of source evaluations on their website. Although each agency uses distinct terms to describe misleading practices, both include keywords such as "Conspiracy," "Fake News," or "Misinformation." Therefore, we flagged any source with one of these terms as a misinformation source. Agreement between both providers was strong, with only 33 domains showing inconsistencies, which we resolved by applying the misinformation label.

Our dataset comprised domains listed by Media Bias Fact Check and NewsGuard, totaling 4,149 domains. Of these, 2,309 domains were classified as factual sources, 467 as misinformation sources, with the remaining domains uncategorized in our analysis.

B.2 Determining the political bias of news sources

To label the political bias of news sources, we relied on NewsGuard and Media Bias Fact Check. NewsGuard provides political bias classifications for 2,939 news sites, categorizing them as Far Left, Slightly Left, Center, Slightly Right, and Far Right. Media Bias Fact Check provides a similar scale for 1,711 news sources, ranging from extreme-left to extreme-right. We normalized these evaluations to align with NewsGuard's scale by mapping:

- Extreme-right and extreme-left as far right and far left, respectively,
- Right as far right and left as far left,
- Right-center as slightly right and left-center as slightly left.

For the 41 domains where the two sources provided differing evaluations, we chose the classification from Media Bias Fact Check.

B.3 Computing time spent on news-related posts on Facebook

To compute the time users spent on each news-related post encountered on Facebook, we recorded which posts were visible on a user's screen every 0.5 seconds, defining a post as "visible" if at least 30% of it remained on the screen. Given the 0.5-second interval, the computed visibility times have a possible accuracy margin of +/-1 second.

B.4 Most frequent news sources in our dataset

Our dataset comprises 70,587 news article URLs visited by users across 1,882 news domains, along with 123,995 news-related posts received by users on Facebook. Table 7 presents the top 20 pages from which users receive news-related posts, while Table 8 displays the top 20 news sources frequently visited by users.

Table 7. The top 20 frequent Facebook pages from which users receive posts with partisanship.

Facebook Page	Frequency	Partisanship
facebook.com/tiphero/	3493	Center
facebook.com/palmbeachpost/	2647	Left
facebook.com/NPR/	2126	Left
facebook.com/NEWS9/	2055	Center
facebook.com/BuzzFeed/	1417	Right
facebook.com/kfor4/	1386	Left
facebook.com/TheRawStory/	1380	Left
facebook.com/nytimes/	1198	Left
facebook.com/nytopinion/	1185	Left
facebook.com/TheBabylonBee/	1178	Right
facebook.com/washingtonpost/	1138	Left
facebook.com/georgehtakei/	1003	Left
facebook.com/cameron.kelley.75	1001	Right
facebook.com/wjhlTV11/	904	Center
facebook.com/WJTV12/	900	Center
facebook.com/thescarymommy/	833	Left
facebook.com/WRBL3/	821	Center
facebook.com/GoodMorningAmerica/	790	Center
facebook.com/NewsOn6/	763	Center
facebook.com/TheOnion/	738	Center

Table 8. The top 20 frequently visited domains with partisanship.

Domain	Frequency	Partisanship
yahoo.com	19974	Center
sports.yahoo.com	9213	Center
finance.yahoo.com	4609	Center
nytimes.com	2431	Left
cnn.com	2251	Left
espn.com	2220	Center
quora.com	2030	–
dailykos.com	1999	Left
247sports.com	1681	Center
healthline.com	1681	Center
msn.com	1679	Center
telemundo.com	1159	Left
news.google.com	940	Center
buzzfeed.com	933	Left
washingtonpost.com	892	Left
cnbc.com	836	Left
people.com	802	Left
medium.com	771	Left
businessinsider.com	695	Left
ncbi.nlm.nih.gov	666	Center

C Tested models

Table 9. Characteristics of the tested models.

Dependent variable	**Distribution**	**AIC**	**Breusch-Pagan**	**KS test**	**Dispersion test**	**Outlier test**
Exposure: fraction mis-information	**Gaussian**	-63	p=0.67	p=0.29	p=0.98	p=0.73
	Beta	**-1068**	p=0.20	p=0.12	p=0.10	p=0.73
	Beta Binomial	1566	p=0.07	p=0.03	p=0.62	p=0.38
	Zero inflated Beta	-198	p=0.20	p=0.04	p=0.05	p=0.47

Table 10. Characteristics of the tested models .

Dependent variable	AIC	Breusch-Pagan	KS test	Dispersion test	Outlier test
Exposure: fraction misinformation	-1068	p=0.20	p=0.12	p=0.10	p=0.73
Consumption: fraction misinformation	-337	p=0	p=0.06	p=0.26	p=0.01
Exposure: political balance	238	p=0.20	p=0.72	p=0.97	p=1
Misinformation: political balance	183	p=0.46	p=0.29	p=0.78	p=1
Time on misinformation on Facebook	473	p=0.35	p=0.14	p=0.22	p=0.73
Time on misinformation off Facebook	-302	p=0	p=0.06	p=0.26	p=0.45
Exposure: fraction misinformation (from consumption)	-260	p=0.29	p=0.95	p=0.87	p=0.35

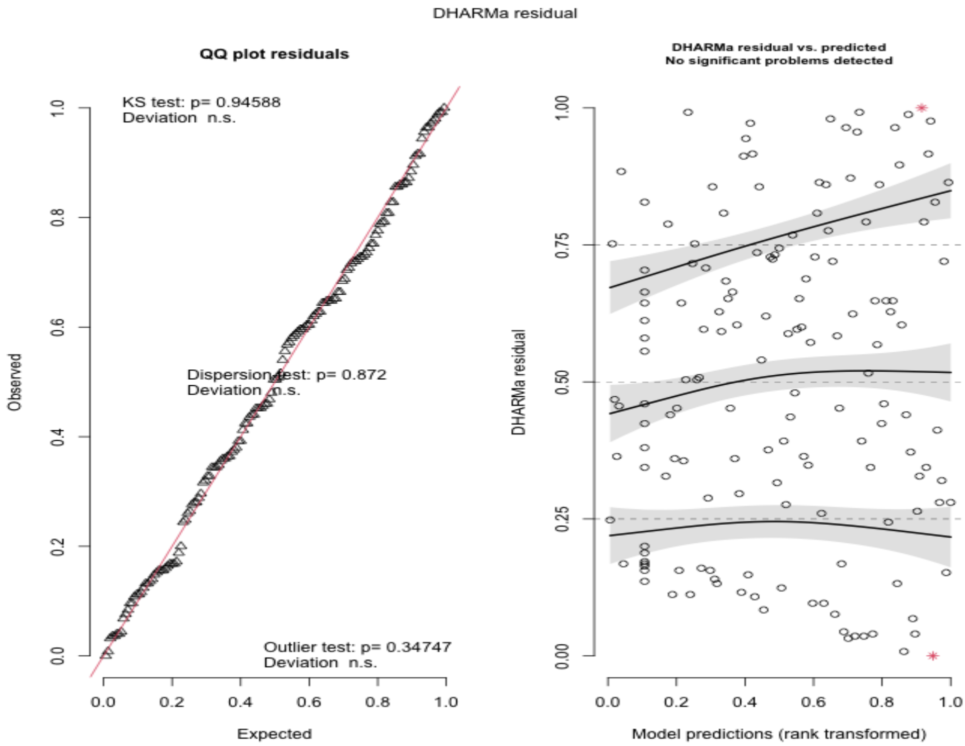

Fig. 8. Diagnostic plots from DHARMa residual analysis for misinformation consumption off Facebook and misinformation exposure on Facebook. The left panel shows a QQ plot of scaled residuals, indicating no significant deviation from uniformity based on the Kolmogorov-Smirnov, dispersion, and outlier tests. The right panel plots DHARMa residuals against model predictions (rank-transformed), with no substantial deviation patterns, suggesting no major violations of model assumptions.

D Predictors of misinformation using website scores

NewsGuard provides trust and credibility ratings for news and information websites, assigning each site a score ranging from 0 to 100. This score reflects the overall trustworthiness and transparency of the news source, based on a set of journalistic criteria. A potential limitation of our approach lies in the application of misinformation labels at the source level rather than at the level of individual news items. This choice is primarily due to the lack of a widely accepted, scalable methodology for reliably identifying misinformation at the article level. To address this limitation, we incorporate website-level ratings from NewsGuard into our dataset. Rather than categorically labeling all content from a particular website as misinformation, we use NewsGuard's credibility score to assign a probabilistic measure. Specifically, for each data entry, we define the likelihood of misinformation as $1 - \frac{\text{NewsGuard Score}}{100}$.

When considering the classification of misinformation at a content level, using NewsGuard ratings in Table 11, we find that most statistically significant predictors remain consistent with those identified using source-level classification in Table 2, though the strength of the associations varies. Republican partisanship, high school education level, and authoritarian attitudes continue to significantly predict exposure to misinformation on Facebook across both approaches. However, two notable differences emerge. First, being aged 25–34, which was a significant predictor of

Table 11. Coefficients for correlates of the prevalence of content from sources known for spreading misinformation on and off Facebook.

		On Facebook: misinformation	Off Facebook: misinformation
Ethnicity			
	White	0.04(0.12)	**0.27* (0.11)**
Religion			
	Non-religious	0.10(0.11)	-0.12 (0.11)
	Other religion	-0.05(0.15)	-0.02(0.15)
Partisanship			
	Republican	**0.43**(0.17)**	**0.41*(0.19)**
	Other	0.21(0.22)	0.07(0.17)
Community			
	Suburban	0.14(0.10)	-0.07(0.12)
	Rural	**0.27†(0.14)**	-0.07(0.14)
Education			
	Undergraduate	**0.31**(0.11)**	0.05(0.11)
	High school	**0.48**(0.15)**	-0.08(0.17)
	Community college	0.04(0.18)	0.24(0.22)
Age			
	[25-34]	0.19(0.13)	-0.26(0.15)
	[35-44]	0.10(0.16)	-0.06(0.18)
	[45-54]	0.23(0.18)	0.002(0.23)
	More than 55	0.09(0.20)	-0.15(0.24)
Gender			
	Male	-0.05(0.10)	0.06(0.11)
Psychologica			
	Authoritarian	**-0.17*(0.09)**	-0.08(0.09)
	Need for closure	0.06(0.10)	0.04(0.11)
	Stress	-0.29(0.20)	-0.30(0.22)
	Threat manag	0.07(0.10)	0.05(0.13)
R2 score		17%	19%

*** p<0.001; ** p<0.01; * p<0.05; † p<0.10

misinformation exposure under source-level classification (β = 0.50*, SE = 0.29), is no longer significant when misinformation is measured at the content level. Second, suburban community status, previously significant in the source-level model (β = 0.34*, SE = 0.14), also loses significance under the content-based approach. For misinformation consumption off Facebook, the set of significant predictors remains the same across both classification strategies. However, the coefficients are larger when using the content-level method.

Received October 2024; revised April 2025; accepted August 2025