

The Puzzle of Misinformation: Exposure to Unreliable Content in the U.S. is Higher among the Better Informed

forthcoming in *New Media & Society*, <https://doi.org/10.1177/14614448231196863>

Alvin Zhou¹, Tian Yang², Sandra González-Bailón^{3*}

Abstract: Healthy news consumption requires limited exposure to unreliable content and ideological diversity in the sources consumed. There are two challenges to this normative expectation: the prevalence of unreliable content online; and the prominence of misinformation within individual news diets. Here, we assess these challenges using an observational panel tracking the browsing behavior of $N \approx 140,000$ individuals in the United States for 12 months (January–December 2018). Our results show that panelists who are exposed to misinformation consume more reliable news and from a more ideologically diverse range of sources. In other words, exposure to unreliable content is higher among the better informed. This association persists after we control for partisan leaning and consider inter- and intra-person variation. These findings highlight the tension between the positive and negative consequences of increased exposure to news content online.

Keywords: political Information, media environment, misinformation, news exposure, news use, web data

¹ Hubbard School of Journalism & Mass Communication, University of Minnesota

² School of Journalism and Communication, Chinese University of Hong Kong

³ Annenberg School for Communication, University of Pennsylvania

* Corresponding author: Annenberg School for Communication, University of Pennsylvania, 3620 Walnut Street, PA 19104, Philadelphia, U.S. e-mail: sgonzalezbailon@asc.upenn.edu

Acknowledgements: we are grateful to Harmony Labs for facilitating access to the data and for crucial research support.

THE PUZZLE OF MISINFORMATION

The Puzzle of Misinformation: Exposure to Unreliable Content in the U.S. is Higher among the Better Informed

The circulation of misinformation through online networks is one of the most salient challenges confronting researchers and policy makers today. The current media environment allows the unmonitored publication of content that online networks then help diffuse and amplify. The visibility of information online results from a combination of social and algorithmic processes that often introduce biases in the content people see (Huszár et al., 2022). Curation mechanisms act as mediators channeling traffic to sources that would have a smaller audience otherwise (Nielsen & Fletcher, 2022; Scharkow et al., 2020). On the one hand, misinformation benefits from the amplification effects of social and algorithmic curation, and from the use of clickbait and emotional triggers (Benklr et al., 2018; Chen et al., 2015; Vosoughi et al., 2018). On the other hand, the pockets of people actively circulating and consuming misinformation are actually very small (at least, as assessed using data from the U.S., e.g., Allen et al., 2020; Grinberg et al., 2019; Guess et al., 2020; Nelson & Taneja, 2016). Most of the information people consume online still comes from reliable news sources – which are, in any case, not the main destination for the majority of the online population (Allen et al., 2020; Altay et al., 2022; Yang et al., 2020). Misinformation, in other words, seems contained within small clusters of specific groups.

The question of reach, however, is separate from the question of impact. If, for instance, the pathways that lead to misinformation reveal partisan divides (i.e., if Republicans are more likely to engage with unreliable content, as prior research shows, Allcott & Gentzkow, 2017; Guess et al., 2019), then the small group of people consuming misinformation may still have a disproportionate impact on the politics of attention and polarizing group dynamics. Likewise, if

THE PUZZLE OF MISINFORMATION

age is the main demographic factor underlying engagement with unreliable content (e.g., Guess et al., 2019), and older groups are also those exhibiting more intense forms of political involvement (Krupnikov & Ryan, 2022), the impact of misinformation is also likely to be farther reaching. Thus, the main issue is not just how prevalent misinformation is but how it fits into people's broader news diets. If unreliable content is the main source of news, its impact is likely to be higher than if it is a small part of a broader, more reliable and ideologically diverse news inventory.

Existing studies have shown that people consuming unreliable content also consume a large amount of reliable news (e.g., Guess et al., 2021; Nelson & Taneja, 2016). Past research also shows that exposure to misinformation is fairly concentrated, with a small number of people driving most of that exposure (e.g., Allen et al., 2020; Grinberg et al., 2019; Watts et al., 2021). Here we provide additional analyses that support these past findings while casting novel light on the relationship between the diversity of news diets and exposure to unreliable content.

The literature on selective exposure, along with prior findings on the hyper-partisan nature of misinformation (e.g., Eady et al., 2023; Guay et al., 2022), seem to suggest that the consumers of unreliable content seek ideologically congruent information and, therefore, should have less ideologically diverse news diets. On the other hand, research also suggests that misinformation consumers often qualify as having higher political interest (e.g., Nelson & Taneja, 2016; Pennycook & Rand, 2021), with the implication that they are likely to be more voracious in their news consumption and visit a wider range of information sources, potentially feeding into a more ideologically diverse inventory of news content.

Our analyses aim to identify which of these alternative scenarios receives empirical support. To do so, we use uniquely rich panel data tracking the web browsing behavior of more

THE PUZZLE OF MISINFORMATION

than a hundred thousand panelists based in the U.S. for a period of twelve months (January to December 2018). The data contains the URLs visited by the panelists as well as their demographic attributes, which allows us to identify behavioral and demographic profiles associated with exposure to misinformation.

Before introducing our data, the following section offers a more detailed discussion of prior research and what it tells us about the prevalence of unreliable content and the audiences more likely to engage with it. We pay special attention to observational studies that analyze the relative volume of misinformation compared to reliable news; and studies that focus on the individual-level correlates of observed exposure to that content. We then describe our panel and how it alleviates some of the limitations of prior work, especially when it comes to answering our main guiding question: what other news do people engaging with misinformation consume? In answering this question, we discuss the puzzle of misinformation. The consumption of unreliable content is just one piece in the larger picture of broader news consumption; and, when analyzed, this larger picture reveals a paradox: that exposure to unreliable information is higher among the better informed. In the context of our data, ‘better informed’ is defined as having a higher engagement with reliable news and a more ideologically diverse news diet.

The puzzle we uncover is similar in nature to the tension between positive and negative engagement that has been studied on social media, i.e., the fact that beneficial behaviors, like increased participation in political talk, are also associated with dysfunctional consequences, like sharing more misinformation and relying on partisan outlets (Fletcher et al., 2021; Rossini et al., 2021; Valenzuela et al., 2019). It is also similar to the observed contradiction between the cynical attitudes towards news media that older adults declare and their tendency to still read and share news they distrust (Munyaka et al., 2022). Unpacking this puzzle, we argue, is important to

THE PUZZLE OF MISINFORMATION

inform interventions designed to curtail the effects of misinformation. For instance, approaches focused on increasing literacy or promoting cross-cutting, diverse exposure will not produce the theorized effects if a higher engagement with diverse, reliable news is positively associated with exposure to unreliable content. We conclude by advocating for a systemic approach to addressing the problem of misinformation (as opposed to just an individual approach centered on psychological mechanisms). We also revisit the concept of informed citizenry, and what it means in the context of our current news ecosystem and its information disorders.

The Study of Misinformation

The analysis of misinformation and so-called ‘fake news’ has been of special interest to academics and journalists since 2016. That year, two high-profile political events – the U.S. Presidential Election and the Brexit referendum – made patent the risks that online networks create in allowing the unmonitored dissemination of false and inaccurate content (Lazer et al., 2018). These risks were also seen and denounced in many other political contexts around the world, especially after the COVID-19 pandemic created new fronts on which to fight misinformation (Zarocostas, 2020). A common fear among observers and analysts is the impact that misinformation may have on election results, compliance with policy recommendations, and, more generally, the legitimacy and stability of the democratic process. Democracies need informed citizens to ensure that accountability mechanisms work. Misinformation not only weakens this foundation; it also threatens to aggravate conflict through misrepresentations and inaccurate portrayals of political and social realities.

The number of empirical articles that have been published on the topic of misinformation has grown substantially in the past six years, but the reported results fall in three broad categories: descriptive statistics measuring the prevalence and reach of misinformation (i.e.,

THE PUZZLE OF MISINFORMATION

Allen et al., 2020; Altay et al., 2022; Grinberg et al., 2019; Guess et al., 2020; Nelson & Taneja, 2018; Vosoughi et al., 2018; Watts, 2021); demographic characteristics of misinformation's likely audiences (e.g., Allcott & Gentzkow, 2017; Guess et al., 2019; Guess et al., 2021); and the impact of interventions designed to curtail or correct misinformation (e.g., Aslett et al., 2022; Pennycook & Rand, 2019; Walter et al., 2020; Vraga, et al., 2022). The studies differ in their empirical focus (i.e., Twitter, Facebook, YouTube, the web) and their measures of engagement with misinformation (e.g., they track exposure to web domains or engagement in the form of sharing on social media platforms). But, despite the different data frames and research designs, all these studies are consistent in their main findings: first, that misinformation is less prevalent, in terms of volume, than the amount of public attention to this problem would suggest; second, that key demographic variables like ideology and age seem to be driving the sharing of misinformation on social media; and third that, on average, people can accurately distinguish between lower- and higher-quality sources but flagging false content is not very effective in reducing misperceptions.

There is a fourth category of studies that use experimental designs outside the natural environment of media platforms, created solely for research purposes, i.e., studies that rely on stylized environments that aim to emulate specific features of real platforms (e.g., Jennings & Stroud, 2023; Porter et al., 2018; Rhodes, 2022; Thorson, 2016). This research aims to test the effects of different correction strategies on misperceptions and how alternative approaches to content labeling can help counterbalance the impact of misinformation on perceptions and beliefs. These studies help uncover important psychological and platform-design mechanisms that can amplify the impact of misinformation, but they are limited by concerns about the external validity of their results. For instance, the results of these lab experiments do not always

THE PUZZLE OF MISINFORMATION

line up with evidence from experiments conducted in the field (e.g., Aslett et al., 2022).

Observational data collected from natural environments, on the other hand, are correlational and less powerful to substantiate causal effects. In addition, most past observational research has focused on sharing behavior, which is an imperfect proxy for actual exposure, especially if we consider that social media often act as mediators that refer traffic elsewhere, i.e., the web (Lazer et al., 2021). Measuring who retweets what on Twitter, or who shares posts on Facebook, does not really measure who is seeing which content.

Beyond limitations about measurement and research designs, past work also leaves open theoretical questions about sources of heterogeneity in exposure to misinformation, especially in terms of which subpopulations are at higher risk of encountering it and engaging with it (Freelon & Wells, 2020). Age and ideology emerge from past work as two of the main predictors of sharing misinformation: older people and people that identify as Republican or Conservative are more likely to engage with unreliable content. However, it is unclear if these divides also hold for other forms of exposure on platforms other than social media. For instance, incidental exposure is less prominent on the web than on social media, i.e., most news consumption on the web results from intentional news-seeking behavior (Yang et al., 2020). Past research also raises the question of whether measures of party affiliation (Republican vs Democrat) are enough to characterize the actual ideological diversity of news diets, or the degree of heterogeneity in the news consumed by individuals that share the same party affiliation. Party affiliation captures political identity; the ideological diversity of news diets captures actual behavior and revealed content preferences.

Our analyses contribute new evidence to address some of these questions. We analyze observed exposure to misinformation on the web with two objectives. The first is to characterize

THE PUZZLE OF MISINFORMATION

the demographic profile of the group of people that consume misinformation. Specifically, we ask: are the demographic divides in exposure to misinformation consistent with what past research has observed on social media? One key difference between our research and past work is that the latter uses measures of engagement (e.g., shares) rather than measures of exposure (i.e., actual impressions or views). In addition, the web creates a very different news ecosystem because social and algorithmic curation are less prominent in determining exposure to content. Even if some of the traffic is driven by social media referrals, the web is a platform that relies on more intentional news seeking and is less impacted by the network effects of social media. The second objective is to determine how the population consuming unreliable information are engaging other news content: are they doing so at the expense of more reliable news, what are the characteristics of their broader news diets, and are they siloed within echo chambers and only reading ideologically congruent content? The answer to this question is particularly important to understand the broader impacts of misinformation and how to successfully address it with interventions.

High levels of political interest are associated to higher levels of education so if people consuming misinformation rank also high in their level of interest, designing programs to increase digital literacy is unlikely to have much impact: the main audiences of unreliable content already have the skills. Likewise, the effects of misinformation are likely to vary across people with different levels of ideological diversity in their news diets. Survey experiments suggest that cross-cutting exposure is one of the mechanisms that can strengthen individuals' ability to be critical in their evaluation of news (Rhodes, 2022). But if people consuming most misinformation are also those with the most diverse diets, in ideological terms, then encouraging more cross-cutting exposure is unlikely to have the theorized effect.

THE PUZZLE OF MISINFORMATION

These, of course, are all empirical possibilities that need to be substantiated with actual data from observed behavior. In the following section, we introduce the data we use to discriminate between these possible scenarios and reconstruct the news behavior of those who engage with unreliable content. We propose a novel measure to capture ideological diversity in news diets, which we use in conjunction with measures of overall interest in news content and demographic correlates to predict exposure to unreliable information. The following section gives more details on our data and methods, including how we operationalize our main dependent variable: exposure to unreliable content. Unlike most research on social media (which, again, uses measures of engagement, e.g., shares, rather than exposure), we measure exposure to unreliable content both as the count of unreliable pages visited and as the time spent on those pages.

Data and Methods

Panel Data

Our data tracks web browsing behavior for $N \sim 140,000$ unique users in the U.S. for the period of January to December of 2018. The data is provided by Nielsen, the media measurement company, and it is weighted to be representative of the U.S. population. One of the strengths of this data source is that it offers URL-level granularity, which means that we can analyze exposure to all the web pages the panelists visited during the observation window, as well as the time (in seconds) they spent visiting those pages and the parent domains. This data source also offers individual-level demographic attributes, such as gender, age, education, race, ethnicity, and income, covering more demographic dimensions than prior studies. Another strength is that the data offer repeated observations for the same individuals, which allows us to control for within-person temporal variability. As we show in figure 1A, half of the panelists in our data are

THE PUZZLE OF MISINFORMATION

active for at least four months, with $N \sim 16,000$ panelists active for the full 12 months we analyze. Compared to other data sources more restricted in sample size and temporal resolution, our data allow us to model individual exposure to unreliable information controlling for longitudinal trends and both within and between-person variability. In other words, the analysis of these data allows us to detect and measure statistical effects that other approaches, based on aggregated time series or cross-sectional data, cannot measure.

-- Figure 1 about here --

Variables

News Exposure. Our main variable of interest is exposure to unreliable information. We identify unreliable sources through the merging of two lists. The first is the list of domains classified as misinformation by Grinberg et al. (2019). This list, which includes $N = 510$ domains, was itself compiled using existing lists produced by journalistic outlets (i.e., BuzzFeed, Politifact, FactCheck.org, Snopes.com) and by prior research (Allcott & Gentzkow, 2017; Guess et al., 2020). Only 199 of these 510 domains (39%) were visited by our panelists at least once. The second list is provided by NewsGuard, a journalism and technology organization that rates the credibility of news and information websites. Each site receives a trust score on a 0-100 scale based on nine criteria, five related to credibility (i.e., the site does not repeatedly publish false content; gathers and presents information responsibly; regularly corrects or clarifies errors; handles the difference between news and opinion responsibly; avoids deceptive headlines) and four related to transparency (i.e., the website discloses ownership and financing; clearly labels advertising; reveals who's in charge, including possible conflicts of interest; and the site provides names of content creators, along with either contact or biographical information). Websites with a score below 60 points receive a red rating, which means that they are deemed to generally fail

THE PUZZLE OF MISINFORMATION

to meet basic standards of credibility and transparency and can be considered as unreliable sources, consistent with prior research (e.g., Aslett et al., 2022). Among the 69 misinformation outlets that are identified by Grinberg et al. (2019), visited by our panelists, and rated by NewsGuard, 58 of them (84%) have credibility scores below 60, signaling high recall rate and correlation between the two lists. We add these websites to our list of unreliable sources using the scores produced in 2018 (the same year for which we have the panel data). This brings the total of news domains classified as unreliable that were visited at least once by our panelists to $N = 504$. For each panelist we then count the number of pages from unreliable sources visited in each month, as well as the time spent on those pages (in seconds).

Another important variable in our analyses is exposure to news. To identify web domains that classify as reliable news (separate from the unreliable sources discussed in the previous paragraph), we merged the lists of news sources used in five previous published studies (i.e., Bakshy et al., 2015; Budak et al., 2016; Grinberg et al., 2019; Peterson et al., 2021; Yang et al., 2020). This merged list has a total of $N = 813$ domains, but only $N = 707$ (87%) was visited at least once by our panelists. In figure 1B we show the cumulative distribution function (CCDF) for the number of pages visited classified as news (blue curve) and those classified as unreliable news (red curve). These functions allow us to determine what percentage of the panelists visit at least x number of pages in each category. In line with prior research, the curves show that ~41% of the panelists do not visit any news page (i.e., a large fraction of the online population opts out of news, Yang et al., 2020); and that the vast majority, i.e., ~79%, do not visit any page classified as unreliable (again, confirming the finding that misinformation affects a small fraction of the online population, e.g., Allen et al., 2020). Still, $N \sim 1,400$ panelists (about 1%) visit at least 100 unreliable news pages during our observation period. This 1% of the panelists are responsible for

THE PUZZLE OF MISINFORMATION

visiting 65.3% of all unreliable pages in our data; by contrast, 1% of the panelists make up for 43.2% of all the reliable news pages visited during our observation window. Exposure to unreliable information is thus more concentrated and skewed than exposure to reliable news content.

These two variables, i.e., the count of news pages and the count of unreliable pages that panelists visit, are moderately correlated, as we show in Figure 1C. As mentioned, few panelists visit many news pages, and even fewer visit many unreliable sources; but for those who have higher counts, there is clearly a positive association, which offers the first piece of evidence supporting what we call the puzzle of misinformation. One of the main questions we want to address is whether this association remains statistically significant once we control for demographic covariates known to predict news consumption and engagement with unreliable content. In addition to measuring individual-level engagement with the news using the count of pages accessed, we also keep track of the time spent on those pages. In Figure 1D we plot the distribution of the average time panelists spent reading news and unreliable sources (in this figure, the measure, in seconds, is aggregated for the full year). In line with what the CCDF curves in panel 1B show, there is also less engagement with unreliable news compared to reliable sources when engagement is measured as time spent on those domains.

Ideological Scores of News Domains. To measure the ideological diversity of news diets, we first assign a media bias label to news domains based on an analyst-based crowdsourcing system provided by Ad Fontes, which samples prominently featured articles on news domains' websites and employs an ideologically balanced panel of experts to rate each news article's ideological slant (e.g., Aslett et al., 2022). This label ranges from -38.5 (most extreme liberal bias) to +38.5 (most extreme conservative bias). In the online Appendix, we report a robustness

THE PUZZLE OF MISINFORMATION

check employing audience-based metrics of ideological scores used in past research (e.g., Tyler et al., 2021; Yang et al., 2020). Our findings remain robust with both sets of ratings.

Demographic Variables. In addition to behavioral indicators of exposure to content on the web, we also have demographic information about the panelists: we have data about their gender, age, education, employment status, race, ethnicity, and income. Table 1 offers details of the proportion of panelists that fall in each demographic category, estimated with and without the weights provided by Nielsen (the statistical analyses that follow use the weighted data). Prior research on online misinformation has highlighted the positive impact that age has on sharing false content (i.e., Guess et al., 2019), but prior research has also shown that other attributes, like education and gender are, in general, also important predictors to understand who consumes news and political content (i.e., Mak, 2021; Scharkow et al., 2020; Shehata & Strömbäck, 2011). For instance, this research suggests that there is a clear gender divide: compared to men, women engage much less with the news. Yet more research is needed to document whether these gaps vary depending on the channel used to access news (i.e., social media vs the web), or if they also appear in the consumption of unreliable content.

In line with our theoretical discussion above, we change the analytical treatment of partisan identity used in past work: instead of focusing on self-disclosed party affiliation, we extrapolate the partisan leaning of our panelists by averaging ideological scores of the news pages they visited each month. This is a more flexible and granular measure than party affiliation because it can change from month to month (if news diets change) and allow us to differentiate individuals that would otherwise look identical in their party affiliation (i.e., two self-disclosed Democrats may have a very different news diet in terms of diversity).

THE PUZZLE OF MISINFORMATION

Modeling Approach

Like all panel data, our observations are nested: news sessions are nested within individual panelists who are nested within temporal aggregations (in our case, months). This data structure is rich enough to allow us to model intra- and inter-person variability. One common assumption when applying statistical models to observational data is that individuals with different characteristics behave differently; but it is also true that the same individuals may exhibit different news seeking behaviors over time. Events exogenous to the data (i.e., political affairs) may also drive overall levels of interest in the news. We account for all these sources of variability using linear mixed-effects models (Bates et al., 2015; Gelman & Hill, 2007). We use panelist ID and month as random effects and the demographic variables described in the previous section as fixed effects (again, using weights). Our main output variable is exposure to misinformation, which we make operational using the number of unreliable pages accessed and the time spent on those pages.

Our two main explanatory variables are interest in the news, which we measure as the number of news pages accessed; and the ideological diversity of news exposure, which we measure as the average distance in the ideological slants of news pages visited. The ideological score of the news page was the score of its news domain. Compared to prior research, we have more controls and more granular measurements, especially when it comes to measuring exposure to misinformation: instead of sharing behavior, we track actual exposure in the form of pages accessed and time spent on those pages. And because of the size of our panel data, which is orders of magnitude larger than the sample sizes used in most prior work, we can also model rare behavior (i.e., exposure to misinformation) with more statistical power. Our panel data also

THE PUZZLE OF MISINFORMATION

allow us to leverage fixed effects model to investigate intra-person changes, as we detail below and in our online Appendix.

Results

The first step in our analyses is to measure the diversity of news diets. In Figure 2A we show the distribution of ideological scores for the news domains in our data, with a small subset of labels shown for illustration, i.e., The Daily Kos (dailykos.com) is predominantly visited by Democrats; The Rush Limbaugh Show (rushlimbaugh.com) is predominantly accessed by Republicans. Using these scores, we calculate the pairwise ideological distance for all news pages visited by the panelists, and then average those distances to assign each panelist a monthly score that we use as a measure of their news diversity. For instance, if panelist i accessed three pages in month m from the Activist Post, BuzzFeed, and the New York Times, we first calculate the ideological distance between each pair of these outlets, and we then average the pairwise distances to a summary statistic for panelist i in month m . This average is smaller for panelists that access ideologically similar domains – so the larger it is in magnitude, the more diverse we can consider their news diet to be. In figure 2B we show the distribution of these scores (the inset shows the log-transformed version of the measure). In general, most panelists have a homogenous news diet, but there is some heterogeneity: because of the way in which the ideology scores are calculated, scores that seem small in magnitude are still signaling meaningful diversity in the information accessed by panelists. This measure allows us to differentiate panelists that would otherwise look identical if we were using their self-disclosed party affiliation.

-- Figure 2 about here --

THE PUZZLE OF MISINFORMATION

The ideological composition of news diets and the overall levels of engagement with news are our two key variables. As we show in figure 3, the most important predictor of exposure to unreliable content is overall interest in the news (as measured by the number of news pages accessed, log-transformed); controlling for this, panelists with news diets that show higher levels of ideological diversity are also exposed to more misinformation (note that the variables have been normalized using the mean value and standard deviation, so they are all on the same scale). As expected, given prior research, the most important demographic covariate is age: older panelists clearly engage more with misinformation than their younger counterparts. But within this age category, people with higher levels of interest and more ideologically diverse diets are consuming more unreliable content.

-- Figure 3 about here --

In the Appendix, we additionally show that this puzzle holds when we control the partisan leaning of panelists (Figure A1) or when we divide our panelists into five quantiles based on their partisan leaning (Table A3 and Table A4). In other words, for liberals and conservatives alike, ideologically diverse news diets consistently predict more exposure to unreliable information. Table A2 adopts fixed-effects models to control time-invariant demographic variables and event shocks. These models show that the puzzle holds when we examine intra-person effects, e.g., if the same person's news diet becomes more diverse from one month to the next, their exposure to unreliable information also becomes higher. Section B in the Appendix reports another set of robustness checks using an alternative measure of ideology, yielding similar findings.

Table A1 and Table B1 in the Appendix summarize outputs of alternative models with different specifications. Higher education levels, for instance, are associated with higher

THE PUZZLE OF MISINFORMATION

exposure to misinformation, but the association disappears when overall interest in news and the diversity of news diets are added to the model. Women, Asian, Black, and Hispanic panelists are also consuming less misinformation than white, male panelists. This is also consistent with what prior research shows in the context of social media (i.e., Guess et al., 2019). In summary, our results suggest that misinformation is not crowding out more reliable sources; instead, panelists with larger and more diverse news diets are also more likely to access unreliable content.

Discussion

Recent research has established that misinformation amounts to a small fraction of all the information circulating online, and that only a small number of people engage actively with it. Our findings are consistent with this pattern: only 1% of our panelists visit at least 100 unreliable pages, and they spend less time on those sources than on reliable news. But our findings also illuminate an important pattern not discussed in prior work: it is people with higher levels of news exposure *and* more diverse news diets that are also more likely to access unreliable sources. According to these findings, the mechanism that seems to be driving exposure to misinformation is overall political interest, which results in more expansive news diets, i.e., nets cast wider to gather more diverse news.

We know that political interest is correlated with education, and we also know that this is one of the main factors associated with political involvement: people exhibiting more intense news-seeking behavior are also more likely to be deeply involved in politics (Krupnikov & Ryan, 2022). This group of people are far from representing the population at large but they have a disproportionate influence in the distribution of political attention. The fact that they are also the main drivers of traffic to misinformation has consequences for how we think about effects and interventions. Most of the existing literature evaluating interventions to counteract the

THE PUZZLE OF MISINFORMATION

effects of misinformation predominantly focus on psychological processes and individual skills, such as levels of literacy (e.g., Guess et al., 2020; Vraga et al., 2022), accuracy prompts (e.g., Pennycook et al., 2021), and fact-checking (e.g., Walter et al., 2020). Our results suggest that those who encounter most unreliable information (on the web, in the U.S.) are also those with broader, more diverse news diets who already have literacy skills and access to reliable and accurate information. Seen through this light, the problem of misinformation is not a problem of literacy or correction. It is more systemic and it requires contextualizing exposure to unreliable context within the larger landscape of evolving news habits and the choices people make (often prompted by technological affordances and algorithmic curation) on how to navigate the larger information environment.

Our results are also constrained by some limitations that are likely to affect generalizability to other contexts. First, the web is an information ecosystem on its own, different from other widely studied contexts such as social media. Exposure to unreliable content on social media may be more incidental and therefore less driven by political interest. There is very little research offering estimates of exposure on social media platforms (cf., Allen et al., 2020; Guess et al., 2021; Moretto et al., 2022); most studies of social media focus on engagement (i.e., sharing or commenting), a measure that undercounts the actual number of people exposed. It is likely that political interest still plays a role in the posting of unreliable content: higher levels of involvement in politics are, after all, associated with higher activity on social media (Krupnikov & Ryan, 2022). Given current measurement limitations, we cannot compare exposure to unreliable content on the web with exposure to unreliable content on social media, but based on past research, we have reasons to believe that the patterns we identify on the web are likely to also translate into exposure to misinformation on social media. Meanwhile, social media

THE PUZZLE OF MISINFORMATION

platforms such as Instagram, YouTube, TikTok, and closed Facebook/WhatsApp groups are also major sources of political information, and misinformation in those channels is published through various modalities, not just through URLs (the basic unit of analysis in our study). For example, Yang et al. (2023), Peng et al., (2023), and Brennen et al. (2021) highlight the prevalence of visual misinformation in the forms of memes, videos, and photographs. These types of misinformation are more casual and light-hearted, which may attract audiences with lower levels of political interest. Thus, visual misinformation may reduce or reverse the positive association we identify here between political interest and news diet diversity, and exposure to unreliable content.

A second limitation is that we do not consider effects. Our analyses are observational, and they center on identifying patterns in news-seeking behavior. We can determine whether misinformation is consumed at the expense of reliable news (it is not); whether those exposed to more unreliable content have news diets that are ideologically siloed (they do not); and which groups of people are most likely to consume unreliable information (older white males with higher levels of political interest). However, on their own, these analyses cannot tell us if exposure to this content has any other behavioral or psychological effects – for instance, a higher engagement in the discussion of politics. This scenario would turn this unrepresentative group into the opinion leaders that could serve as credibility assessors (Wagner & Boczkowski, 2019); but it could also expand the potential impact of misinformation to other audiences. Future research needs to evaluate this possibility, which will require linking observational trails across media channels, i.e., tracking the same group of people as they use the web and other platforms. Our analyses are also agnostic about the persuasive impact of unreliable content, especially given that misinformation's most likely audiences are also the better informed. The exposure measures

THE PUZZLE OF MISINFORMATION

we analyze do not offer any insights about perceptions or attitudes. As qualitative research shows, consuming news is not necessarily an indicator of their perceived credibility or their impact on political beliefs (Munyaka et al., 2022; Weeks et al., 2021). Reading and sharing news (which is what we can observe with observational data) does not, on its own, tell us how much political learning or updating is happening as a consequence of that exposure.

Third, our results are suggestive of the nested puzzles that determine how misinformation finds an audience. Those with higher levels of news consumption (and higher political knowledge) are also those more likely to spend more time on unreliable news; and if, as past research suggests, more intense information seeking is correlated with deeper political involvement, then this sector of the population is also more likely to spread that content through their discussion networks, online and offline, thus amplifying its reach. Whether this amplification is accompanied with corrective statements or not is an empirical question that, again, our data cannot illuminate; it is possible that the better informed are in fact contributing to give more visibility (and credibility) to unreliable content. This possibility is supported by the fact that political engagement often serves as a key antecedent of misinformation sharing (Valenzuela et al., 2019). Unpacking these puzzles is, therefore, central to determine the best point of entry to implement interventions designed to curtail the spread of misinformation.

And fourth, as emphasized throughout the study, the findings reported here are bounded by the regional context of our data (i.e., the U.S. during a specific 12-month period). Considering the high level of partisan sorting, political polarization, and asymmetrical misinformation supply that characterize the U.S. media ecosystem, we are agnostic about the generalizability of our findings to other contexts and the universality of the puzzle we discuss. Our findings suggest the problem of misinformation requires a more systemic approach than individual-oriented studies,

THE PUZZLE OF MISINFORMATION

focused on psychological mechanisms, can afford. The implication is that different media and political systems can shape drastically the information-seeking patterns of news consumers. As web tracking panel data become more accessible in other political contexts, scholars should try to replicate our study to determine whether the paradox appears in those contexts as well.

Ultimately, our study sheds lights on the normative tensions underlying epistemic theories of democracy and how we think about the “informed citizenry” in the current media environment. Healthy information-seeking habits involve, in expectation, (1) consuming more news to keep up with current affairs; (2) contrasting ideologically-diverse sources; and (3) limiting exposure to unreliable information. And yet, our results suggest that when (1) and (2) are in place, (3) seems more difficult to achieve. This is still something that affects only a small fraction of the population, i.e., the group of people with high political interest and high levels of news engagement. Yet it begs the normative question of how to encourage news consumption while shielding users from unreliable sources – sources they are more likely to see the more intense and diverse their news diets become. The answer to this question requires thinking about information architectures and how people navigate them as a function of their choices and the opportunities the architecture offers. Offering a conclusive answer is beyond the scope of this study but the results offer a stepping stone in that direction.

References

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *The Journal of Economic Perspectives*, 31(2), 211-235. <http://www.jstor.org/stable/44235006>
- Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. <https://doi.org/10.1126/sciadv.aay3539>
- Altay, S., Kleis Nielsen, R., & Fletcher, R. (2022). Quantifying the “infodemic”: People turned to trustworthy news outlets during the 2020 coronavirus pandemic. *Journal of Quantitative Description: Digital Media*, 2. <https://doi.org/10.51685/jqd.2022.020>
- Aslett, K., Guess, A. M., Bonneau, R., Nagler, J., & Tucker, J. A. (2022). News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions. *Science Advances*, 8(18), eabl3844. <https://doi.org/doi:10.1126/sciadv.abl3844>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press. <https://books.google.com/books?id=6hhnDwAAQBAJ>

THE PUZZLE OF MISINFORMATION

- Brennen, J. S., Simon, F. M., & Nielsen, R. K. (2021). Beyond (mis)representation: Visuals in COVID-19 misinformation. *The International Journal of Press/Politics*, 26(1), 277–299. <https://doi.org/10.1177/1940161220964780>
- Budak, C., Goel, S., & Rao, J. M. (2016). Fair and balanced? Quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, 80(S1), 250–271. <https://doi.org/10.1093/poq/nfw007>
- Chen, Y., Conroy, N. J., & Rubin, V. L. (2015). *Misleading online content: Recognizing clickbait as "false news"*. Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, Seattle, Washington, USA. <https://doi.org/10.1145/2823465.2823467>
- Eady, G., Paskhalis, T., Zilinsky, J., Bonneau, R., Nagler, J., & Tucker, J. A. (2023). Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 U.S. election and its relationship to attitudes and voting behavior. *Nature Communications*, 14(1), 62. <https://doi.org/10.1038/s41467-022-35576-9>
- Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2021). More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the United Kingdom. *New Media & Society*. <https://doi.org/10.1177/14614448211027393>
- Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, 37(2), 145–156. <https://doi.org/10.1080/10584609.2020.1723755>
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

THE PUZZLE OF MISINFORMATION

- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374-378.
<https://doi.org/10.1126/science.aau2706>
- Guay, B., Berinsky, A., Pennycook, G., & Rand, D. G. (2022). *Examining partisan asymmetries in fake news sharing and the efficacy of accuracy prompt interventions* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/y762k>
- Guess, A., Aslett, K., Tucker, J., Bonneau, R., & Nagler, J. (2021). Cracking open the news feed: Exploring what U.S. Facebook users see and share with large-scale platform data. *Journal of Quantitative Description: Digital Media*, 1.
<https://doi.org/10.51685/jqd.2021.006>
- Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27), 15536–15545. <https://doi.org/10.1073/pnas.1920498117>
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), eaau4586.
<https://doi.org/10.1126/sciadv.aau4586>
- Guess, A. M., Nyhan, B., Lyons, B., & Reifler, J. (2018). *Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think*. Knight Foundation.
- Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 U.S. election. *Nature Human Behaviour*, 4(5), 472–480. <https://doi.org/10.1038/s41562-020-0833-x>

THE PUZZLE OF MISINFORMATION

- Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022). Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences*, 119(1), e2025334119. <https://doi.org/10.1073/pnas.2025334119>
- Jennings, J., & Stroud, N. J. (2023). Asymmetric adjustment: Partisanship and correcting misinformation on Facebook. *New Media & Society*, 25(7), 1501–1521. <https://doi.org/10.1177/14614448211021720>
- Krupnikov, Y., & Ryan, J. B. (2022). *The other divide: Polarization and disengagement in American politics*. Cambridge University Press.
- Lazer, D., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094-1096. <https://doi.org/10.1126/science.aao2998>
- Lazer, D., Hargittai, E., Freelon, D., Gonzalez-Bailon, S., Munger, K., Ognyanova, K., & Radford, J. (2021). Meaningful measures of human society in the twenty-first century. *Nature*, 595(7866), 189-196. <https://doi.org/10.1038/s41586-021-03660-7>
- Mak, M. K. F. (2021). Explaining the gender gap in news access across thirty countries: Resources, gender-bias signals, and societal development. *The International Journal of Press/Politics*, 0(0), 19401612211049444. <https://doi.org/10.1177/19401612211049444>
- Moretto, M., Ortellado, P., Kessler, G., Vommaro, G., Rodriguez-Raga, J. C., Luna, J. P., Heinen, E., Cely, L. F., & Toro, S. (2022). People are more engaged on Facebook as they get older, especially in politics: Evidence from users in 46 countries. *Journal of Quantitative Description: Digital Media*, 2. <https://doi.org/10.51685/jqd.2022.018>

THE PUZZLE OF MISINFORMATION

- Munyaka, I., Hargittai, E., & Redmiles, E. (2022). The misinformation paradox: Older adults are cynical about news media, but engage with it anyway. *Journal of Online Trust and Safety*, 1(4). <https://doi.org/10.54501/jots.v1i4.62>
- Nelson, J. L., & Taneja, H. (2018). The small, disloyal fake news audience: The role of audience availability in fake news consumption. *New Media & Society*, 20(10), 3720-3737. <https://doi.org/10.1177/1461444818758715>
- Nielsen, R. K., & Fletcher, R. (2022). Concentration of online news traffic and publishers' reliance on platform referrals: Evidence from passive tracking data in the UK. *Journal of Quantitative Description: Digital Media*, 2. <https://doi.org/10.51685/jqd.2022.015>
- Peng, Y., Lu, Y., & Shen, C. (2023). An agenda for studying credibility perceptions of visual misinformation. *Political Communication*, 40(2), 225–237. <https://doi.org/10.1080/10584609.2023.2175398>
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), 590–595. <https://doi.org/10.1038/s41586-021-03344-2>
- Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, 116(7), 2521-2526. <https://doi.org/10.1073/pnas.1806781116>
- Pennycook, G., & Rand, D. G. (2021). Research note: Examining false beliefs about voter fraud in the wake of the 2020 Presidential Election. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-51>

THE PUZZLE OF MISINFORMATION

- Peterson, E., Goel, S., & Iyengar, S. (2021). Partisan selective exposure in online news consumption: evidence from the 2016 presidential campaign. *Political Science Research and Methods*, 9(2), 242-258. <https://doi.org/10.1017/psrm.2019.55>
- Porter, E., Wood, T. J., & Kirby, D. (2018). Sex trafficking, Russian infiltration, birth certificates, and pedophilia: A survey experiment correcting fake news. *Journal of Experimental Political Science*, 5(2), 159-164. <https://doi.org/10.1017/XPS.2017.32>
- Rhodes, S. C. (2022). Filter bubbles, echo chambers, and fake news: How social media conditions individuals to be less critical of political misinformation. *Political Communication*, 39(1), 1-22. <https://doi.org/10.1080/10584609.2021.1910887>
- Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2021). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society*, 23(8), 2430-2451. <https://doi.org/10.1177/1461444820928059>
- Scharkow, M., Mangold, F., Stier, S., & Breuer, J. (2020). How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences*, 117(6), 2761-2763. <https://doi.org/10.1073/pnas.1918279117>
- Shehata, A., & Strömbäck, J. (2011). A matter of context: A comparative study of media environments and news consumption gaps in Europe. *Political Communication*, 28(1), 110-134. <https://doi.org/10.1080/10584609.2010.543006>
- Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, 33(3), 460-480. <https://doi.org/10.1080/10584609.2015.1102187>

THE PUZZLE OF MISINFORMATION

Tyler, M., Grimmer, J., & Iyengar, S. (2021). Partisan enclaves and information bazaars:

Mapping selective exposure to online news. *The Journal of Politics*, 84(2), 1057-1073.

<https://doi.org/10.1086/716950>

Valenzuela, S., Halpern, D., Katz, J. E., & Miranda, J. P. (2019). The paradox of participation

versus misinformation: Social media, political engagement, and the spread of

misinformation. *Digital Journalism*, 7(6), 802-823.

<https://doi.org/10.1080/21670811.2019.1623701>

Vraga, E. K., Bode, L., & Tully, M. (2022). Creating news literacy messages to enhance expert

corrections of misinformation on Twitter. *Communication Research*, 49(2), 245–267.

<https://doi.org/10.1177/0093650219898094>

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*,

359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>

Wagner, M. C., & Boczkowski, P. J. (2019). The reception of fake news: The interpretations and

practices that shape the consumption of perceived misinformation. *Digital*

Journalism, 7(7), 870-885. <https://doi.org/10.1080/21670811.2019.1653208>

Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: A meta-analysis of

what works and for whom. *Political Communication*, 37(3), 350–375.

<https://doi.org/10.1080/10584609.2019.1668894>

Watts, D. J., Rothschild, D. M., & Mobius, M. (2021). Measuring the news and its impact on

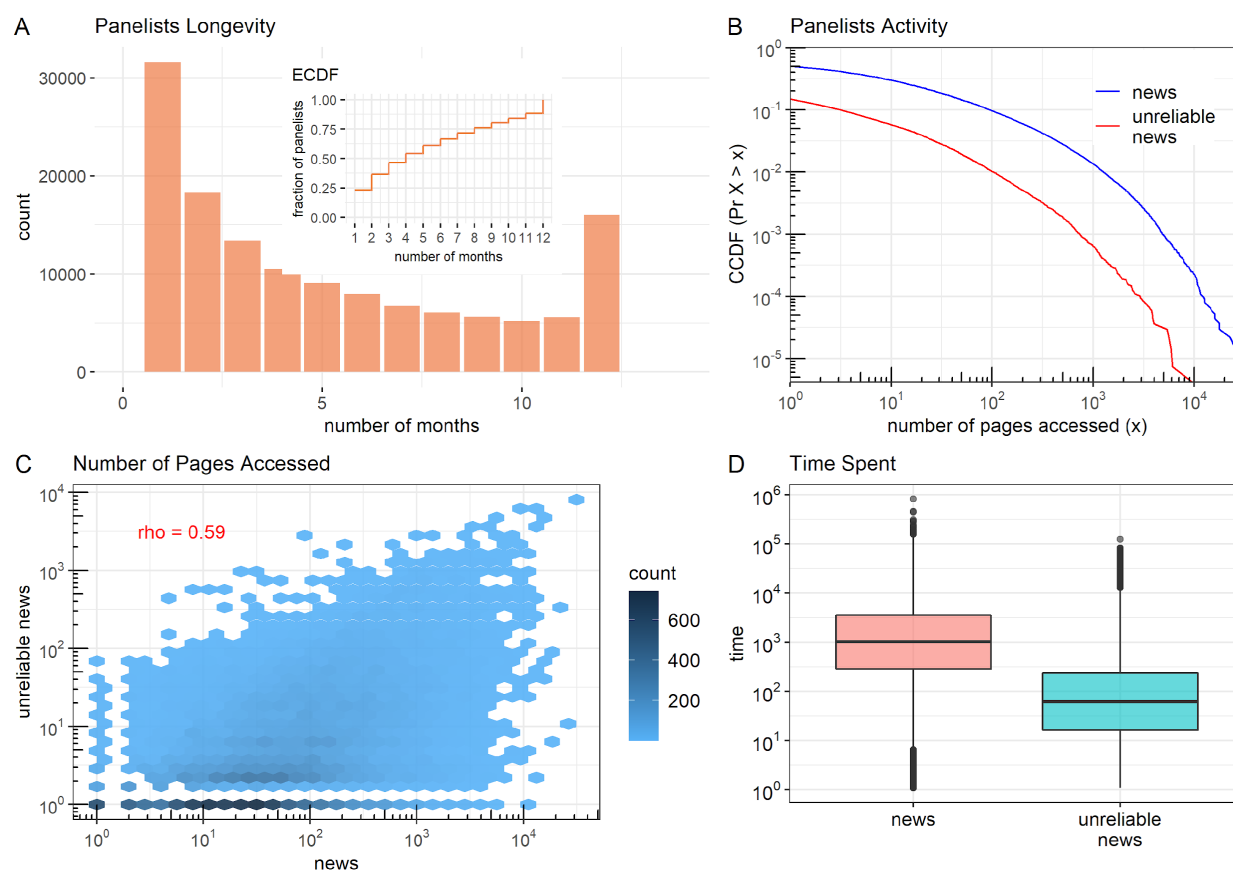
democracy. *Proceedings of the National Academy of Sciences*, 118(15), e1912443118.

<https://doi.org/10.1073/pnas.1912443118>

THE PUZZLE OF MISINFORMATION

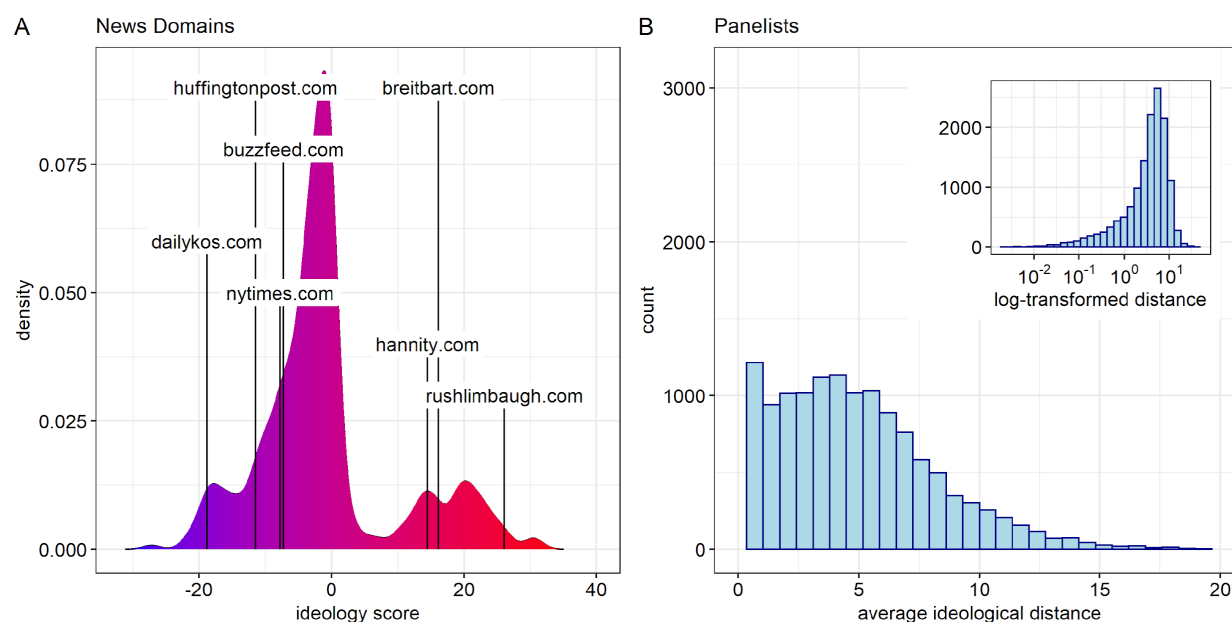
- Weeks, B. E., Menchen-Trevino, E., Calabrese, C., Casas, A., & Wojcieszak, M. (2021). Partisan media, untrustworthy news sites, and political misperceptions. *New Media & Society*, 146144482110333. <https://doi.org/10.1177/14614448211033300>
- Yang, Y., Davis, T., & Hindman, M. (2023). Visual misinformation on Facebook. *Journal of Communication*, jqac051. <https://doi.org/10.1093/joc/jqac051>
- Yang, T., Majó-Vázquez, S., Nielsen, R. K., & González-Bailón, S. (2020). Exposure to news grows less fragmented with an increase in mobile access. *Proceedings of the National Academy of Sciences*, 117(46), 28678-28683. <https://doi.org/10.1073/pnas.2006089117>
- Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, 395(10225).

THE PUZZLE OF MISINFORMATION

Figure 1. Description of the Data.

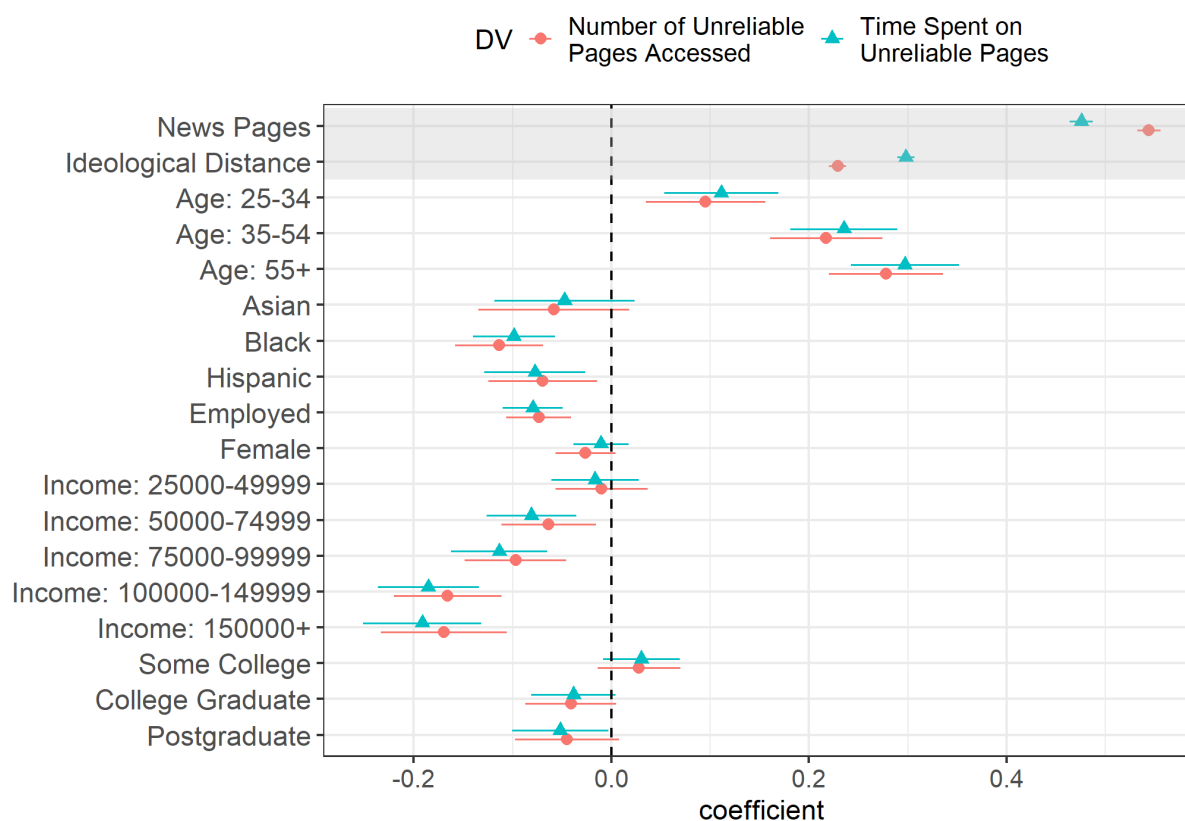
Note. (A) Longevity of panelists measured as the number of months in which they are active in the data (inset shows the empirical cumulative distribution function). (B) Complementary cumulative distribution function of the number of news pages (blue curve) and unreliable news pages (red curve) visited by all panelists. 1% of the panelists are responsible for visiting 43.2% of news pages but, for unreliable news, 1% of the panelists are responsible for visiting 65.3% of the pages. Exposure to unreliable information is thus more concentrated/skewed than exposure to news content. (C) Association between the number of news and unreliable news pages among panelists that visited both types of sites. (D) Distribution of time spent consuming news and unreliable news among panelists that visited both types of sites.

THE PUZZLE OF MISINFORMATION

Figure 2. Ideological Diversity of News Diets.

Note. (A) Ideological scores of the domains visited by the panelists. (B) Ideological diversity of news diets, measured as the average pairwise distance between the ideology of the news pages visited (inset shows the log-transformed version of the measure).

THE PUZZLE OF MISINFORMATION

Figure 3. Predictors of Exposure to Misinformation.

Note. These estimates result from a linear mixed effects model with panelist ID and month as random effects. Variables have been scaled (i.e., normalized using the mean value and standard deviation) and log-transformed where appropriate. See Table A1 and A2 in the Appendix for additional specifications and details.

THE PUZZLE OF MISINFORMATION

Table 1. Demographic Variables

Variables	%	
	weighted	unweighted
Gender		
Male	42	35
Female	58	65
Age		
18-24	10	14
25-34	18	21
35-54	36	35
55+	36	31
Education		
High School or Less	25	22
Some College	41	42
College Graduate	22	23
Postgraduate	12	12
Employment		
Working	72	35
Not Working	28	65
Race		
White	76	74
Black	13	16
Asian	4	4
Hispanic Origin		
Hispanic	14	10
Not Hispanic	86	90
Income		
< 24,999	11	23
25,000-49,999	23	28
50,000-74,999	24	21
75,000-99,999	17	13
100,000-149,999	15	10
> 150,000	9	5
<i>N</i>	136,054	136,054

Note. Percentage of panelists that fall in each demographic group, according to weighted and unweighted data. All statistical analyses use the weighted data. *N* refers to the number of panelists.

Appendix for**The Puzzle of Misinformation: Exposure to Unreliable Information is Higher among the
Better Informed**

Section A. Robustness Check: Alternative Statistical Models

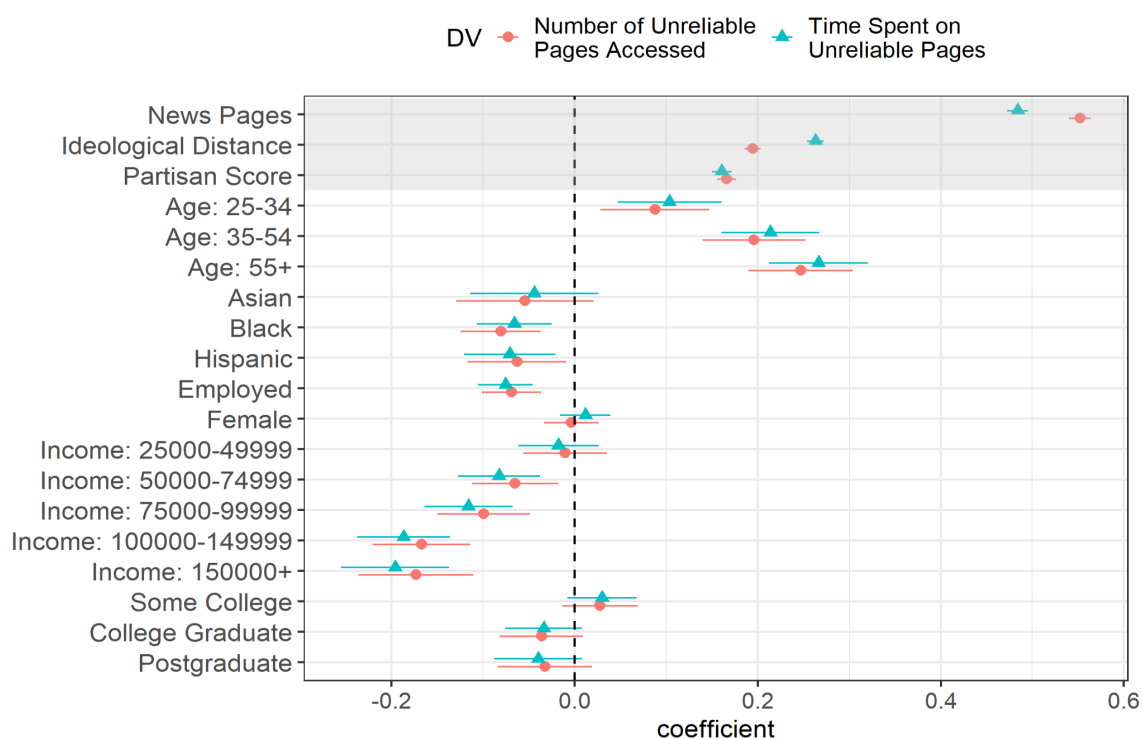
In this section, we report alternative statistical models accompanying findings in our main text.

Figure A1 replicates the Figure 3 in main text but adds partisan score, measured by average ideological scores of news pages visited, as a control variable which, as we described in the main text, equals the average of news sources' ideological ratings visited by the panelist for that month. This measure is more flexible and granular than self-reported categorical variables and helps us differentiate individuals that would otherwise look identical in their party affiliation. A higher partisan score means that the panelist leans more conservative/Republican. Figure A1 shows that conservative/Republican panelists have higher exposure to unreliable information, and that — controlling this partisan difference — ideologically diverse news diets still positively predict misinformation exposure.

Table A1 contains results visualized in Figure 3 in the main text and reports additional statistical models, which uses linear mixed-effects models (nested under panelist ID and month). Table A2 is similar to Table A1 but uses fixed-effects models.

Table A3 (using number of unreliable pages accessed as DV) and Table A4 (using time spent on unreliable pages as DV) report another robustness check. We split our panelists into five equal-size quantiles (ranging from very liberal to liberal to moderate to conservative to very conservative) based on their partisan scores, then ran the same model separately on each partisan group. The puzzle revealed by our main text holds for each partisan group: Diverse news diets positively predicts exposure to unreliable information.

THE PUZZLE OF MISINFORMATION

Figure A1. Predictors of Exposure to Misinformation (with Partisan Score).

Note. These estimates result from a linear mixed effects model with panelist ID and month as random effects. Variables have been scaled (i.e., normalized using the mean value and standard deviation) and log-transformed where appropriate. See Table A3 and A4 for additional specifications and details.

THE PUZZLE OF MISINFORMATION

Table A1. Linear Mixed-Effects Regression Models (1/2)

	<i>Dependent variable:</i>			
	Number of Unreliable Pages		Time on Unreliable Pages	
	(1)	(2)	(3)	(4)
Age: 25-34	0.026*** (0.008)	0.021*** (0.007)	0.095*** (0.023)	0.111*** (0.022)
Age: 35-54	0.112*** (0.008)	0.063*** (0.007)	0.217*** (0.022)	0.235*** (0.021)
Age: 55+	0.246*** (0.008)	0.102*** (0.007)	0.278*** (0.022)	0.297*** (0.021)
Female	-0.022*** (0.005)	-0.012*** (0.004)	-0.026** (0.012)	-0.011 (0.011)
Black	-0.103*** (0.007)	-0.055*** (0.006)	-0.114*** (0.017)	-0.099*** (0.016)
Asian	-0.037*** (0.012)	-0.031*** (0.011)	-0.058* (0.030)	-0.047* (0.028)
Other race	-0.015* (0.009)	0.0001 (0.008)	0.025 (0.025)	0.033 (0.024)
Some College	0.067*** (0.006)	0.022*** (0.005)	0.028* (0.016)	0.030** (0.015)
College Graduate	0.111*** (0.007)	0.011* (0.006)	-0.041** (0.018)	-0.038** (0.017)
Postgraduate	0.151*** (0.009)	0.018** (0.008)	-0.045** (0.020)	-0.052*** (0.019)
Income: 25000-49999	-0.0004 (0.007)	-0.004 (0.006)	-0.010 (0.018)	-0.017 (0.017)
Income: 50000-74999	-0.020*** (0.007)	-0.019*** (0.006)	-0.063*** (0.019)	-0.081*** (0.018)
Income: 75000-99999	-0.041*** (0.008)	-0.037*** (0.007)	-0.097*** (0.020)	-0.113*** (0.019)
Income: 100000-149999	-0.065*** (0.008)	-0.058*** (0.007)	-0.166*** (0.021)	-0.185*** (0.020)
Income: 150000+	-0.060*** (0.010)	-0.052*** (0.009)	-0.169*** (0.025)	-0.191*** (0.023)

THE PUZZLE OF MISINFORMATION

Table A1. Linear Mixed-Effects Regression Models (2/2)

Hispanic	-0.076*** (0.008)	-0.031*** (0.007)	-0.069*** (0.021)	-0.077*** (0.020)
Employed	-0.037*** (0.005)	-0.026*** (0.005)	-0.073*** (0.013)	-0.079*** (0.012)
News Pages		0.357*** (0.001)	0.544*** (0.005)	0.475*** (0.005)
Ideological Distance			0.229*** (0.003)	0.298*** (0.003)
Constant	-0.176 (0.544)	-0.034 (0.926)	-0.234 (0.425)	-0.112*** (0.032)
Random Effects				
Number of Panelists	108120	108120	45730	45730
Panelists Standard Deviation	0.654	0.547	0.904	0.778
Number of Months	12	12	12	12
Months Standard Deviation	1.885	3.208	1.471	0.064
Observations	579,070	579,070	170,895	170,895
Log Likelihood	-756,541.300	-730,014.300	-307,370.100	-310,896.100
Akaike Inf. Crit.	1,513,125.000	1,460,073.000	614,786.300	621,838.200
Bayesian Inf. Crit.	1,513,361.000	1,460,321.000	615,017.400	622,069.400

Note:

* p<0.1; ** p<0.05; *** p<0.01

Note. This table accompanies Figure 3 in the main text and provides more model specifications.

Column 3 and Column 4 are visualized in Figure 3 in the main text.

THE PUZZLE OF MISINFORMATION

Table A2. Fixed-Effects Regression Models

	<i>Dependent variable:</i>					
	Number of Unreliable Pages			Time on Unreliable Pages		
	(1)	(2)	(3)	(4)	(5)	(6)
News Pages	0.208*** (0.009)		0.120*** (0.005)	0.543*** (0.022)		0.386*** (0.013)
Ideological Distance	0.108*** (0.005)	0.130*** (0.006)		0.433*** (0.017)	0.488*** (0.019)	
Observations	208,078	208,078	695,664	208,078	208,078	695,664
R ²	0.781	0.768	0.750	0.697	0.686	0.665
Adjusted R ²	0.695	0.677	0.689	0.579	0.563	0.583
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01					

Note. These estimates result from a fixed-effects model with panelist ID and month as two fixed effects.

THE PUZZLE OF MISINFORMATION

Table A3. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, Number of Unreliable Pages as DV) (1/2)

	<i>Dependent variable:</i>						
	Number of Unreliable Pages						
	All (1)	All (2)	Very Liberal (3)	Liberal (4)	Moderate (5)	Conservative (6)	Very Conservative (7)
Age: 25-34	0.088*** (0.023)	0.087*** (0.023)	0.093*** (0.033)	0.109*** (0.031)	0.078*** (0.028)	0.064** (0.032)	0.031 (0.031)
Age: 35-54	0.196*** (0.022)	0.186*** (0.021)	0.140*** (0.031)	0.210*** (0.029)	0.156*** (0.026)	0.143*** (0.030)	0.062** (0.029)
Age: 55+	0.247*** (0.022)	0.213*** (0.022)	0.161*** (0.031)	0.200*** (0.029)	0.184*** (0.027)	0.176*** (0.031)	0.119*** (0.029)
Female	-0.004 (0.012)	0.004 (0.011)	0.002 (0.016)	0.073*** (0.016)	0.086*** (0.014)	0.046*** (0.015)	-0.002 (0.013)
Black	-0.081*** (0.017)	-0.079*** (0.017)	-0.006 (0.020)	-0.023 (0.023)	-0.030 (0.021)	-0.077*** (0.023)	-0.063*** (0.023)
Asian	-0.054* (0.029)	-0.038 (0.029)	-0.011 (0.044)	-0.048 (0.037)	0.004 (0.033)	-0.064* (0.038)	-0.001 (0.036)
Other race	0.033 (0.025)	0.032 (0.024)	0.048 (0.032)	0.039 (0.034)	0.006 (0.032)	-0.029 (0.034)	-0.011 (0.030)
Some College	0.028* (0.016)	0.025 (0.016)	0.019 (0.020)	-0.030 (0.023)	0.008 (0.020)	0.007 (0.021)	0.010 (0.018)
College Graduate	-0.036** (0.018)	-0.035** (0.017)	-0.008 (0.023)	-0.078*** (0.025)	-0.050** (0.022)	-0.075*** (0.023)	-0.048** (0.020)
Postgraduate	-0.033 (0.020)	-0.035* (0.020)	0.028 (0.026)	-0.111*** (0.027)	-0.074*** (0.024)	-0.107*** (0.027)	-0.047** (0.023)
Income: 25000-49999	-0.011 (0.018)	-0.017 (0.017)	0.006 (0.023)	-0.044* (0.025)	-0.009 (0.023)	-0.028 (0.024)	-0.018 (0.021)
Income: 50000-74999	-0.065*** (0.018)	-0.074*** (0.018)	-0.022 (0.024)	-0.054** (0.026)	-0.060** (0.024)	-0.052** (0.025)	-0.069*** (0.022)
Income: 75000-99999	-0.100*** (0.020)	-0.104*** (0.019)	-0.025 (0.025)	-0.076*** (0.028)	-0.077*** (0.025)	-0.085*** (0.027)	-0.092*** (0.023)
Income: 100000-149999	-0.168*** (0.021)	-0.175*** (0.020)	-0.056** (0.027)	-0.145*** (0.029)	-0.098*** (0.026)	-0.079*** (0.028)	-0.145*** (0.024)
Income: 150000+	-0.173*** (0.024)	-0.176*** (0.024)	-0.069** (0.033)	-0.085*** (0.033)	-0.131*** (0.030)	-0.139*** (0.033)	-0.192*** (0.028)

THE PUZZLE OF MISINFORMATION

Table A3. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, Number of Unreliable Pages as DV) (2/2)

Hispanic	-0.063*** (0.021)	-0.059*** (0.021)	-0.006 (0.028)	-0.046* (0.028)	-0.010 (0.025)	-0.007 (0.028)	-0.023 (0.026)
Employed	-0.069*** (0.013)	-0.059*** (0.012)	-0.044*** (0.016)	-0.033* (0.017)	-0.019 (0.016)	-0.037** (0.017)	-0.063*** (0.014)
News Pages	0.552*** (0.005)	0.541*** (0.004)	0.258*** (0.006)	0.380*** (0.007)	0.394*** (0.006)	0.417*** (0.007)	0.476*** (0.006)
Ideological Distance	0.194*** (0.003)	0.253*** (0.004)	0.224*** (0.005)	0.153*** (0.006)	0.153*** (0.005)	0.167*** (0.006)	0.133*** (0.004)
Partisan Score	0.166*** (0.004)	-0.026*** (0.005)	-0.231*** (0.006)	-0.003 (0.005)	-0.017*** (0.004)	0.013** (0.005)	0.221*** (0.005)
Partisan Score ²		0.133*** (0.002)					
Constant	-0.247 (0.397)	-0.346*** (0.032)	-0.108** (0.046)	-0.103 (0.074)	-0.146*** (0.040)	-0.083 (0.087)	-0.005 (0.038)
Observations	170,895	170,895	36,725	31,620	36,889	31,667	33,994
Log Likelihood	-306,581.600	-304,958.200	-44,221.900	-48,800.440	-57,332.650	-48,845.220	-41,284.620
Akaike Inf. Crit.	613,211.100	609,966.400	88,491.790	97,648.890	114,713.300	97,738.450	82,617.240
Bayesian Inf. Crit.	613,452.300	610,217.600	88,696.060	97,849.570	114,917.700	97,939.160	82,819.660

Note:

* p<0.1; ** p<0.05; *** p<0.01

THE PUZZLE OF MISINFORMATION

Table A4. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, Time Spent on Unreliable Pages as DV) (1/2)

	<i>Dependent variable:</i>						
	Time on Unreliable Pages						
	All (1)	All (2)	Very Liberal (3)	Liberal (4)	Moderate (5)	Conservative (6)	Very Conservative (7)
Age: 25-34	0.103*** (0.022)	0.103*** (0.022)	0.127*** (0.034)	0.109*** (0.030)	0.083*** (0.028)	0.038 (0.032)	0.045 (0.033)
Age: 35-54	0.214*** (0.021)	0.205*** (0.020)	0.174*** (0.032)	0.222*** (0.028)	0.162*** (0.026)	0.138*** (0.030)	0.102*** (0.030)
Age: 55+	0.266*** (0.021)	0.237*** (0.021)	0.186*** (0.032)	0.214*** (0.029)	0.188*** (0.026)	0.170*** (0.030)	0.173*** (0.030)
Female	0.012 (0.011)	0.019* (0.010)	0.004 (0.016)	0.077*** (0.015)	0.074*** (0.013)	0.047*** (0.014)	0.007 (0.014)
Black	-0.066*** (0.016)	-0.065*** (0.016)	0.017 (0.021)	-0.005 (0.022)	-0.016 (0.021)	-0.078*** (0.022)	-0.095*** (0.024)
Asian	-0.044 (0.027)	-0.029 (0.027)	0.022 (0.045)	-0.051 (0.036)	0.031 (0.032)	-0.048 (0.036)	-0.036 (0.038)
Other race	0.040* (0.023)	0.037 (0.023)	0.055* (0.033)	0.051 (0.034)	0.040 (0.031)	-0.008 (0.033)	-0.029 (0.032)
Some College	0.030** (0.015)	0.027* (0.015)	0.025 (0.020)	-0.034 (0.023)	0.006 (0.020)	0.009 (0.020)	0.018 (0.019)
College Graduate	-0.034** (0.016)	-0.033** (0.016)	-0.005 (0.023)	-0.091*** (0.024)	-0.047** (0.021)	-0.060*** (0.022)	-0.039* (0.021)
Postgraduate	-0.040** (0.019)	-0.042** (0.018)	0.018 (0.027)	-0.111*** (0.027)	-0.071*** (0.024)	-0.095*** (0.026)	-0.057** (0.024)
Income: 25000-49999	-0.018 (0.017)	-0.024 (0.017)	0.011 (0.024)	-0.045* (0.025)	-0.004 (0.023)	-0.042* (0.024)	-0.009 (0.022)
Income: 50000-74999	-0.083*** (0.017)	-0.091*** (0.017)	-0.028 (0.024)	-0.053** (0.026)	-0.058** (0.023)	-0.080*** (0.024)	-0.093*** (0.023)
Income: 75000-99999	-0.116*** (0.019)	-0.120*** (0.018)	-0.040 (0.026)	-0.067** (0.027)	-0.087*** (0.025)	-0.104*** (0.026)	-0.122*** (0.024)
Income: 100000-149999	-0.187*** (0.020)	-0.194*** (0.019)	-0.059** (0.028)	-0.148*** (0.028)	-0.109*** (0.026)	-0.108*** (0.028)	-0.178*** (0.026)
Income: 150000+	-0.196*** (0.023)	-0.198*** (0.022)	-0.096*** (0.034)	-0.074** (0.032)	-0.157*** (0.029)	-0.175*** (0.032)	-0.226*** (0.030)

THE PUZZLE OF MISINFORMATION

Table A4. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, Time Spent on Unreliable Pages as DV) (2/2)

Hispanic	-0.071*** (0.019)	-0.067*** (0.019)	0.006 (0.028)	-0.052* (0.027)	-0.035 (0.024)	-0.022 (0.027)	-0.039 (0.027)
Employed	-0.076*** (0.012)	-0.067*** (0.011)	-0.053*** (0.016)	-0.046*** (0.017)	-0.040*** (0.015)	-0.049*** (0.016)	-0.068*** (0.015)
News Pages	0.484*** (0.004)	0.472*** (0.004)	0.225*** (0.006)	0.346*** (0.007)	0.379*** (0.006)	0.383*** (0.007)	0.396*** (0.006)
Ideological Distance	0.263*** (0.004)	0.317*** (0.004)	0.254*** (0.005)	0.194*** (0.006)	0.183*** (0.005)	0.231*** (0.006)	0.189*** (0.005)
Partisan Score	0.161*** (0.004)	-0.013** (0.005)	-0.198*** (0.006)	-0.014*** (0.005)	-0.014*** (0.004)	0.017*** (0.005)	0.225*** (0.006)
Partisan Score ²		0.120*** (0.002)					
Constant	-0.124*** (0.031)	-0.213*** (0.032)	-0.137*** (0.039)	-0.087** (0.039)	-0.121*** (0.036)	-0.035 (0.265)	-0.004 (0.064)
Observations	170,895	170,895	36,725	31,620	36,889	31,667	33,994
Log Likelihood	-310,152.200	-308,872.400	-47,324.890	-48,949.350	-57,250.700	-49,368.880	-44,470.260
Akaike Inf. Crit.	620,352.400	617,794.900	94,697.780	97,946.710	114,549.400	98,785.760	88,988.520
Bayesian Inf. Crit.	620,593.600	618,046.100	94,902.050	98,147.380	114,753.800	98,986.470	89,190.930

Note:

* p<0.1; ** p<0.05; *** p<0.01

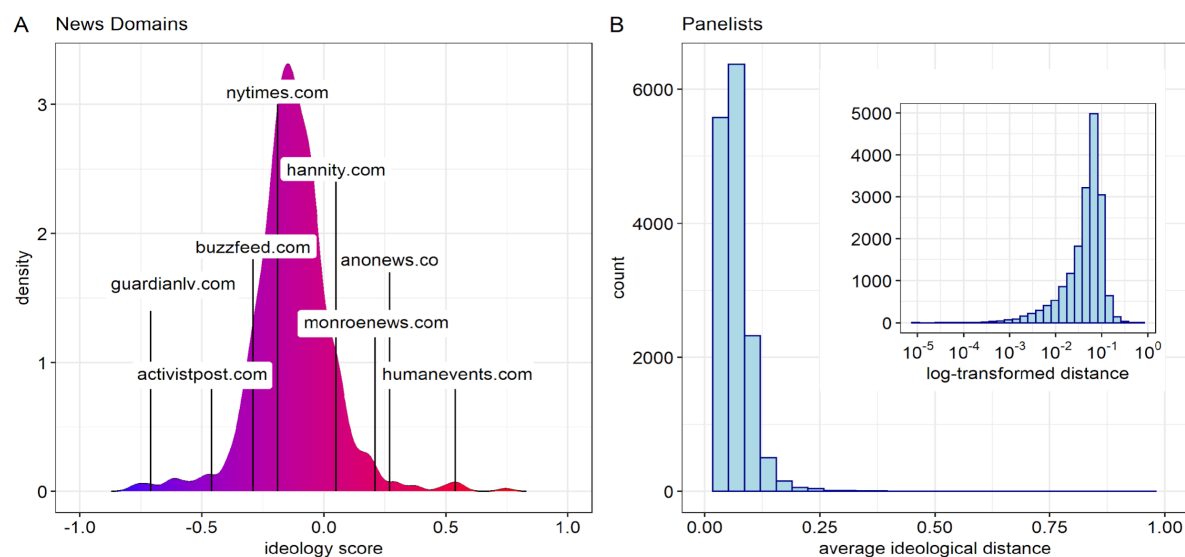
Section B. Robustness Check: Audience-Based Ideological Ratings

In this robustness check, we employ the audience-based metric of ideological slant used in past research (e.g., Tyler et al., 2021; Yang et al., 2020), which they call “favorability score,” to label our news domains. This measure is defined as the ratio $\frac{R_n - D_n}{R_n + D_n}$, where R_n counts the number of people accessing a domain n that identifies as Republican, and D_n counts the number of people accessing the domain that identifies as Democrat. The measure scales so that the closer the score is to -1, the more predominantly Democrat the audience of a domain is, and vice versa: the closer the score is to 1, the more predominantly Republican the audience (values closer to 0 suggest an ideologically mixed audience).

Figure B1 corresponds to Figure 2 in our main text. Figure B2 corresponds to Figure 3 in our main text. Figure B3 corresponds to Figure A1 in the Appendix. Table B1, B2, B3, and B4 correspond to Table A1, A2, A3, and A4.

All findings reported in our main text hold using this alternative measure.

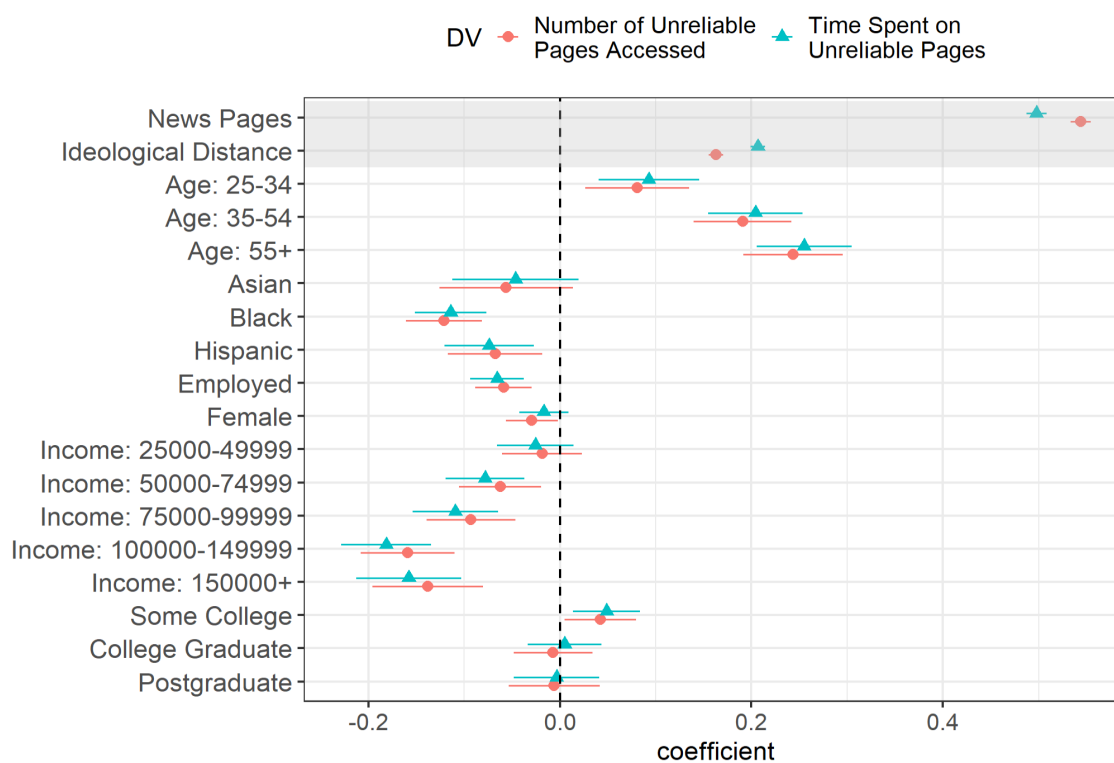
THE PUZZLE OF MISINFORMATION

Figure B1. Ideological Diversity of News Diets (with Audience-Based Ideological Ratings).

Note. This figure corresponds to Figure 2 in the main text but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.

THE PUZZLE OF MISINFORMATION

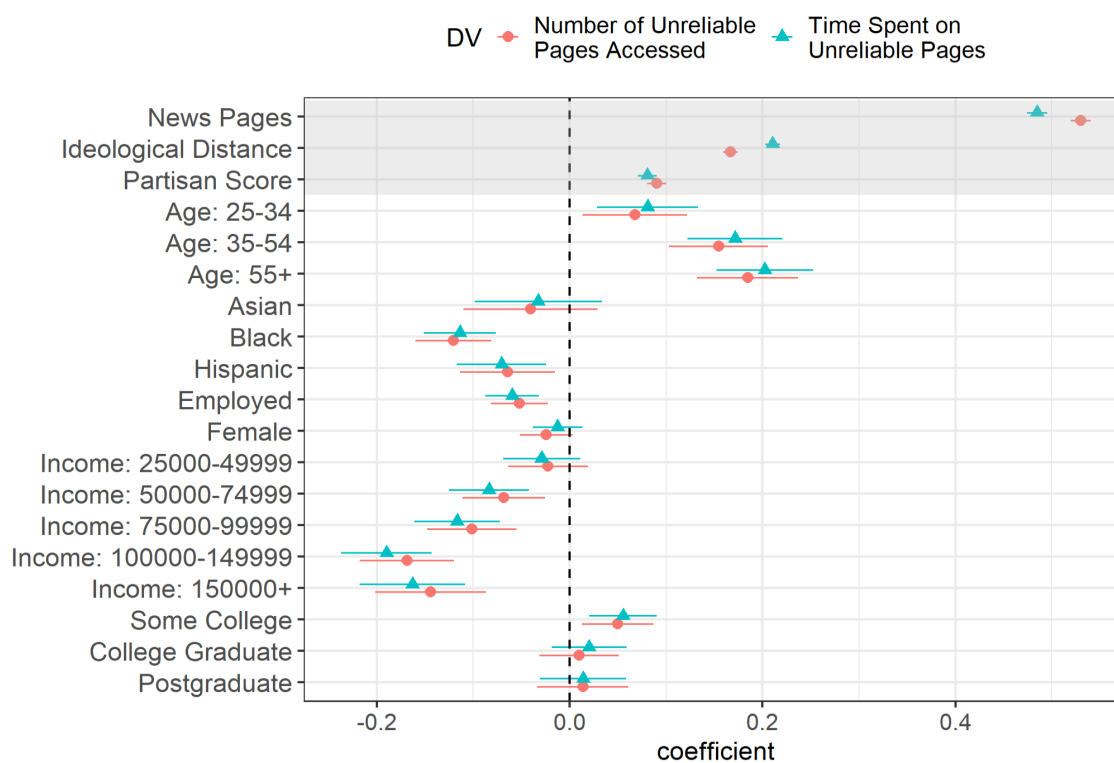
Figure B2. Predictors of Exposure to Misinformation (with Audience-Based Ideological Ratings)



Note. This figure corresponds to Figure 3 in the main text but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.

THE PUZZLE OF MISINFORMATION

Figure B3. Predictors of Exposure to Misinformation (with Audience-Based Ideological Ratings and Partisan Score)



Note. This figure corresponds to Figure A1 in the Appendix but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.

THE PUZZLE OF MISINFORMATION

Table B1. Linear Mixed-Effects Regression Models (with Audience-Based Ideological Ratings) (1/2)

	<i>Dependent variable:</i>			
	Number of Unreliable Pages			Time on Unreliable Pages
	(1)	(2)	(3)	(4)
Age: 25-34	0.026*** (0.008)	0.021*** (0.007)	0.081*** (0.021)	0.093*** (0.020)
Age: 35-54	0.112*** (0.008)	0.063*** (0.007)	0.191*** (0.020)	0.204*** (0.019)
Age: 55+	0.246*** (0.008)	0.102*** (0.007)	0.243*** (0.020)	0.255*** (0.019)
Female	-0.022*** (0.005)	-0.012*** (0.004)	-0.029*** (0.011)	-0.017* (0.010)
Black	-0.103*** (0.007)	-0.055*** (0.006)	-0.121*** (0.015)	-0.114*** (0.015)
Asian	-0.037*** (0.012)	-0.031*** (0.011)	-0.056** (0.027)	-0.046* (0.026)
Other race	-0.015* (0.009)	0.0001 (0.008)	0.025 (0.022)	0.033 (0.022)
Some College	0.067*** (0.006)	0.022*** (0.005)	0.042*** (0.014)	0.049*** (0.014)
College Graduate	0.111*** (0.007)	0.011* (0.006)	-0.007 (0.016)	0.005 (0.015)
Postgraduate	0.151*** (0.009)	0.018** (0.008)	-0.006 (0.018)	-0.003 (0.017)
Income: 25000-49999	-0.0004 (0.007)	-0.004 (0.006)	-0.019 (0.016)	-0.025 (0.016)
Income: 50000-74999	-0.020*** (0.007)	-0.019*** (0.006)	-0.062*** (0.017)	-0.078*** (0.016)
Income: 75000-99999	-0.041*** (0.008)	-0.037*** (0.007)	-0.093*** (0.018)	-0.109*** (0.017)
Income: 100000-149999	-0.065*** (0.008)	-0.058*** (0.007)	-0.159*** (0.019)	-0.181*** (0.018)
Income: 150000+	-0.060*** (0.010)	-0.052*** (0.009)	-0.138*** (0.022)	-0.158*** (0.021)

THE PUZZLE OF MISINFORMATION

Table B1. Linear Mixed-Effects Regression Models (with Audience-Based Ideological Ratings) (2/2)

Hispanic	-0.076*** (0.008)	-0.031*** (0.007)	-0.068*** (0.019)	-0.074*** (0.018)
Employed	-0.037*** (0.005)	-0.026*** (0.005)	-0.059*** (0.011)	-0.066*** (0.011)
News Pages		0.357*** (0.001)	0.544*** (0.004)	0.498*** (0.004)
Ideological Distance			0.163*** (0.003)	0.207*** (0.003)
Constant	-0.176 (0.544)	-0.034 (0.926)	-0.273 (0.425)	-0.185*** (0.036)
Random Effects				
Number of Panelists	108120	108120	50986	50986
Panelists Standard Deviation	0.654	0.547	0.867	0.763
Number of Months	12	12	12	12
Months Standard Deviation	1.885	3.208	1.469	0.095
Observations	579,070	579,070	200,350	200,350
Log Likelihood	-756,541.300	-730,014.300	-352,604.600	-359,051.500
Akaike Inf. Crit.	1,513,125.000	1,460,073.000	705,255.200	718,149.100
Bayesian Inf. Crit.	1,513,361.000	1,460,321.000	705,490.000	718,383.900
<i>Note:</i>			* p<0.1; ** p<0.05; *** p<0.01	

Note. This table corresponds to Table A1 in the Appendix but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.

Table B2. Fixed-Effects Regression Models (with Audience-Based Ideological Ratings)

	<i>Dependent variable:</i>					
	Number of Unreliable Pages			Time on Unreliable Pages		
	(1)	(2)	(3)	(4)	(5)	(6)
News Pages	0.200*** (0.009)		0.120*** (0.005)	0.544*** (0.020)		0.386*** (0.013)
Ideological Distance	1.553*** (0.107)	1.485*** (0.121)		5.900*** (0.361)	5.715*** (0.389)	
Observations	243,775	243,775	695,664	243,775	243,775	695,664
R ²	0.768	0.754	0.750	0.682	0.669	0.665
Adjusted R ²	0.684	0.665	0.689	0.566	0.548	0.583
<i>Note:</i>				*p<0.1; **p<0.05; ***p<0.01		

Note. This table corresponds to Table A2 in the Appendix but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.

THE PUZZLE OF MISINFORMATION

Table B3. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, with Audience-Based Ideological Ratings, Number of Unreliable Pages as DV) (1/2)

	<i>Dependent variable:</i>						
	Number of Unreliable Pages						
	All (1)	All (2)	Very Liberal (3)	Liberal (4)	Moderate (5)	Conservative (6)	Very Conservative (7)
Age: 25-34	0.068*** (0.021)	0.069*** (0.021)	0.092*** (0.025)	0.113*** (0.027)	0.069** (0.029)	0.066* (0.037)	0.032 (0.036)
Age: 35-54	0.154*** (0.020)	0.158*** (0.020)	0.159*** (0.024)	0.178*** (0.026)	0.127*** (0.027)	0.109*** (0.034)	0.104*** (0.033)
Age: 55+	0.185*** (0.020)	0.187*** (0.020)	0.207*** (0.026)	0.156*** (0.026)	0.150*** (0.027)	0.176*** (0.034)	0.156*** (0.033)
Female	-0.024** (0.011)	-0.025** (0.011)	0.031* (0.016)	0.083*** (0.014)	0.067*** (0.013)	0.018 (0.013)	-0.101*** (0.014)
Black	-0.120*** (0.015)	-0.120*** (0.015)	-0.032 (0.022)	0.038* (0.021)	-0.058*** (0.019)	-0.102*** (0.021)	-0.195*** (0.020)
Asian	-0.041 (0.027)	-0.040 (0.027)	-0.027 (0.034)	-0.023 (0.033)	-0.060* (0.033)	-0.027 (0.040)	-0.038 (0.044)
Other race	0.029 (0.022)	0.029 (0.022)	0.019 (0.033)	0.035 (0.031)	0.0004 (0.029)	0.014 (0.031)	0.012 (0.030)
Some College	0.050*** (0.014)	0.050*** (0.014)	-0.007 (0.022)	-0.011 (0.021)	0.008 (0.018)	0.026 (0.018)	0.068*** (0.017)
College Graduate	0.010 (0.016)	0.011 (0.016)	-0.062*** (0.024)	-0.080*** (0.022)	-0.065*** (0.020)	-0.030 (0.020)	0.067*** (0.020)
Postgraduate	0.014 (0.018)	0.015 (0.018)	-0.066** (0.028)	-0.141*** (0.025)	-0.093*** (0.022)	-0.020 (0.024)	0.102*** (0.023)
Income: 25000-49999	-0.022 (0.016)	-0.023 (0.016)	-0.004 (0.024)	-0.048** (0.023)	-0.013 (0.021)	-0.017 (0.022)	0.001 (0.020)
Income: 50000-74999	-0.069*** (0.017)	-0.068*** (0.017)	-0.041 (0.025)	-0.033 (0.024)	-0.049** (0.021)	-0.066*** (0.022)	-0.048** (0.021)
Income: 75000-99999	-0.101*** (0.018)	-0.101*** (0.018)	-0.073*** (0.027)	-0.062** (0.026)	-0.058** (0.023)	-0.079*** (0.024)	-0.083*** (0.022)
Income: 100000-149999	-0.169*** (0.019)	-0.168*** (0.019)	-0.168*** (0.029)	-0.103*** (0.027)	-0.094*** (0.024)	-0.109*** (0.026)	-0.092*** (0.024)
Income: 150000+	-0.144*** (0.022)	-0.143*** (0.022)	0.011 (0.034)	-0.084*** (0.031)	-0.139*** (0.027)	-0.153*** (0.029)	-0.122*** (0.029)

THE PUZZLE OF MISINFORMATION

Table B3. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, with Audience-Based Ideological Ratings, Number of Unreliable Pages as DV) (2/2)

Hispanic	-0.064*** (0.019)	-0.063*** (0.019)	-0.082*** (0.026)	-0.013 (0.025)	-0.010 (0.023)	-0.045* (0.026)	-0.080*** (0.027)
Employed	-0.052*** (0.011)	-0.051*** (0.011)	-0.011 (0.018)	-0.021 (0.016)	-0.016 (0.014)	-0.040*** (0.014)	-0.051*** (0.014)
News Pages	0.530*** (0.004)	0.530*** (0.004)	0.315*** (0.007)	0.379*** (0.007)	0.442*** (0.006)	0.405*** (0.006)	0.310*** (0.005)
Ideological Distance	0.167*** (0.003)	0.168*** (0.003)	0.110*** (0.006)	0.053*** (0.006)	0.127*** (0.005)	0.204*** (0.005)	0.177*** (0.004)
Partisan Score	0.090*** (0.004)	0.089*** (0.004)	0.047*** (0.007)	0.061*** (0.005)	0.003 (0.004)	0.027*** (0.004)	0.098*** (0.005)
Partisan Score ²		0.015*** (0.002)					
Constant	-0.240 (0.427)	-0.257 (0.427)	-0.113*** (0.037)	-0.137** (0.065)	-0.119* (0.069)	-0.085 (0.101)	-0.034 (0.314)
Observations	200,350	200,350	32,051	39,035	42,971	42,403	43,890
Log Likelihood	-352,324.100	-352,301.700	-51,858.520	-61,569.250	-62,699.280	-59,185.060	-48,245.270
Akaike Inf. Crit.	704,696.200	704,653.500	103,765.000	123,186.500	125,446.600	118,418.100	96,538.540
Bayesian Inf. Crit.	704,941.200	704,908.700	103,966.000	123,392.200	125,654.600	118,625.800	96,747.090

Note:

* p<0.1; ** p<0.05; *** p<0.01

Note. This table corresponds to Table A3 in the Appendix but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.

THE PUZZLE OF MISINFORMATION

Table B4. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, with Audience-Based Ideological Ratings, Time Spent on Unreliable Pages as DV) (1/2)

	<i>Dependent variable:</i>						
	Time on Unreliable Pages						
	All (1)	All (2)	Very Liberal (3)	Liberal (4)	Moderate (5)	Conservative (6)	Very Conservative (7)
Age: 25-34	0.081*** (0.020)	0.081*** (0.020)	0.096*** (0.025)	0.103*** (0.026)	0.074** (0.030)	0.086** (0.037)	0.016 (0.037)
Age: 35-54	0.172*** (0.019)	0.172*** (0.019)	0.162*** (0.024)	0.176*** (0.025)	0.133*** (0.028)	0.149*** (0.034)	0.114*** (0.034)
Age: 55+	0.202*** (0.019)	0.203*** (0.019)	0.214*** (0.026)	0.150*** (0.026)	0.158*** (0.028)	0.203*** (0.034)	0.181*** (0.034)
Female	-0.013 (0.010)	-0.013 (0.010)	0.043*** (0.015)	0.082*** (0.014)	0.070*** (0.013)	0.020 (0.013)	-0.108*** (0.014)
Black	-0.114*** (0.015)	-0.114*** (0.015)	-0.020 (0.022)	0.056*** (0.021)	-0.061*** (0.019)	-0.102*** (0.020)	-0.221*** (0.021)
Asian	-0.032 (0.026)	-0.032 (0.026)	-0.004 (0.033)	0.002 (0.032)	-0.054* (0.033)	-0.017 (0.040)	-0.053 (0.045)
Other race	0.035 (0.022)	0.035 (0.022)	0.013 (0.032)	0.048 (0.030)	0.019 (0.029)	0.026 (0.031)	-0.001 (0.031)
Some College	0.055*** (0.014)	0.056*** (0.014)	-0.017 (0.021)	0.004 (0.020)	0.012 (0.018)	0.024 (0.018)	0.078*** (0.018)
College Graduate	0.020 (0.015)	0.020 (0.015)	-0.058** (0.023)	-0.065*** (0.022)	-0.066*** (0.020)	-0.022 (0.020)	0.080*** (0.021)
Postgraduate	0.014 (0.017)	0.014 (0.017)	-0.070*** (0.027)	-0.128*** (0.024)	-0.091*** (0.022)	-0.033 (0.023)	0.118*** (0.024)
Income: 25000-49999	-0.029* (0.016)	-0.029* (0.016)	0.002 (0.023)	-0.060*** (0.023)	-0.017 (0.021)	-0.019 (0.022)	0.004 (0.021)
Income: 50000-74999	-0.084*** (0.016)	-0.084*** (0.016)	-0.048** (0.024)	-0.058** (0.023)	-0.059*** (0.021)	-0.072*** (0.022)	-0.061*** (0.022)
Income: 75000-99999	-0.117*** (0.017)	-0.116*** (0.017)	-0.071*** (0.026)	-0.071*** (0.025)	-0.070*** (0.023)	-0.093*** (0.024)	-0.094*** (0.023)
Income: 100000-149999	-0.190*** (0.018)	-0.190*** (0.018)	-0.200*** (0.028)	-0.113*** (0.026)	-0.110*** (0.024)	-0.119*** (0.025)	-0.104*** (0.025)
Income: 150000+	-0.163*** (0.021)	-0.163*** (0.021)	0.004 (0.033)	-0.127*** (0.030)	-0.164*** (0.027)	-0.146*** (0.029)	-0.160*** (0.030)

THE PUZZLE OF MISINFORMATION

Table B4. Linear Mixed-Effects Regression Models (with Partisan Score and for Various Partisan Groups, with Audience-Based Ideological Ratings, Time Spent on Unreliable Pages as DV) (2/2)

Hispanic	-0.070*** (0.018)	-0.070*** (0.018)	-0.076*** (0.025)	-0.007 (0.024)	-0.023 (0.023)	-0.065** (0.026)	-0.098*** (0.028)
Employed	-0.059*** (0.011)	-0.059*** (0.011)	-0.014 (0.017)	-0.044*** (0.016)	-0.032** (0.014)	-0.050*** (0.014)	-0.061*** (0.014)
News Pages	0.485*** (0.004)	0.485*** (0.004)	0.296*** (0.007)	0.362*** (0.006)	0.426*** (0.006)	0.381*** (0.006)	0.265*** (0.005)
Ideological Distance	0.210*** (0.003)	0.211*** (0.003)	0.134*** (0.006)	0.072*** (0.005)	0.148*** (0.005)	0.248*** (0.005)	0.223*** (0.005)
Partisan Score	0.081*** (0.004)	0.080*** (0.004)	0.043*** (0.007)	0.068*** (0.005)	-0.00002 (0.004)	0.013*** (0.004)	0.099*** (0.005)
Partisan Score ²		0.003 (0.002)					
Constant	-0.155*** (0.036)	-0.159*** (0.036)	-0.112 (0.173)	-0.111*** (0.034)	-0.105 (0.204)	-0.093 (0.080)	-0.034 (0.053)
Observations	200,350	200,350	32,051	39,035	42,971	42,403	43,890
Log Likelihood	-358,835.800	-358,840.100	-52,039.960	-61,239.370	-64,776.020	-60,834.080	-52,401.840
Akaike Inf. Crit.	717,719.600	717,730.200	104,127.900	122,526.700	129,600.000	121,716.200	104,851.700
Bayesian Inf. Crit.	717,964.600	717,985.400	104,328.900	122,732.500	129,808.100	121,923.900	105,060.200

Note:

* p<0.1; ** p<0.05; *** p<0.01

Note. This table corresponds to Table A4 in the Appendix but adopts news slant measures used by past research (e.g., Tyler et al., 2021; Yang et al., 2020) to calculate the ideological distance.