Exposure to (AI-Generated) Untrustworthy Websites in the 2024 U.S. Presidential Election

Ross Dahlke*
Assistant Professor
School of Journalism and Mass Communication
University of Wisconsin–Madison
dahlke2@wisc.edu

Ryan C. Moore*
Assistant Professor
School of Information
The University of Texas at Austin
ryanmoore@utexas.edu

Jeffrey T. Hancock
Harry and Norman Chandler Professor of Communication
Department of Communication
Stanford University
hancockj@stanford.edu

* These authors contributed equally.

Abstract

Generative AI's ability to produce untrustworthy content at scale has heightened concerns about its influence on elections. We provide preregistered evidence of exposure to AI-generated untrustworthy websites during the 2024 U.S. Presidential Election, combining web-browsing logs and surveys from a national sample of American adults (N = 1,069; 6M visits) with expert databases of websites. The share of Americans visiting untrustworthy websites continued the decline witnessed from 2020 and 2016 to reach even lower levels. However, those exposed in 2024 visited more frequently, indicating concentrated engagement. Older adults and Republicans remained more likely to encounter such sites, though gaps narrowed, and Twitter/X's role as a referrer increased. *AI-generated* untrustworthy sites reached more people in 2024 but comprised a small fraction of overall exposure with no pronounced demographic disparities.The results highlight both the persistence of risks and the need for targeted resilience strategies as generative AI tools evolve.

Keywords: 2024 election, untrustworthy websites, AI-generated misinformation, digital trace data, web browsing

Main

Untrustworthy online content–information that contradicts the best available evidence from experts at the time[1]–is a major concern in contemporary society. This concern has been heightened with the advent of generative Artificial Intelligence (AI), which has the capability of producing untrustworthy content on a massive scale with content that is higher-quality, more persuasive, and more personalized[2–5]. Exposure to untrustworthy information online can increase political polarization[6], exacerbate public health crises (e.g., by reducing vaccine uptake)[7–9] and contribute to the decline of trust in democratic institutions[10,11]. Concerns about the negative societal consequences of untrustworthy content continue to grow. The World Economic Forum's 2024 Global Risks Report, which draws on input from 1,400 global risks experts, policy-makers, and industry leaders, ranked misinformation and disinformation as the number one risk facing the planet over the next two years[12]. Those experts' concerns are primarily driven by fears that advances in generative AI technologies will be used to generate and spread mis/disinformation at scales and in forms not seen before.

As concerns over untrustworthy content have mounted, so has scholarly research on documenting people's exposure to such content online[13]. Understanding the extent to which people are exposed to untrustworthy information and the nature of their exposures is important to contextualize the scope of untrustworthy information's reach. Documenting exposure can also suggest directions for effective interventions to build resilience against untrustworthy content, identifying who is most likely to be exposed to dubious information and how the content fits into their broader media experiences.

Exposure to Untrustworthy Information Over Time

Much research on untrustworthy information exposure relies on self-reported exposure measures (e.g., [14–24]). While insightful, these measures are limited by two key factors that make drawing strong conclusions difficult. First, people may not accurately remember exposures to untrustworthy content[25]. Self-reported media use is weakly correlated with actual use[26–29], and

self-reports of political information consumption are particularly inaccurate[30,31]. Second, people may not detect untrustworthy information as such upon exposure[32].

A smaller body of work leverages passively-collected digital trace data of people's day-to-day media use to identify exposures to untrustworthy online sources objectively. These studies rely on software that participants install on their personal devices that logs the websites they visit, which researchers then pair with databases of websites known to disseminate unreliable content to identify exposures. Digital trace studies have identified a general decline in exposure to untrustworthy websites over time[33–36]. The overall incidence of exposure has declined among American adults since the 2016 election. The groups most likely to consume untrustworthy information, older adults and political conservatives, have remained consistent across time[34,35,37–39].

Exposure to AI-Generated Untrustworthy Information

While evidence suggests that exposure to untrustworthy sources has declined over the past two U.S. elections[34], experts are concerned that exposure to untrustworthy content may have surged during the 2024 election[40–47] because of the availability of tools that leverage AI to generate content (e.g., OpenAI's ChatGPT). Beyond increasing the quantity of untrustworthy content that spreads online, generative AI may also create content that is higher-quality, more persuasive, and more personalized[2–5] – factors that may increase user engagement. AI was one of the key drivers behind the World Economic Forum's 2024 Global Risks Report rating disinformation as the number one challenge facing the globe over the next two years[12]. The American public is also concerned that AI will be used to disseminate misinformation during elections[49]. A national survey found that 83.4% of Americans expressed some level of concern about AI's use in election misinformation[50]. Supporting the public's concerns, there is evidence that AI has bolstered state-backed disinformation campaigns[51].

Researchers are now exploring how much untrustworthy content generated by AI appears online. Hanley & Durumeric found that synthetic news articles on untrustworthy

domains rose by 457% from January 2022 to May 2023, increasing exponentially when ChatGPT was released in late 2022[52]. In laboratory settings, individuals often struggle to distinguish AI-generated text from human-generated text[3,53–55], images[56–60], and videos[61,62], prompting concerns that AI could be successfully used to generate unreliable content [3]. Nonetheless,although this is a highly active area of research, with some recent evidence suggesting AI-generated untrustworthy content tends to be perceived as less accurate than human-generated untrustworthy content[63]. Despite evidence that untrustworthy websites are increasingly using AI, and that humans generally struggle to identify AI-generated content, little existing research documents how much people are *actually* exposed to AI-generated untrustworthy sources in their everyday media use.

The Present Study

In this study, we provide evidence addressing the concerns of experts and citizens by documenting Americans' exposure to untrustworthy websites online during the 2024 presidential election, a period when many fear that generative AI would be used to flood the ecosystem with unreliable information[46,49]. In a preregistered study (preregistration document available at https://doi.org/10.17605/OSF.IO/CS3Z6), we collected web browsing data from a national sample of American adults via YouGov Pulse ($N_{participants}$ = 1,069, $N_{WebVisits}$ = 6M), a program where consenting participants install software on their computers, smartphones, and tablets that passively collects the websites that they visit, from October 8, 2024 (four weeks before the election) to November 12, 2024 (one week after the election). Following our conceptual definition of untrustworthy content, we identified databases of domains evaluated by experts that would allow us to operationally identify untrustworthy websites in participants' web browsing data. Specifically, we compiled two distinct lists of website domains to identify general untrustworthy and AI-generated untrustworthy websites. First, to facilitate over-time comparisons, we compiled a list of 2,573 general untrustworthy domains following the general

approach of Moore et al.[34], who compared exposure to untrustworthy websites during the 2020 and 2016 US elections, but made critical updates to the list to reflect changes in the information ecosystem (more details in Methods section). Second, to specifically assess exposure to AI-generated untrustworthy websites, we compiled a separate list using NewsGuard's Unreliable AI-generated News Sites (UAINS) database (for more information about NewsGuard's rating process, see[64]). This list contained 810 domains meeting four criteria: (1) a substantial portion of content produced by AI; (2) strong evidence of publication without significant human oversight (e.g., containing AI error messages or chatbot-specific language); (3) content presented in a way that could be mistaken for human-generated content; and (4) a lack of clear disclosure about AI generation. We then matched these lists against participants' URL-level browsing logs to identify which visits participants made were to domains in either list.

We paired web browsing data with a survey to capture participants' sociodemographics (e.g., age, gender, race/ethnicity, level of education) and political attitudes and beliefs. Our approach (collecting web browsing data via YouGov Pulse paired with a survey) allows us to make apples-to-apples comparisons to prior studies of Americans' exposure to untrustworthy websites during the 2016[35] and 2020[34] elections that used the same data source and analytical strategy. Furthermore, our approach to identifying exposure to untrustworthy sources allows us to distinguish between non-AI-generated and AI-generated misinformation, directly addressing the core concern about the misinformation landscape in 2024.

We focused our analyses on two primary questions: (1) *How did Americans' exposure to untrustworthy websites during the 2024 election compare to exposure during the 2016 and 2020 elections?* (2) *To what extent was the untrustworthy information that people were exposed to during the 2024 election AI-generated as opposed to non-AI-generated?* Our findings reveal that fewer Americans visited untrustworthy websites in 2024 than in 2020 or 2016, marking a continued decline in overall exposure. However, those who did visit such sites tended to do so more frequently, reflecting an intensification of exposure among a smaller group. Older adults

and Republicans remained more likely than younger adults and Democrats to encounter untrustworthy websites, suggesting that existing demographic disparities persisted although gaps narrowed relative to past elections. We observed a significant increase in Twitter/X's role as a referrer to untrustworthy content—consistent with our hypothesis that its influence might grow. This hypothesis was based on work showing that Twitter/X has reduced trust and safety staff and programs[65–67], which in the past effectively curbed the spread of untrustworthy content on the platform[68,69]. We also found that Facebook's share of referrals declined substantially from 2020 and 2016.

Meanwhile, in 2024, 6.5% of Americans encountered AI-generated untrustworthy websites, up from 2.6% in 2020. However, these AI-generated sites accounted for a smaller fraction of the total pool of untrustworthy visits than they did in 2020. Moreover, Americans did not spend more time on AI-generated untrustworthy sites compared to non-AI sites, and we detected no pronounced age or partisan differences in exposure to AI-generated untrustworthy websites. Taken together, these results indicate that although overall exposure to untrustworthy content continues to decline, intensified engagement persists among certain groups. While AI-generated untrustworthy content reaches more people than before, its share remains small and less demographically polarized than exposure to non-AI-generated untrustworthy content.

Results

Total consumption of untrustworthy websites in 2024

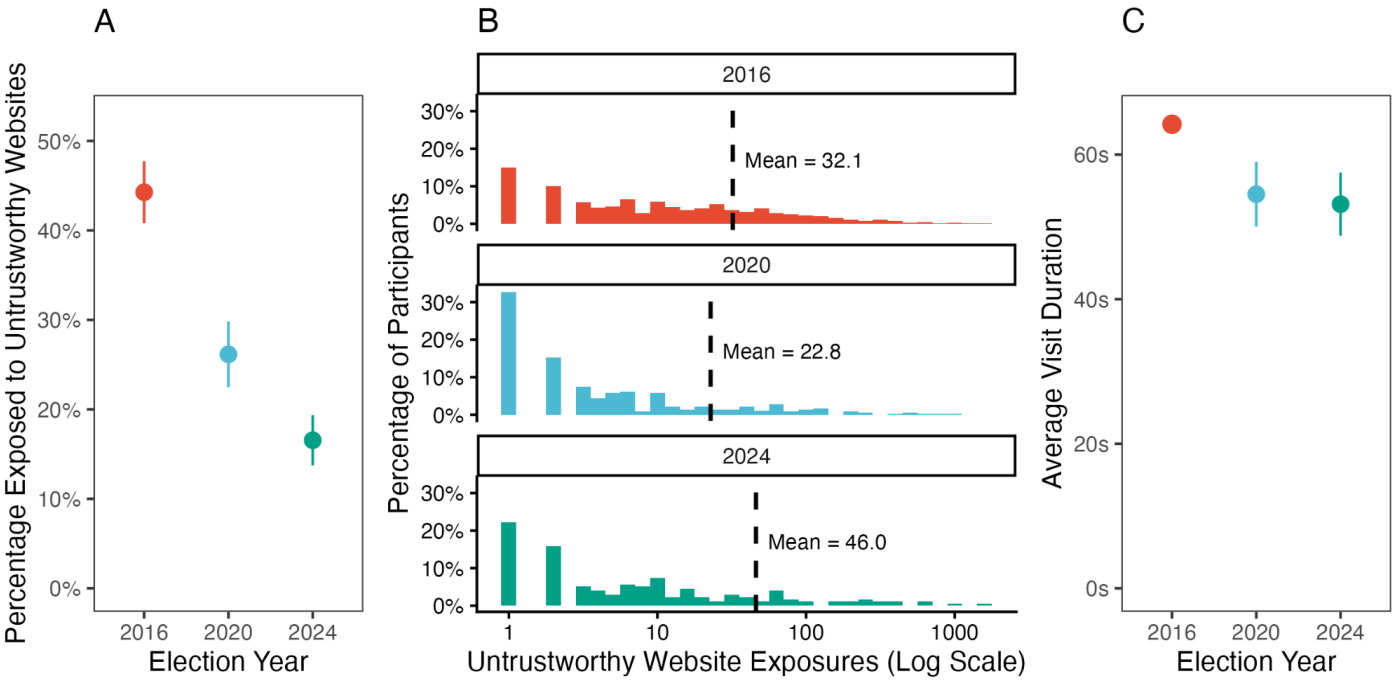
To begin, we analyze our preregistered hypotheses (see Supplementary Information S1). In 2024, 16.5% (95% CI [13.7%, 19.4%]) of Americans visited at least one untrustworthy website, a significant decline from the 26.1% (95% CI [22.5%, 29.8%]) who did so in 2020 (H1, $t = -5.567$, $p < 0.001$, Cohen's $d = -0.236$, one-tailed) and the 44.3% (95% CI [40.8%, 47.7%]) who did so in 2016 (see Figure 1). These results indicate that the proportion of Americans exposed to untrustworthy websites has fallen over time, and this pattern holds at different

thresholds of the number of visits needed to consider someone “exposed”, from one visit to more than 50 (see Supplementary Information S2).

While the number of people visiting untrustworthy websites has fallen, one notable difference compared to 2020 is the intensity of exposures among those who encounter such websites. The average number of visits to untrustworthy websites among those exposed rose from 22.8 (95% CI [15.2, 30.4]) in 2020 to 46.0 (95% CI [23.4, 68.6]) in 2024 (H2, $t = 1.90$, $p = 0.029$, Cohen’s $d = 0.193$, one-tailed; see Figure 1; for a CCDF plot, which similarly shows the heavy-tailed distribution of exposures, see Supplementary Information S3). The average duration per visit was statistically unchanged—53 seconds (95% CI [49, 58]) in 2024 versus 55 seconds (95% CI [50, 59]) in 2020 (H3, $t = -0.42$, $p = 0.661$, Cohen’s $d = -0.006$, one-tailed).

This pattern of results suggest that although the number of overall visitors to untrustworthy websites decreased in 2024, those who did visit untrustworthy sites did so more frequently, with visits still lasting approximately as long as in 2020. This concentration pattern might suggest a shift from incidental or accidental exposure toward more intentional seeking behavior among a dedicated subset of users. We found no significant decline from 2020 to 2024 in the number of Americans accessing credible “hard news” outlets or in their frequency (see Supplementary Information S4), suggesting that the reduction in visits to untrustworthy websites was not merely a byproduct of a reduction in the consumption of news more broadly.

Fig. 1: Untrustworthy website exposure across U.S. presidential elections.



Note: Untrustworthy website exposure during the 2016, 2020, and 2024 presidential elections. Panel A shows overall exposure rates, with the y-axis representing the percentage of participants who visited at least one untrustworthy website. Data are weighted means ± 95% CIs. Panel B shows the distribution of visit frequency among all participants, with the x-axis representing number of visits (log scale) and y-axis representing percentage of participants. Exposure is concentrated among a small subset of participants. Vertical lines represent means.

Who consumed untrustworthy websites in 2024?

Consistent with observations from 2016 and 2020, older adults continued to be more likely to visit untrustworthy websites than younger adults (see Figure 2A). In 2024, 24.8% (95% CI [19.3%, 30.3%]) of adults aged 65 years or older visited at least one untrustworthy site, compared with 13.8% (95% CI [0.6%, 27.0%]) of those aged 18–29 (H4, t = 2.41, p = 0.008, Cohen's d = 0.28, one-tailed). Although both older and younger adults show lower absolute exposure rates relative to 2020, the age gap remains consistent with patterns observed in

previous elections. While this bivariate comparison reveals a significant difference, after controlling for other sociodemographic and political variables examined in prior studies, we do not find a significant difference between those aged 65 and older and those aged 18-29 (the reference group in the regression in Table 1), despite similar regression analyses showing a significant difference between the two age groups in 2016 and 2020. One explanation for this discrepancy is that the variance in older adults' exposure to untrustworthy websites is increasingly explained by other factors, such as partisanship and political interest, rather than age. Indeed, post-hoc auxiliary regressions indicated that other factors significantly associated with untrustworthy website exposure (Republican affiliation, political interest) were also significantly associated with being aged 65+, even more so than in past elections (see Supplementary Information S5). Moreover, while older adults remain more likely to encounter untrustworthy websites, differences in the intensity of engagement across age groups are less pronounced, and such sites continue to make up only a small portion of overall website visits (see Supplementary Information S6).

Given that age groups are made up of different mixtures of generational cohorts (e.g., Baby Boomers, Gen X, Gen Z), it may be instructive to analyze exposure to untrustworthy content across generations. Generations often engage with different media platforms and have been socialized to different media ecosystems (e.g., cable TV dominant eras vs. social media dominant eras)[70–72]. Indeed, breaking down changes in exposure to untrustworthy websites over time by generation (see Figure 2B) shows that, while exposure has fallen across all generational cohorts from 2016 to 2024, the timing and magnitude of those declines varied considerably, with the Silent Generation experiencing the largest decline between the 2020 and 2024 elections while Baby Boomers and Millennials experienced their largest declines in exposure between the 2016 and 2020 elections. Generational differences were less pronounced when considering intensity of engagement, with most groups visiting untrustworthy sites at similar rates once exposed (Supplementary Information S6). The Silent Generation is one notable exception,

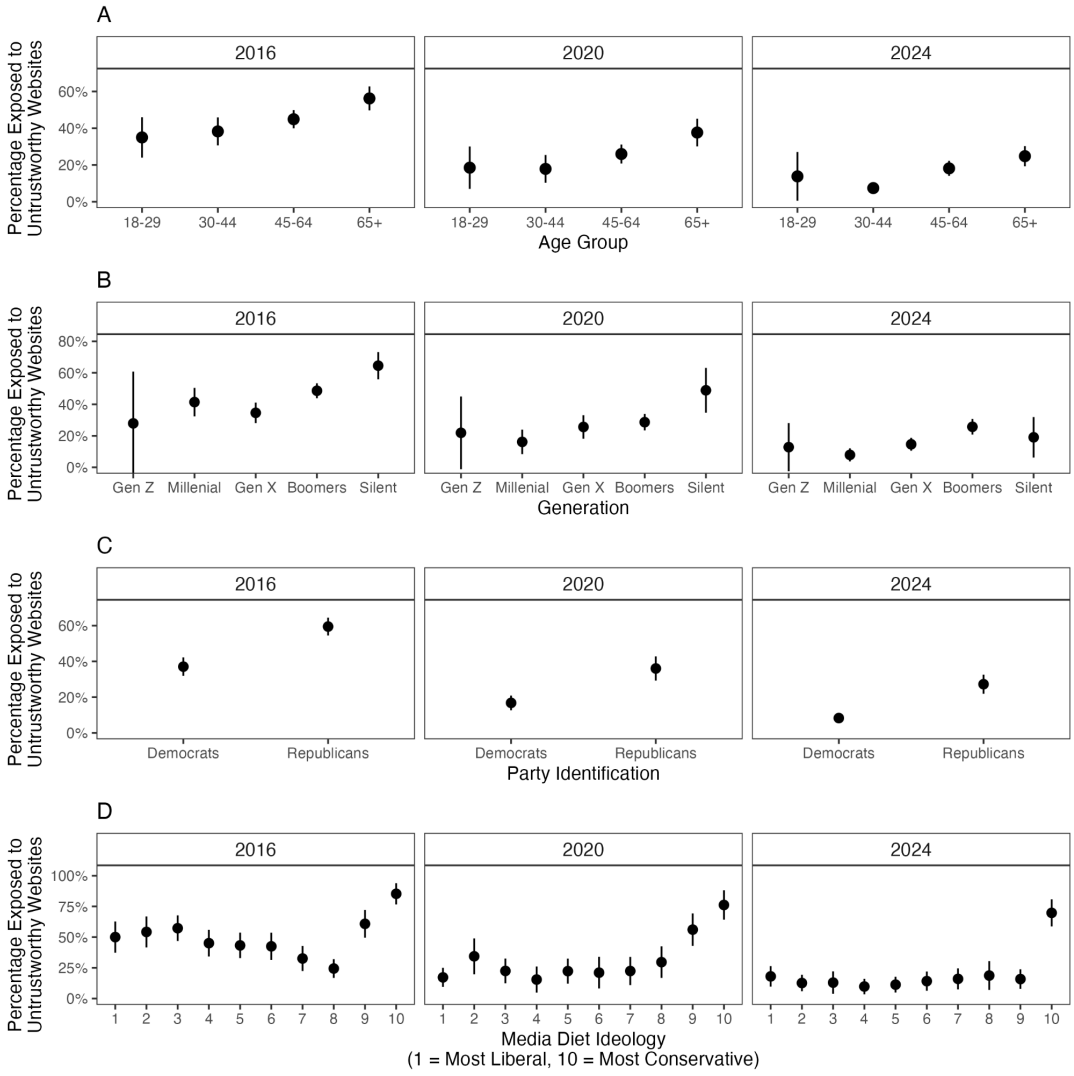
where a small subset of users drove very high average visit counts and a comparatively larger share of their total web activity consisting of such sites, though even for this group untrustworthy websites accounted for less than 1% of web visits.

In terms of partisanship, we observed a continued partisan divide in exposure (Figure 2C). Republicans (27.2%, 95% CI [21.9%, 32.6%]) continued to be more likely than Democrats (8.3%, 95% CI [5.6%, 11.0%]) to visit untrustworthy websites (H5, $t = 7.56$, $p < 0.001$, Cohen's $d = 0.51$, one-tailed). This divide is even more pronounced when examining candidate support, with Trump supporters showing substantially higher exposure rates (29.0%, 95% CI [23.4%, 34.6%]) compared to Harris supporters (8.9%, 95% CI [6.2%, 11.6%) ($t = 7.77$, $p < 0.001$, Cohen's $d = 0.25$). Moreover, exposure appears increasingly concentrated among self-reported ideological conservatives in 2024, with 33.1% (95% CI [24.3%, 41.8%]) of Very Conservative individuals and 29.0% (95% CI [20.5%, 37.5%]) of Conservatives ($t = 0.81$, $p = 0.42$, Cohen's $d = 0.09$) exposed. This concentration reflects a broader temporal pattern: between 2020 and 2024, exposure declined significantly across all liberal and moderate groups (Very Liberal: from 23.8% to 9.9%, $t = -3.33$, $p < 0.001$, Cohen's $d = -0.38$; Liberal: from 23.6% to 9.6%, $t = -4.06$, $p < 0.001$, Cohen's $d = -0.38$; Moderate: from 17.8% to 9.3%, $t = -3.36$, $p < 0.001$, Cohen's $d = -0.25$), while conservative groups showed smaller, non-significant declines (Conservative: from 35.9% to 29.0%, $t = -1.47$, $p = 0.142$, Cohen's $d = -0.15$; Very Conservative: from 42.6% to 33.1%, $t = -1.66$, $p = 0.098$, Cohen's $d = -0.20$), resulting in their persistently elevated exposure rates (see Supplementary Information S7).

These findings regarding self-reported ideology align with our finding that Americans in the most conservative media diet decile have the highest rate of exposure to untrustworthy websites in each of the past three presidential elections (see Figure 2D). Notably, however, the exposure rates among individuals in all other media diet slant deciles declined in 2024 compared to 2020 and 2016. One possible explanation for these results lies in changes in which untrustworthy sites people frequent (see Supplementary Information S8 for discussion of the

most visited untrustworthy websites). Additionally, political interest remained a significant positive predictor of exposure, just as it was in 2016 and 2020, with politically engaged individuals consistently showing higher exposure rates across all three election cycles. However, several demographic predictors that were significant in previous elections showed notable changes in 2024. College education, which was a significant predictor in both 2016 and 2020, was no longer significant in 2024. Similarly, while non-white individuals were less likely to be exposed in both 2016 and 2020, this coefficient was no longer significant in 2024 ($p = 0.084$).

Fig. 2: Untrustworthy website exposure by demographic and political characteristics.



Note: Untrustworthy website exposure during the 2016, 2020, and 2024 presidential elections. Panel A shows exposure rates by age group, with the y-axis representing the percentage of participants who visited at least one untrustworthy website. Panel B shows exposure rates by generation from youngest to oldest (Gen Z, Millennial, Gen X, Boomers, Silent; generational windows taken from[73]). Panel C shows exposure rates by party identification (Democrats vs. Republicans). Panel D shows exposure rates by media diet ideology decile (1 = Most Liberal, 10 = Most Conservative). Each panel displays three years (2016, 2020, 2024) side-by-side for comparison. Data are weighted means ± 95% CIs.

Table 1: Who chooses to visit general untrustworthy news websites

	2016	2020	2024
Republican	$\beta = 0.161$ s.e. 0.021 P<0.001 95% CI 0.119 to 0.203	$\beta = 0.128$ s.e. 0.027 P<0.001 95% CI 0.075 to 0.182	$\beta = 0.152$ s.e. 0.024 P<0.001 95% CI 0.104 to 0.199
Political interest	$\beta = 0.073$ s.e. 0.011 P<0.001 95% CI 0.051 to 0.095	$\beta = 0.059$ s.e. 0.014 P<0.001 95% CI 0.031 to 0.087	$\beta = 0.051$ s.e. 0.012 P<0.001 95% CI 0.027 to 0.074
College	$\beta = 0.055$ s.e. 0.021 P=0.010 95% CI 0.013 to 0.096	$\beta = 0.065$ s.e. 0.027 P=0.017 95% CI 0.012 to 0.119	$\beta = 0.017$ s.e. 0.024 P=0.459 95% CI -0.029 to 0.064
Female	$\beta = -0.037$ s.e. 0.019 P=0.052 95% CI -0.075 to 0.000	$\beta = -0.013$ s.e. 0.025 P=0.623 95% CI -0.062 to 0.037	$\beta = 0.007$ s.e. 0.023 P=0.762 95% CI -0.038 to 0.052
Non-white	$\beta = -0.096$ s.e. 0.021 P<0.001 95% CI -0.138 to -0.054	$\beta = -0.082$ s.e. 0.030 P=0.005 95% CI -0.140 to -0.024	$\beta = -0.045$ s.e. 0.026 P=0.084 95% CI -0.096 to 0.006
Age 30-44 years	$\beta = 0.036$ s.e. 0.032 P=0.270 95% CI -0.028 to 0.099	$\beta = -0.010$ s.e. 0.042 P=0.814 95% CI -0.092 to 0.073	$\beta = -0.053$ s.e. 0.038 P=0.157 95% CI -0.127 to 0.021
Age 45-64 years	$\beta = 0.068$ s.e. 0.029 P=0.021 95% CI 0.011 to 0.126	$\beta = 0.048$ s.e. 0.039 P=0.218 95% CI -0.029 to 0.125	$\beta = 0.019$ s.e. 0.036 P=0.595 95% CI -0.051 to 0.089
Age 65+ years	$\beta = 0.134$ s.e. 0.033 P<0.001 95% CI 0.069 to 0.200	$\beta = 0.138$ s.e. 0.043 P=0.001 95% CI 0.055 to 0.222	$\beta = 0.037$ s.e. 0.038 P=0.336 95% CI -0.038 to 0.112
Constant	$\beta = 0.123$ s.e. 0.046 P=0.007 95% CI 0.033 to 0.213	$\beta = -0.006$ s.e. 0.058 P=0.922 95% CI -0.120 to 0.109	$\beta = -0.053$ s.e. 0.052 P=0.309 95% CI -0.154 to 0.049
R^2	0.088	0.086	0.093
N	2514	1151	1050

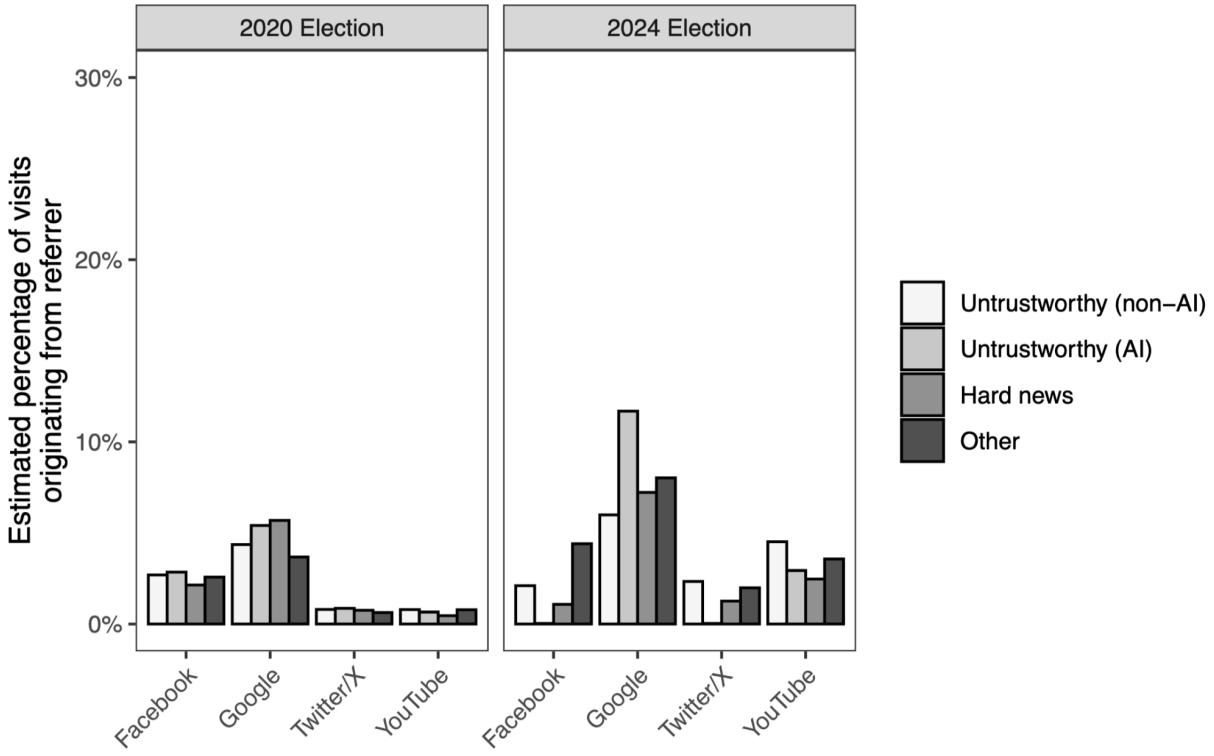
Note: Ordinary least squares regression coefficients are shown with standard errors (s.e.) in parentheses (models estimated using survey weights). The model on the left represents the 2016 election, the model in the middle represents the 2020 election, and the model on the right represents the 2024 election. The dependent variable in both models is a binary variable indicating whether an individual was exposed to at least one general untrustworthy website during the data collection period (1) or not (0). 'Republican' variable: identifying as Republican or Republican-leaning: 1; identifying as Democrat, Independent, or other: 0. 'Political interest' variable: variable ranging from 1 to 4, where 4 represents people who say they pay attention to what is going on in government and politics 'most of the time' and 1 represents those who pay attention 'hardly at all'. 'College' variable: 1: college graduate; 0: not a college graduate. 'Female' variable: 1: indicates identifying as female; 0: did not indicate identifying as female. 'Non-white' variable: 1: indicates identifying as a race other than white; 0: indicates identifying as white. Age variables represent indicator variables for each age group, with ages 18-29 years

serving as the reference category. P values are two-sided and not corrected for multiple comparisons.

How were people exposed to untrustworthy websites in 2024?

Following the approach in previous studies, we examined the “referrers” that directed individuals to untrustworthy websites (i.e., the sites that were within three visits prior and within 30 seconds of the visit; see Figure 3). Aligned with our hypothesis that Twitter/X would expand its role as a gateway, referrals to untrustworthy websites from Twitter/X increased between 2020 (0.8%) and 2024 (2.3%) (H6, $\chi^2 = 94.8$, $p < 0.001$, Cohen’s $h = 0.125$, one-tailed). In exploratory analyses, Facebook’s share dropped from 2.7% to 2.1% ($\chi^2 = 7.2$, $p = 0.007$, Cohen’s $h = -0.039$, two-tailed). In comparison, Google’s referrals rose from 4.4% to 6.0% ($\chi^2 = 29.5$, $p < 0.001$, Cohen’s $h = 0.072$, two-tailed), and YouTube’s referrals rose from 0.8% to 4.5% ($\chi^2 = 367.8$, $p < 0.001$, Cohen’s $h = 0.248$, two-tailed). Collectively, these results suggest that no single platform dominated as an entry point to untrustworthy websites in 2024, but several platforms exhibited notable shifts in their referral shares compared to the previous election.

Fig. 3: Referrers to untrustworthy new websites and other sources



Note: Percentage of visits to different website categories referred by major platforms during the 2020 and 2024 elections. Left panel: 2020 election. Right panel: 2024 election. The y-axis represents the percentage of visits to each website type preceded by visits to each platform (Facebook, Google, Twitter/X, YouTube). Bar colors distinguish website categories: untrustworthy non-AI (white), untrustworthy AI-generated (light gray), hard news (medium gray), and other websites (dark gray).

AI-generated untrustworthy websites in 2024

A new component of this research on untrustworthy website exposure is the identification of *AI-generated* untrustworthy websites. We find that 6.5% (95% CI [4.6%, 8.5%]) of Americans visited at least one such site in 2024, significantly up from 2.6% (95% CI [1.3%, 4.0%]) in 2020 (H7, $t = 4.37$, $p < 0.001$, Cohen's $d = 0.187$, one-tailed). Although this indicates that more Americans are encountering AI-generated untrustworthy sources, these AI-generated sites did not account for a larger proportion of overall untrustworthy website visits in 2024 compared to in 2020 (2024: 1.6%,95% CI [1.3%, 1.8%]; 2020: 2.1%, 95% CI [1.8%, 2.5%]; H8,

χ^2 = 6.8, p = 0.995, Cohen's h = -0.043, one-tailed). Thus, while the absolute number of users visiting AI-generated content is growing, it has not supplanted other untrustworthy sources.

We hypothesized that visits to AI-generated untrustworthy websites would last longer than visits to non-AI untrustworthy websites, owing to AI's potential to make untrustworthy content more engaging[2,3]. However, we did not find support for this hypothesis: The average visit duration for AI-generated untrustworthy sites was not significantly different from the average visit duration for non-AI untrustworthy sites, 44 seconds (95% CI: [30, 59]) for AI, compared to 53 seconds (95% CI: [49, 58]) for non-AI (H9, t = -1.39, p = 0.917, Cohen's d = -0.054, one-sided).

Moreover, demographic factors associated with exposure to AI-generated untrustworthy websites in 2024 diverged from those associated with exposure to non-AI untrustworthy websites (see Table 2). Given higher exposure to non-AI untrustworthy websites, we hypothesized that older adults would be more likely than younger adults to encounter AI-generated untrustworthy content (H10). However, the data did not support this hypothesis: 5.8% of adults aged 65+ (95% CI [3.1%, 8.5%]) were exposed to AI-generated untrustworthy websites, compared to 10.0% (95% CI [–0.2%, 20.2%]) of adults aged 18–29 (H10, t = –1.19, p = 0.882, Cohen's d = –0.16, one-tailed). Similarly, we hypothesized that Republicans would be more likely than Democrats to encounter AI-generated untrustworthy content (H11). While the proportion of Republicans exposed was higher (7.4%, 95% CI [3.9%, 11.0%]) than Democrats (5.0%, 95% CI [3.0%, 7.0%]), the difference was not statistically significant (H11, t = 1.51, p = 0.066, Cohen's d = 0.10, one-tailed). In examining sociodemographic and political variables in a regression, we find that those with lower levels of political knowledge and identifying as white were more likely to visit AI-generated untrustworthy websites. Those aged 30-44 were also less likely to visit such websites than those aged 18-29. Taken together, these findings suggest that although AI-generated untrustworthy websites reached a larger audience in 2024 than in 2020, they still represent a relatively small share of total untrustworthy website exposure and do not

feature the pronounced age and partisan disparities in exposure observed for untrustworthy content more broadly.

Table 2: Who chooses to visit AI-generated untrustworthy news websites

	2020	2024
Republican	$\beta = -0.018$	$\beta = -0.002$
	s.e. 0.010	s.e. 0.017
	P=0.076	P=0.920
	95% CI -0.039 to 0.002	95% CI -0.034 to 0.031
Political interest	$\beta = -0.001$	$\beta = -0.019$
	s.e. 0.005	s.e. 0.008
	P=0.898	P=0.021
	95% CI -0.011 to 0.010	95% CI -0.036 to -0.003
College	$\beta = 0.009$	$\beta = 0.004$
	s.e. 0.010	s.e. 0.016
	P=0.409	P=0.806
	95% CI -0.012 to 0.029	95% CI -0.028 to 0.036
Female	$\beta = 0.018$	$\beta = -0.010$
	s.e. 0.010	s.e. 0.016
	P=0.058	P=0.540
	95% CI -0.001 to 0.037	95% CI -0.041 to 0.021
Non-white	$\beta = -0.026$	$\beta = -0.067$
	s.e. 0.011	s.e. 0.018
	P=0.020	P<0.001
	95% CI -0.048 to -0.004	95% CI -0.103 to -0.032
Age 30-44 years	$\beta = 0.037$	$\beta = -0.062$
	s.e. 0.016	s.e. 0.026
	P=0.020	P=0.017
	95% CI 0.006 to 0.068	95% CI -0.113 to -0.011
Age 45-64 years	$\beta = 0.026$	$\beta = -0.032$
	s.e. 0.015	s.e. 0.025
	P=0.081	P=0.187
	95% CI -0.003 to 0.055	95% CI -0.081 to 0.016
Age 65+ years	$\beta = 0.031$	$\beta = -0.048$
	s.e. 0.016	s.e. 0.027
	P=0.057	P=0.068
	95% CI -0.001 to 0.062	95% CI -0.101 to 0.004
Constant	$\beta = 0.005$	$\beta = 0.189$
	s.e. 0.022	s.e. 0.036
	P=0.813	P<0.001
	95% CI -0.038 to 0.048	95% CI 0.119 to 0.259
R^2	0.015	0.022
N	1151	1050

Note: Ordinary least squares regression coefficients are shown with standard errors (s.e.) in parentheses (models estimated using survey weights). The model on the left represents the 2020 election, and the model on the right represents the 2024 election. The dependent variable in both models is a binary variable indicating whether an individual was exposed to at least one AI-generated untrustworthy website during the data collection period (1) or not (0). 'Republican' variable: identifying as Republican or Republican-leaning: 1; identifying as Democrat, Independent, or other: 0. 'Political interest' variable: variable ranging from 1 to 4, where 4 represents people who say they pay attention to what is going on in government and politics 'most of the time' and 1 represents those who pay attention 'hardly at all'. 'College' variable: 1: college graduate; 0: not a college graduate. 'Female' variable: 1: indicates identifying as female; 0: did not indicate identifying as female. 'Non-white' variable: 1: indicates identifying as a race

other than white; 0: indicates identifying as white. Age variables represent indicator variables for each age group, with ages 18-29 years serving as the reference category. P-values are two-sided and not corrected for multiple comparisons.

Finally, we conducted extensive exploratory analyses to explore the relationship between other constructs and untrustworthy website exposure, including perceptions of untrustworthy website exposure, digital and AI literacy, AI knowledge, usage, and attitudes, the ability to discern AI-generated images, self-rated ability to spot AI-generated content, perceptions of AI's electoral impact, AI usage purposes, and temporal patterns of exposure before and after the election (see Supplementary Information S9-S19). We highlight two noteworthy findings. First, both Democrats and Republicans accurately estimated their untrustworthy website exposure, whereas nonpartisans did not – that is, there was a significant association between partisans' self-reported perceptions of the percentage of their media diet comprised of untrustworthy websites and the actual percentage based on their browsing data. This finding could suggest that partisans are, at least to some extent, aware that they are consuming untrustworthy content, while nonpartisans may have more undetected encounters with such content.

Second, greater digital literacy was associated with greater exposure to untrustworthy websites. While potentially counterintuitive, this finding highlights important nuances in the relationship between digital literacy and exposure to untrustworthy content. Greater digital literacy may actually increase exposure to untrustworthy sites as users navigate a broader variety of sources and platforms, and individuals with higher digital literacy may feel more willing to visit untrustworthy sources because of confidence in their ability to be discerning during those encounters. This finding reiterates that researchers should not interpret exposures to untrustworthy content in digital trace data as representing belief in that content[74].

Discussion

We find that the proportion of Americans visiting untrustworthy websites declined in 2024 compared to 2020, continuing a trend between 2020 and 2016. This drop persists even though

our classification of untrustworthy websites includes a larger and more up-to-date database of domains than prior efforts. Despite the decline in the number of Americans exposed, people who visited untrustworthy sites visited them more frequently in 2024 than in 2020. In other words, untrustworthy content exposure on the web is becoming increasingly concentrated among a smaller group of users, a deviation from past trends where fewer individuals were exposed *and* exposure among those who were exposed also decreased[34].

Certain groups that were more likely to visit untrustworthy websites in previous elections, such as older adults and conservatives, continued to do so in 2024. However, their rates of exposure were generally lower than in earlier cycles. Furthermore, we identified nuances in age differences in exposure that emerged with the 2024 election, suggesting that other sociodemographic factors, such as partisanship and political interest, may increasingly influence exposure to untrustworthy websites beyond the influence of age. Future research should continue to unpack age differences in engagement with untrustworthy content to better understand drivers of exposure and identify pathways for effective intervention (for an example of such work, see[75]). In general, however, the broad persistence of demographic trends in exposure underscores the need for interventions and resources specifically targeting populations more likely to encounter untrustworthy content, especially because these populations may be at heightened risks from the effects of exposure[76].

We hypothesized that social media platforms, particularly Twitter/X, might serve as increasingly important gateways to untrustworthy websites as the platform made sizable cuts to trust and safety infrastructure[65,66]. Our data indicate that Twitter/X “referred” a higher percentage of visits to untrustworthy websites in 2024 than in 2020. Findings about platforms may also be informative for designing and delivering interventions that meet users on platforms where they are more likely to be referred to untrustworthy content. Importantly, however, future work should aim to clarify why referrals from certain platforms (such as Facebook) have declined, whether

due to platform policies, changes in user behavior, or broader shifts in how people consume news.

The role of AI-generated untrustworthy websites expanded modestly in terms of the share of Americans exposed (up from 2.6% in 2020 to 6.5% in 2024), but remains a small fraction of total untrustworthy visits. In addition, our data do not support the notion that AI-generated untrustworthy websites are more “engaging” in terms of visit duration. Our findings corroborate other reports that AI did not dominate the media landscape for the 2024 US election[77–79], but with comprehensive data from the actual online behaviors of a representative sample of American adults during the election. From a societal perspective, however, it is important to note that even a small fraction of AI-generated untrustworthy content can significantly affect public discourse and beliefs, particularly if such content is highly persuasive or targeted at vulnerable groups.

Although the percentage of Americans exposed to untrustworthy websites in 2024 decreased relative to 2020, one should not interpret these findings as meaning that untrustworthy content is no longer a significant social challenge. The absolute number of exposed users remains considerable, and even limited exposures can fuel polarization[6], public health crises[7–9], and distrust in democratic processes[10,11]. Moreover, the fact that more people now see AI-generated content—despite it representing a small share of overall untrustworthy traffic—suggests that AI-based disinformation may be on a growth trajectory, with new tactics possibly emerging in the years ahead as the capabilities of generative AI technologies continue to advance.

Our study offers three main contributions. First, we reinforce the importance of longitudinal comparisons. By examining 2024 exposure using the same data-collection methods and definitions of “untrustworthy websites” employed in 2020 and 2016, we can directly assess how exposure patterns have evolved. The consistency of this methodology enables apples-to-apples comparisons across multiple election cycles. Second, our findings highlight the

persistence of demographic factors, such as age and partisanship, which help focus future attention and resources on higher-risk populations, as well as provide insight into the roles different popular online platforms play in exposing untrustworthy content. Third, we provide a comprehensive assessment of AI-generated untrustworthy website exposure to date in a national sample of American adults.

Why did untrustworthy website exposure continue to decline from 2020 to 2024, even amidst new generative AI capabilities? This decline is particularly intriguing given widespread public expectation of the opposite trend (see Supplementary Information S21). A recent national survey found that most Americans (56%) incorrectly believed that exposure to untrustworthy websites increased from 2020 to 2024, with fewer than one in ten (8%) accurately perceiving a decline in exposure. Several speculative explanations may account for the decline observed in our data. One possibility is that platforms have become more proactive in reducing visibility of content from known untrustworthy domains, thereby diminishing visits via social media referrals. Another explanation is that audiences increasingly consume content through platform feeds and private apps, which do not generate external link visits captured by browser logs. Generative AI might also produce more ephemeral or localized misinformation that circulates within private channels or ephemeral stories, further reducing external domain visits.

Despite falling overall exposure rates, the possibility that AI could eventually yield new forms of more targeted, persuasive, untrustworthy websites underscores the need for policymaking and thoughtful technology design. Policymakers might consider implementing or encouraging guardrails around generative AI tools, such as watermarking or traceability that flags AI-generated text. Media literacy interventions could focus specifically on teaching users to distinguish between synthetic and authentic content. At the same time, our exploratory findings highlight a “digital literacy paradox”: individuals with stronger digital literacy skills are, in some cases, more exposed to untrustworthy websites, likely due to broader online engagement. This paradox underscores the need for interventions that address not only users’ ability to identify

false information but also the overconfidence or exploratory behaviors that may inadvertently increase their exposure. Finally, the ongoing concentration of untrustworthy site consumption suggests that interventions targeting specific high-risk communities may be more efficient than broad, population-level campaigns. Ultimately, it remains important not to overstate the prevalence of AI-generated untrustworthy content at present, as it still accounts for a relatively small proportion of overall untrustworthy exposure. Overemphasizing the scale of AI-generated untrustworthy content may inadvertently fuel generalized distrust in legitimate information sources[80–82].

Despite the strengths of our approach, several limitations exist that future work should address. First, like many other studies, we measure untrustworthiness at the domain level. Although previous research validates this method of classifying entire sites as untrustworthy, some domains might host both untrustworthy and neutral or even reliable articles[9,83]. Similarly, some AI-generated misinformation could appear on otherwise trustworthy websites but would not be captured by our domain-level approach. Second, browser-based logging data do not capture platform-native app usage (for example, within mobile apps, feed scrolling on social media, or messaging platforms). Our data cannot reflect whether users encounter AI-generated or other untrustworthy content without clicking a link to an external site. Given the rise of news consumption via apps, private groups, and encrypted messaging, future research should integrate alternative data sources (for example, screen captures or direct platform data) to obtain a broader view of consumption. Third, although we rely on newly updated classifications of AI-generated websites, new or ephemeral sites may evade detection. Generative AI could potentially create sites that vanish before rating organizations, such as NewsGuard, can catalog them[84]. Fourth, our data cannot speak to how untrustworthy exposure affects attitudes or behavior, such as voting, trust in institutions, or political participation. Addressing these outcomes will require additional experimental or longitudinal designs linking exposure to changes in beliefs or behaviors.

Methods

This research was approved by the Stanford University Institutional Review Board (protocol no. IRB-53941). All participants provided informed consent prior to participating and received incentives from YouGov, the survey company that administered both the survey and web-tracking software. Data processing and statistical analyses were conducted using R (v4.4.2).

Following the analytic approach originally established for the study of the 2016 U.S. election[35] and further developed in the 2020 study[34], we examined participants' web-browsing logs collected using YouGov's Pulse software. Similar to those studies, we used a domain-level method to classify exposures to untrustworthy websites. This approach allows for direct comparisons with exposure patterns documented in both 2016 and 2020. By adopting the same core design, i.e., collecting real-time web traffic data over the four weeks before and one week after election day, we assess how untrustworthy content exposure shifted over these three consecutive elections.

YouGov recruited 1,069 participants for the 2024 study. Participants were enrolled in a two-wave online survey and YouGov Pulse software installed on their smartphones, computers, and/or tablets allowed us to passively collect web-browsing data between October 8 and November 12, 2024 (election day was November 5, 2024). The final sample consisted of 46% males and 54% females, with an age distribution of 8.1% under 30, 28.4% aged 30–44, 38.1% aged 45–64, and 25.4% aged 65 or older. In terms of political preferences, 43.0% reported supporting Kamala Harris, and 38.9% reported supporting Donald Trump. Sample demographic characteristics were comparable across the 2016, 2020, and 2024 studies, ensuring comparability between election years (see Supplementary Information Section S20). YouGov weighted participants to ensure that the final dataset aligns with key demographic benchmarks (such as age, gender, and race/ethnicity) of the U.S. population. All analyses presented here utilize these survey weights to approximate nationally representative estimates. Consistent with

the 2020 study methodology[34]—but in contrast to the 2016 study that focused solely on desktop data[35]—we included both desktop and mobile web-browsing logs in the 2024 analysis. As news consumption continues to shift toward smartphones, it is critical to capture cross-device activity to more accurately represent overall exposure[85]. We identified untrustworthy websites by following the general approach of Moore et al.[34] of combining multiple sources to create as comprehensive a list of untrustworthy websites as possible. First, we began with 490 domains from Grinberg et al.[38] (382 of which were "identified by journalists and fact checkers as notable publishers of false or misleading content", 61 of which were identified via human review as having "little regard for the truth", and 47 of which were identified as being "negligent or deceptive") and 66 domains from Allcott & Gentzkow[86] "frequently publishing dubious information" (the two lists comprising the original list used by Guess et al.[35]). Next, the 2020 study[34] supplemented this base by adding 1,240 domains from NewsGuard that were rated as "repeatedly publishing false content," resulting in a total of 1,796 general untrustworthy domains. For the 2024 study, we updated the NewsGuard data to include an additional 777 domains rated since 2020, yielding a final list of 2,573 general untrustworthy domains..

To capture emerging patterns of AI-driven untrustworthy sources, we incorporated an additional 810 domains from NewsGuard's UAINS (Unreliable AI-generated News Sites) database. According to NewsGuard, UAINS are sites that meet the following four criteria: (1) there is clear evidence that a substantial portion of the site's content is produced by AI; (2) there is strong evidence that the content is being published without significant human oversight—for example, numerous articles might contain error messages or language specific to chatbot responses, indicating that the content was produced by AI tools without adequate editing (news sites that deploy effective human oversight are not classified as UAINS); (3) the site is presented in a way that an average reader could assume that its content is produced by human writers or journalists; and (4) the site does not clearly disclose that its content is produced by AI.

We matched our updated lists against each participant's URL-level browsing logs so that any visit to a domain on these lists was classified as an exposure to untrustworthy content, distinguishing between non-AI and AI-generated sites. In parallel, we constructed a comprehensive list of credible "hard news" sites (including the Bakshy et al.[87] top 500 hard news sites and all domains that NewsGuard rated as not repeatedly publishing false content and had a "quality" score of 60+ out of 100[88]) to capture participants' exposure to mainstream or reliable sources.

We maintained a domain-level strategy to classify both non-AI and AI-generated untrustworthy sites for continuity with our earlier analyses in 2016 and 2020. This approach, validated in previous research, facilitates straightforward comparisons of exposure over time. Nevertheless, it cannot capture article-specific variance within domains (for example, a large site that occasionally posts misleading stories[9]). AI-generated misinformation, in particular, may appear on subdomains or ephemeral sites that elude rapid detection. To mitigate these concerns, we relied on ongoing lists compiled by NewsGuard throughout the 2024 campaign period. We note that the 2020 general untrustworthy statistics were calculated using the 2020 version of the domain list, while the 2024 general untrustworthy statistics were calculated using the 2024 updated version of the list.

We replicated the same data-collection window as the 2016 and 2020 studies, four weeks before and one week after election day. Similar to the 2016 and 2020 studies[34,35], our data distributions were assumed to be normal, although this assumption was not formally tested. Anonymized survey data along with summary web traffic data used for the analyses in the paper are available at: https://github.com/rossdahlke/2024-exposure-public. Full web traffic histories are not available in order to protect subject confidentiality.

References

1. Vraga, E. K. & Bode, L. Defining Misinformation and Understanding its Bounded Nature: Using Expertise and Evidence for Describing Misinformation. *Polit. Commun.* **37**, 136–144 (2020).
2. Simon, F. M., Altay, S. & Mercier, H. Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harv. Kennedy Sch. Misinformation Rev.* https://doi.org/10.37016/mr-2020-127 (2023) doi:10.37016/mr-2020-127.
3. Spitale, G., Biller-Andorno, N. & Germani, F. AI model GPT-3 (dis)informs us better than humans. *Sci. Adv.* **9**, eadh1850 (2023).
4. Dan, V. Deepfakes as a Democratic Threat: Experimental Evidence Shows Noxious Effects That Are Reducible Through Journalistic Fact Checks. *Int. J. Press.* 19401612251317766 (2025) doi:10.1177/19401612251317766.
5. Goldstein, J. A., Chao, J., Grossman, S., Stamos, A. & Tomz, M. How persuasive is AI-generated propaganda? *PNAS Nexus* **3**, pgae034 (2024).
6. Kaiser, J., Vaccari, C. & Chadwick, A. Partisan Blocking: Biased Responses to Shared Misinformation Contribute to Network Polarization on Social Media. *J. Commun.* **72**, 214–240 (2022).
7. Vinck, P., Pham, P. N., Bindu, K. K., Bedford, J. & Nilles, E. J. Institutional trust and misinformation in the response to the 2018–19 Ebola outbreak in North Kivu, DR Congo: a population-based survey. *Lancet Infect. Dis.* **19**, 529–536 (2019).
8. Kricorian, K., Civen, R. & Equils, O. COVID-19 vaccine hesitancy: misinformation and perceptions of vaccine safety. *Hum. Vaccines Immunother.* **18**, 1950504 (2022).
9. Allen, J., Watts, D. J. & Rand, D. G. Quantifying the impact of misinformation and vaccine-skeptical content on Facebook. *Science* **384**, eadk3451 (2024).
10. Ognyanova, K., Lazer, D., Robertson, R. E. & Wilson, C. Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side

is in power. *Harv. Kennedy Sch. Misinformation Rev.* https://doi.org/10.37016/mr-2020-024 (2020) doi:10.37016/mr-2020-024.

11. Dahlke, R. & Hancock, J. The Effect of Online Misinformation Exposure on False Election Beliefs. Preprint at https://doi.org/10.31219/osf.io/325tn (2022).
12. World Economic Forum. Global Risks 2024: Disinformation Tops Global Risks 2024 as Environmental Threats Intensify. *World Economic Forum* https://www.weforum.org/press/2024/01/global-risks-report-2024-press-release/ (2024).
13. Adams, Z., Osman, M., Bechlivanidis, C. & Meder, B. (Why) Is Misinformation a Problem? *Perspect. Psychol. Sci.* **18**, 1436–1463 (2023).
14. Hwang, Y. & Jeong, S.-H. Misinformation Exposure and Acceptance: The Role of Information Seeking and Processing. *Health Commun.* **38**, 585–593 (2023).
15. Jones-Jang, S. M., Kim, D. H. & Kenski, K. Perceptions of mis- or disinformation exposure predict political cynicism: Evidence from a two-wave survey during the 2018 US midterm elections. *New Media Soc.* **23**, 3105–3125 (2021).
16. Kim, H. K. & Tandoc, E. C. Consequences of Online Misinformation on COVID-19: Two Potential Pathways and Disparity by eHealth Literacy. *Front. Psychol.* **13**, (2022).
17. Lee, J. J. *et al.* Associations Between COVID-19 Misinformation Exposure and Belief With COVID-19 Knowledge and Preventive Behaviors: Cross-Sectional Online Study. *J. Med. Internet Res.* **22**, e22205 (2020).
18. Lee, T., Johnson, T., Jia, C. & Lacasa-Mas, I. How social media users become misinformed: The roles of news-finds-me perception and misinformation exposure in COVID-19 misperception. *New Media Soc.* 14614448231202480 (2023) doi:10.1177/14614448231202480.
19. Luk, T. T. *et al.* Exposure to health misinformation about COVID-19 and increased tobacco and alcohol use: a population-based survey in Hong Kong. *Tob. Control* **30**, 696–699 (2021).

20. Neely, S. R., Eldredge, C., Ersing, R. & Remington, C. Vaccine Hesitancy and Exposure to Misinformation: a Survey Analysis. *J. Gen. Intern. Med.* **37**, 179–187 (2022).
21. Vegetti, F. & Mancosu, M. Perceived Exposure and Concern for Misinformation in Different Political Contexts: Evidence From 27 European Countries. *Am. Behav. Sci.* 00027642221118255 (2022) doi:10.1177/00027642221118255.
22. Warner, E. L., Waters, A. R., Cloyes, K. G., Ellington, L. & Kirchhoff, A. C. Young adult cancer caregivers' exposure to cancer misinformation on social media. *Cancer* **127**, 1318–1324 (2021).
23. Xiao, X. Not doomed: Examining the path from misinformation exposure to verification and correction in the context of COVID-19 pandemic. *Telemat. Inform.* **74**, 101890 (2022).
24. Valenzuela, S., Halpern, D. & Araneda, F. A Downward Spiral? A Panel Study of Misinformation and Media Trust in Chile. *Int. J. Press.* **27**, 353–373 (2022).
25. Revilla, M., Ochoa, C. & Loewe, G. Using Passive Data From a Meter to Complement Survey Data in Order to Study Online Behavior. *Soc. Sci. Comput. Rev.* **35**, 521–536 (2017).
26. Ernala, S. K., Burke, M., Leavitt, A. & Ellison, N. B. How Well Do People Report Time Spent on Facebook? An Evaluation of Established Survey Questions with Recommendations. in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* 1–14 (Association for Computing Machinery, New York, NY, USA, 2020). doi:10.1145/3313831.3376435.
27. Naab, T. K., Karnowski, V. & Schlütz, D. Reporting Mobile Social Media Use: How Survey and Experience Sampling Measures Differ. *Commun. Methods Meas.* **13**, 126–147 (2019).
28. Parry, D. A. *et al.* A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nat. Hum. Behav.* **5**, 1535–1547 (2021).
29. Vraga, E., Bode, L. & Troller-Renfree, S. Beyond Self-Reports: Using Eye Tracking to Measure Topic and Style Differences in Attention to Social Media Content. *Commun. Methods Meas.* **10**, 149–164 (2016).

30. Guess, A., Munger, K., Nagler, J. & Tucker, J. How Accurate Are Survey Responses on Social Media and Politics? *Polit. Commun.* **36**, 241–258 (2019).
31. Haenschen, K. Self-Reported Versus Digitally Recorded: Measuring Political Activity on Facebook. *Soc. Sci. Comput. Rev.* **38**, 567–583 (2020).
32. Pfänder, J. & Altay, S. Spotting false news and doubting true news: a systematic review and meta-analysis of news judgements. *Nat. Hum. Behav.* 1–12 (2025) doi:10.1038/s41562-024-02086-1.
33. Guess, A., Lyons, B., Montgomery, J. M., Nyhan, B. & Reifler, J. *Fake News, Facebook Ads, and Misperceptions: Assessing Information Quality in the 2018 U.S. Midterm Election Campaign*. 36 (2019).
34. Moore, R. C., Dahlke, R. & Hancock, J. T. Exposure to untrustworthy websites in the 2020 US election. *Nat. Hum. Behav.* **7**, 1096–1105 (2023).
35. Guess, A. M., Nyhan, B. & Reifler, J. Exposure to untrustworthy websites in the 2016 US election. *Nat. Hum. Behav.* **4**, 472–480 (2020).
36. Allcott, H., Gentzkow, M. & Yu, C. Trends in the diffusion of misinformation on social media. *Res. Polit.* **6**, 2053168019848554 (2019).
37. Zhou, A., Yang, T. & González-Bailón, S. The puzzle of misinformation: Exposure to unreliable content in the United States is higher among the better informed. *New Media Soc.* 14614448231196863 (2023) doi:10.1177/14614448231196863.
38. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. & Lazer, D. Fake news on Twitter during the 2016 U.S. presidential election. *Science* **363**, 374–378 (2019).
39. Allen, J., Howland, B., Mobius, M., Rothschild, D. & Watts, D. J. Evaluating the fake news problem at the scale of the information ecosystem. *Sci. Adv.* **6**, eaay3539 (2020).
40. Scott, L. World Faces 'Tech-Enabled Armageddon,' Maria Ressa Says. *Voice of America* https://www.voanews.com/a/world-faces-tech-enabled-armageddon-maria-ressa-says-/7256196.html (2023).

41. Metz, C. ‘The Godfather of A.I.’ Leaves Google and Warns of Danger Ahead. *The New York Times* (2023).
42. Gold, A. & Fischer, S. Chatbots trigger next misinformation nightmare. *Axios* https://www.axios.com/2023/02/21/chatbots-misinformation-nightmare-chatgpt-ai (2023).
43. Hsu, T. & Thompson, S. A. Disinformation Researchers Raise Alarms About A.I. Chatbots. *The New York Times* (2023).
44. Bond, S. As tech evolves, deepfakes will become even harder to spot. *NPR* (2022).
45. Benson, T. This Disinformation Is Just for You. *Wired* (2023).
46. Sanderson, Z., Messing, S. & Tucker, J. A. Misunderstood mechanics: How AI, TikTok, and the liar’s dividend might affect the 2024 elections. *Brookings* https://www.brookings.edu/articles/misunderstood-mechanics-how-ai-tiktok-and-the-liars-dividend-might-affect-the-2024-elections/ (2024).
47. Kapoor, S. & Narayanan, A. How to Prepare for the Deluge of Generative AI on Social Media. *Knight First Amendment Institute at Columbia University* http://knightcolumbia.org/content/how-to-prepare-for-the-deluge-of-generative-ai-on-social-media (2023).
48. Baig, E. C. Too scary? Elon Musk’s OpenAI company won’t release tech that can generate fake news. *USA TODAY* https://www.usatoday.com/story/tech/2019/02/15/elon-musks-openai-fake-news-generator-too-scary-release/2880790002/ (2019).
49. Morning Consult & Axios. *National Tracking Poll 2308055*. https://pro-assets.morningconsult.com/wp-uploads/2023/09/2308055_topline_AXIOS_AI_Adults_v1_EP-1.pdf (2023).
50. Yan, H. Y., Morrow, G., Yang, K.-C. & Wihbey, J. The origin of public concerns over AI supercharging misinformation in the 2024 U.S. presidential election. *Harv. Kennedy Sch. Misinformation Rev.* https://doi.org/10.37016/mr-2020-171 (2025)

doi:10.37016/mr-2020-171.

51. Wack, M., Ehrett, C., Linvill, D. & Warren, P. Generative propaganda: Evidence of AI's impact from a state-backed disinformation campaign. *PNAS Nexus* **4**, pgaf083 (2025).
52. Hanley, H. W. A. & Durumeric, Z. Machine-Made Media: Monitoring the Mobilization of Machine-Generated Articles on Misinformation and Mainstream News Websites. Preprint at https://doi.org/10.48550/arXiv.2305.09820 (2024).
53. Kreps, S., McCain, R. M. & Brundage, M. All the News That's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation. *J. Exp. Polit. Sci.* **9**, 104–117 (2022).
54. Jakesch, M., Hancock, J. T. & Naaman, M. Human heuristics for AI-generated language are flawed. *Proc. Natl. Acad. Sci.* **120**, e2208839120 (2023).
55. Ramu, D., Jain, R. & Jain, A. Generation Z's Ability to Discriminate Between AI-generated and Human-Authored Text on Discord. Preprint at https://doi.org/10.48550/arXiv.2401.04120 (2023).
56. Lu, Z. *et al.* Seeing is not always believing: Benchmarking Human and Model Perception of AI-Generated Images. *Adv. Neural Inf. Process. Syst.* **36**, 25435–25447 (2023).
57. Partadiredja, R. A., Serrano, C. E. & Ljubenkov, D. AI or Human: The Socio-ethical Implications of AI-Generated Media Content. in *2020 13th CMI Conference on Cybersecurity and Privacy (CMI) - Digital Transformation - Potentials and Challenges(51275)* 1–6 (2020). doi:10.1109/CMI51275.2020.9322673.
58. Ha, A. Y. J. *et al.* Organic or Diffused: Can We Distinguish Human Art from AI-generated Images? Preprint at https://doi.org/10.48550/arXiv.2402.03214 (2024).
59. Bray, S. D., Johnson, S. D. & Kleinberg, B. Testing human ability to detect 'deepfake' images of human faces. *J. Cybersecurity* **9**, tyad011 (2023).
60. Yan, H. Y., Moore, R., Tu, F. & Hancock, J. Detecting Synthetic, Doubting Authentic: AI Attribution Bias for Political Imagery. Preprint at https://doi.org/10.31219/osf.io/w6bzx_v1 (2025).

61. Sundar, S. S., Molina, M. D. & Cho, E. Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps? *J. Comput.-Mediat. Commun.* **26**, 301–319 (2021).
62. Jin, X. *et al.* Assessing the perceived credibility of deepfakes: The impact of system-generated cues and video characteristics. *New Media Soc.* 14614448231199664 (2023) doi:10.1177/14614448231199664.
63. Bashardoust, A., Feuerriegel, S. & Shrestha, Y. R. Comparing the willingness to share for human-generated vs. AI-generated fake news. Preprint at https://doi.org/10.48550/arXiv.2402.07395 (2024).
64. Jussim, L., Mackey, J. L. & Eppard, L. M. Chapter 12: How NewsGuard Works. in *The Poisoning of the American Mind* 104–115 (University of Virginia Press, 2024).
65. The Associated Press. Musk's Twitter has dissolved its Trust and Safety Council. *NPR* (2022).
66. McGuirk, R. X Corp. has slashed 30% of trust and safety staff, an Australian online safety watchdog says. *AP News* https://apnews.com/article/x-corp-musk-australia-staff-safety-bc4772369cab1fe8dd975132fd8d61ed (2024).
67. Hickey, D., Fessler, D. M. T., Lerman, K. & Burghardt, K. X under Musk's leadership: Substantial hate and no reduction in inauthentic activity. *PLOS ONE* **20**, e0313293 (2025).
68. Jhaver, S., Boylston, C., Yang, D. & Bruckman, A. Evaluating the Effectiveness of Deplatforming as a Moderation Strategy on Twitter. *Proc ACM Hum-Comput Interact* **5**, 381:1-381:30 (2021).
69. McCabe, S. D., Ferrari, D., Green, J., Lazer, D. M. J. & Esterling, K. M. Post-January 6th deplatforming reduced the reach of misinformation on Twitter. *Nature* **630**, 132–140 (2024).
70. Ghersetti, M. & Westlund, O. Habits and Generational Media Use. *Journal. Stud.* **19**, 1039–1058 (2018).

71. Bolin, G. *Media Generations: Experience, Identity and Mediatised Social Change*. (Routledge, London, 2016). doi:10.4324/9781315694955.
72. Moore, R. C. Short educational videos improve older adults' digital literacy and resilience to online deception at scale. (Stanford University, Stanford, California, 2025).
73. Dimock, M. Defining generations: Where Millennials end and Generation Z begins. *Pew Research Center* https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/ (2019).
74. Tokita, C. K. *et al.* Measuring receptivity to misinformation at scale on a social media platform. *PNAS Nexus* **3**, pgae396 (2024).
75. Lyons, B., Montgomery, J. M. & Reifler, J. Partisanship and Older Americans' Engagement with Dubious Political News. *Public Opin. Q.* **88**, 962–990 (2024).
76. Kunst, J. R. *et al.* Leveraging artificial intelligence to identify the psychological factors associated with conspiracy theory beliefs online. *Nat. Commun.* **15**, 7497 (2024).
77. Schneier, B. & Sanders, N. The apocalypse that wasn't: AI was everywhere in 2024's elections, but deepfakes and misinformation were only part of the picture. *Ash Center* https://ash.harvard.edu/articles/the-apocalypse-that-wasnt-ai-was-everywhere-in-2024s-elections-but-deepfakes-and-misinformation-were-only-part-of-the-picture/ (2024).
78. Chow, A. R. AI's Underwhelming Impact On the 2024 Elections. *TIME* https://time.com/7131271/ai-2024-elections/ (2024).
79. Hasan, S. & Ahmed, A. Gauging the AI Threat to Free and Fair Elections | Brennan Center for Justice. https://www.brennancenter.org/our-work/analysis-opinion/gauging-ai-threat-free-and-fair-elections (2024).
80. Tulin, M., Pantazi, M., Starke, C., Sivolap, M. & Dobber, T. Generative AI and Disinformation| Echoes of Doubt: Exposure to Information About Generative AI Decreases

Believability of News. *Int. J. Commun.* **19**, 24–24 (2025).

81. Hameleers, M. & Marquart, F. It's Nothing but a Deepfake! The Effects of Misinformation and Deepfake Labels Delegitimizing an Authentic Political Speech. *Int. J. Commun.* **17**, 21–21 (2023).
82. von Sikorski, C. & Hameleers, M. Disinformation in the Age of Artificial Intelligence (AI): Implications for Journalism and Mass Communication. *Journal. Mass Commun. Q.* **102**, 941–957 (2025).
83. Goel, P., Green, J., Lazer, D. & Resnik, P. S. Using co-sharing to identify use of mainstream news for promoting potentially misleading narratives. *Nat. Hum. Behav.* 1–18 (2025) doi:10.1038/s41562-025-02223-4.
84. Dahlke, R., Kumar, D., Durumeric, Z. & Hancock, J. T. Quantifying the Systematic Bias in the Accessibility and Inaccessibility of Web Scraping Content From URL-Logged Web-Browsing Digital Trace Data. *Soc. Sci. Comput. Rev.* 08944393231218214 (2023) doi:10.1177/08944393231218214.
85. González-Bailón, S. & Xenos, M. A. The blind spots of measuring online news exposure: a comparison of self-reported and observational data in nine countries. *Inf. Commun. Soc.* **0**, 1–19 (2022).
86. Allcott, H. & Gentzkow, M. Social Media and Fake News in the 2016 Election. *J. Econ. Perspect.* **31**, 211–236 (2017).
87. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).
88. Bhadani, S. *et al.* Political audience diversity and news reliability in algorithmic ranking. *Nat. Hum. Behav.* **6**, 495–505 (2022).

Supplementary Information for

Exposure to (AI-Generated) Untrustworthy Websites in the 2024 U.S. Election

Ross Dahlke, Ryan C. Moore, Jeffrey T. Hancock

Correspondence to: ross.dahlke@wisc.edu

Contents

S1 Pre-Registered Hypotheses

RQ1: How does Americans' exposure to misinformation websites during the 2024 election compare to exposure during the 2016 and 2020 elections?

H1a*: The percentage of Americans exposed to misinformation will be greater during the 2024 election than during the 2020 election.*

We find no support for H1a; the weighted one-sided t-test showed a 2024 exposure rate of 16.5% versus 26.1% in 2020 (difference = -9.6%, $t = -5.567$, $P = 1.000$, Cohen's d = -0.236).

H1b*: The percentage of Americans exposed to misinformation will be smaller in 2024 than in 2020.*

We find strong support for H1b; the weighted one-sided t-test indicated a significant decrease in exposure from 26.1% in 2020 to 16.5% in 2024 (difference = -9.6%, $t = -5.567$, $P < 0.001$, Cohen's d = -0.236).

H2a*: The average number of exposures to misinformation websites (among those exposed) will be greater during the 2024 election than during the 2020 election.*

We find support for H2a; among exposed users, the one-sided weighted t-test indicates that the average number of untrustworthy website visits in 2024 was 46.0, compared to 22.8 in 2020 (difference = 23.2, $t = 1.90$, $P = 0.029$, Cohen's $d = 0.193$).

H2b*: The average number of exposures to misinformation websites (among those exposed) will be smaller in 2024 than in 2020.*

We find no support for H2b; among exposed users, the one-sided weighted t-test indicates that the average number of untrustworthy website visits in 2024 was 46.0, compared to 22.8 in 2020 (difference = 23.2, $t = 1.90$, $P = 0.971$, Cohen's $d = 0.193$).

H3a*: The average length of time Americans spend on misinformation websites will be longer during the 2024 election than during the 2020 election.*

We find no support for H3a; among visits to untrustworthy websites, the one-sided weighted t-test indicates that the average visit duration in 2024 was 53.15 seconds compared to 54.52 seconds in 2020 (difference = -1.37, $t = -0.42$, $P = 0.661$, Cohen's d = -0.006).

H3b*: The average length of time Americans spend on misinformation websites will be shorter in 2024 than in 2020.*

We find no support for H3b; among visits to untrustworthy websites, the one-sided weighted t-test indicates that the average visit duration in 2024 was 53.15 seconds compared to 54.52 seconds in 2020 (difference = -1.37, $t = -0.42$, $P = 0.339$, Cohen's d = -0.006).

H4*: Older adults (aged 65+) will be more likely than younger adults (aged 18–29) to be exposed to misinformation websites during the 2024 election.*

Our pre-registered weighted linear regression analysis that controls for various other sociodemographic and political variables did not yield a statistically significant effect of being in the age group 65+ on exposure to untrustworthy websites ($\beta = 0.0369$, $P = 0.336$). In an exploratory one-sided weighted t-test — which provides a more direct comparison of group differences without controlling for potential collinear variables — we found that adults aged 65+ had a significantly higher exposure rate (24.8%) than adults aged 18–29 (13.8%), with a difference of 11.0% ($t = 2.41$, $P = 0.008$, Cohen's $d = 0.28$).

H5*: Those identifying as Republicans will be more likely than those identifying as Democrats to be exposed to misinformation websites during the 2024 election.*

Our pre-registered weighted linear regression analysis that controls for various other sociodemographic and political variables indicated that identification as Republican was significantly associated with higher exposure to untrustworthy websites ($\beta = 0.1515$, $P < 0.001$). In an exploratory one-sided weighted t-test — which directly compares the group differences without additional covariates — we found that Republicans had a significantly higher exposure rate (27.2%) compared to Democrats (8.3%), with a difference of 18.9% ($t = 7.56$, $P < 0.001$, Cohen's $d = 0.51$).

H6*: Twitter will be more likely to refer people to misinformation websites in the 2024 election than in 2020.*

We find support for H6; the weighted chi-square test revealed that the proportion of Twitter referrals to untrustworthy websites was significantly higher in 2024 (2.3%) than in 2020 (0.8%), yielding a difference of 1.5% (95% CI [1.2%, 1.7%], $\chi^2 = 94.811$, $P < 0.001$, Cohen's $h = 0.125$). For Twitter referrals to AI-generated untrustworthy sites, the difference between 2024 (0%) and 2020 (0.9%) was not significant ($\chi^2 = 0.076$, $P = 0.783$, Cohen's $h = -0.19$).

RQ2: To what extent is the misinformation that people are exposed to during the 2024 election AI-generated?

H7*: The percentage of Americans exposed to AI-generated misinformation websites will be greater during the 2024 election than during the 2020 election.*

We find support for H7; the weighted t-test shows that the percentage of Americans exposed to AI-generated untrustworthy websites increased significantly from 2.6% in 2020 (95% CI [1.3%, 4.0%]) to 6.5% in 2024 (95% CI [4.6%, 8.5%]), yielding a difference of 3.9% (95% CI [1.5%, 6.3%]), $t = 4.37$, $P < 0.001$, with an effect size of Cohen's $d = 0.187$.

H8*: The proportion of misinformation website exposures that are AI-generated will be greater in 2024 than in 2020.*

Our pre-registered hypothesis predicted that the proportion of untrustworthy website exposures that are AI-generated would be greater in 2024 than in 2020. In a one-sided weighted proportion test, however, we found that only 1.6% (95% CI [1.3%, 1.8%]) of exposures in 2024 were AI-generated compared to 2.1% (95% CI [1.8%, 2.5%]) in 2020,

yielding a difference of -0.6% (95% CI [-1.0%, -0.1%]), a relative ratio of 0.73 (95% CI [0.58, 0.92]), and Cohen's h = -0.043, with a non-significant one-sided $P = 0.995$. In contrast, an exploratory two-sided test indicated a statistically significant difference ($P = 0.009$), suggesting that the proportion of AI-generated exposures was lower in 2024 than in 2020.

H9*: Americans in 2024 will spend more time on AI-generated misinformation websites than on non-AI-generated misinformation websites.*

We find no support for H9; the one-sided weighted t-test indicates that the average visit duration for AI-generated untrustworthy websites was 44.48 seconds (95% CI [30.04, 58.91]) compared to 53.15 seconds (95% CI [48.78, 57.51]) for non-AI untrustworthy websites. The resulting difference of -8.67 s (95% CI [-23.75, 6.41]) was not statistically significant ($t = -1.39$, $P = 0.917$, Cohen's d = -0.054).

H10*: Older adults (aged 65+) will be more likely than younger adults (aged 18–29) to be exposed to AI-generated misinformation websites during the 2024 election.*

Our pre-registered weighted linear regression analysis that adjusts for other sociodemographic and political variables did not yield a statistically significant effect of age on exposure to AI-generated untrustworthy websites (β = -0.0485, $P = 0.0679$). In an exploratory one-sided weighted t-test—which provides a more direct comparison of group means—we found that the mean exposure rate among adults aged 65+ was 5.8% (95% CI [3.1%, 8.5%]), compared to 10.0% (95% CI [-0.2%, 20.2%]) for those aged 18–29, resulting in a difference of -4.2% (95% CI [-11.1%, 2.7%], $t = -1.19$, $P = 0.882$, Cohen's d = -0.156).

H11*: Those identifying as Republicans will be more likely than those identifying as Democrats to be exposed to AI-generated misinformation websites during the 2024 election.*

Our pre-registered weighted linear regression analysis that adjusts for other sociodemographic and political variables indicated that identification as Republican was not significantly associated with exposure to AI-generated untrustworthy websites (β = -0.00166, $P = 0.920$). In contrast, an exploratory one-sided weighted t-test—which directly compares the group means without additional control—revealed that Republicans had a mean exposure rate of 7.4% (95% CI [3.9%, 11.0%]) versus 5.0% (95% CI [3.0%, 7.0%]) for Democrats, yielding a difference of 2.5% (95% CI [-0.6%, 5.6%], $t = 1.51$, $P = 0.066$, Cohen's $d = 0.102$).

Multiple Testing Correction Summary

To address concerns about multiple comparisons, we present the results of our 14 pre-registered hypothesis tests with adjusted p-values using both Holm-Bonferroni and Bonferroni corrections (Table S1):

Hyp.	Description	Original P	Holm P	Bonferroni P	Original Sig.	Holm Sig.	Bonferroni Sig.
H1a	2024 > 2020 exposure	1.000	1.000	1.000	No	No	No
H1b	2024 < 2020 exposure	<0.001	<0.001	<0.001	Yes	Yes	Yes
H2a	2024 > 2020 intensity	0.029	0.291	0.407	Yes	No	No
H2b	2024 < 2020 intensity	0.971	1.000	1.000	No	No	No
H3a	2024 > 2020 duration	0.661	1.000	1.000	No	No	No
H3b	2024 < 2020 duration	0.339	1.000	1.000	No	No	No
H4	Age 65+ > 18–29 exposure	0.336	1.000	1.000	No	No	No
H5	Republicans > Democrats exposure	<0.001	<0.001	<0.001	Yes	Yes	Yes
H6	Twitter refs 2024 > 2020	<0.001	<0.001	<0.001	Yes	Yes	Yes
H7	AI exposure 2024 > 2020	<0.001	<0.001	<0.001	Yes	Yes	Yes
H8	AI proportion 2024 > 2020	0.995	1.000	1.000	No	No	No
H9	AI > non-AI duration	0.917	1.000	1.000	No	No	No
H10	Age 65+ > 18–29 AI exposure	0.068	0.611	0.950	No	No	No
H11	Republicans > Democrats AI exposure	0.920	1.000	1.000	No	No	No

Table S1: Summary of hypothesis tests with multiple testing corrections. Four hypotheses remain significant after Bonferroni correction (H1b, H5, H6, H7). We applied both Holm-Bonferroni and traditional Bonferroni methods to maintain familywise error rate control.

S2 Exposure Threshold

To address concerns that our topline exposure findings might be driven by incidental single-visit encounters, we examine the percentage of participants exposed at various threshold levels. This analysis tests whether the declining exposure patterns observed across election years persist when requiring progressively higher minimum visit counts to classify someone as "exposed."

Figure S1 presents exposure rates across six different thresholds: $\geq$1, $\geq$2, $\geq$5, $\geq$10, $\geq$25, and $\geq$50 visits to untrustworthy websites.

Several key patterns emerge from this analysis. First, the fundamental pattern of declining exposure from 2016 to 2020 to 2024 holds consistently across all thresholds. Second, despite the general pattern holding, the difference between the elections gets smaller at each threshold, meaning that the percentage of American adults who were exposed at higher thresholds is more similar than smaller number of minimum exposures. Third, meaningful exposure persists even at higher thresholds.

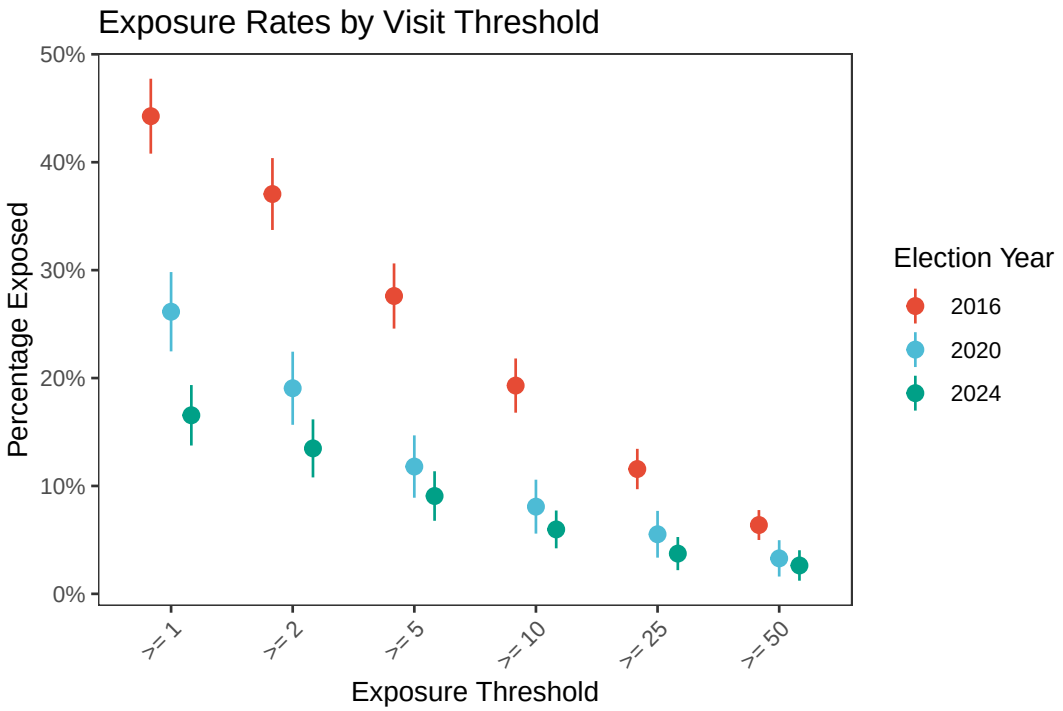


Figure S1: Exposure Rates by Visit Threshold Across Election Years. Percentage of Americans exposed to untrustworthy websites at different minimum visit count thresholds during the 2016, 2020, and 2024 elections. Each point represents the weighted percentage of participants who visited untrustworthy websites at least the specified number of times, with 95% confidence intervals. The consistent pattern of 2016 $>$ 2020 $>$ 2024 exposure rates across all thresholds demonstrates that the observed decline in exposure is not driven by incidental single-visit encounters but reflects meaningful changes in engagement with untrustworthy content.

S3 Exposure Intensity Distribution (CCDF)

To further analyze the intensity of untrustworthy website exposures, we also visualized exposed participants' number of exposures via a complementary cumulative distribution function (CCDF) (see Figure S2). This visual analysis suggests that the long-tailed distribution of the number of visits was strongest in 2024, contributing to the uptick in the average number of exposures among the exposed that we found.

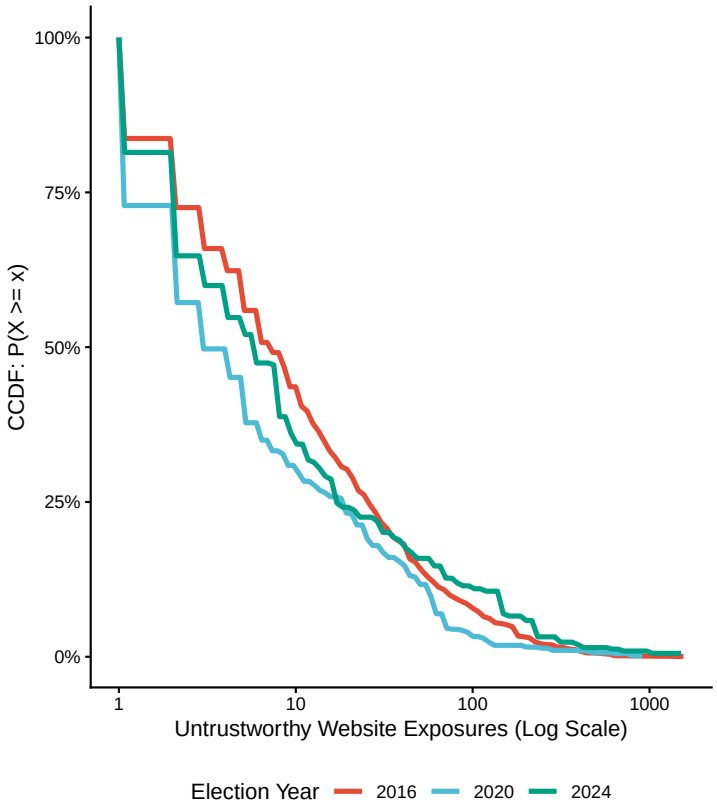


Figure S2: Complementary Cumulative Distribution Function of Untrustworthy Website Exposure Intensity. The CCDF shows the probability that exposure count exceeds a given value among individuals who visited at least one untrustworthy website. The heavy-tailed distribution across all election years demonstrates that while most exposed individuals have relatively few visits, a small subset accounts for disproportionately high exposure levels. The 2016 election shows the longest tail, consistent with higher overall exposure rates in that year.

S4 Credible News Consumption

One potential explanation for the decline in exposure to untrustworthy websites between the 2020 and 2024 elections is that Americans may be consuming less online news in general, leading to lower overall exposure to all types of content, including untrustworthy sources. To investigate this possibility, we examined participants' consumption of credible news sources. We used the same database approach for credible news domains employed in the 2020 study (i.e., we analyzed visits to website classified as hard news by [1] and added domains that were rated by NewsGuard as not repeatedly publishing false content and having a quality score of 60+ out of 100, following past work (e.g., [2, 3]), allowing us to directly compare how many people accessed credible outlets and how frequently they did so in 2020 versus 2024.

As shown in Figure S3, the proportion of Americans who visited at least one credible news site remained statistically unchanged from 2020 (89.7%, 95% CI [86.6%, 92.7%]) to 2024 (89.8%, 95% CI [87.4%, 92.2%]), with an absolute difference of only 0.2% (95% CI [-3.7%, 4.1%], $t = 0.13$, $P = 0.893$). Among those who did visit credible news sites, the average number of visits was 270 in 2024 (95% CI [221.2, 318.8]) compared to 312.5 in 2020 (95% CI [255.9, 369]), a nonsignificant difference of -42.5 visits (95% CI [-117.1, 32.2], $t = -1.22$, $P = 0.223$). Taken together, these results suggest that the overall level of online engagement with reputable news sources did not materially decline during the 2024 election period.

This stability in credible news consumption reinforces our interpretation of the decline in exposure to untrustworthy websites. It indicates that the drop in visits to untrustworthy sites is not merely a byproduct of people consuming less online information overall. Instead, the data point to a more targeted reduction in visits to sources classified as untrustworthy, while engagement with credible news outlets remains steady. This pattern suggests that the observed decrease in exposure to untrustworthy websites likely reflects changes in how Americans navigate online information ecosystems, rather than a blanket reduction in news consumption.

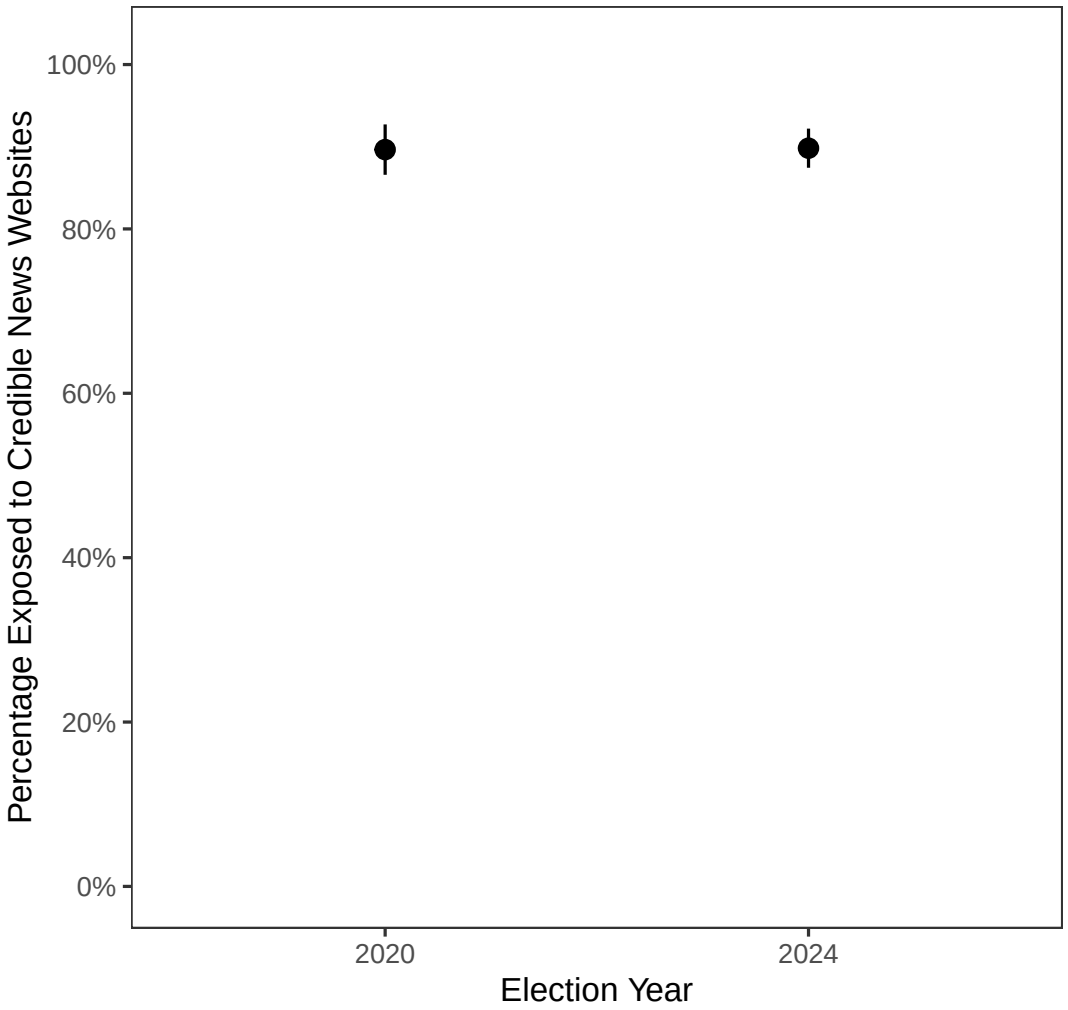


Figure S3: Credible News Consumption Proportion.

S5 Age Variance Analysis

In our primary models predicting exposure to at least one untrustworthy website, individuals aged 65 years and older were significantly more likely than 18–29-year-olds to have been exposed in both 2016 ($\beta = 0.134$, SE $= 0.024$, $P < 0.001$) and 2020 ($\beta = 0.138$, SE $= 0.043$, $P = 0.001$). However, by 2024, the association between being aged 65 years or older and untrustworthy website exposure was no longer statistically significant ($\beta = 0.037$, SE $= 0.038$, $P = 0.336$). This attenuation occurred despite a simple weighted t-test continuing to show a significant bivariate association. To investigate why the multivariable effect of age disappeared in 2024, we conducted a comparative analysis of the underlying variance structures in both the 2020 and 2024 models.

First, we evaluated the extent of multicollinearity among the covariates in both years by calculating the Variance Inflation Factor (VIF) for each predictor. In both 2020 and 2024, all VIF values were well below conventional thresholds for concern (in 2024, the highest was 2.37 for the Age 45 to 64 indicator; in 2020, it was 2.33 for the same indicator). These results suggest that multicollinearity was not a major problem in either model and that standard errors were not being unduly inflated due to redundancy among predictors.

To more directly assess variance overlap, we conducted auxiliary regressions in both years, predicting the Age 65 and older indicator from the other covariates. The models revealed a substantial and remarkably stable degree of variance overlap across elections: 56.6% of the variance in the Age 65+ indicator was explained by other covariates in 2024 ($R^2 = 0.566$), compared to 55.7% in 2020 ($R^2 = 0.557$). In both years, Republican affiliation was a significant, positive predictor of being aged 65 or older (2024: $\beta = 0.065$, $P < 0.001$; 2020: $\beta = 0.059$, $P = 0.002$). This result indicates that older age was tightly linked to other predictive characteristics in both election cycles.

The critical difference between the two years, however, lies in the extent to which partisanship statistically absorbed the variance otherwise attributed to age. To formally test this, we re-estimated the models for both years, excluding Republican affiliation as a control. In 2020, removing partisanship only modestly increased the coefficient for Age 65+, from $\beta = 0.138$ to $\beta = 0.157$, with the effect remaining highly significant. In stark contrast, removing partisanship from the 2024 model caused the age coefficient to increase by nearly 70%, from $\beta = 0.037$ to $\beta = 0.062$, demonstrating a much stronger mediating effect. Furthermore, a nested model comparison reveals that including Republican affiliation significantly improved model fit in both years; however, the effect was substantially larger in 2024 ($F(1, 1041) = 39.9$) than in 2020 ($F(1, 1142) = 22.2$).

These comparative results support a more nuanced conclusion. While individuals aged 65 and older continued to be exposed to more untrustworthy websites at a descriptive level, the statistical relationship between age and exposure changed between 2020 and 2024. In 2020, age was a significant predictor of exposure even when accounting for partisanship. By 2024, its independent predictive power was fully captured by other covariates in the model, particularly Republican affiliation. This pattern reflects an evolving information

environment where the demographic and political correlates of misinformation exposure have intensified, making partisanship a statistically dominant predictor that now more fully accounts for age-related differences.

S6 Generation and Age Analysis

While the main text focuses on age group differences in exposure to untrustworthy websites, here we provide a generational breakdown to better understand which cohorts contributed most to the overall changes between 2016, 2020, and 2024. These analyses reveal important nuances in the evolution of exposure patterns over time.

Across all five generations, exposure to untrustworthy websites declined between 2016 and 2024, but the magnitude and timing of these changes varied (see Figure S4 and Table S2). The Silent Generation experienced substantial declines across both periods: a significant decrease from 2016 to 2020 (Diff = -15.6%, $t = -2.66$, $P = 0.009$) followed by an even sharper reduction from 2020 to 2024 (Diff = -29.8%, $t = -3.33$, $P = 0.001$). These results suggest that the oldest adults not only began with very high exposure rates but also underwent the most pronounced improvements over time.

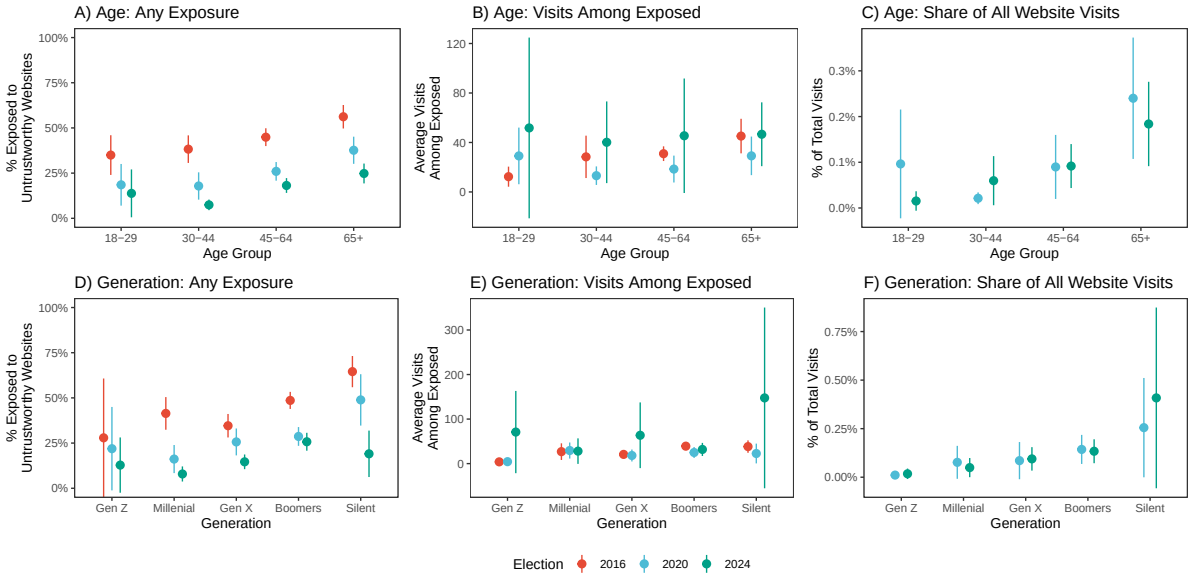


Figure S4: Age and Generational Exposure Patterns Across Elections. Comprehensive comparison showing exposure to untrustworthy websites across age groups (top row) and generational cohorts (bottom row). Panels A/D: percentage with any exposure. Panels B/E: average visits among those exposed. Panels C/F: untrustworthy visits as percentage of total visits. Note the persistent age and generational gradients across all measures, with older groups maintaining higher exposure rates despite overall declines from 2016 to 2024. Data are weighted means ± 95% confidence intervals.

In contrast, Boomers exhibited a different pattern. Boomers showed a large initial decline between 2016 and 2020 (Diff = -19.9%, $t = -8.59$, $P < 0.001$), but no statistically significant change between 2020 and 2024 (Diff = -2.9%, $t = -0.98$, $P = 0.325$). This

plateau suggests that much of the improvement among older adults occurred early in the study period, and that further reductions among Boomers may be more challenging to achieve without additional interventions.

Generation X demonstrated a steady and consistent improvement across both periods. Exposure among Gen X declined from 2016 to 2020 (Diff = -9.0%, $t = -2.72$, $P = 0.007$) and again from 2020 to 2024 (Diff = -11.0%, $t = -3.42$, $P = 0.001$). Millennials similarly showed a sharp initial decline (Diff = -25.3%, $t = -5.97$, $P < 0.001$) and a smaller, but still statistically significant, additional drop between 2020 and 2024 (Diff = -8.3%, $t = -2.47$, $P = 0.014$). These findings suggest that Gen X and Millennials steadily reduced their exposure to untrustworthy content across multiple election cycles.

In contrast, Gen Z maintained relatively low exposure levels throughout the three election cycles, with no statistically significant changes from 2016 to 2020 (Diff = -6.0%, $t = -0.35$, $P = 0.729$) or from 2020 to 2024 (Diff = -9.1%, $t = -1.00$, $P = 0.317$). This suggests that Gen Z entered adulthood during a period of lower exposure and largely remained at these low levels over time.

Taken together, these generational analyses reveal that the overall reductions in untrustworthy website exposure observed in the main paper were driven disproportionately by the Silent Generation, Generation X, and Millennials. Boomers' exposure rates declined sharply initially but leveled off by 2024, while Gen Z exhibited consistently low but relatively stable exposure rates. These findings add nuance to the broader age-group analyses, highlighting that improvements were not uniform across cohorts and that certain generations contributed more strongly to the patterns of change observed at the aggregate level.

Table S2: Generational Differences in Untrustworthy Website Exposure Over Time.

Generation	2016	2020	2024	2020 versus 2016	2024 versus 2020
Silent	64.5% [55.9, 73.2]	48.9% [34.7, 63.1]	19.1% [6.2, 31.9]	Diff=-15.6%, t=-2.66, p=0.009	Diff=-29.8%, t=-3.33, p=0.001
Boomers	48.6% [43.9, 53.3]	28.7% [23.5, 33.9]	25.7% [20.8, 30.7]	Diff=-19.9%, t=-8.59, p=0.000	Diff=-2.9%, t=-0.98, p=0.325
Gen X	34.6% [28.1, 41.1]	25.6% [18.2, 33.1]	14.6% [10.6, 18.7]	Diff=-9.0%, t=-2.72, p=0.007	Diff=-11.0%, t=-3.42, p=0.001
Millennial	41.4% [32.4, 50.5]	16.2% [8.4, 23.9]	7.9% [3.7, 12.1]	Diff=-25.3%, t=-5.97, p=0.000	Diff=-8.3%, t=-2.47, p=0.014
Gen Z	27.9% [-4.9, 60.7]	21.9% [-1.2, 45.0]	12.8% [-2.5, 28.1]	Diff=-6.0%, t=-0.35, p=0.729	Diff=-9.1%, t=-1.00, p=0.325

S7 Vote Choice and Political Ideology Analysis

This section provides a comprehensive analysis of how vote choice, voter registration status, and personal ideology relate to exposure to untrustworthy websites in 2024. These analyses address questions about which specific segments of the political spectrum drive observed partisan differences in exposure.

S7.1 Vote Choice Analysis

When examining exposure by 2024 presidential vote choice, stark differences emerge between supporters of the two major candidates. Trump supporters had substantially higher exposure rates (29.0%, 95% CI [23.4%, 34.6%]) compared to Harris supporters (8.9%, 95% CI [6.2%, 11.6%]). A weighted t-test confirmed this difference was statistically significant ($t = 7.77$, $P < .001$, Cohen's $d = 0.25$), representing a 20.1 percentage point gap in exposure rates.

This vote choice analysis reveals that the partisan divide in untrustworthy website exposure extends beyond general party identification to specific candidate support, with Trump supporters showing exposure rates more than three times higher than Harris supporters.

S7.2 Voter Registration Analysis

Voter registration status also significantly predicted exposure to untrustworthy websites. Registered voters were more likely to be exposed (17.6%, 95% CI [14.6%, 20.6%]) compared to non-registered individuals (7.5%, 95% CI [0.4%, 14.7%]). This difference was statistically significant ($t = 3.48$, $P < .001$, Cohen's $d = 0.416$), representing a 10.1 percentage point difference.

The higher exposure rates among registered voters may reflect greater overall political engagement and information-seeking behavior, though this engagement unfortunately extends to untrustworthy sources as well as credible ones.

S7.3 Interaction Analyses

To better understand how partisanship and voter registration interact, we examined exposure rates across combinations of these factors. Table S3 presents the results of this analysis.

Table S3: Exposure to Untrustworthy Websites by Partisanship and Voter Registration Status

Party Group	Registration Status	Exposure %	95% CI
Democrat/Lean Dem	Not Registered	2.0	[-1.9, 5.8]
Democrat/Lean Dem	Registered	8.6	[5.7, 11.5]
Independent	Not Registered	4.7	[-1.9, 11.4]
Independent	Registered	11.3	[6.0, 16.6]
Republican/Lean Rep	Not Registered	9.6	[-3.0, 22.3]
Republican/Lean Rep	Registered	28.6	[23.0, 34.2]

The interaction reveals that the partisan gap in exposure is primarily driven by registered Republicans, who show exposure rates of 28.6% (95% CI [23.0%, 34.2%]) compared to 8.6% (95% CI [5.7%, 11.5%]) for registered Democrats. Among non-registered individuals, partisan differences are much smaller: Republicans at 9.6% (95% CI [-3.0%, 22.3%]) versus Democrats at 2.0% (95% CI [-1.9%, 5.8%]).

Similarly, when examining Trump support and voter registration (Table S4), the pattern becomes even more pronounced.

Table S4: Exposure to Untrustworthy Websites by Trump Support and Voter Registration Status

Trump Support	Registration Status	Exposure %	95% CI
Non-Trump Supporter	Not Registered	8.3	[0.4, 16.1]
Non-Trump Supporter	Registered	8.6	[6.3, 11.0]
Trump Supporter	Not Registered	0.0	[0.0, 0.0]
Trump Supporter	Registered	29.8	[24.0, 35.5]

Notably, among non-Trump supporters, voter registration status made little difference in exposure rates (8.3%, 95% CI [0.4%, 16.1%] vs 8.6%, 95% CI [6.3%, 11.0%]). However, among Trump supporters, registered voters showed substantially higher exposure (29.8%, 95% CI [24.0%, 35.5%]) compared to non-registered Trump supporters (0.0%, 95% CI [0.0%, 0.0%]).

S7.4 Political Ideology Analysis

To examine whether exposure patterns reflect ideological extremism beyond party identification, we analyzed exposure rates across a five-point personal ideology scale. Figure S5 and Table S5 present these results.

Table S5: Exposure to Untrustworthy Websites by Personal Ideology (2024)

Ideology	**Exposure %**	**95% CI**
Very Liberal	9.9	[4.6, 15.2]
Liberal	9.6	[4.8, 14.5]
Moderate	9.3	[6.2, 12.3]
Conservative	29.0	[20.5, 37.5]
Very Conservative	33.1	[24.3, 41.8]
Not Sure	4.3	[-3.9, 12.4]

The results reveal a clear ideological gradient, with conservative and very conservative individuals showing substantially higher exposure rates than liberal, moderate, or uncertain respondents. Notably, the gap between Conservative (29.0%, 95% CI [20.5%, 37.5%]) and Very Conservative (33.1%, 95% CI [24.3%, 41.8%]) individuals suggests that exposure is concentrated among the most ideologically committed conservatives rather than being broadly distributed across the political right.

When comparing Very Conservative individuals to all others combined, the difference is substantial: Very Conservative individuals had exposure rates of 33.1% (95% CI [24.3%, 41.8%]) compared to 13.6% for all other groups combined. This concentration among ideological extremists suggests that untrustworthy website consumption in 2024 is increasingly limited to the most committed conservative partisans.

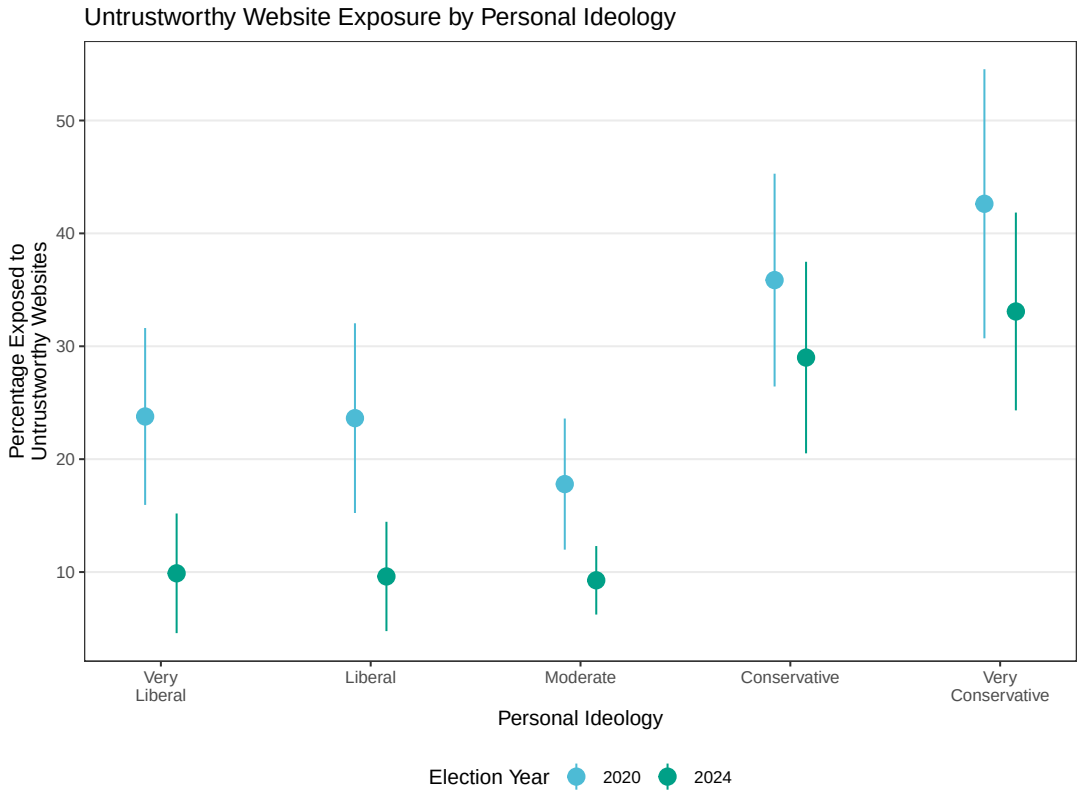


Figure S5: Exposure to Untrustworthy Websites by Personal Ideology (2020 vs 2024). Percentage of Americans exposed to untrustworthy websites across different ideological categories during the 2020 and 2024 elections. The figure shows exposure rates with 95% confidence intervals for each ideological group. While all groups experienced declines from 2020 to 2024, the decreases were most pronounced among liberal and moderate groups, with conservative groups showing smaller and non-significant reductions.

S7.5 Political Ideology Exposure by Election

To understand how ideological patterns changed over time, we compared 2024 exposure rates with those from 2020 across all ideological groups. Table S6 presents these comparisons.

Table S6: Temporal Comparison of Exposure by Personal Ideology (2020 vs 2024)

Ideology	**2020 %**	**2024 %**	**Change**	**t-value**	**p-value**	**Cohen's d**
Very Liberal	23.8 [15.9, 31.6]	9.9 [4.6, 15.2]	-13.9	-3.33	< 0.001	-0.38
Liberal	23.6 [15.2, 32.0]	9.6 [4.8, 14.5]	-14.0	-4.06	< 0.001	-0.38
Moderate	17.8 [12.0, 23.6]	9.3 [6.2, 12.3]	-8.5	-3.36	< 0.001	-0.25
Conservative	35.9 [26.4, 45.3]	29.0 [20.5, 37.5]	-6.9	-1.47	0.142	-0.15
Very Conservative	42.6 [30.7, 54.5]	33.1 [24.3, 41.8]	-9.5	-1.66	0.098	-0.20

The temporal analysis reveals that exposure declined significantly across liberal and moderate groups between 2020 and 2024. Very liberal individuals showed a 13.9 percentage point decline from 23.8

These findings reinforce the conclusion that untrustworthy website exposure in 2024 is increasingly concentrated among conservative partisans, particularly those with the strongest ideological commitments and attachment to Trump.

S8 Most Visited Untrustworthy Websites

Supplemental Table S7 presents the five untrustworthy domains with the highest overall visitation rates in 2024, alongside each site's rank in 2020 and its standing within the top decile of liberal and conservative media diets.

Table S7: Most visited untrustworthy websites overall and by media-diet decile in 2024.

Rank (2024 All)	2024 Most Visited Sites	Rank 2020	Rank 2024 Most Liberal Media Decile	Rank 2024 Most Conservative Media Decile
1	infowars.com	37	–	1
2	dailykos.com	1	1	–
3	theepochtimes.com	12	3	2
4	newsmax.com	4	–	4
5	thegatewaypundit.com	2	2	5

infowars.com exhibits the most dramatic upward shift, climbing from 37th overall in 2020 to 1st in 2024. This rise appears to be driven by participants in the most conservative media decile, who also visited the most untrustworthy website. In contrast, not a single participant in the most liberal decile visited *infowars.com*. This trend reveals how InfoWars has reemerged as a central node within the conservative media ecosystem, possibly fueled by recent legal controversies and high-profile political commentary surrounding its founder, while remaining largely insulated from opposing audiences. This finding underscores both the volatility of media outlets and the deep ideological segregation of today's news environment.

Two domains, *theepochtimes.com* and *thegatewaypundit.com*, bridge both audiences: the Epoch Times ranks 3rd overall and appears as 3rd in the liberal decile and 2nd in the conservative decile. At the same time, The Gateway Pundit is 5th overall, 2nd among liberals, and 5th among conservatives. Their dual presence suggests that these sites, for whatever reason, draw relatively equal traffic across the partisan divide. Meanwhile, *newsmax.com* holds steady at 4th overall (also 4th in 2020) and ranks 4th in the conservative decile, but was not visited by any participants in the most liberal media decile.

Together, these patterns reveal both strong ideological anchors—Infowars and Newsmax for conservatives, and Daily Kos for liberals—and a subset of "crossover" domains that reach multiple segments. The rise of Infowars and the sustained prominence of Daily Kos illustrate the persistence of distinct partisan information ecosystems. At the same time, the cross-decile reach of the Epoch Times and The Gateway Pundit highlights the potential for certain untrustworthy sources to span ideological boundaries.

S9 Untrustworthy Website Exposure Perceptions

This section explores whether participants' perceptions of untrustworthy website exposure correspond to their actual browsing behavior. We consider three separate independent variables: self-assessment (individuals' beliefs about how much they themselves encounter untrustworthy content), others-assessment (individuals' beliefs about how much other people encounter untrustworthy content), and the third-person gap (the difference between how much participants believe others are exposed compared to themselves). We are particularly interested in examining the moderating relationship of partisan identity and perceptions. In other words, we investigate whether partisanship influences the perception of exposure to untrustworthy websites and whether individuals are actually exposed to such websites.

S9.1 Self-Assessment

The self-assessment measure indicates how much respondents believe they personally see untrustworthy content online, regardless of whether they think others encounter the same or different amounts. Specifically, participants were asked, "In the last week, how many visits do you think you have made to websites that repeatedly share untrustworthy news?" and were able to respond with an integer.

Table S8 includes two models for self-assessment. Model 1 incorporates partisan identification and self-assessment along with their interactions, while Model 2 adds demographic and political interest controls. In both models, participants identifying as Republican exhibit higher levels of actual visits to untrustworthy websites than Independents, but the self-assessment variable alone is not statistically significant. Once controls are added in Model 2, a significant interaction emerges between Republican and self-assessment, suggesting that Republican respondents who believe they personally see more untrustworthy content do, in fact, visit such sites at higher rates than those who report lower self-assessment.

Figure S6 displays the relationship between self-assessment and actual exposure across all three outcome measures in a three-panel format. Panel A shows the binary exposure relationship, Panel B displays visit counts, and Panel C presents percentage exposure. Across all panels, Republican respondents (red) exhibit higher baseline exposure and stronger positive associations between self-assessment and actual exposure. Democrats (blue) and Independents (gray) show flatter relationships and lower overall exposure levels. The consistency across outcome measures strengthens confidence in the robustness of partisan differences in perception-behavior alignment.

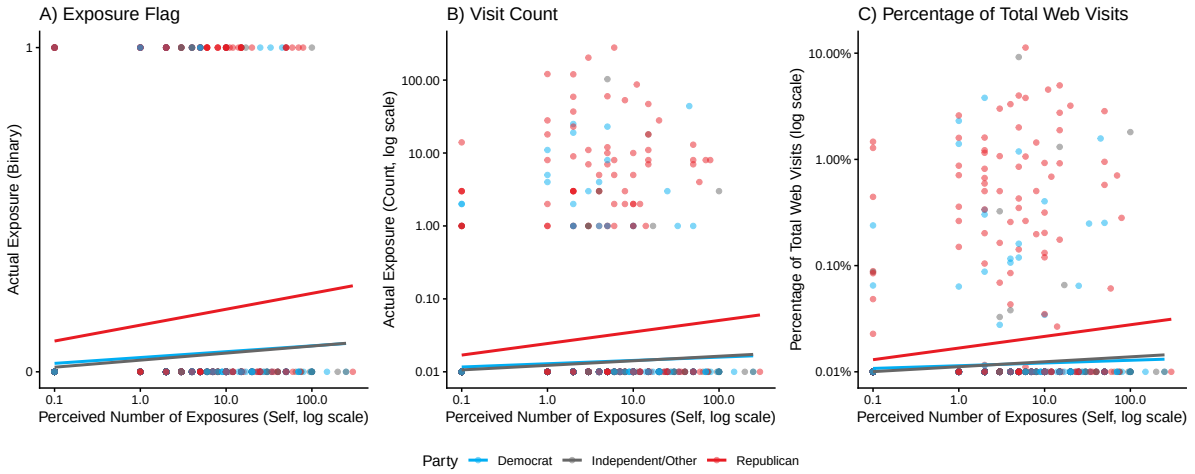


Figure S6: Self-Assessment and Actual Exposure Across Three Outcome Measures (7-Day Period). Panel A shows binary exposure flag, Panel B shows visit counts, and Panel C shows percentage exposure, all plotted against self-assessed exposure perceptions with partisan identification indicated by color. All measures reflect the 7 days prior to participants taking the survey.

Table S8 displays two models for self-assessment, labeled Model 1 (basic) and Model 2 (full). Republicans show greater untrustworthy site exposure overall, and the full model reveals a significant interaction for Republicans, indicating that Republican respondents who perceive higher self-exposure also tend to have higher actual exposure relative to Republicans who perceive lower self-exposure.

Table S8: Regression Models Predicting Actual Untrustworthy Website Exposure by Self-Assessment (7-Day Period).

	Dependent variable:	
	Actual Untrustworthy Website Exposure	
	Model 1 (Basic)	Model 2 (Full)
Republican	0.137***	0.080**
	(0.026)	(0.029)
Democrat	0.017	-0.028
	(0.026)	(0.029)
Self-Assessment	0.0001	-0.0002
	(0.001)	(0.001)
Republican × Self-Assessment		0.045***
		(0.010)
Democrat × Self-Assessment		0.010
		(0.019)
Political Interest		-0.033
		(0.018)
College		-0.026
		(0.021)
Female		-0.046
		(0.030)
Non-White		-0.018
		(0.029)
Age 30–44		0.038
		(0.031)
Age 45–64		0.001
		(0.001)
Age 65+		0.001
		(0.001)
Constant	0.028	-0.035
	(0.022)	(0.044)
Observations	1,069	1,050
R^2	0.049	0.093
Adjusted R^2	0.044	0.082

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

S9.2 Others Assessment and Partisan Perceptions

The others-assessment measure gauges how much participants believe other people encounter untrustworthy content, independent of their own exposure. Additionally, we examine partisan-specific perceptions by analyzing how respondents assess Democrats' and Republicans' untrustworthy website exposure. Participants were asked: "In the last week, how many visits do you think the average American made to websites that repeatedly share untrustworthy news?", "In the last week, how many visits do you think the average Democrat made to websites that repeatedly share untrustworthy news?", and "In the last week, how many visits do you think the average Republican made to websites that repeatedly share untrustworthy news?"

Figure S7 presents a comprehensive 3×3 visualization showing the relationship between different partisan perception measures (rows) and actual exposure across three outcome

measures (columns). The rows display: (1) general others assessment, (2) Democrat-specific assessments, and (3) Republican-specific assessments. The columns show: (A) binary exposure flag, (B) visit counts, and (C) percentage exposure. All measures reflect the 7-day period prior to participants taking the survey.

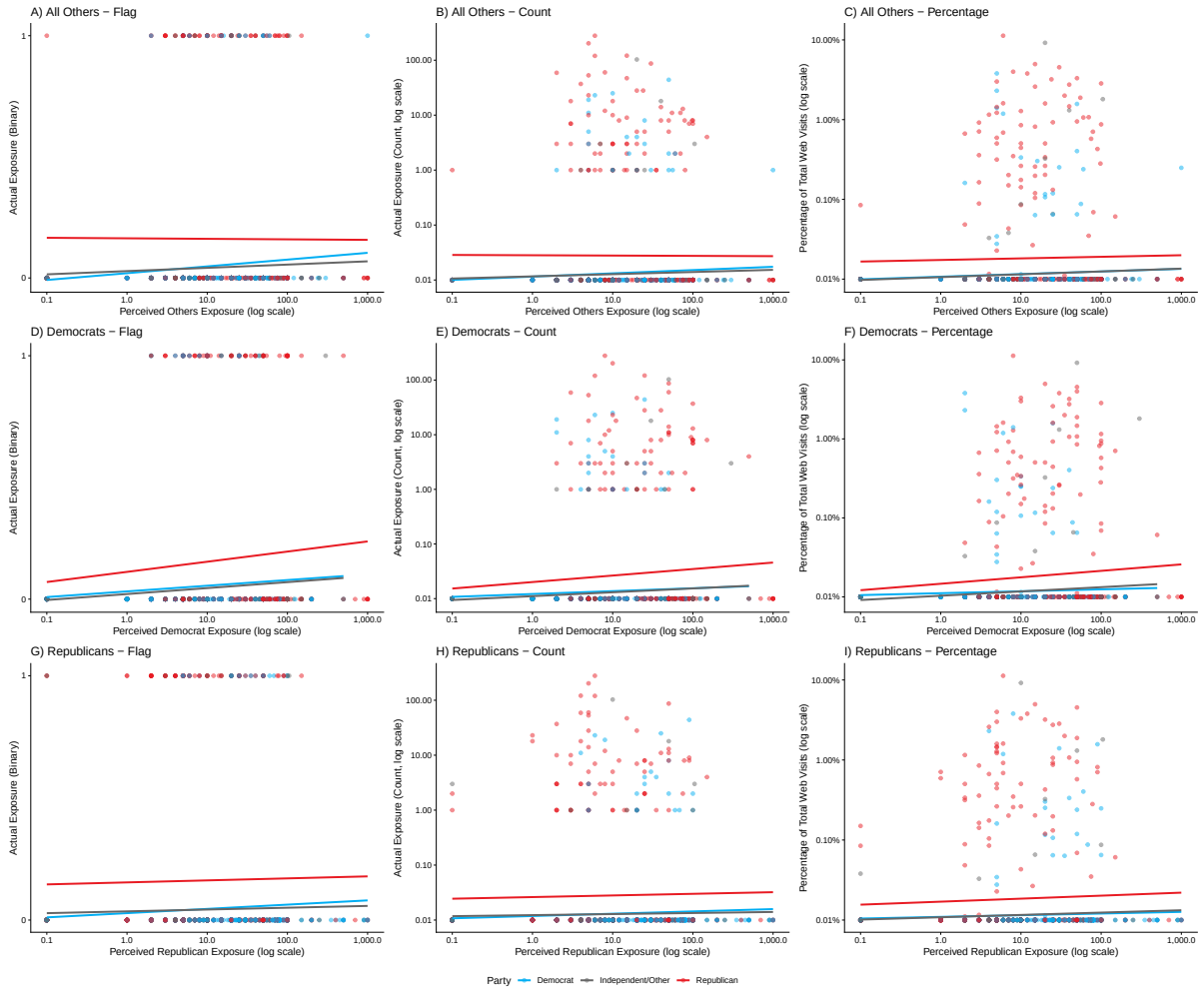


Figure S7: Partisan Perception and Actual Exposure Across Multiple Measures (7-Day Period). The 3×3 grid shows relationships between partisan perceptions (rows: Others, Democrats, Republicans) and actual exposure (columns: Binary Flag, Visit Count, Percentage). All measures reflect the 7 days prior to participants taking the survey.

Table S9 displays six models examining different partisan perception measures. Models 1-2 focus on others-assessment, Models 3-4 examine Republican-specific assessments, and Models 5-6 analyze Democrat-specific assessments. Each pair includes a basic model (partisan ID + perception + interactions) and a full model with demographic controls.

Consistent with the self-assessment findings, Republicans show stronger alignment between their perceptions and actual behavior across all three perception targets.

Table S9: Regression Models Predicting Actual Untrustworthy Website Exposure by Partisan Perceptions (7-Day Period).

	Dependent variable:					
	Actual Untrustworthy Website Exposure					
	Others Assessment		Republican Assessment		Democrat Assessment	
	Basic	Full	Basic	Full	Basic	Full
	(1)	(2)	(3)	(4)	(5)	(6)
Republican	0.150***	0.097***	0.148***	0.094**	0.158***	0.108***
	(0.026)	(0.029)	(0.026)	(0.029)	(0.027)	(0.030)
Democrat	0.023	-0.016	0.020	-0.020	0.029	-0.008
	(0.027)	(0.029)	(0.027)	(0.029)	(0.028)	(0.030)
Others-Assessment	-0.00001	-0.00004	–	–	–	–
	(0.0002)	(0.0002)				
Rep-Assessment	–	–	-0.00002	-0.00005	–	–
			(0.0002)	(0.0002)		
Dem-Assessment	–	–	–	–	0.0004	0.0005
					(0.0004)	(0.0004)
Rep × Others	-0.0002	-0.0001	–	–	–	–
	(0.0003)	(0.0003)				
Dem × Others	-0.00003	-0.0001	–	–	–	–
	(0.0002)	(0.0002)				
Rep × Rep-Assess	–	–	-0.0002	-0.0001	–	–
			(0.0003)	(0.0003)		
Dem × Rep-Assess	–	–	-0.00002	-0.0001	–	–
			(0.0002)	(0.0002)		
Rep × Dem-Assess	–	–	–	–	-0.001	-0.001
					(0.0004)	(0.0005)
Dem × Dem-Assess	–	–	–	–	-0.0004	-0.001
					(0.001)	(0.001)
Political Interest		0.044***		0.044***		0.045***
		(0.010)		(0.010)		(0.010)
College		0.011		0.011		0.011
		(0.019)		(0.019)		(0.019)
Female		-0.033		-0.033		-0.032
		(0.018)		(0.018)		(0.018)
Non-White		-0.024		-0.024		-0.023
		(0.021)		(0.021)		(0.021)
Age 30–44		-0.042		-0.042		-0.043
		(0.030)		(0.030)		(0.030)
Age 45–64		-0.013		-0.013		-0.015
		(0.028)		(0.028)		(0.028)
Age 65+		0.040		0.040		0.040
		(0.031)		(0.031)		(0.031)
Constant	0.030	-0.041	0.030	-0.041	0.019	-0.055
	(0.022)	(0.043)	(0.022)	(0.043)	(0.023)	(0.044)
Observations	1,069	1,050	1,069	1,050	1,069	1,050
R^2	0.056	0.097	0.050	0.093	0.050	0.093
Adjusted R^2	0.051	0.086	0.045	0.082	0.045	0.083

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

S9.3 Third-Person Gap

The third-person gap is defined as the difference between the perceived exposure of others and one's own perceived exposure. A larger gap indicates that a participant believes others are substantially more exposed to untrustworthy content than they personally are. The third-person gap is computed using each participant's pair of responses to the perceived own exposure and perceived others' exposure survey questions.

Table S10 presents two corresponding models. Model 1 shows that Republicans have higher average untrustworthy exposure, although the gap measure itself is not statistically significant. In Model 2, the addition of demographic and political interest controls yields a notable interaction for Republican respondents, implying that a larger third-person gap is associated with higher untrustworthy site visits among this group.

Table S10: Regression Models Predicting Actual Untrustworthy Website Exposure by Third-Person Gap.

	Dependent variable:	
	Actual Untrustworthy Website Exposure	
	(1)	(2)
Republican	0.180***	0.110***
	(0.032)	(0.035)
Democrat	-0.029	-0.078**
	(0.032)	(0.035)
Third-Person Gap	-0.0002	-0.0002
	(0.0003)	(0.0003)
Rep × Third-Person Gap	-0.0001	0.054***
	(0.0003)	(0.012)
Dem × Third-Person Gap	0.001*	0.022
	(0.0004)	(0.024)
Political Interest		0.012
		(0.024)
College		-0.042
		(0.026)
Female		-0.050
		(0.038)
Non-White		0.020
		(0.036)
Age 30–44		0.039
		(0.038)
Age 45–64		-0.0001
		(0.0003)
Age 65+		0.001
		(0.0004)
Constant	0.103***	-0.020
	(0.026)	(0.054)
Observations	1,069	1,050
R^2	0.067	0.102
Adjusted R^2	0.063	0.091

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

Figure S8 plots the third-person gap on the horizontal axis—how much participants think others see untrustworthy content compared to themselves—and actual untrustwor-

thy exposure (as a share of total site visits) on the vertical axis. The Republican line (red) sits higher overall, suggesting that Republicans tend to have a larger fraction of untrustworthy visits than Democrats (blue) or Independents (gray). Moreover, Republicans show a modestly upward slope, indicating that those who perceive a larger gap between others' exposure and their own also exhibit somewhat greater actual untrustworthy exposure. By contrast, Democrats and Independents have flatter or slightly negative slopes, with lower overall levels of untrustworthy browsing. The broader takeaway is that Republicans' beliefs about how much more susceptible others are (relative to themselves) align more closely with their own site visits, whereas Democrats and Independents exhibit less connection between these perceptions and actual behavior.

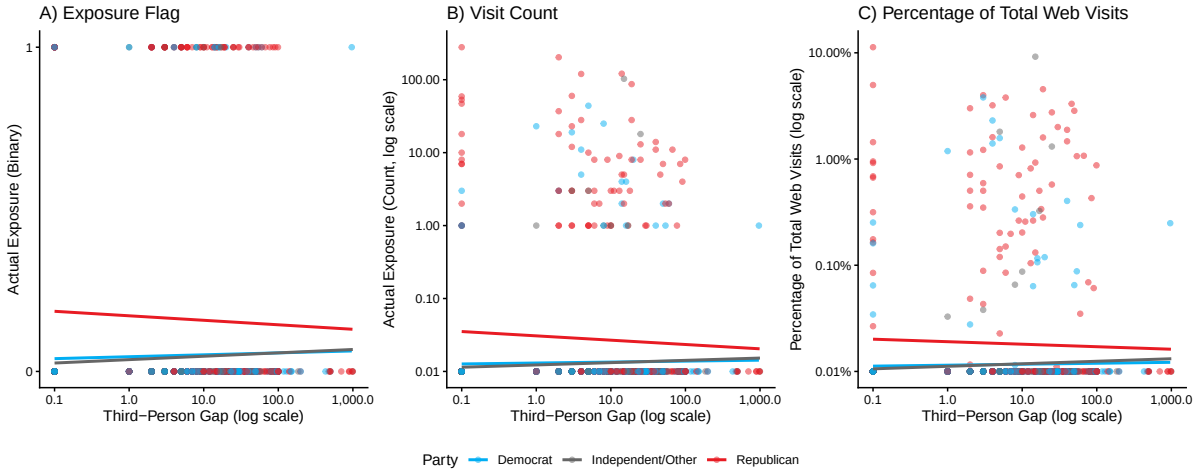


Figure S8: Third-Person Gap Assessment.

S10 Technology Attitudes

S10.1 Digital Literacy

This section examines the relationship between Americans' digital literacy and technological familiarity and their exposure to untrustworthy websites. Digital literacy was measured using Hargittai and Hsieh [4]'s succinct measure of web-use skills.

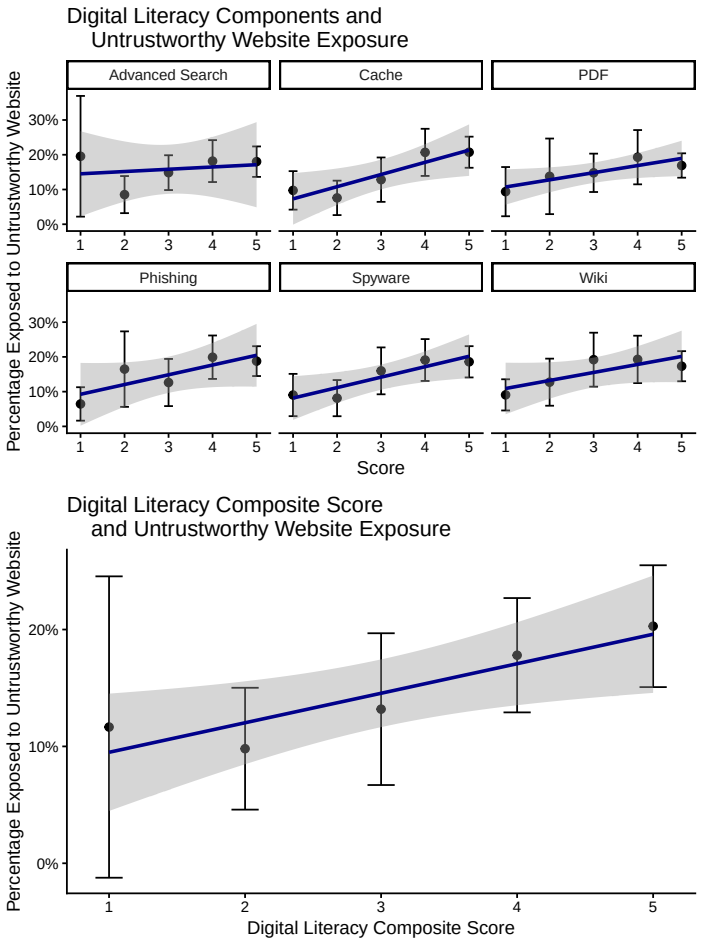


Figure S9: Digital Literacy Components (General Untrustworthy Exposure).

Regression analyses (Table S11) revealed significant positive associations between exposure to untrustworthy websites and higher scores in Spyware ($\beta = 0.026$, $P < 0.01$), Cache ($\beta = 0.036$, $P < 0.001$), Phishing ($\beta = 0.027$, $P < 0.01$), and Wiki ($\beta = 0.018$, $P < 0.05$). Directional, but not significant, associations were found for Advanced Search and PDF.

Table S11: Digital Literacy Predicting General Untrustworthy Exposure.

	Dependent variable:							
	Untrustworthy Exposure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Term: Adv. Search	0.013							
	(0.009)							
Term: PDF		0.015						
		(0.010)						
Term: Spyware			0.026**					
			(0.009)					
Term: Wiki				0.018*				
				(0.008)				
Term: Cache					0.036***			
					(0.009)			
Term: Phishing						0.027**		
						(0.009)		
Digital Term Knowledge (Composite)							0.035**	0.027*
							(0.011)	(0.012)
Republican								0.105***
								(0.033)
Political Interest								0.041**
								(0.013)
College								0.011
								(0.024)
Female								0.011
								(0.023)
Non-White								−0.045
								(0.026)
Age 30–44								−0.052
								(0.038)
Age 45–64								0.025
								(0.036)
Age 65+								0.050
								(0.039)
Constant	0.115**	0.105*	0.066	0.100**	0.037	0.063	0.034	−0.133*
	(0.037)	(0.041)	(0.036)	(0.031)	(0.033)	(0.035)	(0.043)	(0.063)
Observations	1,069	1,069	1,069	1,069	1,069	1,069	1,069	1,050
R^2	0.002	0.002	0.008	0.005	0.015	0.009	0.009	0.097
Adjusted R^2	0.001	0.001	0.007	0.004	0.014	0.008	0.008	0.090

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Moreover, the aggregated Digital Term Knowledge Composite Score positively predicted exposure ($\beta = 0.035$, $P < 0.01$), and this relationship remained significant, though slightly attenuated, after adding demographic and political controls ($\beta = 0.027$, $P < 0.05$). Although initially counterintuitive, these findings provide an important substantive insight into the relationship between digital term knowledge and exposure to misinformation. Specifically, higher digital term knowledge does not necessarily shield individuals from encountering untrustworthy websites; instead, those who are more knowledgeable about digital terms, particularly those familiar with concepts like cache, phishing, and spyware, may browse a broader range of online sources, inadvertently increasing their chances of encountering untrustworthy content. Alternatively, digitally knowledgeable individuals might feel overly confident in their abilities, leading them to engage more

freely with questionable sources. Overall, these results suggest that efforts to enhance digital term knowledge should be accompanied by strategies emphasizing cautious navigation online, highlighting the nuanced relationship between digital skills and exposure to misinformation.

S10.2 Digital Literacy and AI Untrustworthy Exposure

We next examined whether digital term knowledge was associated with Americans' exposure to AI-generated misinformation websites during the 2024 presidential election. Digital term knowledge was assessed across six technical terms (Advanced Search, PDF, Spyware, Wiki, Cache, and Phishing) as well as through an overall composite score (see Figure S10).

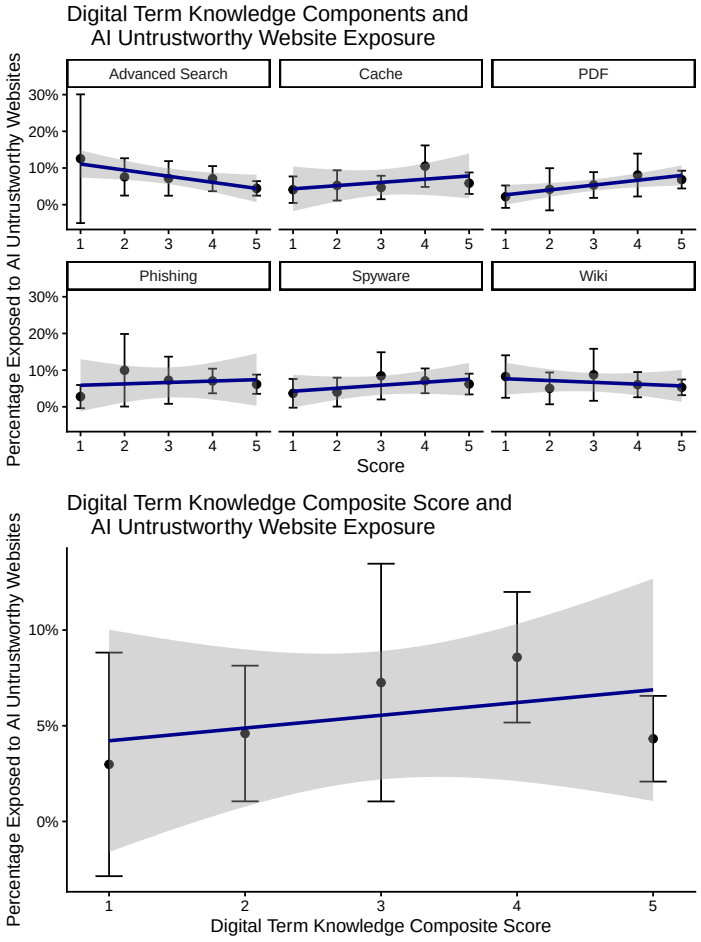


Figure S10: Digital Literacy Components (AI Exposure).

Regression results (Table S12) revealed minimal relationships between digital term

knowledge and exposure to AI-generated misinformation. Specifically, Advanced Search was the only dimension significantly associated with exposure, showing a modest protective effect; participants who reported higher knowledge of advanced search techniques were less likely to encounter AI-generated misinformation (β = -0.015, $P < 0.05$). Other digital term knowledge dimensions—such as PDF, Spyware, Wiki, Cache, and Phishing – showed no statistically significant associations. Furthermore, the Digital Term Knowledge Composite Score also did not significantly predict exposure to AI-generated misinformation, neither when tested alone (β = -0.0002) nor when controlling for demographic and political characteristics ($\beta = 0.002$).

Table S12: Digital Literacy Predicting AI Untrustworthy Exposure.

	Dependent variable:							
	AI-generated untrustworthy exposure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Term: Adv. Search	−0.015*							
	(0.006)							
Term: PDF		0.010						
		(0.006)						
Term: Spyware			0.004					
			(0.006)					
Term: Wiki				−0.007				
				(0.005)				
Term: Cache					0.007			
					(0.006)			
Term: Phishing						0.002		
						(0.006)		
Digital Term Knowledge (Composite)							−0.0002	0.002
							(0.007)	(0.008)
Republican								−0.002
								(0.017)
Political Interest								−0.020*
								(0.009)
College								0.003
								(0.016)
Female								−0.009
								(0.016)
Non-White								−0.067***
								(0.018)
Age 30–44								−0.062*
								(0.026)
Age 45–64								−0.032
								(0.025)
Age 65+								−0.047
								(0.027)
Constant	0.122***	0.025	0.051*	0.090***	0.040	0.058*	0.066*	0.182***
	(0.025)	(0.027)	(0.024)	(0.021)	(0.022)	(0.023)	(0.029)	(0.043)
Observations	1,069	1,069	1,069	1,069	1,069	1,069	1,069	1,050
R^2	0.005	0.002	0.000	0.002	0.001	0.000	0.000	0.023
Adjusted R^2	0.004	0.001	−0.001	0.001	0.0004	−0.001	−0.001	0.014

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Substantively, these findings indicate that traditional digital term knowledge has lim-

ited predictive power regarding exposure to AI-generated misinformation. Unlike more general forms of misinformation, where digital term knowledge correlated positively with exposure, AI-generated misinformation exposure seems less connected to an individual's knowledge of digital terminology. The notable exception regarding advanced search knowledge suggests that individuals with better search skills might be more effective at avoiding AI-generated misinformation sources. Overall, the results suggest that AI-generated misinformation may propagate differently, perhaps targeting or reaching individuals irrespective of their general digital term knowledge. Consequently, addressing AI-generated misinformation may require targeted approaches beyond conventional digital literacy training.

S11 AI Skills

We examined whether Americans' self-reported practical abilities with AI technologies relate to their exposure to untrustworthy content online. AI skills were measured using an abbreviated version of Wang et al. [5]'s measure of AI literacy.

S11.1 AI Skills and General Untrustworthy Exposure

Next, we explored whether Americans' AI skills were associated with their exposure to untrustworthy websites during the 2024 presidential election. AI skills comprised six distinct practical abilities: Skill: Distinguish Devices, Skill: Identify AI Tech, Skill: Use AI for Work, Skill: Improve Efficiency w/ AI, Skill: Choose AI Solution, and Skill: Choose AI App. We also assessed AI skills using a composite score aggregating these individual skill measures (see Figure S11).

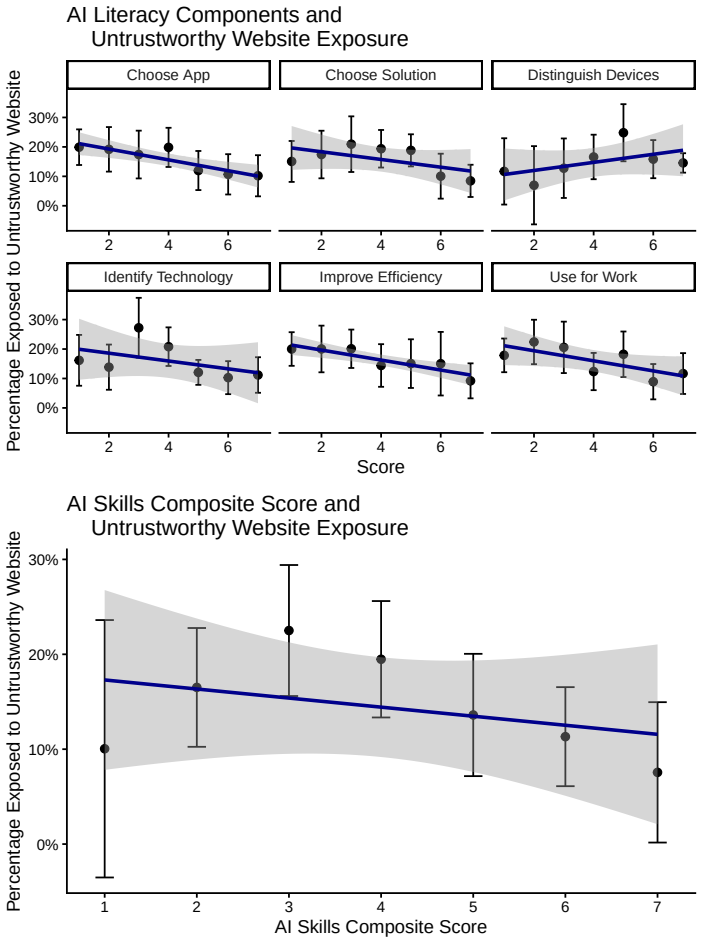


Figure S11: AI skills components and general untrustworthy website exposure.

Regression analyses (Table S13) indicated clear negative relationships between several AI skills and exposure to untrustworthy websites. Specifically, individuals with greater skills in Identify AI Tech ($\beta = -0.018$, $P < 0.01$), Use AI for Work ($\beta = -0.015$, $P < 0.05$), Improve Efficiency w/ AI ($\beta = -0.017$, $P < 0.01$), Choose AI Solution ($\beta = -0.014$, $P < 0.05$), and Choose AI App ($\beta = -0.018$, $P < 0.01$) all experienced significantly lower levels of misinformation exposure. In contrast, Distinguish Devices was not significantly associated with exposure. The aggregated AI Skills Composite Score similarly showed a significant negative relationship initially ($\beta = -0.022$, $p < 0.01$), but this relationship became non-significant when controlling for demographic and political variables ($\beta = -0.010$, $p = 0.264$).

Table S13: AI Skills Predicting General Untrustworthy Exposure.

	Dependent variable:							
	Untrustworthy Exposure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skill: Distinguish Devices	−0.004							
	(0.008)							
Skill: Identify AI Tech		−0.018**						
		(0.007)						
Skill: Use AI for Work			−0.015*					
			(0.006)					
Skill: Improve Efficiency w/ AI				−0.017**				
				(0.006)				
Skill: Choose AI Solution					−0.014*			
					(0.007)			
Skill: Choose AI App						−0.018**		
						(0.006)		
AI Skill Score (Composite)							−0.022**	−0.010
							(0.008)	(0.009)
Republican								0.105***
								(0.033)
Democrat								−0.068**
								(0.034)
Political Interest								0.058***
								(0.012)
College								0.020
								(0.023)
Female								0.006
								(0.023)
Non-White								−0.039
								(0.027)
Age 30–44								−0.053
								(0.038)
Age 45–64								0.013
								(0.036)
Age 65+								0.024
								(0.040)
Constant	0.186***	0.244***	0.220***	0.228***	0.224***	0.230***	0.258***	0.016
	(0.045)	(0.031)	(0.025)	(0.025)	(0.030)	(0.026)	(0.035)	(0.063)
Observations	1,069	1,069	1,069	1,069	1,069	1,069	1,069	1,050
R^2	0.0002	0.007	0.006	0.008	0.004	0.007	0.007	0.098
Adjusted R^2	−0.001	0.006	0.005	0.007	0.003	0.006	0.006	0.089

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings offer important substantive implications for misinformation research and interventions. Higher self-reported skill with AI, particularly in applying it to specific tasks, was associated with lower exposure to untrustworthy content. For example, those who reported being able to skillfully use AI for work or choose the right AI solution for a task had significantly lower exposure. This suggests that practical, hands-on experience with AI may equip individuals to better evaluate or avoid untrustworthy content online, perhaps by fostering a more critical stance. The attenuation of the composite AI skill score effect after controlling for demographics and political attitudes further suggests these skills intersect with broader sociopolitical characteristics linked to misinformation vulnerability.

S11.2 AI Skills and AI Untrustworthy Exposure

We further examined whether Americans' AI skills were specifically associated with their exposure to AI-generated misinformation websites during the 2024 presidential election. AI skills encompassed six separate practical abilities (Skill: Distinguish Devices, Skill: Identify AI Tech, Skill: Use AI for Work, Skill: Improve Efficiency w/ AI, Skill: Choose AI Solution, and Skill: Choose AI App) and was additionally assessed through an aggregate composite score (see Figure S12).

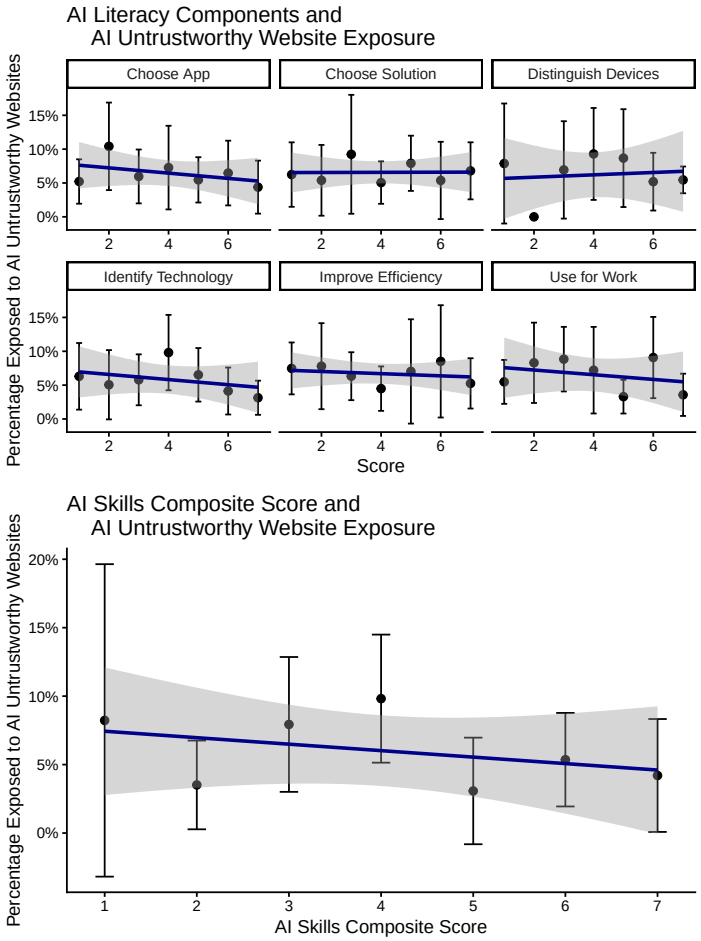


Figure S12: AI skills components and AI-generated untrustworthy website exposure.

Regression analyses (Table S14) revealed that none of the AI skills individually predicted significant differences in exposure to AI-generated misinformation, though each showed small negative relationships. The aggregated AI Skill Score Composite also failed to significantly predict exposure, either in initial models or after controlling for demographic and political variables.

Table S14: AI Skills Predicting AI Untrustworthy Exposure.

	Dependent variable:							
	AI-generated untrustworthy exposure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skill: Distinguish Devices	−0.007							
	(0.005)							
Skill: Identify AI Tech		−0.005						
		(0.005)						
Skill: Use AI for Work			−0.004					
			(0.004)					
Skill: Improve Efficiency w/ AI				−0.002				
				(0.004)				
Skill: Choose AI Solution					−0.0001			
					(0.004)			
Skill: Choose AI App						−0.003		
						(0.004)		
AI Skill Score (Composite)							−0.005	−0.001
							(0.005)	(0.006)
Republican								−0.002
								(0.017)
Political Interest								−0.019**
								(0.008)
College								0.004
								(0.016)
Female								−0.010
								(0.016)
Non-White								−0.067***
								(0.018)
Age 30–44								−0.062**
								(0.026)
Age 45–64								−0.033
								(0.025)
Age 65+								−0.050
								(0.028)
Constant	0.104***	0.086***	0.078***	0.073***	0.066***	0.077***	0.086***	0.193***
	(0.030)	(0.021)	(0.017)	(0.017)	(0.020)	(0.017)	(0.023)	(0.043)
Observations	1,069	1,069	1,069	1,069	1,069	1,069	1,069	1,050
R^2	0.002	0.001	0.001	0.000	0.000	0.000	0.001	0.022
Adjusted R^2	0.001	0.0001	−0.0002	−0.001	−0.001	−0.0004	−0.001	0.014

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings offer meaningful insights within the broader context of misinformation research. Despite the protective role of AI skills against general misinformation exposure, these same skills did not significantly influence exposure to AI-generated misinformation websites. Given that AI-generated misinformation remains relatively uncommon, this lack of significant association may indicate that exposure is still largely random or indiscriminate. Alternatively, it may suggest that recognizing AI-generated misinformation requires different competencies beyond self-reported practical skills. Thus, educational interventions aiming to combat AI-generated misinformation specifically may need to move beyond general AI awareness and towards more targeted strategies.

S12 AI Knowledge

This section investigates how familiarity with specific AI technologies affects exposure to untrustworthy websites. Our questions gauging participants' knowledge of AI tools and technologies were based on Yan et al. [6]'s measures of generative AI knowledge but were expanded in scope to cover different aspects and applications of generative AI.

S12.1 AI Knowledge and General Untrustworthy Exposure

We next examined whether Americans' familiarity with specific AI technologies—such as GPT-4, Large Language Models (LLMs), Image Generation tools, and Speech Recognition software—was associated with their exposure to untrustworthy websites during the 2024 presidential election. Specifically, we asked participants to identify the latest publicly available version of ChatGPT, to define what the acronym "LLM" stands for (large language model), to identify what Midjourney is known to be used for (image generation), and to identify what the primary function of OpenAI's Whisper model is (speech recognition). Additionally, we assessed an aggregate AI Knowledge Composite, combining familiarity scores across these different AI technologies (see Figure S13).

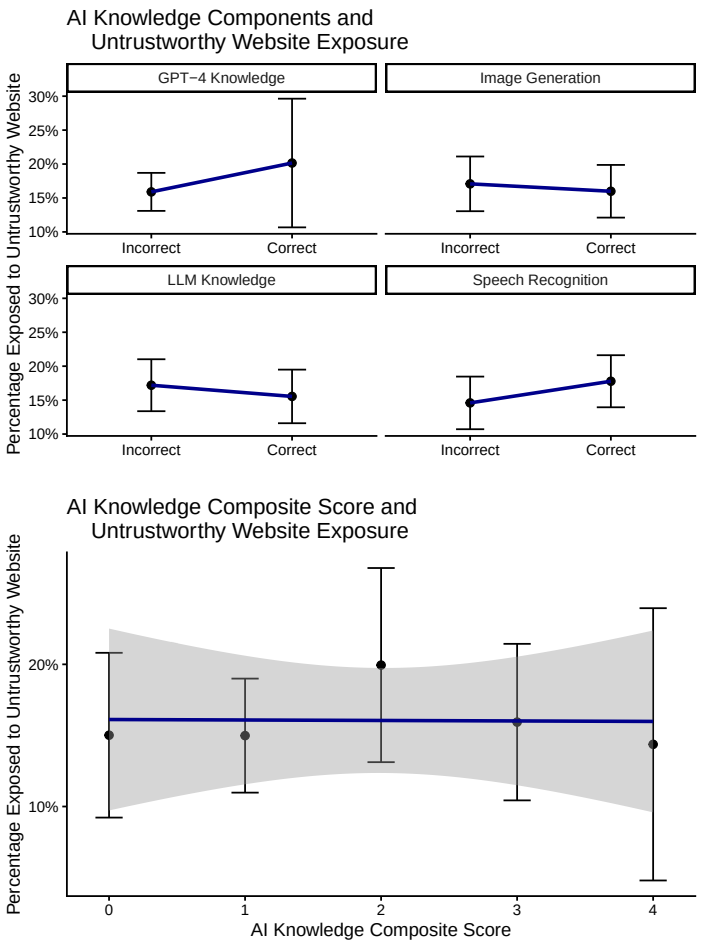


Figure S13: AI Knowledge Components (General Exposure).

Regression analyses (Table S15) indicated that none of the specific AI technology familiarity measures significantly predicted misinformation exposure. While familiarity with GPT-4 and Speech Recognition showed small positive associations with untrustworthy website exposure, familiarity with LLMs and Image Generation had small negative relationships; however, none of these individual associations reached statistical significance. Similarly, the aggregated AI Knowledge Composite measure was not significantly associated with exposure, either independently or when demographic and political controls were considered.

Table S15: AI Knowledge Predicting General Untrustworthy Exposure.

	Dependent variable:					
	Untrustworthy Exposure					
	(1)	(2)	(3)	(4)	(5)	(6)
GPT-4 Knowledge	0.043					
	(0.031)					
LLM Knowledge		-0.016				
		(0.023)				
Image Generation			-0.011			
			(0.023)			
Speech Recognition				0.032		
				(0.023)		
AI Knowledge (Composite)					0.006	0.013
					(0.011)	(0.011)
Republican						0.105***
						(0.033)
Democrat						-0.068**
						(0.034)
Political Interest						0.058***
						(0.012)
College						0.013
						(0.024)
Female						0.010
						(0.023)
Non-White						-0.046
						(0.026)
Age 30–44						-0.048
						(0.038)
Age 45–64						0.027
						(0.036)
Age 65+						0.047
						(0.039)
Constant	0.159***	0.172***	0.171***	0.146***	0.157***	-0.075
	(0.012)	(0.015)	(0.016)	(0.018)	(0.021)	(0.055)
Observations	1,068	1,069	1,069	1,069	1,068	1,049
R^2	0.002	0.0005	0.0002	0.002	0.0003	0.098
Adjusted R^2	0.001	-0.0005	-0.001	0.001	-0.001	0.089

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

These findings provide valuable insights into the role of technological awareness in susceptibility to misinformation. Specifically, mere familiarity with prominent AI technologies like GPT-4 or generative image models does not appear sufficient, by itself, to meaningfully influence an individual's likelihood of encountering misinformation online. This suggests that possessing baseline knowledge about AI tools may not directly confer protective or risk-enhancing effects regarding misinformation exposure. It may also reflect that misinformation exposure occurs through channels and mechanisms largely independent of individuals' specific technological knowledge. Consequently, effective interventions to mitigate misinformation exposure likely require more than just increasing public awareness about specific AI technologies; they must encourage deeper critical reasoning and evaluative skills tailored explicitly toward digital content credibility and authenticity.

S12.2 AI Knowledge and AI Untrustworthy Exposure

Next, we examined whether Americans' familiarity with specific AI technologies—GPT-4, Large Language Models (LLMs), Image Generation tools, and Speech Recognition software—was related specifically to their exposure to AI-generated misinformation websites during the 2024 presidential election. Additionally, we assessed a composite measure of AI technology familiarity, combining responses across these individual technologies (see Figure S14).

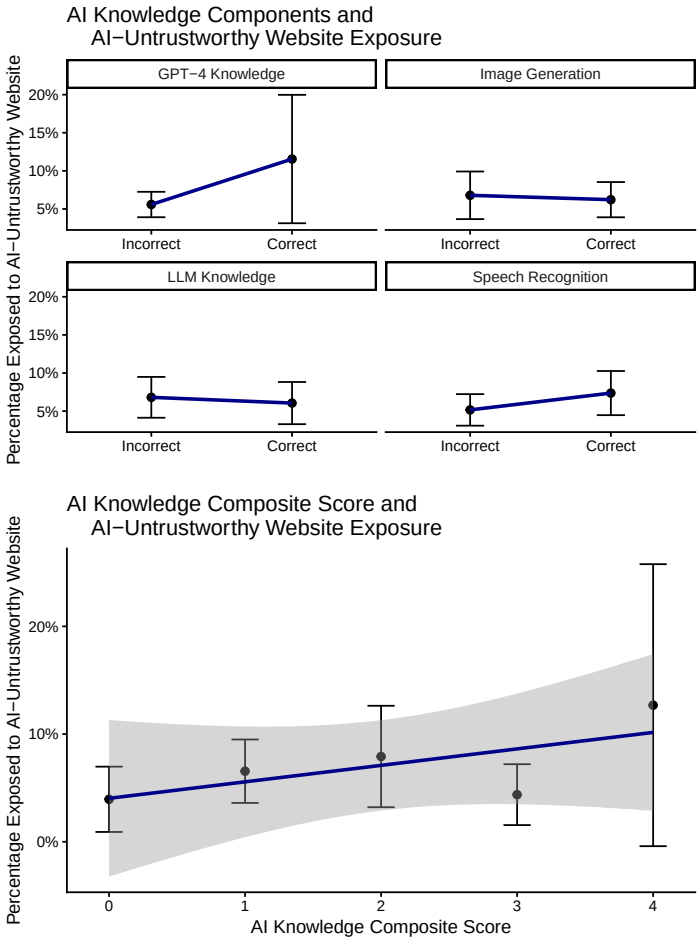


Figure S14: AI Knowledge Components (AI Exposure).

Regression analyses (Table S16) demonstrated a significant positive relationship between familiarity with GPT-4 and exposure to AI-generated misinformation, indicating that participants who were more familiar with GPT-4 experienced greater exposure to this type of untrustworthy content. Familiarity with other AI technologies—LLMs, Image Generation, and Speech Recognition—did not exhibit significant associations with AI

misinformation exposure. Similarly, the aggregate AI Knowledge Composite score was not significantly associated with exposure, either in initial analyses or after controlling for demographic and political characteristics.

Table S16: AI Knowledge Predicting AI Untrustworthy Exposure.

	Dependent variable:					
	AI Untrustworthy Exposure					
	(1)	(2)	(3)	(4)	(5)	(6)
GPT-4 Knowledge	0.060**					
	(0.021)					
LLM Knowledge		-0.008				
		(0.015)				
Image Generation			-0.006			
			(0.015)			
Speech Recognition				0.022		
				(0.016)		
AI Knowledge (Composite)					0.008	0.012
					(0.007)	(0.008)
Republican						-0.0003
						(0.017)
Political Interest						-0.021*
						(0.008)
College						0.0004
						(0.016)
Female						-0.007
						(0.016)
Non-White						-0.068***
						(0.018)
Age 30–44						-0.058*
						(0.026)
Age 45–64						-0.025
						(0.027)
Age 65+						-0.040
						(0.027)
Constant	0.056***	0.068***	0.068***	0.052***	0.051***	0.171***
	(0.008)	(0.010)	(0.011)	(0.012)	(0.014)	(0.038)
Observations	1,068	1,069	1,069	1,069	1,068	1,049
R^2	0.008	0.0002	0.0001	0.002	0.001	0.025
Adjusted R^2	0.007	-0.001	-0.001	0.001	0.0004	0.016

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

These findings provide valuable insights into the complexity of misinformation exposure in an AI-driven information environment. The specific association between GPT-4 familiarity and increased AI-generated misinformation exposure might suggest that individuals who know more about GPT-4 actively interact with content related to advanced generative AI, inadvertently increasing their exposure risk. It could also reflect greater confidence among GPT-4-aware individuals in their capacity to differentiate genuine from artificial content, causing them to engage more readily with sources that turn out to be untrustworthy. The absence of similar effects for other AI technologies emphasizes GPT-4's distinctive public prominence and highlights how targeted familiarity with particular technologies might inadvertently amplify exposure to misinformation. This finding

underscores that addressing AI-generated misinformation likely requires nuanced strategies that consider how familiarity with prominent AI technologies can both empower and inadvertently expose individuals to new misinformation risks.

S13 AI Usage

We analyzed whether individuals' usage patterns of AI tools influence their exposure to untrustworthy content.

S13.1 AI Usage and General Untrustworthy Exposure

We next investigated whether Americans' use of AI tools was associated with exposure to untrustworthy websites during the 2024 presidential election. AI usage was assessed through a single measure capturing frequency of AI tool usage across various applications (7-point scale ranging in frequency from never to multiple times a day) (see Figure S15).

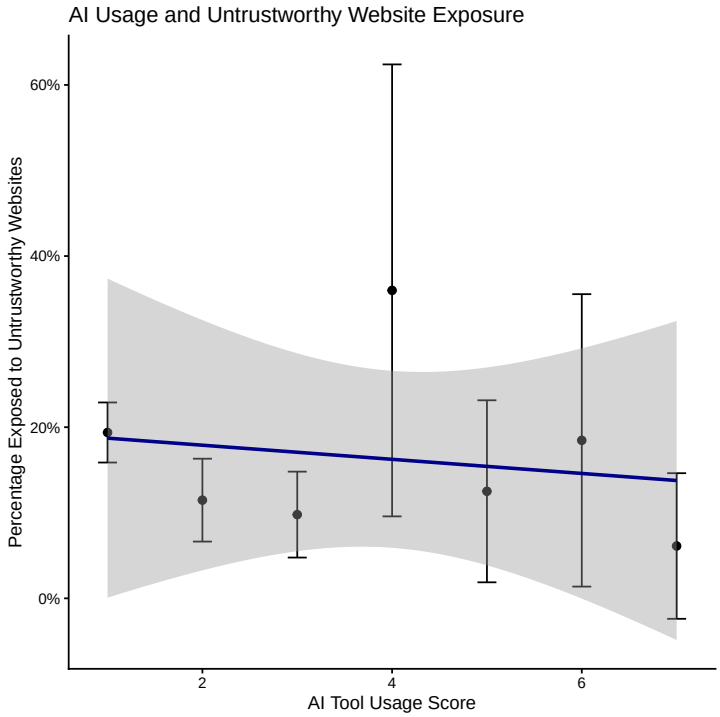


Figure S15: AI Usage/Attitude Components (General Exposure).

Regression analyses (Table S17) revealed that AI Tool Usage was not significantly associated with exposure to untrustworthy websites ($\beta = -0.012$, $P = .099$). This suggests that the frequency of AI tool usage does not meaningfully predict an individual's likelihood of encountering untrustworthy content online.

Table S17: AI Usage Predicting General Untrustworthy Exposure.

	Dependent variable:	
	Untrustworthy Exposure	
	(1)	(2)
AI Tool Usage	−0.012	−0.002
	(0.007)	(0.008)
Republican		0.105***
		(0.033)
Democrat		−0.068**
		(0.034)
Political Interest		0.058***
		(0.012)
College		0.014
		(0.024)
Female		0.009
		(0.023)
Non-White		−0.045
		(0.026)
Age 30–44		−0.049
		(0.038)
Age 45–64		0.027
		(0.036)
Age 65+		0.048
		(0.039)
Constant	0.192***	−0.076
	(0.019)	(0.055)
Observations	1,069	1,050
R^2	0.003	0.094
Adjusted R^2	0.002	0.086

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings provide substantive insights into how AI tool usage relates to misinformation exposure. The absence of a significant relationship highlights that merely interacting with AI technologies does not confer inherent advantages or disadvantages regarding misinformation exposure. This suggests that interventions designed to mitigate misinformation exposure might benefit from focusing on critical evaluation skills rather than simply increasing AI tool familiarity.

S13.2 AI Usage and AI Untrustworthy Exposure

We next investigated whether Americans' usage of AI tools was specifically associated with exposure to AI-generated misinformation websites during the 2024 presidential election (see Figure S16).

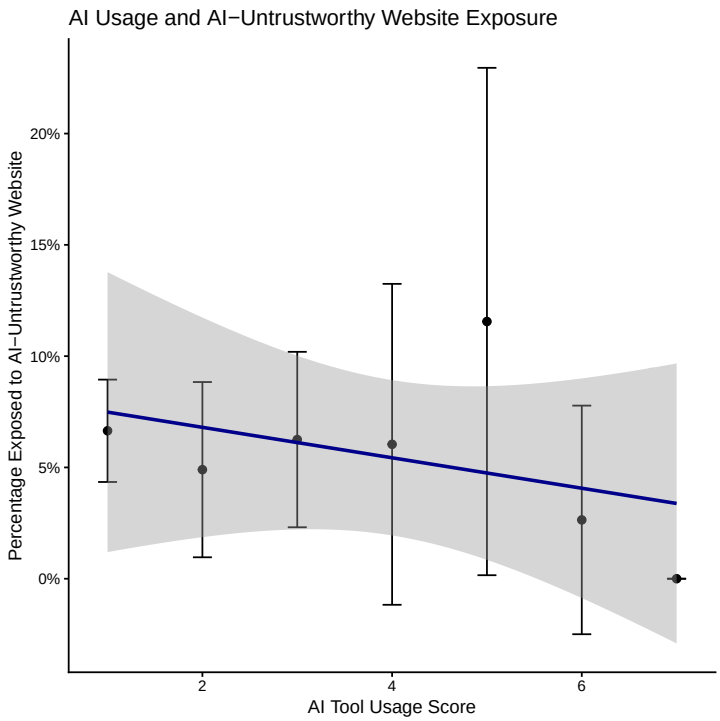


Figure S16: AI Usage/Attitude Components (AI Exposure).

Regression analyses (Table S18) indicated that AI Tool Usage was not significantly associated with exposure to AI-generated misinformation websites ($\beta = -0.005$, $P = .474$). This relationship remained non-significant even after controlling for demographic and political characteristics.

Table S18: AI Usage Predicting AI-Generated Untrustworthy Exposure.

	Dependent variable:	
	AI-generated untrustworthy exposure	
	(1)	(2)
AI Tool Usage	−0.005	−0.001
	(0.005)	(0.005)
Republican		−0.001
		(0.017)
Political Interest		−0.020*
		(0.008)
College		0.001
		(0.016)
Female		−0.008
		(0.016)
Non-White		−0.067***
		(0.018)
Age 30–44		−0.061*
		(0.026)
Age 45–64		−0.031
		(0.025)
Age 65+		−0.048
		(0.027)
Constant	0.076***	0.179***
	(0.013)	(0.038)
Observations	1,069	1,050
R^2	0.0005	0.022
Adjusted R^2	−0.0004	0.013

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Substantively, these results suggest that AI tool usage does not appear to vary based on exposure to AI-generated misinformation. This form of misinformation might still be too novel or infrequent to reflect clear relationships with individual-level AI usage patterns. Consequently, addressing the threat of AI-generated misinformation likely requires developing targeted strategies focused explicitly on the distinctive patterns through which such misinformation propagates, rather than relying solely on AI usage patterns.

S14 AI Attitudes

This section explores how Americans' attitudes toward artificial intelligence relate to their exposure to untrustworthy websites. We measured attitudes toward AI using French and Hancock [7]'s semantic differential approach to assessing attitudes toward technology. Seven attitude dimensions measured participants' perceptions of AI as: Effective vs. Ineffective, Useful vs. Useless, Informative vs. Uninformative, Interesting vs. Uninteresting, Important vs. Unimportant, Meaningful vs. Meaningless, and Powerful vs. Powerless.

S14.1 AI Attitudes and General Untrustworthy Exposure

We examined whether Americans' attitudes toward AI across seven dimensions, as well as a composite AI attitude score, were associated with exposure to untrustworthy websites during the 2024 presidential election. Additionally, we created a composite AI attitude score by averaging across all seven dimensions (see Figure S17 and Figure S18).

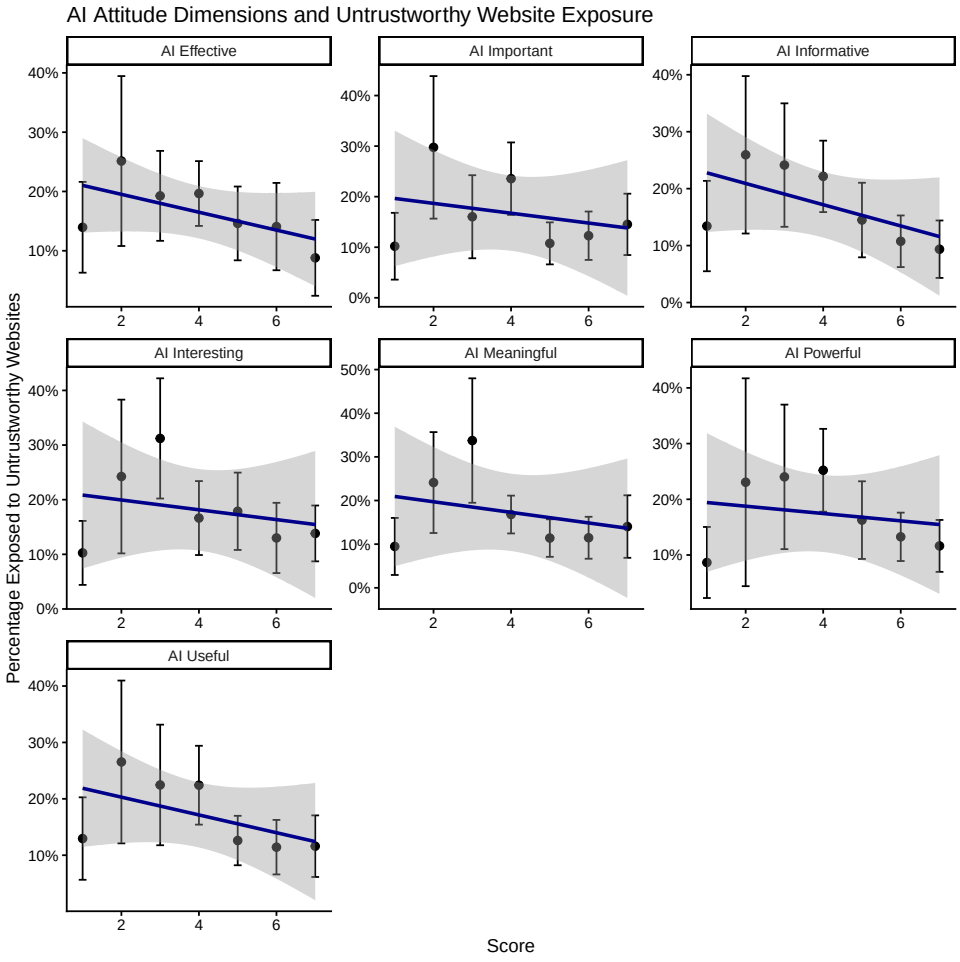


Figure S17: AI attitude dimensions and general untrustworthy website exposure.

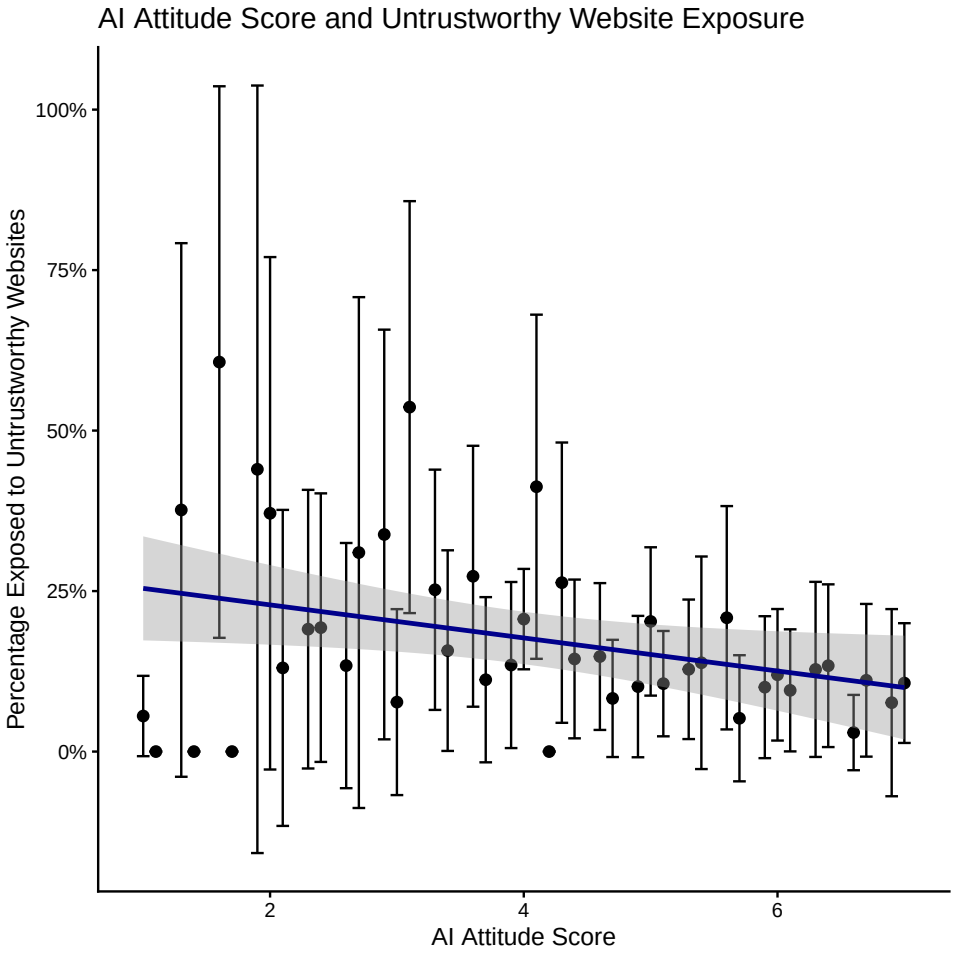


Figure S18: AI attitude score and general untrustworthy website exposure.

Regression analyses revealed significant associations between several AI attitude dimensions and exposure to untrustworthy websites. Specifically, more positive attitudes toward AI as Effective (β = -0.017, $P < 0.05$), Useful (β = -0.019, $P < 0.01$), and Informative (β = -0.023, $P < 0.001$) were significantly associated with lower exposure to untrustworthy websites. Additionally, more positive attitudes toward AI as Meaningful (β = -0.018, $P < 0.01$) and Powerful (β = -0.014, $P < 0.05$) were also protective against untrustworthy website exposure. The composite AI attitude score showed a similar protective pattern (β = -0.020, $P < 0.01$).

Table S19: AI Attitudes Predicting General Untrustworthy Website Exposure

	Dependent variable: Untrustworthy Exposure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Effective	−0.017*							
	(0.007)							
Useful		−0.019**						
		(0.007)						
Informative			−0.023***					
			(0.007)					
Interesting				−0.008				
				(0.006)				
Important					−0.012			
					(0.007)			
Meaningful						−0.018**		
						(0.007)		
Powerful							−0.014*	
							(0.007)	
AI Attitude Score								−0.020**
								(0.008)
Constant	0.238***	0.248***	0.263***	0.202***	0.218***	0.219***	0.185***	0.250***
	(0.029)	(0.030)	(0.030)	(0.027)	(0.026)	(0.026)	(0.025)	(0.025)
Observations	1,069	1,069	1,069	1,069	1,069	1,069	1,069	1,069
R^2	0.005	0.008	0.011	0.002	0.003	0.007	0.004	0.006
Adjusted R^2	0.004	0.007	0.010	0.001	0.002	0.006	0.003	0.005

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings suggest that more positive attitudes toward AI are generally protective against exposure to untrustworthy websites. The protective effects were strongest for attitudes related to AI's informational value (Informative, Useful) and effectiveness, suggesting that individuals who view AI as a valuable and reliable information tool may be more discerning in their online information consumption. The composite AI attitude score's significant negative association reinforces this pattern, indicating that overall positive AI attitudes are associated with reduced misinformation exposure.

S14.2 AI Attitudes and AI Untrustworthy Exposure

We also examined whether AI attitudes were specifically associated with exposure to AI-generated misinformation websites during the 2024 presidential election (see Figure S19 and Figure S20).

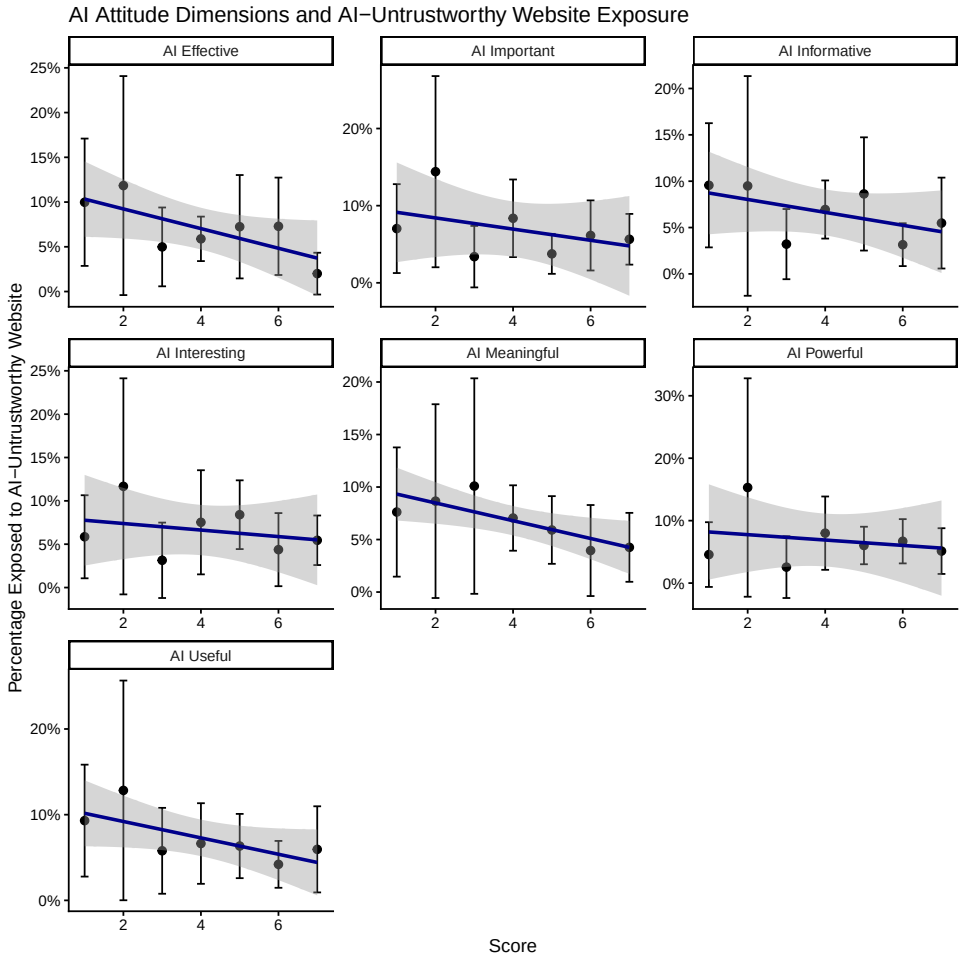


Figure S19: AI Attitude Dimensions and AI-Untrustworthy Website Exposure.

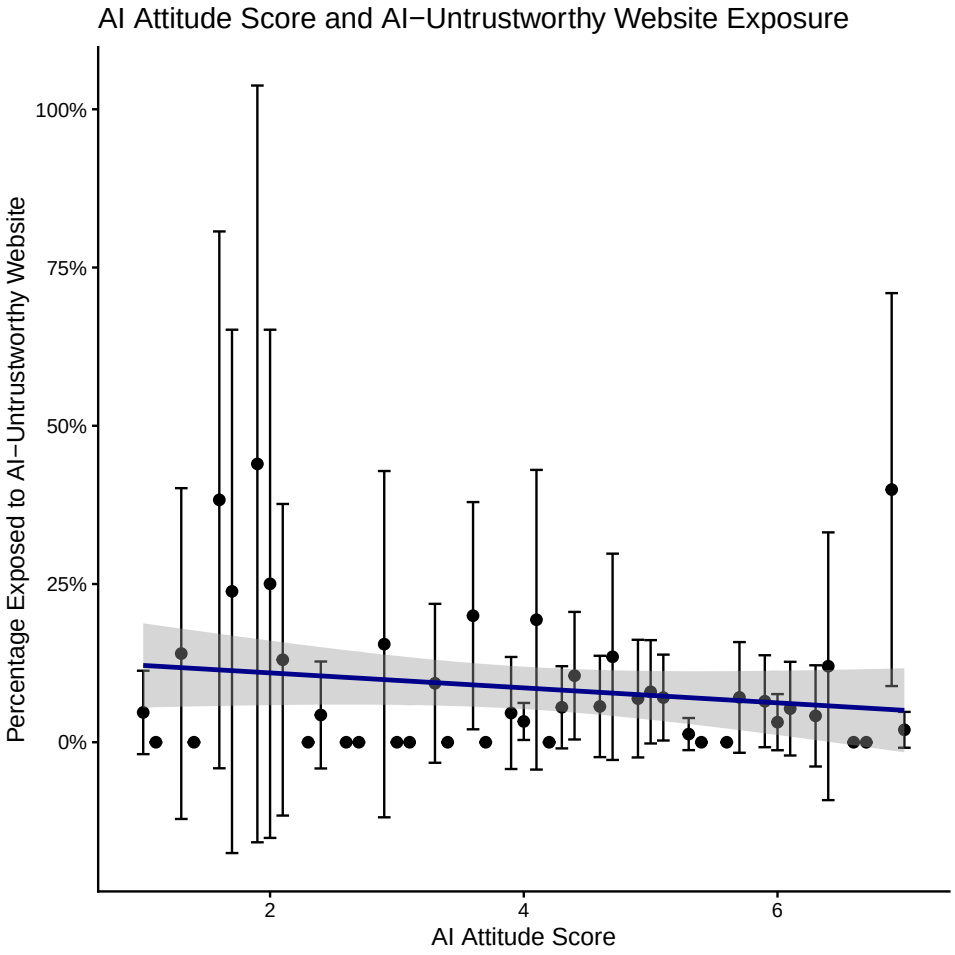


Figure S20: AI Attitude Score and AI-Untrustworthy Website Exposure.

Regression analyses revealed that none of the individual AI attitude dimensions significantly predicted exposure to AI-generated misinformation websites. Similarly, the composite AI attitude score was not significantly associated with AI-generated misinformation exposure.

Table S20: AI Attitudes Predicting AI-Generated Untrustworthy Website Exposure

	Dependent variable: AI-generated untrustworthy exposure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Effective	−0.003							
	(0.005)							
Useful		−0.007						
		(0.005)						
Informative			−0.008					
			(0.005)					
Interesting				−0.003				
				(0.004)				
Important					−0.005			
					(0.004)			
Meaningful						−0.006		
						(0.004)		
Powerful							−0.001	
							(0.004)	
AI Attitude Score								−0.006
								(0.004)
Constant	0.078***	0.095***	0.103***	0.076***	0.088***	0.090***	0.068***	0.091***
	(0.020)	(0.021)	(0.021)	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
Observations	1,069	1,069	1,069	1,069	1,069	1,069	1,069	1,069
R^2	0.0004	0.002	0.003	0.0005	0.001	0.002	0.00003	0.002
Adjusted R^2	−0.0005	0.001	0.002	−0.0004	0.0004	0.001	−0.001	0.001

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings indicate that while AI attitudes protect against general misinformation exposure, they do not significantly influence exposure to AI-generated misinformation specifically. This pattern suggests that AI-generated misinformation may operate through different mechanisms than traditional misinformation, potentially bypassing the protective effects of positive AI attitudes. This could reflect the novelty of AI-generated content or indicate that current AI attitudes, while protective against general misinformation, have not yet adapted to address the specific risks posed by AI-generated misinformation.

S15 AI Discernment

We investigated whether individuals' ability to distinguish AI-generated from human-generated content affects their exposure to untrustworthy websites. We assessed discernment by showing participants a series of images circulating on social media during the election campaign – both authentic and AI-generated – and asking them to identify whether images were either real or generated by AI. More details on the stimuli and task can be found in Yan et al. [8].

S15.1 AI Discernment and General Untrustworthy Exposure

We next examined whether individuals' ability to discern AI-generated content from human-generated content was associated with their exposure to untrustworthy websites during the 2024 presidential election. We computed participants' discernment abilities using accuracy measures for distinguishing AI-generated from real content overall (Total Accuracy), as well as separately for AI-generated content (AI Accuracy) and human-generated (Real Accuracy). Additionally, we explored discernment accuracy specifically for politically partisan content (Democratic-leaning and Republican-leaning) and analyzed interactions between participants' own partisan identities and their discernment skills (see Figure S21).

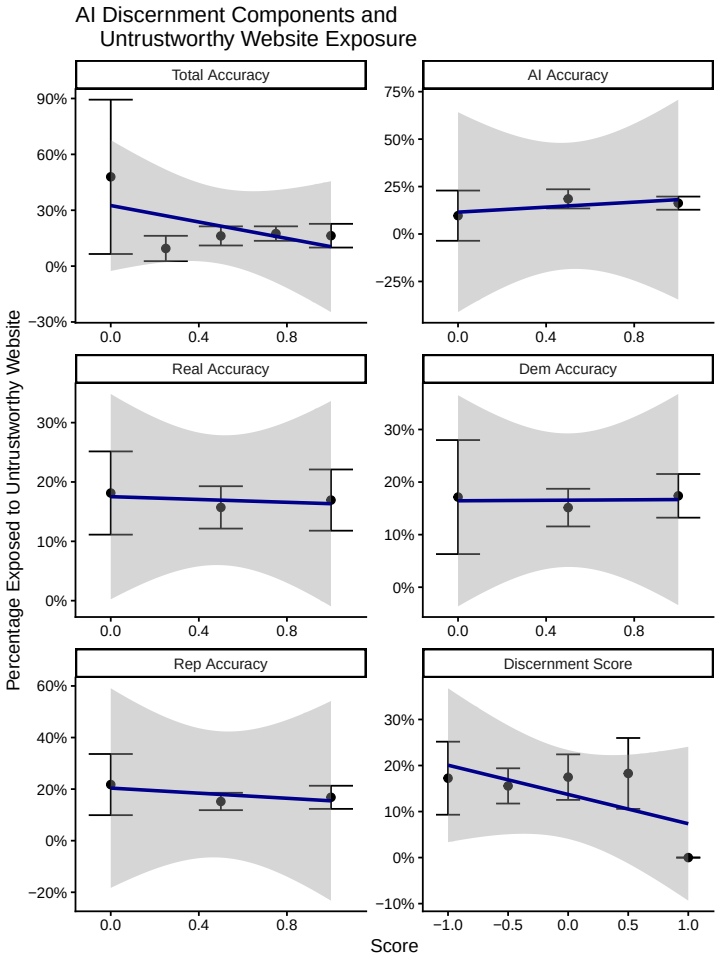


Figure S21: AI Discernment Components (General Exposure).

Regression analyses (Table S21) revealed no significant associations between general discernment abilities – measured by Total Accuracy, AI Accuracy, Real Accuracy, or an overall Discernment Score – and exposure to untrustworthy websites. However, nuanced partisan-specific analyses identified a significant interaction between participants' partisan identities and their accuracy discerning Democratic-oriented content. Specifically, Republicans who were better able to accurately identify Democratic-leaning content as AI-generated or human-generated experienced significantly higher exposure to untrustworthy websites relative to Republicans with lower accuracy. Such a relationship did not emerge for Democratic participants or for accuracy discerning Republican-oriented content.

Table S21: AI Discernment Predicting General Untrustworthy Exposure.

	Dependent variable:					
	Untrustworthy Exposure					
	(1)	(2)	(3)	(4)	(5)	(6)
Total Accuracy	0.003					
	(0.048)					
AI Accuracy		0.009				
		(0.039)				
Real Accuracy			−0.003			
			(0.033)			
Republican				0.034	0.185**	
				(0.074)	(0.068)	
Democrat				−0.113	0.051	
				(0.080)	(0.070)	
Dem Accuracy				−0.089		
				(0.079)		
Dem Accuracy × Republican				0.196*		
				(0.095)		
Dem Accuracy × Democrat				0.133		
				(0.100)		
Rep Accuracy					0.018	
					(0.078)	
Rep Accuracy × Republican					−0.016	
					(0.094)	
Rep Accuracy × Democrat					−0.095	
					(0.096)	
Discernment Score						−0.006
						(0.027)
Constant	0.163***	0.158***	0.167***	0.162**	0.086	0.164***
	(0.036)	(0.034)	(0.023)	(0.062)	(0.054)	(0.012)
Observations	1,069	1,069	1,069	1,069	1,069	1,069
R^2	0.000	0.000	0.000	0.066	0.062	0.000
Adjusted R^2	−0.001	−0.001	−0.001	0.061	0.058	−0.001

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Substantively, these results suggest a critical nuance in how discernment relates to misinformation exposure. Broadly accurate discernment skills in distinguishing AI-generated from human-generated content do not universally translate to reduced exposure. In-

stead, a counterintuitive pattern emerges for politically incongruent content. Republicans with higher discernment accuracy regarding Democratic-oriented content actually show increased exposure to untrustworthy websites. This suggests that skill in identifying politically oppositional AI-generated content may paradoxically be associated with greater engagement with potentially problematic online information sources. This could reflect that Republicans who are adept at discerning Democratic AI content are more likely to seek out and engage with diverse information sources, including untrustworthy ones, perhaps to critically evaluate opposing political content. These findings indicate that discernment skills alone may not be protective against misinformation exposure and may sometimes be associated with increased exposure to untrustworthy sources.

S15.2 AI Discernment and AI Untrustworthy Exposure

Next, we examined whether participants' discernment skills – defined as their ability to distinguish AI-generated from human-generated content – were associated with their exposure specifically to AI-generated misinformation websites during the 2024 presidential election. Discernment was assessed using measures of overall accuracy (Total Accuracy), accuracy specifically identifying AI-generated content (AI Accuracy), accuracy identifying human-generated content (Real Accuracy), and discernment accuracy scores for politically partisan AI-generated content (Democratic-leaning and Republican-leaning). We also considered interactions between these partisan discernment scores and participants' own partisan identities (see Figure S22).

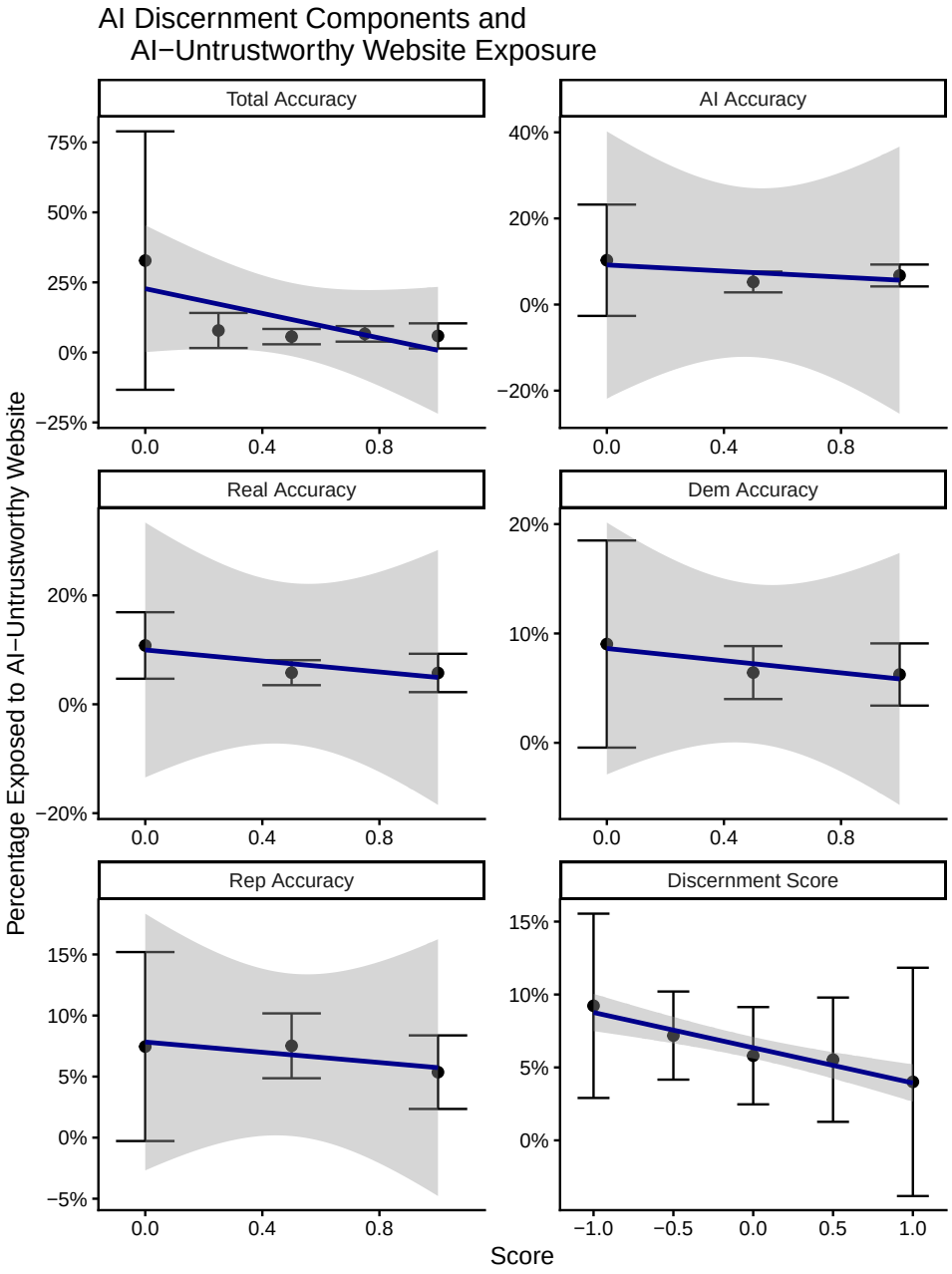


Figure S22: AI discernment and AI-generated untrustworthy website exposure.

Regression analyses (Table S22) indicated that general discernment measures – Total Accuracy, AI Accuracy, Real Accuracy, and an overall Discernment Score – were not significantly associated with exposure to AI-generated misinformation. However, a nuanced pattern emerged when examining discernment of politically partisan AI-generated content. Specifically, accuracy discerning Democratic-oriented AI content significantly interacted

with participants' partisan identity: Republicans who exhibited greater accuracy discerning Democratic-oriented AI-generated content had significantly higher exposure to AI-generated misinformation. Additionally, higher accuracy discerning Republican-oriented AI content was negatively associated with AI misinformation exposure overall; however, this relationship was moderated by partisan identity, such that Republicans who were particularly adept at discerning Republican-oriented AI-generated content experienced higher exposure compared to Republicans with lower discernment skills.

Table S22: AI Discernment Predicting AI Untrustworthy Exposure.

	Dependent variable:					
	AI-generated untrustworthy exposure					
	(1)	(2)	(3)	(4)	(5)	(6)
Total Accuracy	−0.042					
	(0.032)					
AI Accuracy		−0.002				
		(0.026)				
Real Accuracy			−0.039			
			(0.022)			
Republican				−0.108*	−0.102*	
				(0.051)	(0.046)	
Democrat				−0.023	−0.089	
				(0.055)	(0.047)	
Dem Accuracy				−0.081		
				(0.054)		
Dem Accuracy × Republican				0.147*		
				(0.065)		
Dem Accuracy × Democrat				0.003		
				(0.068)		
Rep Accuracy					−0.136*	
					(0.053)	
Rep Accuracy × Republican					0.158*	
					(0.064)	
Rep Accuracy × Democrat					0.105	
					(0.066)	
Discernment Score						−0.025
						(0.018)
Constant	0.095***	0.067**	0.089***	0.135**	0.160***	0.060***
	(0.024)	(0.022)	(0.016)	(0.042)	(0.037)	(0.008)
Observations	1,069	1,069	1,069	1,069	1,069	1,069
R^2	0.002	0.000	0.003	0.011	0.009	0.002
Adjusted R^2	0.001	−0.001	0.002	0.006	0.005	0.001

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings provide substantive insights into the complex role that discernment skills play in shaping exposure to AI-generated misinformation, particularly highlighting the influence of partisan dynamics. Broadly accurate discernment of AI-generated content

does not necessarily reduce misinformation exposure. Instead, discernment skills specifically targeted toward politically oppositional AI-generated content appear to increase exposure risk for Republicans. This suggests that Republicans who are skilled at recognizing politically challenging AI content may paradoxically engage more with potentially problematic online information sources, thereby increasing their exposure risk. This could reflect active information-seeking behavior where discerning individuals deliberately seek out diverse content sources, including untrustworthy ones, perhaps to monitor or critique opposing viewpoints. Conversely, general accuracy discerning politically congruent AI-generated content is also associated with increased exposure when partisan moderation is considered. Taken together, these findings indicate that discernment skills alone may not be protective against AI-generated misinformation exposure and highlight the complex interplay between political identity, discernment abilities, and information-seeking behavior.

S16 AI Self-Perceived Detection

This final section examines whether individuals' self-perceived ability to detect AI-generated content influences their exposure to untrustworthy websites.

S16.1 AI Self-Perceived Detection and General Untrustworthy Exposure

We examined whether Americans' perceptions of their own (self-assessed) and others' abilities to detect AI-generated content relate to exposure to untrustworthy websites during the 2024 election (Figure S23). We also analyze the AI Detection Gap (self minus others).

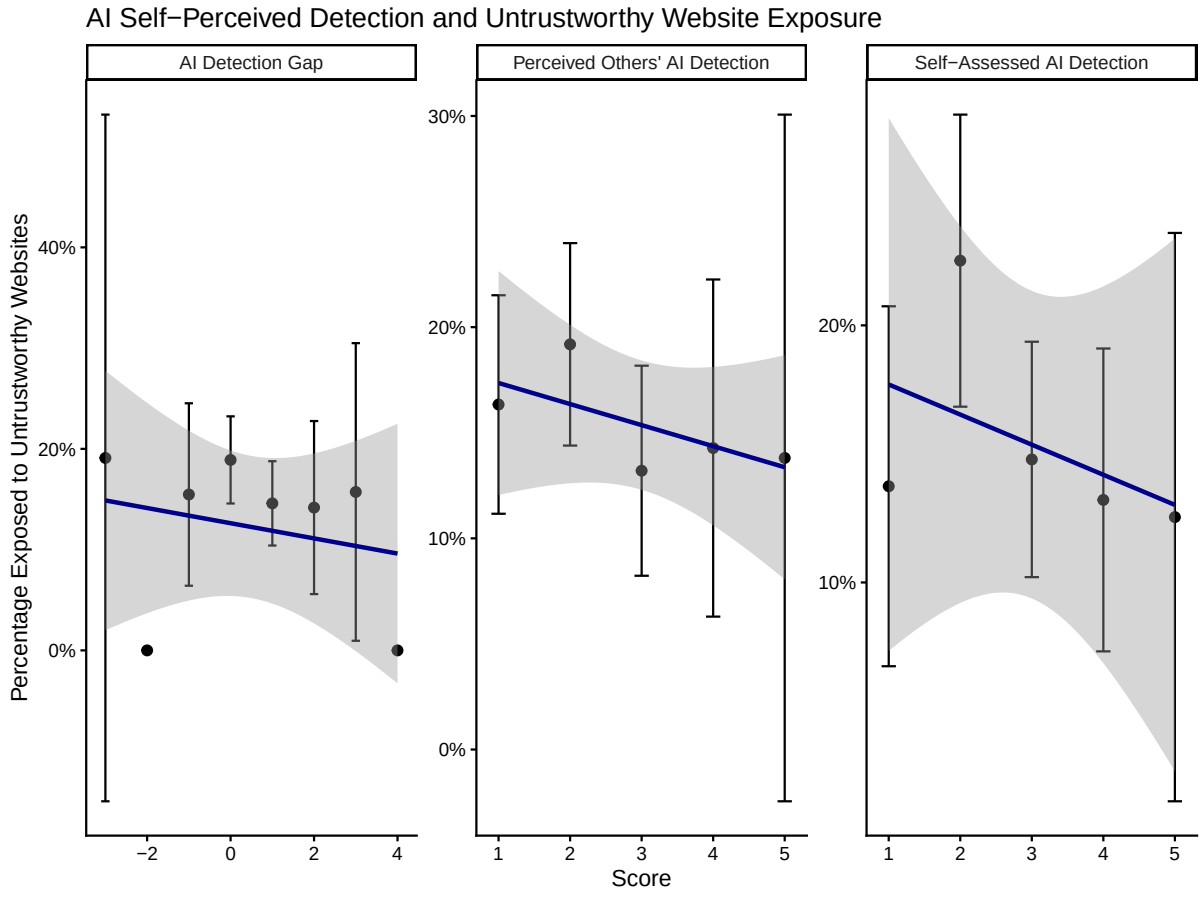


Figure S23: AI self-perceived detection and general untrustworthy website exposure.

Table S23: AI Self-Perceived Detection Predicting General Untrustworthy Exposure.

	(1)	(2)	(3)
Self-Assessed Detection	-0.022^{*}		
	(0.011)		
Others Detection		-0.016	
		(0.012)	
Detection Gap			-0.010
			(0.012)
Constant	0.226^{***}	0.201^{***}	0.171^{***}
	(0.033)	(0.030)	(0.013)
Observations	1,069	1,069	1,069
R^2	0.004	0.002	0.001

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Regression analyses reveal that self-assessed AI detection confidence provides modest protection against general untrustworthy website exposure ($\beta = -0.022$, $P = 0.049$), while perceptions of others' detection abilities and the detection gap show no significant associations. This suggests that personal confidence in one's ability to identify AI-generated content may encourage more cautious online behavior, though the protective effect is relatively small.

S16.2 AI Self-Perceived Detection and AI Untrustworthy Exposure

Figure S24 shows the relationships for exposure specifically to AI-generated misinformation.

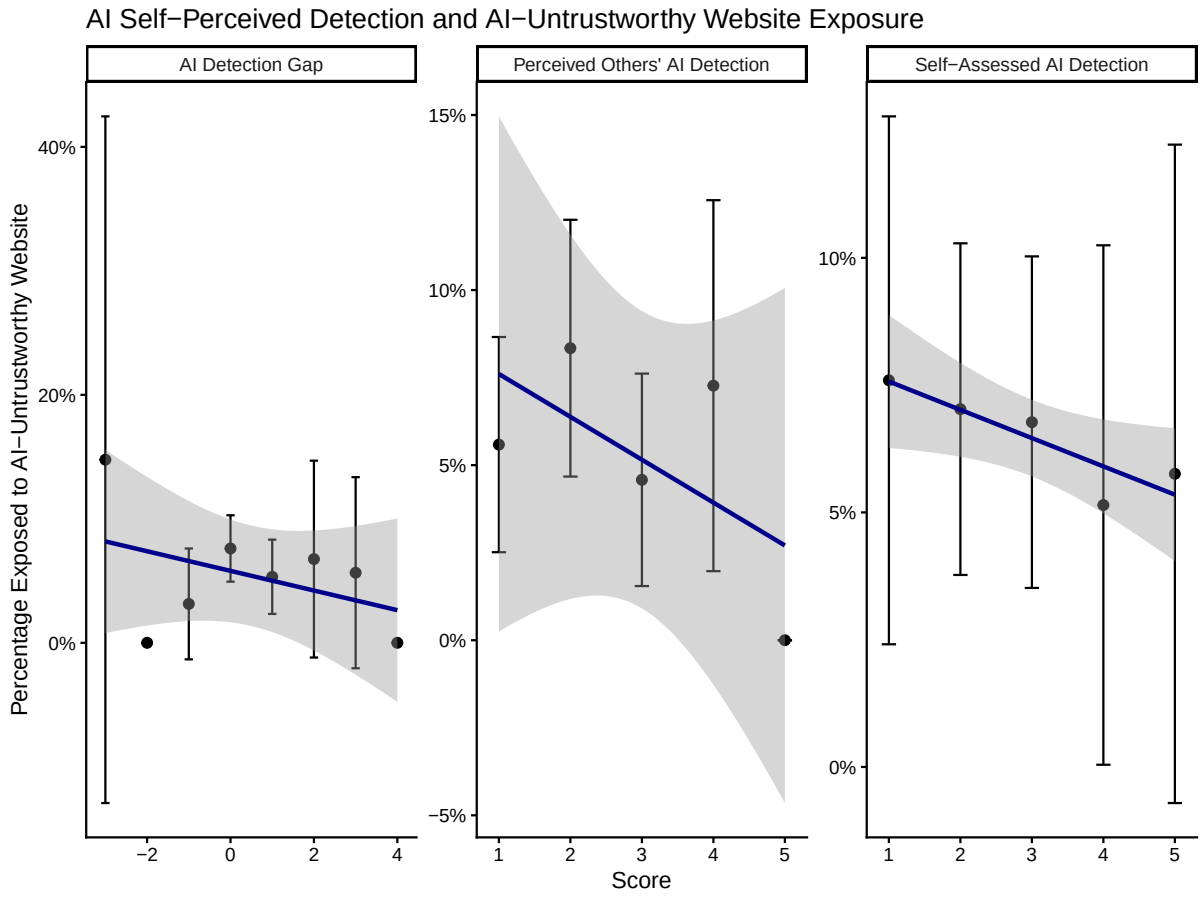


Figure S24: AI self-perceived detection and AI-generated untrustworthy website exposure.

Table S24: AI self-perceived detection predicting AI-generated untrustworthy exposure

	(1)	(2)	(3)
Self-Assessed Detection	−0.007		
	(0.007)		
Others Detection		−0.008	
		(0.008)	
Detection Gap			−0.001
			(0.008)
Constant	0.085***	0.082***	0.066***
	(0.022)	(0.020)	(0.009)
Observations	1,069	1,069	1,069
R^2	0.001	0.001	< 0.001

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

In contrast to general untrustworthy exposure, none of the self-perceived AI detection measures significantly predict exposure to AI-generated misinformation websites. This indicates that current self-assessments of AI detection abilities may not be well-calibrated to the specific challenges of identifying sophisticated AI-generated content, suggesting that as AI-generated misinformation becomes more prevalent and sophisticated, individuals' confidence in their detection abilities may need recalibration.

S17 AI Impact Perceptions

We examined whether individuals' perceptions about AI's impact on elections influence their exposure to untrustworthy websites. Participants rated their concerns about AI's impact on elections (AI Election Concern) via the question "How concerned are you about the potential for AI to affect the outcome of this election?" (answer options: Not at all concerned – Extremely concerned) (based on the questionnaire from Yan et al. [6]). Participants were also asked how much they believed AI would influence their own voting decisions (AI Influence on Self) via the question "To what extent do you believe AI-generated information has influenced your opinions in this election?", as well as others' voting decisions (AI Influence on Others) via the question "To what extent do you believe AI-generated information has influenced the opinions of the average voter in this election?" (answer options for both questions: Not at all – A great deal). Each item was measured on a 5-point scale.

S17.1 AI Impact Perceptions and General Untrustworthy Exposure

We investigated whether perceptions about AI's impact on elections were associated with exposure to untrustworthy websites during the 2024 presidential election. We examined three components measuring different aspects of AI impact perceptions, as well as a composite AI impact score (see Figure S25).

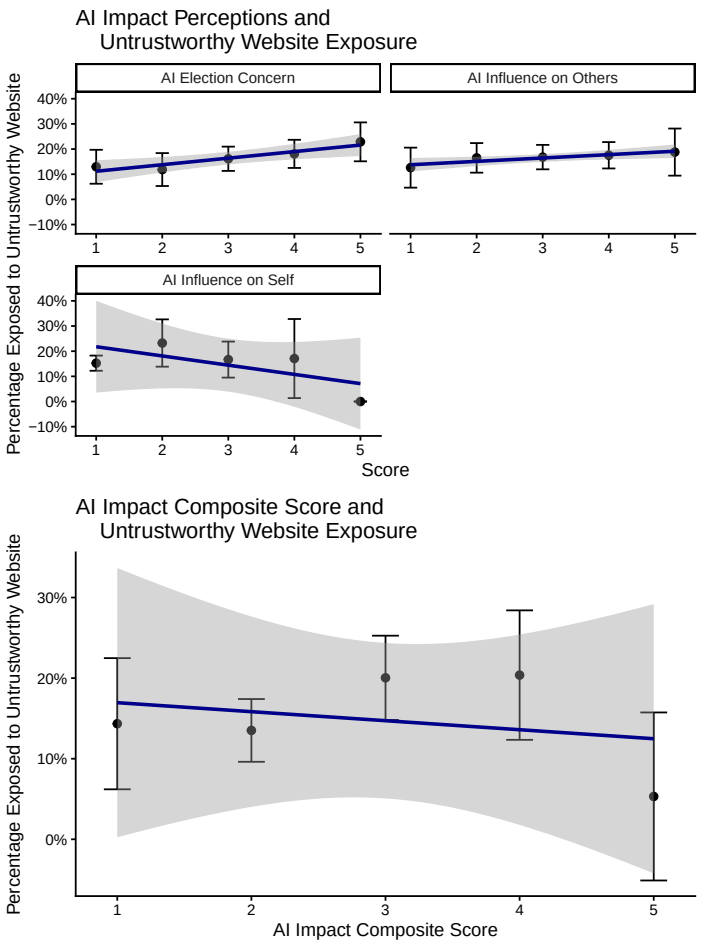


Figure S25: AI impact perceptions and general untrustworthy website exposure.

Regression analyses (Table S25) revealed mixed findings. Greater concern about AI's impact on elections was significantly associated with higher exposure to untrustworthy websites ($\beta = 0.027$, $P < 0.01$). However, beliefs about AI's influence on one's own voting ($\beta = 0.003$, $P = 0.819$) and on others' voting ($\beta = 0.012$, $P = 0.263$) were not significant predictors. The composite AI impact score initially showed a significant positive association ($\beta = 0.031$, $P < 0.05$), but this relationship became non-significant when controlling for demographic and political variables ($\beta = 0.011$, $P = 0.445$).

Table S25: AI Impact Perceptions Predicting General Untrustworthy Exposure.

	Dependent variable:				
	Untrustworthy Exposure				
	(1)	(2)	(3)	(4)	(5)
AI Election Concern	0.027**				
	(0.009)				
AI Influence on Self		0.003			
		(0.013)			
AI Influence on Others			0.012		
			(0.011)		
AI Impact (Composite)				0.031*	0.011
				(0.014)	(0.014)
Republican					0.151***
					(0.024)
Political Interest					0.049***
					(0.012)
College					0.017
					(0.024)
Female					0.006
					(0.023)
Non-White					−0.045
					(0.026)
Age 30–44					−0.054
					(0.038)
Age 45–64					0.019
					(0.036)
Age 65+					0.037
					(0.038)
Constant	0.081**	0.160***	0.131***	0.090*	−0.074
	(0.029)	(0.023)	(0.033)	(0.036)	(0.058)
Observations	1,069	1,069	1,069	1,069	1,050
R^2	0.009	0.0001	0.001	0.005	0.094
Adjusted R^2	0.008	−0.001	0.0002	0.004	0.086

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These findings reveal a nuanced relationship between AI impact perceptions and untrustworthy website exposure. Individuals with greater concern about AI's impact on elections showed higher exposure to untrustworthy content, suggesting that anxiety about

AI's electoral influence may correlate with broader patterns of online information consumption. However, beliefs about AI's direct influence on voting behavior (either one's own or others') did not predict exposure. The attenuation of the composite AI impact score effect after controlling for political and demographic characteristics suggests that concerns about AI intersect with broader sociopolitical factors linked to misinformation vulnerability.

S17.2 AI Impact Perceptions and AI Untrustworthy Exposure

We next examined whether AI impact perceptions were specifically associated with exposure to AI-generated misinformation websites during the 2024 presidential election (see Figure S26).

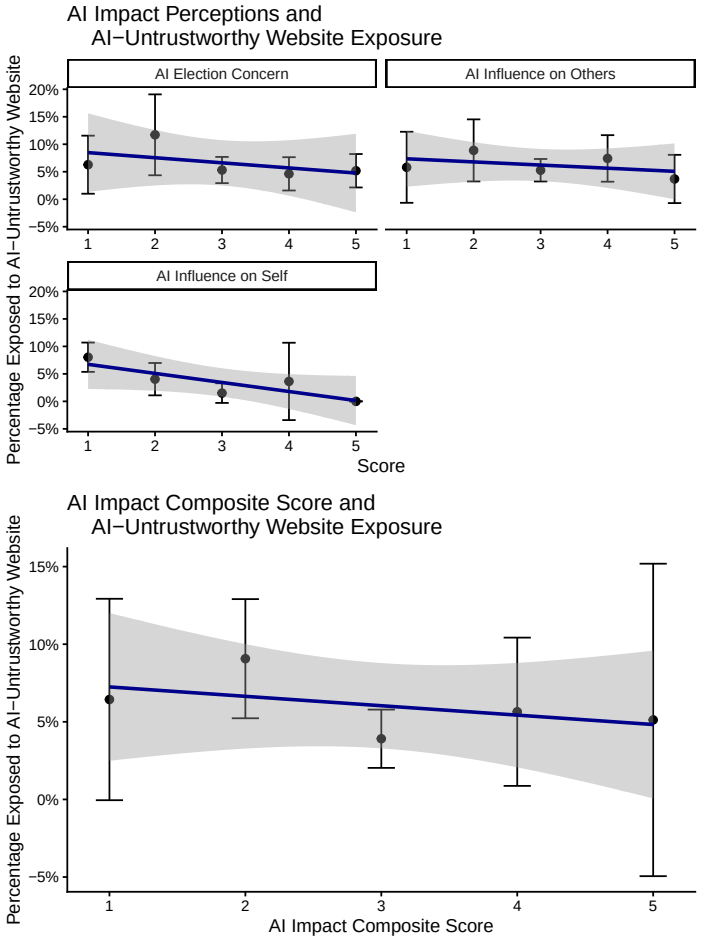


Figure S26: AI impact perceptions and AI-generated untrustworthy website exposure.

Regression analyses (Table S26) revealed significant protective associations between

certain AI impact perceptions and exposure to AI-generated misinformation websites. Specifically, believing that AI would influence one's own voting decisions was significantly associated with lower exposure to AI-generated untrustworthy content ($\beta = -0.027$, $P < 0.01$). Concerns about AI's impact on elections ($\beta = -0.011$, $P = 0.072$) and beliefs about AI influencing others' voting ($\beta = -0.006$, $P = 0.392$) were not significant. The composite AI impact score showed a significant negative association that remained robust even after controlling for demographic and political variables ($\beta = -0.021$, $P < 0.05$ without controls; $\beta = -0.022$, $P < 0.05$ with controls).

Table S26: AI Impact Perceptions Predicting AI-Generated Untrustworthy Exposure.

	Dependent variable:				
	AI-generated untrustworthy exposure				
	(1)	(2)	(3)	(4)	(5)
AI Election Concern	−0.011				
	(0.006)				
AI Influence on Self		−0.027**			
		(0.009)			
AI Influence on Others			−0.006		
			(0.007)		
AI Impact (Composite)				−0.021*	−0.022*
				(0.009)	(0.010)
Republican					−0.001
					(0.017)
Political Interest					−0.015
					(0.008)
College					0.005
					(0.016)
Female					−0.007
					(0.016)
Non-White					−0.069***
					(0.018)
Age 30–44					−0.062*
					(0.026)
Age 45–64					−0.033
					(0.025)
Age 65+					−0.048
					(0.026)
Constant	0.098***	0.105***	0.082***	0.118***	0.230***
	(0.019)	(0.015)	(0.022)	(0.024)	(0.040)
Observations	1,069	1,069	1,069	1,069	1,050
R^2	0.003	0.009	0.001	0.005	0.027
Adjusted R^2	0.002	0.008	−0.0003	0.004	0.019

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

These results reveal an intriguing protective pattern: individuals who believe AI will influence their own voting decisions show significantly lower exposure to AI-generated misinformation. This suggests that personal awareness of AI's potential electoral influ-

ence may promote more cautious evaluation of AI-generated content. The robust negative association of the composite AI impact score, even after controlling for political and demographic factors, indicates that overall concern about AI's electoral impact may equip individuals to better identify or avoid AI-generated untrustworthy content. This contrasts with general misinformation exposure, where AI concerns were associated with higher exposure, suggesting different vulnerability mechanisms for AI-generated versus traditional misinformation.

S18 AI Usage Purposes

We examined whether the specific purposes for which individuals use AI tools are associated with their exposure to untrustworthy websites. Participants indicated how frequently they use AI for twelve different purposes: work/school tasks, learning, creative projects, entertainment, personal advice, social interaction, consuming news, producing news, gathering election information, discussing elections, creating election-related text content, and creating election-related visual content. Specifically, participants were asked "When you have used generative AI tools, how often did you use them for the following?" and for each of the twelve aforementioned purposes participants indicated their frequency of use on a 5-point scale ranging from Not at all – All the time. This question is similar to others asking about the frequency of generative AI use for different purposes (e.g., [9]).

S18.1 AI Usage Purposes and General Untrustworthy Exposure

We investigated whether the purposes for which Americans use AI tools were associated with exposure to untrustworthy websites during the 2024 presidential election. We examined twelve specific usage purposes as well as a composite AI usage purposes score (see Figure S27 and Figure S28).

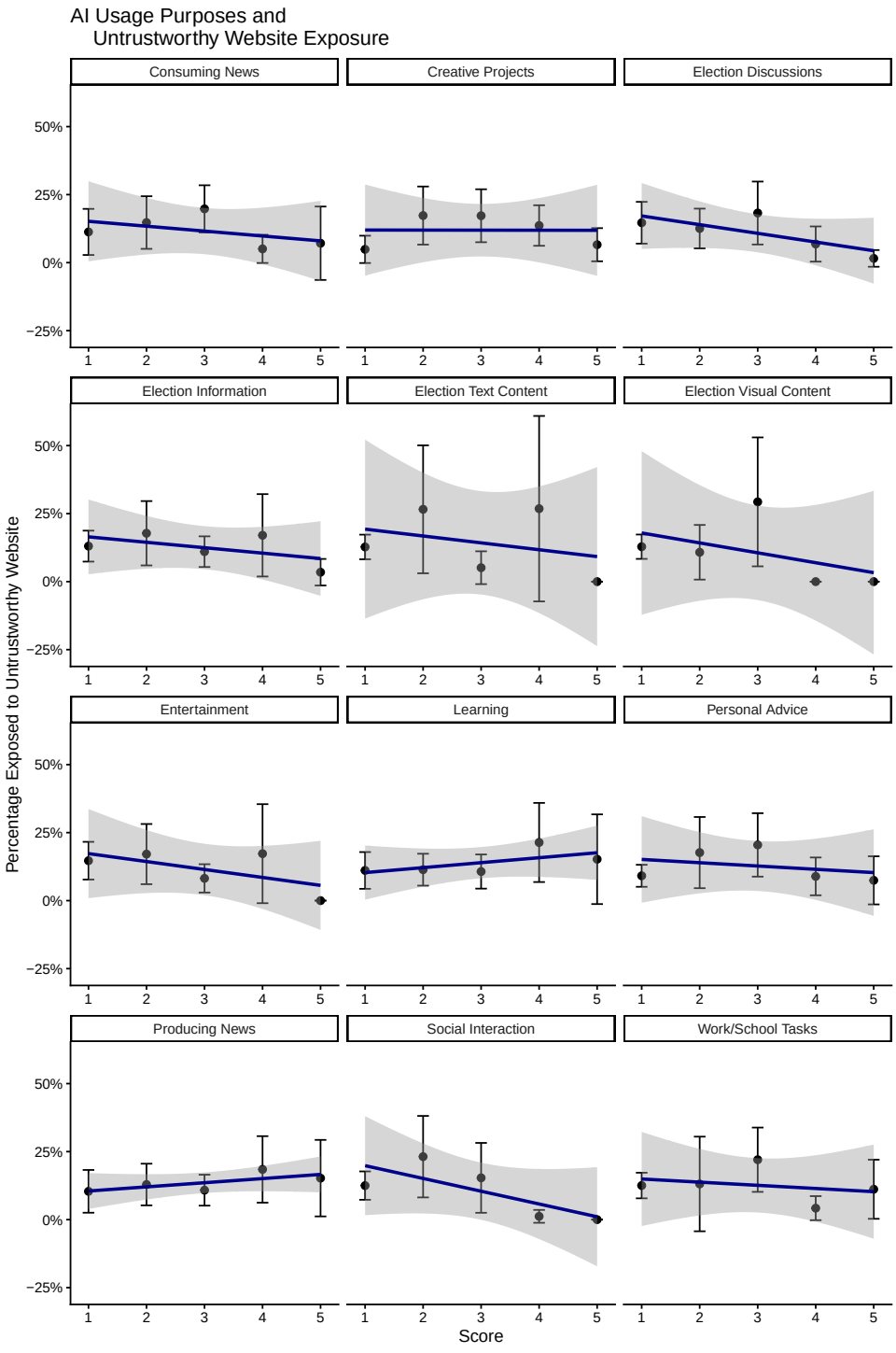


Figure S27: AI usage purposes dimensions and general untrustworthy website exposure.

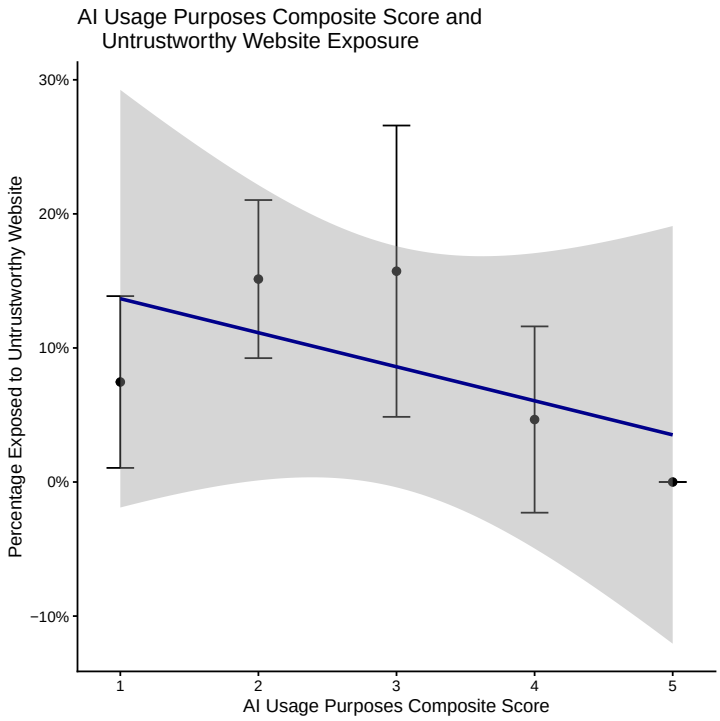


Figure S28: AI usage purposes composite score and general untrustworthy website exposure.

Regression analyses (Table S27) focused on four election-related AI usage purposes: gathering election information, discussing elections online, creating election-related text, and creating election-related visual content. None of these election-specific AI usage purposes significantly predicted exposure to untrustworthy websites. The composite AI usage purposes score similarly showed no significant association with untrustworthy website exposure, even after controlling for demographic and political characteristics.

Table S27: AI Usage Purposes Predicting General Untrustworthy Exposure.

	Dependent variable:					
	Untrustworthy Exposure					
	(1)	(2)	(3)	(4)	(5)	(6)
Election Info	−0.012					
	(0.012)					
Election Discuss		−0.018				
		(0.012)				
Election Text			−0.003			
			(0.015)			
Election Visual				−0.001		
				(0.015)		
Usage Purposes (Composite)					−0.006	0.008
					(0.019)	(0.020)
Republican						0.202***
						(0.032)
Political Interest						0.035*
						(0.017)
College						0.005
						(0.031)
Female						0.010
						(0.030)
Non-White						−0.028
						(0.034)
Age 30–44						−0.100*
						(0.042)
Age 45–64						−0.037
						(0.043)
Age 65+						−0.044
						(0.049)
Constant	0.167***	0.175***	0.139***	0.136***	0.148**	−0.021
	(0.036)	(0.032)	(0.027)	(0.027)	(0.047)	(0.078)
Observations	503	503	503	503	503	495
R^2	0.002	0.004	0.0001	0.00000	0.0002	0.113
Adjusted R^2	−0.0001	0.002	−0.002	−0.002	−0.002	0.097

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Note: AI usage purposes items were only asked of respondents who reported using generative AI tools, resulting in a smaller sample size.

These findings indicate that the specific purposes for which individuals use AI do not predict their exposure to untrustworthy websites. Even election-related AI usage—such as gathering political information, engaging in political discussions, or creating political content with AI assistance—shows no significant relationship with misinformation exposure. This suggests that it is not what people use AI for, but rather how discerningly they evaluate information, that matters for misinformation vulnerability.

S18.2 AI Usage Purposes and AI Untrustworthy Exposure

We next examined whether AI usage purposes were specifically associated with exposure to AI-generated misinformation websites during the 2024 presidential election (see Figure S29 and Figure S30).

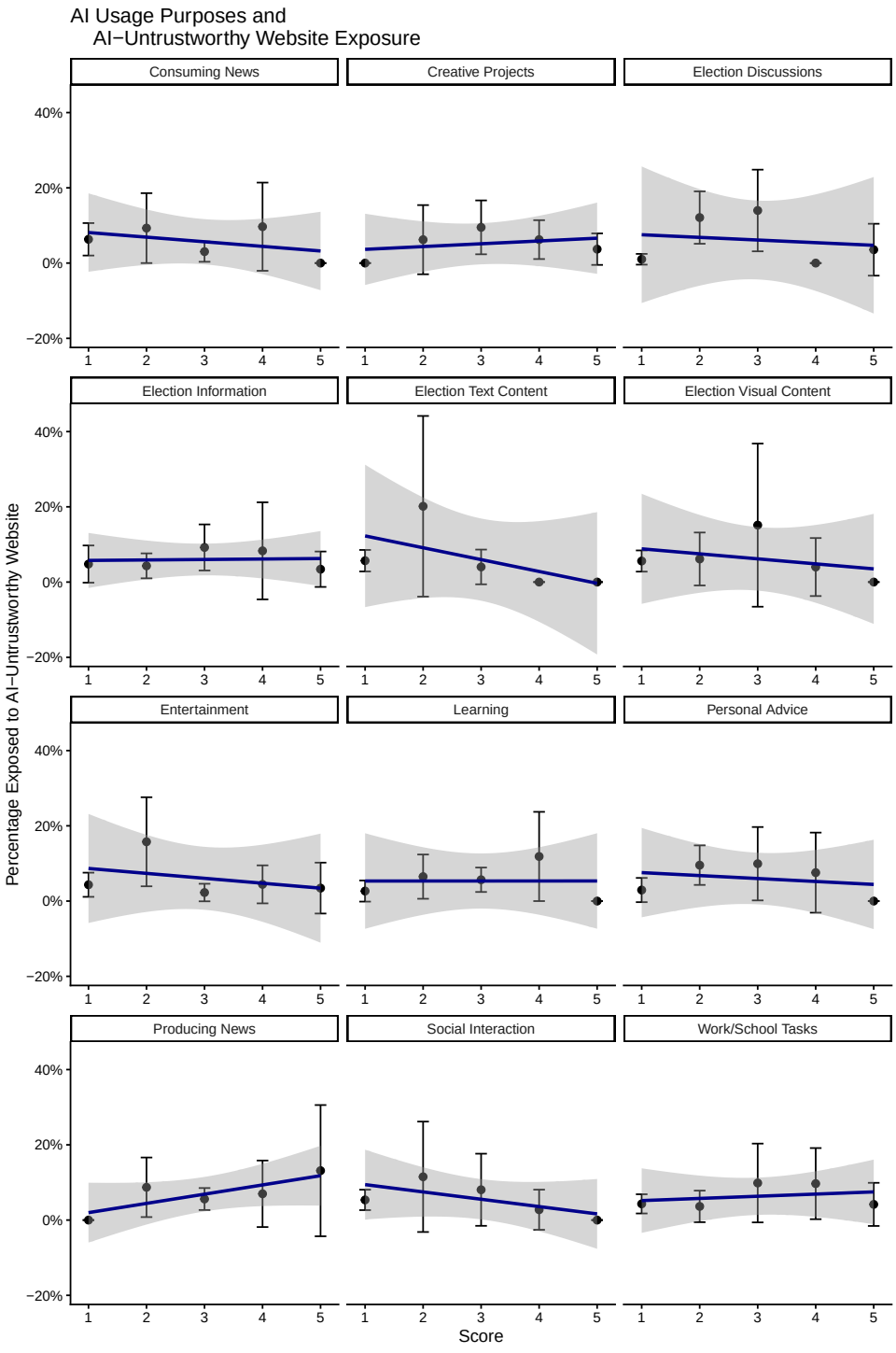


Figure S29: AI usage purposes dimensions and AI-generated untrustworthy website exposure.

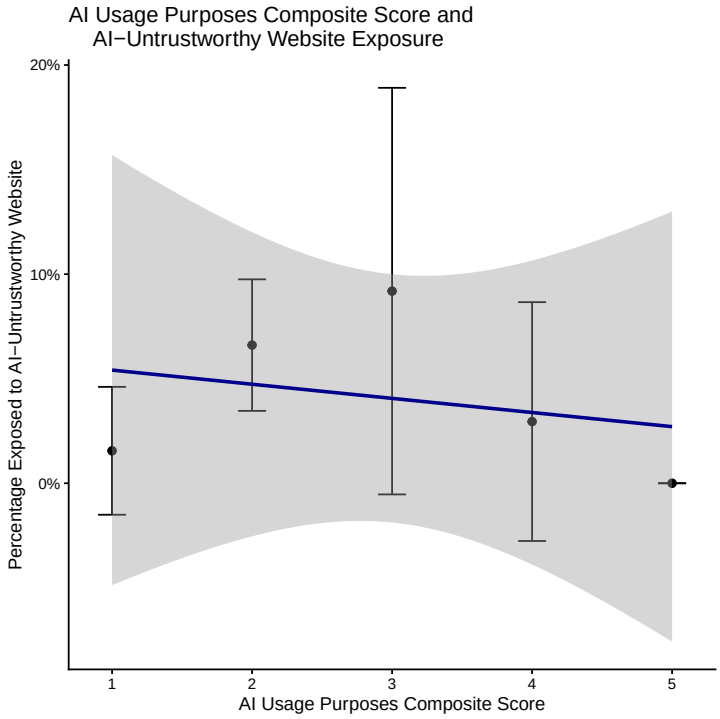


Figure S30: AI usage purposes composite score and AI-generated untrustworthy website exposure.

Regression analyses (Table S28) revealed no significant associations between election-related AI usage purposes and exposure to AI-generated misinformation websites. Specifically, using AI to gather election information, discuss elections, create election-related text, or create election-related visuals were all non-significant predictors. The composite AI usage purposes score similarly did not predict AI-generated misinformation exposure.

Table S28: AI Usage Purposes Predicting AI-Generated Untrustworthy Exposure.

	Dependent variable:					
	AI-generated untrustworthy exposure					
	(1)	(2)	(3)	(4)	(5)	(6)
Election Info	0.007					
	(0.009)					
Election Discuss		0.012				
		(0.009)				
Election Text			−0.008			
			(0.011)			
Election Visual				0.008		
				(0.011)		
Usage Purposes (Composite)					0.013	0.018
					(0.014)	(0.015)
Republican						0.017
						(0.024)
Political Interest						−0.022
						(0.012)
College						0.009
						(0.023)
Female						−0.052*
						(0.022)
Non-White						−0.062*
						(0.025)
Age 30–44						−0.090**
						(0.031)
Age 45–64						−0.079*
						(0.032)
Age 65+						−0.096**
						(0.037)
Constant	0.046	0.036	0.075***	0.052**	0.034	0.197***
	(0.025)	(0.023)	(0.019)	(0.020)	(0.034)	(0.058)
Observations	503	503	503	503	503	495
R^2	0.001	0.004	0.001	0.001	0.002	0.054
Adjusted R^2	−0.001	0.002	−0.001	−0.001	−0.0003	0.036

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Note: AI usage purposes items were only asked of respondents who reported using generative AI tools, resulting in a smaller sample size.

These results parallel the findings for general untrustworthy exposure, indicating that the purposes for which people use AI do not predict exposure to AI-generated misinformation specifically. Even politically-engaged AI usage does not differentiate who encounters AI-generated untrustworthy content. This suggests that as AI-generated misinformation proliferates, exposure risk may be more universal than previously anticipated, affecting users regardless of their specific AI usage patterns.

S19 Temporal Analysis: Pre-Election vs. Post-Election Exposure Patterns

To examine whether exposure to untrustworthy websites varies across election periods, we conducted a comprehensive temporal analysis comparing exposure patterns during the four weeks before and one week after both the 2024 and 2020 presidential elections. This analysis addresses questions about how election timing and outcomes might influence information-seeking behavior across different partisan groups.

Our approach involved collecting daily exposure data from October 8 to November 12, 2024 (totaling 38,484 person-days across 1,069 participants) and comparing these patterns with equivalent data from the 2020 election period (October 6 to November 10, 2020; 41,436 person-days across 1,151 participants). We calculated daily exposure rates as the percentage of participants who visited at least one untrustworthy website on each day, then aggregated these into pre-election (28 days before Election Day) and post-election (7 days after Election Day) periods for statistical comparison.

When examining overall exposure patterns without considering partisan differences, neither election year showed significant changes between pre- and post-election periods. In 2024, overall exposure remained stable from pre-election (3.27%, 95% CI [2.72%, 3.83%]) to post-election (3.22%, 95% CI [2.96%, 3.48%]) periods ($t = 0.229$, $P = 0.819$). Similarly, 2020 showed no significant overall change from pre-election (5.62%, 95% CI [4.84%, 6.40%]) to post-election (5.65%, 95% CI [5.26%, 6.03%]) periods ($t = -0.106$, $P = 0.916$). These null findings for overall patterns, however, mask important partisan differences in temporal exposure patterns.

The most striking findings emerge when examining partisan groups separately. In 2024, Republicans showed significantly higher exposure during the pre-election period (24.31%, 95% CI [19.25%, 29.37%]) that declined substantially following the election (16.47%, 95% CI [11.99%, 20.96%]), representing a statistically significant decrease ($t = -2.876$, $P = 0.004$, Cohen's $d = 1.673$). In contrast, Democratic exposure remained relatively stable across the election period, declining only modestly from 5.32% (95% CI [3.20%, 7.45%]) pre-election to 4.18% (95% CI [2.36%, 6.00%]) post-election ($t = -0.804$, $P = 0.422$, Cohen's $d = 1.273$).

This partisan asymmetry in 2024 stands in notable contrast to patterns observed during the 2020 election. In 2020, both parties showed significant post-election declines: Democrats decreased from 18.40% (95% CI [14.17%, 22.63%]) to 8.42% (95% CI [4.96%, 11.88%]) ($t = -5.21$, $P < 0.001$, Cohen's $d = 1.785$), while Republicans declined from 36.22% (95% CI [29.38%, 43.07%]) to 22.49% (95% CI [16.90%, 28.09%]) ($t = -4.264$, $P < 0.001$, Cohen's $d = 1.771$). The 2020 pattern suggests that post-election decreases were more universal, affecting both winning and losing partisan groups.

To formally test these partisan differences in temporal patterns, we estimated weighted linear regression models with party identification, election period, and their interaction as predictors of exposure. The 2024 analysis revealed a significant partisan $\times$ period

interaction ($\beta = -0.067$, $t = -2.161$, $P = 0.031$), indicating that the post-election change in exposure differed significantly between Republicans and Democrats. In contrast, the equivalent 2020 model showed no significant interaction ($\beta = -0.037$, $t = -1.067$, $P = 0.286$), consistent with both parties experiencing similar post-election declines during that election cycle.

Figure S31 presents the daily exposure patterns throughout both election periods, illustrating the detailed temporal dynamics around Election Day. The event study visualization reveals that partisan differences in exposure were consistently maintained throughout both study periods, with Republican exposure consistently exceeding Democratic exposure across virtually all days. However, the post-election patterns differ notably between years: in 2020, both parties show visible decreases following Election Day, while in 2024, the Republican decrease is more pronounced and the Democratic pattern remains relatively flat.

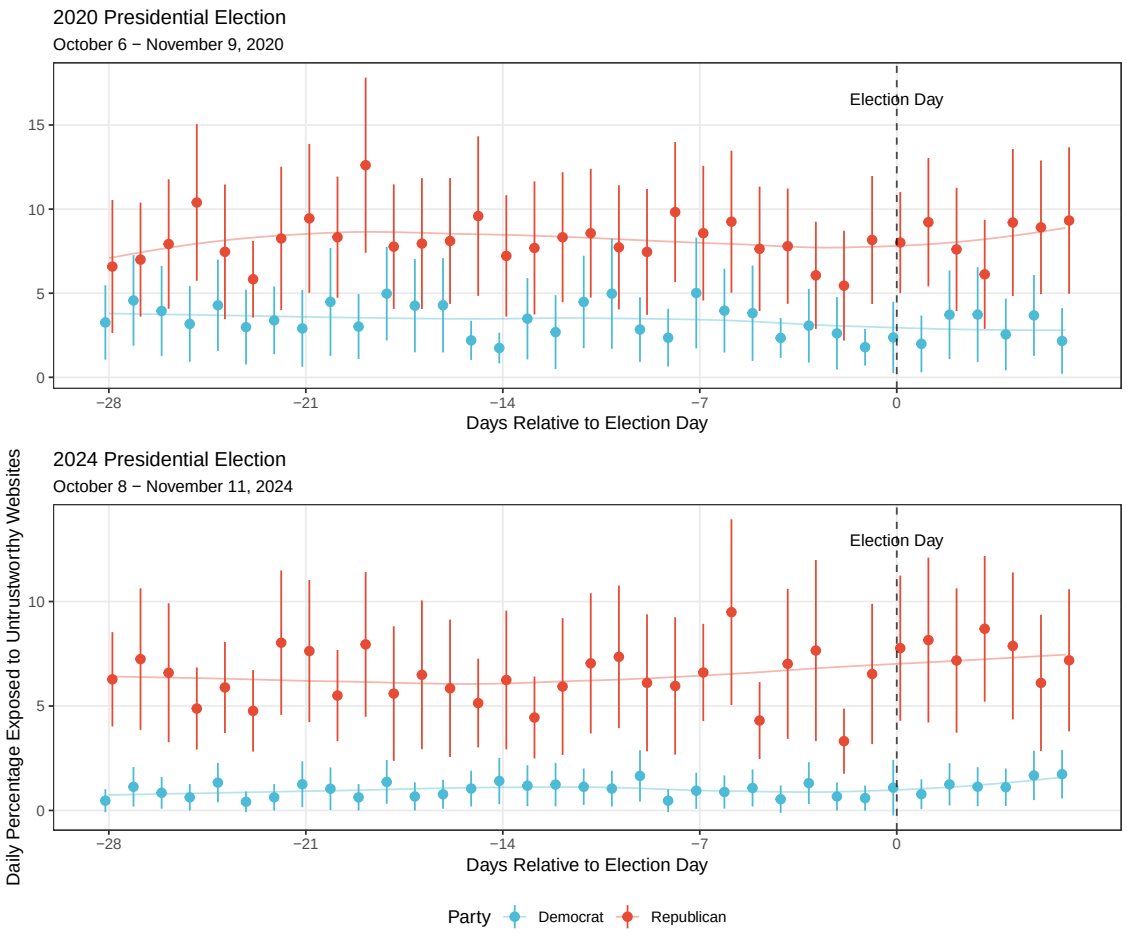


Figure S31: Daily Exposure Patterns Around Election Day (2020 vs 2024). Daily percentage of participants exposed to untrustworthy websites during the four weeks before and one week after Election Day in both 2020 and 2024. Each point represents the weighted percentage of participants who visited at least one untrustworthy website on that day, with 95% confidence intervals. The vertical dashed line indicates Election Day (Day 0). The plots illustrate consistent partisan differences across both elections, with notable post-election dynamics that differ between the two election years.

Figure S32 summarizes the pre-election versus post-election comparison across both years, highlighting the key interaction effects. The visualization clearly demonstrates the asymmetric post-election response in 2024, where Republicans show a substantial decline while Democrats remain stable, contrasted with the more symmetric declines observed in 2020.

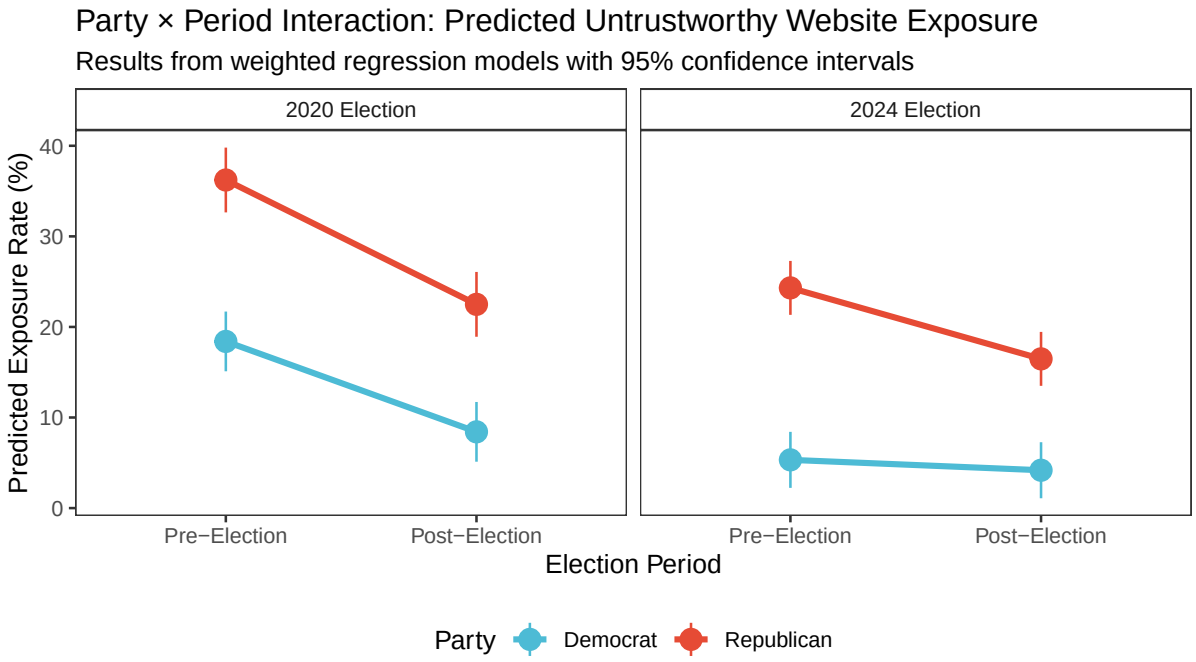


Figure S32: Pre-Election vs. Post-Election Exposure by Party and Year. Predicted exposure rates from weighted regression models showing the interaction between partisan identification and election period for both 2020 and 2024 elections. Points represent predicted exposure percentages with 95% confidence intervals. The plot illustrates the significant partisan × period interaction in 2024 ($p = 0.031$), where Republicans show substantial post-election declines while Democrats remain stable, contrasted with the more symmetric post-election patterns observed in 2020.

These temporal patterns have important implications for understanding how election outcomes influence information-seeking behavior across partisan groups. The 2024 findings suggest that supporters of the winning candidate (Republicans) reduced their consumption of untrustworthy websites following their candidate's victory, while supporters of the losing candidate (Democrats) maintained relatively stable consumption patterns. This contrasts with 2020, when supporters of both the winning candidate (Democrats) and losing candidate (Republicans) showed post-election decreases in exposure.

The asymmetric response in 2024 may reflect different information needs and motivations following electoral victory versus defeat. Winning party supporters may feel less need to seek alternative information sources or challenge mainstream narratives once their preferred outcome is achieved. Conversely, losing party supporters may maintain their information-seeking patterns as they process unexpected results or seek explanations for their candidate's defeat. However, these interpretations remain speculative and would benefit from additional research examining the motivations underlying temporal exposure

patterns.

These findings also highlight the importance of considering temporal dynamics when studying political information consumption. Static measurements taken at single time points may miss important fluctuations in exposure patterns that relate to political events and outcomes. Future research on untrustworthy information exposure should consider incorporating temporal analysis to better understand how political contexts influence information-seeking behavior.

S20 Public Perception of Change in Untrustworthy-Website Exposure

Survey design and fieldwork

To gauge public expectations about the 2024 election information environment, we embedded a single question in a nationally representative online survey conducted by Verasight, fielded from April 9 to 15, 2025 ($N = 1{,}000$).[S1]

Sampling quotas were set to match the February 2025 benchmarks from the Current Population Survey for age, race/ethnicity, sex, income, education, region, and metropolitan status. Additional post-stratification aligned the sample to a three-year rolling average of partisanship distributions from Pew Research Center's National Public Opinion Reference Survey (NPORS) and population-level estimates of the 2024 presidential vote. The design-adjusted margin of sampling error for the society module is ±3.5 percentage points.

Respondents were recruited into the Verasight Community via a combination of address-based sampling, randomized SMS outreach, and dynamic online targeting. To verify participant identity, all panelists completed multi-factor authentication, including confirmation via SMS from a Tier-1 U.S. mobile carrier. In-survey quality control included a Google reCAPTCHA v3 check, embedded attention and straight-lining items, and flagging of respondents who completed the module in less than 30% of the median completion time. Post-field procedures removed duplicates, verified U.S. IP geolocation, and excluded speeders or bot-like response patterns. Final weights were applied after all quality control measures to ensure representativeness.

Question wording. *"In the weeks around the 2020 U.S. Presidential Election, a past study found that 26.2% of American adults visited at least one 'untrustworthy' website (i.e. websites that repeatedly publish false content). Do you think the percentage of American adults that visited at least one such website in the weeks around the 2024 U.S. Presidential Election was more, less, or the same?"*

Response options were presented in a randomized order per participant; no midpoint or "don't know" option was offered.

Results

Table S29 reports weighted national results. A clear majority of adults (55.6%) believed that exposure to untrustworthy websites *increased* between 2020 and 2024. Only 8.2% thought exposure *decreased*, while 36.3% expected it had stayed *the same*.

[S1] Full project documentation and results are available at https://verasight-mpsa-2025.tiiny.co/#key-takeaways.

Table S29: Perceived change in untrustworthy-website exposure, overall sample.

Response	%
More	56
Less	8
Same	36

Note: Weighted percentages.

We observe clear heterogeneity in perceived change (Tables S30–S34). Partisanship shows the widest split: 70% of Democrats or Democratic leaners answered "more," whereas only 45% of Republicans or Republican leaners and 47% of independents did so; Republicans were almost twice as likely as Democrats to say exposure stayed the same (44% vs 25%).

Table S30: Perceived change by partisan identification.

	More	Less	Same
Democrat / lean Dem	70	5	25
Independent	47	10	43
Republican / lean Rep	45	11	44

Age differences follow a U-curve. Pessimism peaks among 18–29-year-olds (65% "more"), dips for 30–49-year-olds (45% "more," 46% "same"), and rises again for adults aged 50–64 (57% "more") and 65+ (61% "more") (Table S31).

Table S31: Perceived change by age group.

	More	Less	Same
18–29	65	8	28
30–49	45	9	46
50–64	57	9	33
65+	61	6	33

Note: Weighted percentages.

Socio-economic indicators also show a divide. The share expecting "more" exposure climbs from 44% among households earning $<$ \$50,000 to 69% among those earning $\geq$ \$150,000 (Table S32).

Table S32: Perceived change by household income.

	More	Less	Same
\$ $<$ 50,000	44	10	46
\$50,000–\$99,999	57	8	35
\$100,000–\$149,999	68	5	28
\$ $\geq$ 150,000	69	8	23

A similar pattern emerges for education: 45% of respondents with a high-school education or less chose "more," rising steadily to 65% among those holding a four-year or postgraduate degree (Table S33).

Table S33: Perceived change by educational attainment.

	More	Less	Same
High school or less	45	10	45
Some college / 2-yr degree	58	8	34
4-yr / postgraduate degree	65	6	29

Gender gaps are modest: 58% of men versus 53% of women predicted an increase, and women were slightly more likely to attribute no change (39% vs 33%) (Table S34).

Table S34: Perceived change by gender.

	More	Less	Same
Male	58	9	33
Female	53	8	39

Patterns by race and ethnicity show modest variation but still reflect a broad belief in increased exposure (Table S35). A majority of each racial/ethnic group expected exposure to rise, including 56% of White, 50% of Black, 58% of Hispanic, and 57% of respondents identifying as another racial or ethnic background. Notably, Hispanic (11%) and respondents in the "Other" category (13%) were more likely than others to believe exposure declined, compared to just 6% of White and 8% of Black respondents. Black respondents were the most likely to say levels remained the same (41%), suggesting somewhat greater skepticism toward claims of rising exposure.

Table S35: Perceived change by race/ethnicity.

	More	Less	Same
White	56	6	38
Black	50	8	41
Hispanic	58	11	31
Other	57	13	30

Taken together, the survey highlights a widespread, yet, as we find, empirically inaccurate, belief that exposure to untrustworthy websites rose in 2024 compared to 2020. This misperception is strongest among groups that are younger (65% of 18–29 year olds), higher-income (69% among those earning $\geq$ \$150,000), college-educated (65%), and Democratic-leaning (70%): demographics that tend to consume more political news online and are often more attentive to misinformation warnings. Thus, objects that may warn of widespread, untrustworthy content, such as misinformation identification interventions and media coverage, may inadvertently reinforce exaggerated perceptions of online misinformation prevalence. Future interventions may therefore need to address not only factual exposure levels but also public narratives that amplify the perceived scale of the problem.

S21 Sample Demographics Across Election Years

To address concerns about whether changes in exposure patterns between election years could be attributable to differences in sample composition, we present comprehensive demographic comparisons across all three election years (2016, 2020, and 2024). These unweighted sample characteristics demonstrate that our study samples maintained consistent demographic profiles across election cycles, ensuring that observed differences in exposure patterns reflect genuine temporal changes rather than compositional artifacts.

Table S36 presents the demographic breakdown for each election year. The samples show remarkable consistency across key demographic dimensions. Age distributions remained relatively stable, with the most notable change being a modest increase in the 30-44 age group representation in 2024 (28.4%) compared to 2016 (14.9%) and 2020 (14.9%), partially offset by decreases in older age groups. Gender representation remained consistent across all years, with females comprising approximately 51-54% of each sample.

Educational attainment showed minor variations, with college degree holders representing 49.7% in 2016, 39.2% in 2020, and 41.0% in 2024. While these differences exist, they fall within reasonable bounds for population sampling variation and are unlikely to drive the substantial changes in exposure patterns observed in our analyses. Race and ethnicity composition showed the expected demographic trends, with non-white or Hispanic participants increasing from 16.5% in 2016 to 27.6% in 2024, reflecting broader U.S. demographic changes over this period.

Political affiliation distributions varied somewhat across election years, which is consistent with documented changes in party identification during this period. Republican identification increased from 37.1% in 2016 to 40.6% in 2024, while Democratic identification decreased from 48.8% to 41.8%. However, these partisan shifts are accounted for in our analyses through the use of survey weights and explicit modeling of partisan differences.

Table S36: Sample Demographics Across Election Years (Unweighted)

Demographic Group	Category	2016	2020	2024
Age	18-29	5.7%	8.2%	8.1%
Age	30-44	14.9%	14.9%	28.4%
Age	45-64	47.6%	46.4%	38.1%
Age	65+	31.8%	30.6%	25.4%
Gender	Female	51.3%	54.2%	54.0%
Gender	Male	48.7%	45.8%	46.0%
Education	College degree or higher	49.7%	39.2%	41.0%
Education	Less than college	50.3%	60.8%	59.0%
Race/Ethnicity	Non-white or Hispanic	16.5%	19.9%	27.6%
Race/Ethnicity	White, non-Hispanic	83.5%	80.1%	72.4%
Political Affiliation	Democrat	48.8%	53.9%	41.8%
Political Affiliation	Independent/Other	14.1%	12.1%	17.6%
Political Affiliation	Republican	37.1%	34.1%	40.6%

Importantly, all analyses presented in this study employ YouGov's survey weights, which adjust for any sampling differences and ensure that results reflect nationally representative estimates rather than raw sample characteristics. The demographic consistency across samples, combined with the use of appropriate weights, provides confidence that our findings reflect genuine changes in online information consumption patterns rather than artifacts of sample composition.

The robustness of our findings is further supported by the replication of key patterns across multiple demographic subgroups within each election year. For instance, the documented decline in overall exposure from 2020 to 2024 is observed consistently across age groups, education levels, and political affiliations, suggesting that demographic composition changes cannot account for the primary findings.

References

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

[2] Saumya Bhadani, Shun Yamaya, Alessandro Flammini, Filippo Menczer, Giovanni Luca Ciampaglia, and Brendan Nyhan. Political audience diversity and news reliability in algorithmic ranking. *Nature Human Behaviour*, 6(4):495–505, 2022.

[3] Ryan C Moore, Ross Dahlke, and Jeffrey T Hancock. Exposure to untrustworthy websites in the 2020 us election. *Nature Human Behaviour*, 7(7):1096–1105, 2023.

[4] Eszter Hargittai and Yuli Patrick Hsieh. Succinct survey measures of web-use skills. *Social Science Computer Review*, 30(1):95–107, 2012.

[5] Bingcheng Wang, Pei-Luen Patrick Rau, and Tianyi Yuan. Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & information technology*, 42(9):1324–1337, 2023.

[6] Harry Yaojun Yan, Garrett Morrow, Kai-Cheng Yang, and John Wihbey. The origin of public concerns over ai supercharging misinformation in the 2024 us presidential election. *Harvard Kennedy School Misinformation Review*, 2025.

[7] Megan French and Jeff Hancock. What's the folk theory? reasoning about cyber-social systems. *Reasoning About Cyber-Social Systems (February 2, 2017)*, 2017.

[8] Harry Yaojun Yan, Ryan C Moore, Fangjing Tu, and Jeffery T Hancock. Detecting synthetic, doubting authentic: Ai attribution bias for political imagery. 2025.

[9] Clare Baek, Tamara Tate, and Mark Warschauer. "chatgpt seems too good to be true": College students' use and perceptions of generative ai. *Computers and Education: Artificial Intelligence*, 7:100294, 2024.