Political Communication

ISSN: 1058-4609 (Print) 1091-7675 (Online) Journal homepage: www.tandfonline.com/journals/upcp20

Multi-Platform Referrers of Misinformation: A Comparative Analysis of Misinformation Visits Referred by Facebook, Twitter, Instagram, Reddit, YouTube, Snapchat, and TikTok

Ross Dahlke, Ryan C. Moore, Danya Adib-Azpeitia, Johan Ugander & Jeffrey T. Hancock

To cite this article: Ross Dahlke, Ryan C. Moore, Danya Adib-Azpeitia, Johan Ugander & Jeffrey T. Hancock (09 Jun 2026): Multi-Platform Referrers of Misinformation: A Comparative Analysis of Misinformation Visits Referred by Facebook, Twitter, Instagram, Reddit, YouTube, Snapchat, and TikTok, Political Communication, DOI: 10.1080/10584609.2026.2679492

To link to this article: https://doi.org/10.1080/10584609.2026.2679492

View supplementary material

Published online: 09 Jun 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=upcp20

POLITICAL COMMUNICATION
https://doi.org/10.1080/10584609.2026.2679492

Routledge
Taylor & Francis Group

Check for updates

Multi-Platform Referrers of Misinformation: A Comparative Analysis of Misinformation Visits Referred by Facebook, Twitter, Instagram, Reddit, YouTube, Snapchat, and TikTok

Ross Dahlke[a], Ryan C. Moore[b], Danya Adib-Azpeitia[c], Johan Ugander[d], and Jeffrey T. Hancock[e]

[a]School of Journalism and Mass Communication, University of Wisconsin-Madison, WI, USA; [b]School of Information, The University of Texas at Austin, Austin, USA; [c]Symbolic Systems Program, Stanford University, Stanford, CA, USA; [d]Department of Statistics & Data Science, Yale University, New Haven, CT, USA; [e]Department of Communication, Stanford University, USA, Stanford, USA

ABSTRACT

Journalistic and scholarly accounts alike often depict social media platforms as key players in exposing people to misinformation. Unfortunately, existing research examining social media's role in exposing people to misinformation online implicitly operates under what we call the *direct referrer paradigm*, focusing only on visits to misinformation that directly follow social media. We introduce the *embedded referrer paradigm*, which considers social media's more comprehensive role in leading people to visit misinformation throughout their online journeys and operationalize it via counterfactual simulations. We estimate the effects of social media in exposing people to misinformation websites under both paradigms using 3 months of a large sample of Americans' ($N = 1{,}240$) web browsing behavior (21 million website visits). We explore seven different platforms' (Facebook, Instagram, Reddit, Snapchat, TikTok, Twitter, YouTube) effects in exposing people to misinformation sites with various ideological leanings. Overall, our results suggest that analyses under the direct referrer paradigm underestimate the role of social media in exposing people to misinformation, with estimated effects under the embedded referrer paradigm ranging from 1.5 to 15+ times larger than effects under the direct referrer paradigm. The difference in estimated effects under both paradigms is generally consistent across platforms, misinformation types (liberal, conservative, non-ideological), and individuals' partisanship. However, for some combinations of platforms, types of misinformation, and mobile vs. desktop devices, the discrepancy between the two paradigms is even greater.

KEYWORDS
Facebook; Twitter; YouTube; Reddit; Instagram; Snapchat; TikTok; misinformation; web browsing; simulation; clickstream

The spread of political misinformation online poses fundamental challenges to democratic processes and institutions. During recent elections, tens of millions of Americans were exposed to false political claims online (A. Guess et al., 2020; Dahlke et al., 2026 R. C. Moore et al., 2023), with consequences ranging from decreased voter turnout (Green et al., 2022) to misperceptions of electoral outcomes (Dahlke & Hancock, 2025). A central question for political communication scholars is understanding how citizens

CONTACT Ross Dahlke ross.dahlke@wisc.edu School of Journalism and Mass Communication, University of Wisconsin, Vilas Communication Hall, 5115, 821 University Ave,-Madison, WI 53706, USA

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10584609.2026.2679492

© 2026 Taylor & Francis Group, LLC

encounter this political misinformation in their media environments. Do voters actively seek out false political content? Do they stumble upon misinformation while using search engines or social media platforms? These questions are crucial for theories of selective and incidental exposure to political information, as well as for designing interventions to protect democratic discourse. Both journalists and researchers have identified social media platforms as key vectors for political misinformation exposure (Allcott & Gentzkow, 2017; Allen et al., 2020; Dahlke et al., 2026; Fox, 2018; Frenkel, 2021; A. Guess et al., 2020; Jones-Jang et al., 2021; Lee et al., 2023; Luk et al., 2021; R. C. Moore et al., 2023, 2024; Myers, 2022; Neyazi et al., 2022; Warner et al., 2021). However, existing research methods provide an incomplete picture of how social media shapes citizens' exposure to political misinformation.

Extant research primarily conceives of social media's role as a direct referrer. For example, Allcott and Gentzkow (2017) use web browsing data to identify how many visits to misinformation websites were *directly preceded* by visits to social media platforms (e.g., 41.8% of misinformation website visits during the 2016 U.S. election were preceded by a visit to a social media platform). This approach is commonplace, with several scholars using it to examine pathways to misinformation websites (Allen et al., 2020; Guess, Lyons, et al., 2019; A. Guess et al., 2020; R. C. Moore et al., 2023), and to mainstream news sites (Cardenal et al., 2019; Flaxman et al., 2016; Möller et al., 2020; Scharkow et al., 2020; Stier et al., 2022; Wojcieszak et al., 2021). Analyses tend to collapse across visits, not allowing scholars to easily examine the influence of individual differences (e.g., partisanship) on social media's role in exposing people to misinformation. We call this approach – in which uses of social media directly precede visits to misinformation – the *direct referrer paradigm*.

Another way to conceptualize platforms' influence is to imagine a counterfactual online world in which an individual cannot visit a given platform. For instance, we could examine what would happen to visits to misinformation sites if Facebook were not accessible, such as a world where Facebook did not exist.[1] Would people be more or less likely to visit misinformation?

There are myriad ways in which a social media platform might influence the likelihood that an individual would visit a misinformation site. One possibility is that using social media platforms causes the activation of goals and concepts that lead people to visit misinformation later in their browsing (Knoll et al., 2020). For example, someone may encounter information about a news event while scrolling through their Facebook Feed and then, after they have finished using Facebook, conduct a search where they then visit misinformation about that news event (e.g., Facebook Feed → curiosity → other website → misinformation website). In this example, Facebook did not directly expose people to misinformation, but other content that later led people to misinformation. Another possibility is that social media influences people's visits to misinformation because it leads them to other websites that serve as intermediaries (e.g., Facebook → blog post → misinformation).

The direct referrer paradigm, although popular, provides an incomplete picture of social media's role, as the examples above illustrate. Moreover, in Figure 1A, this paradigm would entirely credit Facebook for the misinformation website visit. However, in Figure 1B, it would not consider Facebook to have played any role. We instead propose an alternative paradigm, which relies on this type of counterfactual reasoning to remove people's ability to visit a social media platform, accounting for social media's role when directly preceding

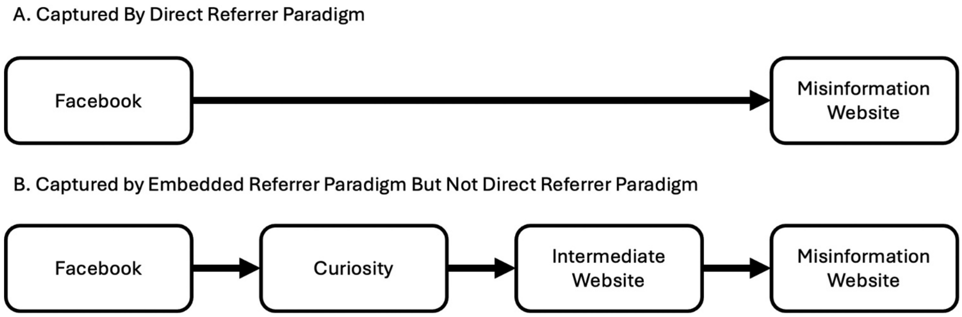

Figure 1. Example path from social media to misinformation sources captured under the direct and embedded referrer paradigms. Caption: Illustrative example of the limitations of the direct referrer paradigm in considering the role of social media platforms on misinformation exposure. Each bubble represents a website that an individual is at, and the arrows represent an individual's path from one website to another. In Panel A, which represents the direct referrer paradigm, Facebook would receive all of the "credit" for referring the individual to the misinformation website. However, in Panel B, the direct referrer paradigm would not attribute any credit to Facebook for the visits to the misinformation website, despite the possibility that Facebook indirectly influenced the misinformation visit by leading the individual to an intermediate website, which led them to the misinformation website. Our proposed embedded referrer paradigm accounts for Facebook's influence in this path by simulating a counterfactual scenario which estimates what an individual's path would be if they could not visit Facebook.

information sources *and* when embedded further back in users' online journeys. We call this alternative approach the *embedded referrer paradigm*.

The approach of removing an element from a system to identify that element's impact on the system features throughout the natural and social sciences. For example, a commonly used technique in biology and genetics is a complete gene knockout, where researchers remove a target gene from an organism (Hall et al., 2009; Kieffer & Gavériaux-Ruff, 2002). By comparing organisms with and without that specific gene, researchers learn about that gene's role in the biological system and its effects on the organism. Our counterfactual world without social media is conceptually similar to gene knockouts and other types of "ablation studies" – by removing a social media platform from the online ecosystem and observing where people go in its absence, we can learn about the effects of that platform on people's visits to information sources of interest; in our case, misinformation sites (defined as websites rated by professional journalists, editors, or fact-checkers as repeatedly publishing false content, more detail about the operationalization is contained in the Methods & Measures section).

We draw further inspiration from advertising attribution research in marketing science. Online retailers traditionally measure media's impact on purchases via "last-click attribution" (Gordon et al., 2021), examining advertisements individuals see directly before a purchase and assigning the totality of the purchase to that advertisement (Li et al., 2016). Thus, last-click attribution conceptually resembles the direct referrer paradigm. However, marketing scholars find that last-click attribution does not adequately capture advertising's effect on purchasing decisions, missing important steps in consumer journeys and thus causing inefficient allocations of firms' marketing budgets (Lee, 2010; Li et al., 2016). One alternative method uses consumers' likelihood of moving from one website to another given their prior websites to estimate advertising effect and thus offers holism that inspires our embedded referrer paradigm (Anderl et al., 2016; de Haan et al., 2016). Another area of social science research that provides inspiration for the present study is work that combines clickstream data with Markov chains to understand the sequence of user transitions throughout pages on a website (e.g., Vermeer & Trilling, 2020; Vermeer et al., 2020).

How will social media's role differ under the embedded referrer paradigm? It could be larger than previously thought, as the embedded referrer paradigm accounts for instances in which social media does not directly precede people's visits to misinformation but still influences them (e.g., Facebook → blog post → misinformation website). Alternatively, it could be smaller because social media platforms occupy people's time and attention online, limiting people's visits to misinformation they would otherwise spend time on if social media did not exist. Our counterfactual approach allows us to examine these possibilities.

In this paper, we analyze 3 months of web browsing data collected during the 2020 U.S. presidential election from a large sample of American adults ($N = 1,240$, 21 million website visits) to estimate seven social media platforms' roles in leading voters to misinformation websites under both the direct referrer paradigm and our novel embedded referrer paradigm. This electoral context is crucial, as it represents a period of intense political information seeking when citizens are most actively forming opinions about candidates and issues. Our findings reveal that the direct referrer paradigm dramatically underestimates social media platforms' role in political misinformation exposure. Under our embedded referrer paradigm, social media's estimated effect on misinformation visits is 1.5 to 15+ times larger than effects estimated under the direct referrer paradigm. This underestimation generally holds across platforms, with some heterogeneity across participants' partisan identities and misinformation websites with different ideological leanings. We argue that the embedded referrer paradigm provides a more comprehensive account of how social media platforms shape citizens' exposure to political misinformation during critical democratic moments, advancing our understanding of selective exposure to electoral information.

Introducing the embedded referrer paradigm is our paper's first contribution to political communication. This new framework helps us understand how social media platforms shape citizens' information environments beyond simple direct exposure. Developing an analytical approach to operationalize the embedded referrer paradigm is our second contribution. While we demonstrate its value in the context of political misinformation, our approach could be applied to study social media's role in exposing people to various political information sources (e.g., partisan news websites, election-related content) as well as other content (e.g., mainstream news websites, websites hosting hateful or problematic content).

Our third contribution is empirical, revealing how different social media platforms distinctly shape exposure to political misinformation across partisan groups. While underestimation is consistent across the social media platforms we examine (Facebook, Instagram, Reddit, Snapchat, TikTok, Twitter, YouTube), the magnitude varies in ways that challenge existing theories about platform effects and selective exposure. For example, under the direct referrer paradigm, Twitter directly precedes approximately 2% of all misinformation website visits. However, under the embedded referrer paradigm, the counterfactual scenario without Twitter has 12% fewer visits to misinformation sites than the observed world, with particularly strong effects among Republicans compared to

Democrats and Independents, suggesting that Twitter plays a substantially larger role in political information exposure than previously considered. Reddit exhibits a similar discrepancy, preceding fewer than 0.2% of misinformation visits under the direct referrer paradigm but under the embedded referrer paradigm, the world with Reddit had nearly 1.8% more visits to misinformation sites than in the world without Reddit. For certain combinations of platforms and specific types of political misinformation (e.g., YouTube and liberal misinformation, Twitter and non-ideological misinformation) and partisans, the discrepancies are even starker, with the embedded referrer paradigm suggesting that platforms have roles 15+ times larger than estimates under the existing direct referrer paradigm. These partisan asymmetries in platform effects demonstrate how different social media sites distinctly shape citizens' exposure to various types of political misinformation, advancing our understanding of how digital media environments may reinforce or challenge partisan information diets.

Overall, our findings call for a re-interpretation of existing knowledge about social media's role in leading people to misinformation. They indicate new mechanisms that future research should explore that may explain how social media platforms lead people to visit online content, positioning our study as especially relevant to the literatures in communication on selective (Knobloch-Westerwick, 2014) and incidental exposure (Schäfer, 2023). Furthermore, our analysis of seven social media platforms underscores the importance of examining platforms individually rather than collectively when studying misinformation exposure (Matassi & Boczkowski, 2021). Integrating a multi-platform perspective with the embedded referrer paradigm clarifies that each platform may uniquely shape misinformation exposure, highlighting the need for targeted theoretical and empirical exploration of platform-specific dynamics.

Existing Research on Social Media's Role in Exposing People to Misinformation

Work examining social media's effect on exposing people to misinformation online has taken two primary forms. First, many studies draw on self-reported data about people's internet browsing (Hwang & Jeong, 2023; Jones-Jang et al., 2021; Kim & Tandoc, 2022; Lee et al., 2020, 2023; Luk et al., 2021; Neely et al., 2022; Neyazi et al., 2022; Vegetti & Mancosu, 2022; Warner et al., 2021; Xiao, 2022). For example, by asking people to self-report the amount they use different social media platforms as well as if they have been exposed to misinformation, Neyazi et al. (2022) found that greater usage of WhatsApp and Instagram to consume political content was positively associated with exposure to misinformation, while usage of Facebook and Twitter to consume political content was not associated with exposure to misinformation.

Self-reported use of social media is, at best, modestly correlated with actual use, both the overall amount of one's use (Ernala et al., 2020; Naab et al., 2019; Parry et al., 2021; Vraga et al., 2016) and with specific behaviors on platforms (Haenschen, 2020; Staddon et al., 2013). Self-reports of political information consumption are particularly inaccurate (Guess, Munger, et al., 2019; Haenschen, 2020), largely because people cannot remember all their media use (Revilla et al., 2017).

To overcome these limitations, another body of work uses individuals' web browsing data to infer social media's effect on visits to misinformation websites (Allcott & Gentzkow, 2017; Allen et al., 2020; Altay et al., 2022; Guess, Lyons, et al., 2019; A. Guess et al., 2020;

R. C. Moore et al., 2023; Weeks et al., 2021; Zhou et al., 2023), including conspiratorial (Moore et al., 2024) and ones mimicking journalism but built for financial or political purposes (R. C. Moore et al., 2023). Data collection occurs through passive tracking via web browser plugins (e.g., A. Guess et al., 2020) or data donations participants make to researchers (e.g., Weeks et al., 2021). Using web browsing data, A. Guess et al. (2020) found that 15% of visits Americans made to misinformation websites during the 2016 election were directly preceded by Facebook, while Twitter preceded just 1%. Allcott and Gentzkow (2017) also studied the 2016 US election, finding that social media (Facebook and Twitter) was a larger source of visits to misinformation websites than to high-quality news sites. During the 2020 US election, R. C. Moore et al. (2023) found that Facebook directly preceded 5.6% of visits to misinformation websites, a significant reduction compared to Facebook's role in referring people to misinformation in 2016. While these studies leveraging web browsing data more accurately measure people's media use behaviors compared to self-reports, nearly all using web-browsing data operate under the direct referrer paradigm.

We argue the direct referrer paradigm, while insightful, does not measure social media platforms' total role in people's visits to misinformation. Our embedded referrer paradigm more adequately captures this total role because it accounts for instances where platforms may influence people's trajectories through the internet, even when they do not directly send people to misinformation. One way social media can influence trajectories is by exposing people to intermediary websites that then expose them to misinformation sites (Houidi et al., 2019; Scharkow et al., 2020). They may also influence subsequent browsing behaviors by causing the psychological activation of concepts or goals that influence visits long after leaving a platform (Knoll et al., 2020).

A few studies have leveraged objective web use data and do not necessarily operate under the strictest definition of the direct referrer paradigm. These studies go beyond the simple "one-step-back" approach of other studies. One such study is Vermeer et al. (2020), which examines how people move through the internet to arrive at news websites. Rather than limiting themselves to which website users visited immediately before a news website, the authors analyze the past five websites that users visited before visiting a news site. Vermeer and Trilling (2020) analyzed a sample of Dutch users' behaviors on news websites using Markov Chains, estimating the probability of individuals going from one part of a news website (e.g., homepage) to another part (e.g., business section), taking into account the users' activity. While these papers use Markov chains to consider transitions between webpages, they cannot speak to social media's role in contributing to news consumption, as it is limited to participants' navigation on news websites. Furthermore, for our purposes, neither of these studies examines misinformation exposure. Although Vermeer et al. (2020) and Vermeer and Trilling (2020) more comprehensively account for social media's role in exposing people to online information sources, they are still limited in scope. We draw inspiration from this work to propose a more holistic theoretical and methodological framework, namely through a novel method of counterfactual simulations, to understand indirect referral effects.

How social media use might facilitate exposure to and engagement with misinformation can be understood through selective and incidental exposure mechanisms, two prominent theoretical perspectives in political communication research. Selective exposure generally refers to the tendency for people to be exposed to news and information that aligns with existing beliefs, while incidental exposure refers to the tendency for people to be exposed to news when they are not necessarily seeking it out due to processes like algorithmic curation (Stroud, 2008, 2010). In terms of selective exposure, social media can (1) selectively expose people to misinformation that is likely to be congenial with their preexisting beliefs via personalization and network effects, which research has shown makes people more likely to believe information to be true (Sultan et al., 2024) and (2) selectively expose people to a feed of congenial information in general, which can condition people to be less critical of misinformation when it is encountered (Rhodes, 2022). In terms of incidental exposure, social media can incidentally expose people to news and political topics when they are otherwise not actively seeking them out (Oeldorf-Hirsch, 2018), around which they may be exposed to misinformation about those topics.

Unfortunately, existing research exploring selective and incidental processes on social media (not just in terms of exposure to misinformation, but exposure to news in general) is limited in its ability to characterize the extent of selective and incidental exposure in three key ways. First, most studies rely on self-reported measures of exposure (e.g., Kwon et al., 2015; Weeks et al., 2017), which can be inaccurate as discussed previously. Second, these measures often ask participants about "social media" (e.g., Malinen et al., 2018), or only assess one platform (e.g., Facebook; Mukerjee & Yang, 2021, Twitter; Shin, 2020) and do not disaggregate selective/incidental exposure by specific platforms, which likely play different roles given differences between platforms in content, algorithms, networks, and other affordances. Third, even when objective measures of media use are used to examine selective exposure, this work is typically either focused only on one platform (e.g., Bakshy et al., 2015) or done under the direct referrers paradigm (e.g., Cardenal et al., 2019; Flaxman et al., 2016), which as was previously discussed limits the scope of this work's ability to elucidate the role of social media in selective exposure processes, particularly if one conceptualizes selective exposure as having potentially both immediate and delayed effects on people's journeys through their individual information ecosystems.

Finally, one important additional limitation of existing research on social media's role in misinformation exposure, covering studies that rely on self-report and those that use web browsing activity, is that most studies in the literature examine misinformation exposure via only two platforms, Facebook and Twitter (A. M. Guess & Lyons, 2020; Yang et al., 2023), a problem plaguing social media research more broadly (Lukito et al., 2023). While these platforms undoubtedly have large user bases and play important roles in the social media ecosystem, there is growing concern about misinformation's spread on other social media platforms whose user bases are large and growing rapidly, such as TikTok (Basch et al., 2021; Southwick et al., 2021).

Methods & Measures

Participants

Data was collected from participants recruited by the survey organization YouGov. All participants consented to the terms of the research and were compensated by YouGov for their participation. These participants ($N = 1{,}240$; 550 males, 683 females) completed a survey during the 2020 election and provided passively gathered web-browsing data (i.e., every time a consenting participant visits a new URL, it gets automatically recorded in our dataset without the participant needing to take another step) through YouGov's Pulse software from 24 August 2020 to 7 December 2020. Twenty-nine percent of participants ($n = 358$) were aged 65+, 48% ($n = 599$) were 45–64, 14% ($n = 175$) were 30–44, and 9% ($n = 106$) were under 30. Of the participants in this study, 59% ($n = 726$) supported Joe Biden in the 2020 election and 36% ($n = 445$) supported Donald Trump. The web-browsing data was collected across desktop, mobile, and tablet devices, though not evenly. Of our 1,240 participants, 924 (75%) contributed desktop browsing sessions, 335 (27%) contributed mobile browsing sessions, and 65 (5%) contributed tablet browsing sessions, with participants using multiple devices. This distribution reflects both the technological capabilities of browser tracking in 2020 and participants' device preferences for installing tracking software. The average desktop user produced approximately 315 browsing paths, while the average mobile user produced approximately 272 paths, and the average tablet user produced approximately 123 paths. The data that support the findings of this study are openly available in OSF at https://osf.io/fegy3/.

Procedure

The participants visited approximately 21 million web pages, at over 270k unique domains, throughout our data collection period. Following previous academic (Vermeer & Trilling, 2020; Vermeer et al., 2020) and industry (Google, 2022) definitions, we define a web-browsing session as "ending" if a participant did not make any new visits to websites within 30 minutes. Each session contains at least one visit to a website, and our data contain 395,936 sessions.

Defining Social Media, Misinformation, and News Websites

Following Aridor (2023), we operationalize a visit to a social media platform as a visit to Facebook, Instagram, Reddit, Snapchat, TikTok, Twitter, or YouTube (see Supplemental Materials A [Table S1] for the full list of domains for each platform). After identifying domains which were one of these social media platforms, 270,072 domains were left.

To identify which domains were misinformation websites, we take a similar approach to R. C. Moore et al. (2023), who operationalize misinformation websites as domains rated by professional journalists, editors, and fact-checkers as repeatedly publishing false content. Specifically, misinformation domains in our study consisted of websites rated by the organization NewsGuard[2] as "repeatedly publishing false content" (a binary criteria used in prior work to define misinformation websites, e.g., Ahmad et al., 2024), as well as websites identified by fact-checkers and journalists (e.g., PolitiFact, Snopes) as misinformation sites and used in prior academic studies (Allcott & Gentzkow, 2017; Grinberg et al., 2019).

To examine social media platforms' effects on selective exposure, that is when platforms are more likely to expose individuals to ideologically congenial misinformation, we also divide misinformation domains by their ideological slant. To accomplish this task, we first classify domains' ideology via NewsGuard. Domains that NewsGuard categorized as "Far Left" or "Slightly Left," we label as liberal. Domains categorized as "Far Right" or "Slightly Right," we label as conservative, and domains categorized as "N/A," we label as non-ideological. For domains that do not have an ideological rating from NewsGuard, we turn to scores developed by Robertson et al. (2018). These scores range from −1 (the most liberal) to 1 (the most conservative). We divide the domains into terciles based on these scores (to see the distribution of scores, see Supplemental Materials B [Figure S1]). We label domains in the bottom tercile as liberal, the middle tercile as non-ideological, and the top tercile as conservative. Fourteen domains were identified as misinformation that did not have an ideology label after drawing on the NewsGuard and Robertson et al. (2018) ratings. Two research team members visited and hand-labeled the 14 remaining domains, discussing domains with disagreements to finalize ideological labels.

There were 269,823 domains that remained uncategorized as either social media or misinformation. To categorize these domains, we turn to quality scores provided by Lin et al. (2023), the largest academic list detailing domain credibility to our knowledge. Specifically, we divide the remaining domains found in Lin et al. (2023) into two categories: credibility scores below the median score are categorized as low-quality domains, and scores above the median are categorized as high-quality domains (to see a distribution of scores, see Supplemental Materials B [Figure S2]). Lin et al. (2023) calculated domain quality scores through a principle component analysis (PCA) of six different lists of website domain quality ratings, including NewsGaurd, Ad Fontes Media, Media Bias/Fact-Check, Fact-Checker, the Iffy index of unreliable sources, and another academic study (Lasser et al., 2022). This PCA results in a comprehensive list of domain quality ratings that are highly correlated with the underlying ratings (see Lin et al., 2023 for more details). We repeat the process to subcategorize these domains by ideology (conservative, liberal, non-ideological) using the same approach as above for the misinformation domains (hand coding 657 domains). We categorize the remaining 266,172 domains that participants visited as "other." "Other" domains largely consisted of internet utilities (e.g., e-mail clients, banking, and shopping websites).

To summarize, we proceed with our analysis with domains categorized as social media (Facebook, Instagram, Reddit, Snapchat, TikTok, Twitter, YouTube), misinformation (conservative, liberal, non-ideological), low-quality (conservative, liberal, non-ideological), high-quality (conservative, liberal, non-ideological), or other.

Data Analysis

Direct Referrer Paradigm

Under the direct referrer paradigm, social media's effect on visits is typically measured via a *referral analysis*, which is simply analyzing the percentage of visits to a given website that were directly preceded by a social media website (e.g., Allcott & Gentzkow, 2017; Möller et al., 2020). Formally, this analysis is calculated by the function $ReferralPercentage = \frac{r_{is}}{t_i} \times 100$ (1) where r_{is} is the number of visits to misinformation website i that were directly preceded by social media s and t_i, is the total number of visits to misinformation websites i. Then, this number is converted to a percentage. We calculate this effect for each social media platform s.

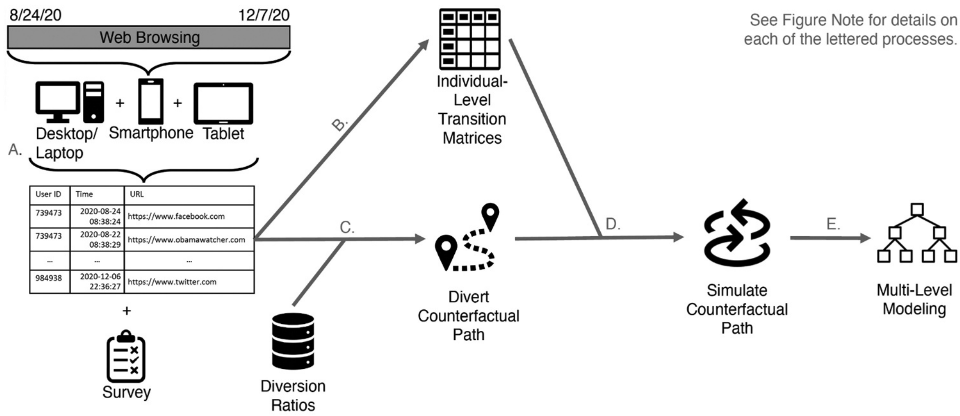

Figure 2. Study diagram visually displaying the analysis steps. Caption: The provided diagram represents the conceptual framework of the study, detailing the data collection and analysis process from web browsing to multi-level modeling. A. *Data Collection*: Web browsing data was captured across Desktop/ Laptop Computers, Smartphones, and Tablets from 24 August 2020 to 7 December 2020. The data includes timestamps and URLs visited by each user, uniquely identified by a User ID. Survey data was also collected on each individual to capture information like sociodemographics (e.g., political partisanship). B. *Individual-Level Transition Matrices*: Transition matrices were computed for each individual, capturing transition probabilities (i.e., probability of moving from one website to another). C. *Divert Counterfactual Path*: If an observed path was about to arrive at a given social media platform of interest (e.g., Facebook), the path is diverted to a different social media platform using experimentally derived diversion ratios (Aridor, 2023). D. *Simulate Counterfactual Path*: If diverted in step C, the counterfactual path is simulated to completion via the individuals' transition matrix. E.*Multi-Level Modeling*: Multi-level models were fit on the counterfactual data to assess the "effect" of counterfactually removing a given social media platform on visits to misinformation websites compared to the observed world.

Embedded Referrer Paradigm

We operationalize the embedded referrer paradigm by modeling and simulating counterfactual browsing sessions and observing how the number of visits to misinformation websites changes in these counterfactual paths compared to actual paths (see Figure 2). This process consists of three main steps: 1. Diverting counterfactual paths once a social media platform we want to create counterfactual paths for is visited, 2. Continuing paths based on individuals' observed transition probabilities, 3. Fitting multi-level models to estimate the change in the number of visits to misinformation websites. Below, we describe each of these three steps in greater detail.

Diverting Counterfactual Path

For each of our 395,936 paths of browsing sessions, we simulate counterfactual paths in scenarios in which each of the seven social media platforms cannot be visited. We sequentially proceed along each step (i.e., visited state) in the path. If the path does not include the social media platform we want to remove, the path continues as it did in the observed world. If a path attempts to move to the state we seek to remove, we divert the path by sampling from a distribution of diversion ratios for each platform. Social media diversion ratios are estimates of what social media individuals would use in the absence of another platform (e.g., if YouTube access is restricted, what proportion of those visits go to different platforms). Our diversion ratios come from Aridor (2023), who conducted an experiment in which users could not access a social media platform (e.g., Instagram) and observed what platforms they visited instead. Aridor (2023) provides diversion ratios for each of the platforms (Facebook, Twitter, Instagram, Reddit, YouTube, Snapchat, TikTok). By using experimentally derived diversion ratios to redirect participants in the absence of social media platforms, we follow recent calls from scholars to inform simulations of human behavior with experimentally derived estimates of behaviors foundational to conducting those simulations (Hammond & Barkin, 2024; Schwabl et al., 2024).

For example, to create counterfactual paths without Facebook, we sequentially step through each website visit in a session. When a path's next step is Facebook, we prevent the path from continuing to Facebook and instead determine the path's next step by sampling from Aridor's (2023) estimates of Facebook's diversion ratios. In the next step, the path is continued by simulating a path from the individuals' estimated transition probabilities (that excludes Facebook) until the path naturally arrives at a "path end" state or the path reaches the maximum number of steps in a path for the given person's observed data.

Individual-Level Transition Matrices

For each individual, we also estimate their transition probabilities via a multinomial logit demand model. This is a demand model in that it predicts choices individuals make as they move across the internet. It is a logit model because the choice probabilities all have the logit functional form; the model arises as a random utility model with error terms following a Gumbel distribution (Luce, 1959). It is multinomial because there are multiple possible outcomes (i.e., websites) that individuals can move to. These individual-level matrices, opposed to a single transition matrix for the entire sample, allows each individual's counterfactual paths to be unique to them. In other words, if an individual has a propensity to visit misinformation websites via multiple platforms, these transition matrices consider this individual-level factor that the person may find the misinformation website regardless of a single platform's role.

Let M denote a choice set consisting of (i) the seven social media platforms, (ii) the three misinformation website categories, and (iii) the seven remaining website categories. Let s and e denote additional nodes we refer to as "path start" and "path end" which represent the beginning and the ending of a browsing session, respectively. Then the *transition probability* of moving from $j \in M \cup stok \in M \cup e \, for \, individuals \, is$

$$Pr_i(j \to k) = \frac{e^{u_{ijk}}}{\sum_{m \in M \cup e} e^{u_{ijm}}} (2) where_{u_{ijk}} = \alpha_{ijk}$$

We normalize $u_{ijk} \, for \, k = e$ to zero. Thus, we are estimating the probability of transitioning from j to k for each individual i

We form an estimate $\hat{\alpha}_{jk}$ that matches predicted probabilities to the individuals' averages corresponding to these probabilities between all pairs of website categories we observe in our data. We use these probabilities to fill in a *transition matrix* T_i between the possible "states" of the user. In this matrix, each row represents a website category j and the columns represent a website category k. This transition matrix represents the probabilities of how the individual i moves from website category to website category. The transition matrix T_i is a representation of the transitions between websites in the world as we observed it in our data.

These transition probabilities are used when the counterfactual paths are diverted. Once a counterfactual path is diverted, the path is "continued" via the individual i's transition matrix T_i. For example, in a scenario in which a person may not visit Facebook, if the counterfactual path is diverted to Instagram, the path continues from the Instagram state via the transition probabilities in the individuals' transition matrix. Importantly, the social media platform of interest is removed from the participants' transition matrix, ensuring their counterfactual path cannot visit the platform later in the path. Steps along the counterfactual paths are simulated until the path naturally hits the "path end" state. Alternatively, we also end paths when the user has reached the maximum number of steps observed in their actual browsing behavior. Paths are also ended if an individual is diverted to a platform not in their transition matrix. This process of creating counterfactual paths, diverting the path if and when the path is about to arrive at the social media platform of interest and continues based on the individual's transition matrix is completed for each of the 395,936 paths for each of the seven social media platforms, resulting in 3,167,488 total paths (including the original paths).

Multi-Level Modeling

To estimate the difference in the number of visits to misinformation websites in the counterfactual paths compared to the actual paths, we fit a multi-level regression model at the path level, with all 3,167,488 observed and counterfactual paths in the model. To account for the hierarchical nature of the data, we cluster by participants and original paths (i.e., the original path by which the seven counterfactual paths were derived). Thus, each observation represents either the actually observed paths or one of the seven counterfactual scenarios in which a given social media platform cannot be visited. These scenarios are represented as binary variables. For example, the *NoFacebook* scenario variable is a "1" if the path represents the counterfactual scenario in which Facebook may not be visited and a "0" in all other instances. Formally, our outcome model is

$$\begin{aligned} Y_{ip} = {} & (\beta_0 + u_{0i} + u_{1p}) + \beta_1 \text{NoFacebook} + \beta_2 \text{NoInstagram} + \beta_3 \text{NoReddit} \\ & + \beta_4 \text{NoSnapchat} + \beta_5 \text{NoTikTok} + \beta_6 \text{NoTwitter} + \beta_7 \text{NoYouTube} + \varepsilon_{ip} \end{aligned} \tag{3}$$

This model predicts Y_{ip}, the number of visits to misinformation websites (we estimate separate models for all misinformation websites, as well as conservative, liberal, and non-ideological misinformation websites) in a path p for individual i. The model is fit with Maximum Likelihood (ML), meaning that the parameter estimates are obtained by finding the values that maximize the probability of the observed data given the model. The term β_0 is the overall intercept, representing the expected value of Y_{ip} in the actually observed paths, since the observed paths do not have any of the binary scenario variables, and thus serves as the reference for variable comparisons. The random effects, i.e., the individual- and path-specific deviations, $u_{0i}\,and\,u_{1p}$, allow the intercept to vary by individual $U1_P$ allows the intercept to vary by individual i and path p. These random effects are assumed to follow a normal distribution with zero means, estimated variances, and a covariance of

$$u_{0i}, u_{1p} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{0i}^2 & \sigma_{01,ip} \\ \sigma_{01,ip} & \sigma_{1p}^2 \end{bmatrix}\right).$$

β_1 *through* β_7 are the fixed effect coefficients for the binary variables representing the counterfactual scenarios in which a given social media platform cannot be visited. For example, $\beta_1 NoFacebook$ represents the scenarios in which Facebook cannot be visited. Therefore, each of the coefficients are the main coefficients of interest and represent the "effect" of removing each of the platforms on the outcome Y_{ip} (in this case, the number of visits in a path). The term ε_{ip} is the residual error term for each observation, capturing the deviation of the observed outcome from the predicted value based on the model.

To make the output of these models comparable to the output of referral analyses, for each of the social media coefficients we calculate what proportion of the intercept it represents, giving us the change in the number of misinformation visits in each counterfactual scenario without a given social media platform compared to the observed world. We also rerun the same analysis but break participants down by Republicans, Democrats, and Independents. We present the non-translated model outputs in Supplemental Materials C. One benefit to our approach compared to referral analyses is that it also allows us to control for other variables, such as sociodemographics, which we present as a robustness check in Supplemental Materials D.

Another benefit of our approach is that it allows us to translate the counterfactual paths to measure the effect of the counterfactual scenarios on other metrics of interest to extant research. For example, two frequently estimated quantities in the literature on misinformation exposure are the percentage of people exposed to at least one misinformation website and the average number of exposures among the exposed (A. Guess et al., 2020; R. C. Moore et al., 2023). Under our methodological approach, we calculate these numbers under the observed world and each of the counterfactual scenarios by summarizing to the individual level and fitting a model similar to Equation (3):

$$\begin{aligned} Y_i = (\beta_0 + u_{0i}) + \beta_1 NoFacebook + \beta_2 NoInstagram + \beta_3 NoReddit + \beta_4 NoSnapchat \\ + \beta_5 NoTikTok + \beta_6 NoTwitter + \beta_7 NoYouTube + \varepsilon_i \end{aligned} \quad (4)$$

This model operates at the individual-scenario level (i.e., not at the path level) and similar to Equation (3) is fit using Maximum Likelihood (ML). We estimate one set of models where Y_i is a binary variable "1" or "0" representing whether the individual was exposed to at least one misinformation website across *all* the paths in their counterfactual scenarios. Thus, the coefficients here represent the change in the proportion of individuals who were exposed to at least one misinformation website under the counterfactual scenario without the given social media platform. We fit another set of models via (ML) where we only include individuals who were exposed to at least one misinformation website and fit the model with Y_i being the number of visits to misinformation websites. Therefore, the coefficients represent the change in the average number of visits to misinformation websites among those who visited at least one misinformation website when the given social media platform is inaccessible.

Results

Exposure to Misinformation Under the Direct and Embedded Referrer Paradigms

We present the results of the direct and embedded referrer paradigms in Table 1, finding across platforms that the direct referrer paradigm consistently underestimates the effects of social media removal on misinformation exposure compared to estimates from the embedded referrer paradigm. In examining visits to all misinformation websites, the direct referrer paradigm suggests that Facebook directly preceded 9.25% of visits to all misinformation sites. In contrast, the Embedded Referrer Paradigm suggests a larger impact, with 14.07% (95% CI = 10.96, 17.19) more visits to all misinformation websites in a world with Facebook compared to the simulated world without Facebook. The magnitude of the difference is even greater for Twitter, with the direct referrer analysis indicating that Twitter directly preceded 2.17% of visits to misinformation websites, but the embedded referrer paradigm reveals that the world with Twitter results in 12.27% (95% CI = 9.16, 15.39) more visits to misinformation websites compared to the simulated world without Twitter. For YouTube, the direct referrer paradigm indicates that YouTube directly preceded 2.08% of visits to misinformation sites. In contrast, the embedded referrer paradigm shows that the world with YouTube had 4.92% (95% CI = 1.81, 8.04) more visits to misinformation websites than the simulated world without YouTube.

The results generally hold when examining conservative, liberal, and non-ideological misinformation websites. For example, Facebook directly preceded 11.43% of visits to liberal misinformation websites, but the world with Facebook had 58.56% (95% CI = 40.07, 77.04) more visits to liberal misinformation websites than the simulated world without Facebook. Examining Twitter and YouTube also revealed large discrepancies between the direct and embedded referrer analyses for liberal misinformation. Reddit only directly preceded 0.20% of visits to non-ideological misinformation. However, the observed world with Reddit had 18.22% (95% CI = 6.02, 30.42) more visits to non-ideological misinformation than the simulated world without Reddit, showing that the embedded referrer analysis can reveal how removing different platforms may differentially impact visits to various types of misinformation websites in a way that is not apparent under the direct referrer paradigm. Across partisans, results show differences between the direct and embedded referrer paradigms for both congenial and cross-cutting misinformation. For example, among Republicans (Table 2), Twitter directly preceded 1.99% of visits to congenial conservative misinformation, but the world with Twitter had 11.14% (95% CI = 7.55, 14.72) more visits to congenial conservative misinformation than the world without Twitter. Or, among Democrats (Table 3), YouTube directly preceded 2.06% of visits to congenial liberal misinformation websites. However, there were 32.09% (95% CI = 11.58, 52.59) more visits to congenial liberal misinformation websites in the world with YouTube compared to the simulated world without YouTube. Given these substantial differences, previous research operating under the direct referrers paradigm (e.g., Cardenal et al., 2019; Flaxman et al., 2016) to assess selective exposure would underestimate the effects of removing these different platforms on partisans' exposure to congenial misinformation. There are also notable differences when examining cross-cutting misinformation. For example, among Democrats, Facebook directly preceded 11.87% of visits to conservative misinformation websites, but the embedded referrer paradigm suggests that the world with Facebook results in 21.51% (95% CI = 14.68, 28.34) more visits to this cross-cutting misinformation than the world without Facebook.

Among Independents (Table 4), Reddit directly preceded 0.40% of visits to non-ideological misinformation websites, but the world with Reddit had a staggering 55.14%

Table 1. Direct versus embedded referrer paradigm results.

	All Misinformation			Conservative Misinformation			Liberal Misinformation			Non-Ideological Misinformation		
Platform	Direct Referrers Est.	Embedded Referrers Est.	CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI
Facebook	−9.25%	−14.07%	[−17.19%, −10.96%]	−9.40%	−10.68%	[−13.82%, −7.54%]	−11.43%	−58.56%	[−77.04%, −40.07%]	−3.96%	−18.79%	[−30.99%, −6.59%]
Instagram	−0.14%	−0.88%	[−4.00%, 2.23%]	−0.14%	−1.10%	[−4.24%, 2.04%]	0.00%	0.02%	[−18.47%, 18.50%]	−0.26%	−0.38%	[−12.58%, 11.82%]
Reddit	−0.16%	−1.82%	[−4.94%, 1.29%]	−0.18%	−0.96%	[−4.10%, 2.18%]	0.00%	−0.18%	[−18.67%, 18.30%]	−0.20%	−18.22%	[−30.42%, −6.02%]
Snapchat	0.00%	0.21%	[−2.91%, 3.32%]	0.00%	−0.02%	[−3.16%, 3.12%]	0.00%	0.09%	[−18.39%, 18.58%]	0.00%	−0.27%	[−12.47%, 11.93%]
TikTok	0.00%	0.12%	[−2.99%, 3.24%]	0.00%	0.06%	[−3.08%, 3.20%]	0.00%	0.02%	[−18.46%, 18.51%]	0.00%	−0.77%	[−12.97%, 11.43%]
Twitter	−2.17%	−12.27%	[−15.39%, −9.16%]	−2.06%	−10.70%	[−13.84%, −7.56%]	−3.59%	−24.22%	[−42.71%, −5.74%]	−1.52%	−23.80%	[−36.00%, −11.60%]
YouTube	−2.08%	−4.92%	[−8.04%, −1.81%]	−2.16%	−3.50%	[−6.64%, −0.36%]	−1.93%	−29.30%	[−47.79%, −10.82%]	−1.32%	−2.32%	[−14.52%, 9.88%]

Note: $N_{participants}$= 1,240. $N_{sessions}$ = 395,936. Values represent the implied percentage reduction of visits to misinformation websites if a given platform were removed. Est. = estimate. CI = Confidence Interval (95% confidence interval).

Table 2. Direct versus embedded referrer paradigm results among Republicans.

	All Misinformation			Conservative Misinformation			Liberal Misinformation			Non-Ideological Misinformation		
Platform	Direct Referrers Est.	Embedded Referrers Est.	CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI
Facebook	−9.66%	−9.45%	[−12.98%, −5.91%]	−9.70%	−9.53%	[−13.12%, −5.95%]	−15.15%	−10.87%	[−35.43%, 13.70%]	−7.61%	−4.71%	[−16.30%, 6.88%]
Instagram	−0.15%	−1.03%	[−4.56%, 2.51%]	−0.14%	−1.08%	[−4.67%, 2.50%]	0.00%	−5.41%	[−29.97%, 19.16%]	−0.41%	2.11%	[−9.48%, 13.70%]
Reddit	−0.15%	−0.83%	[−4.36%, 2.71%]	−0.15%	−0.97%	[−4.55%, 2.62%]	0.00%	0.06%	[−24.50%, 24.63%]	−0.21%	5.80%	[−5.79%, 17.39%]
Snapchat	0.00%	−0.01%	[−3.55%, 3.53%]	0.00%	−0.02%	[−3.61%, 3.56%]	0.00%	2.79%	[−21.77%, 27.36%]	0.00%	0.28%	[−11.31%, 11.87%]
TikTok	0.00%	0.04%	[−3.49%, 3.58%]	0.00%	0.05%	[−3.53%, 3.64%]	0.00%	0.06%	[−24.50%, 24.63%]	0.00%	−0.51%	[−12.10%, 11.08%]
Twitter	−1.99%	−11.11%	[−14.65%, −7.58%]	−1.99%	−11.14%	[−14.72%, −7.55%]	0.00%	−24.51%	[−49.08%, 0.05%]	−2.06%	−8.13%	[−19.72%, 3.45%]
YouTube	−2.22%	−3.92%	[−7.46%, −0.39%]	−2.21%	−3.79%	[−7.37%, −0.21%]	−3.03%	0.08%	[−24.49%, 24.64%]	−2.47%	−10.47%	[−22.06%, 1.12%]

Note: $N_{republicans}$ = 415. $N_{sessions}$= 129,606. Values represent the implied percentage reduction of visits to misinformation websites if a given platform were removed. Est. = estimate. CI = Confidence Interval (95% confidence interval).

Table 3. Direct versus embedded referrer paradigm results among Democrats.

	All Misinformation			Conservative Misinformation			Liberal Misinformation			Non-Ideological Misinformation		
Platform	Direct Referrers Est.	Embedded Referrers Est.	CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI
Facebook	−10.25%	−43.00%	[−54.37%, −31.63%]	−11.87%	−21.51%	[−28.34%, −14.68%]	−11.67%	−62.51%	[−83.02%, −42.00%]	−2.48%	−21.86%	[−34.50%, −9.22%]
Instagram	−0.15%	−1.29%	[−12.66%, 10.08%]	−0.43%	−2.63%	[−9.45%, 4.20%]	0.00%	−1.14%	[−21.65%, 19.37%]	−0.38%	−2.01%	[−14.65%, 10.63%]
Reddit	−0.27%	−0.88%	[−12.25%, 10.49%]	−1.30%	−2.00%	[−8.82%, 4.83%]	0.00%	−0.68%	[−21.19%, 19.83%]	0.00%	−2.01%	[−14.65%, 10.62%]
Snapchat	0.00%	−0.06%	[−11.43%, 11.31%]	0.00%	0.04%	[−6.78%, 6.87%]	0.00%	−0.79%	[−21.30%, 19.72%]	0.00%	0.04%	[−12.59%, 12.68%]
TikTok	0.00%	−0.28%	[−11.65%, 11.09%]	0.00%	−0.27%	[−7.09%, 6.56%]	0.00%	−0.91%	[−21.42%, 19.60%]	0.00%	−0.70%	[−13.34%, 11.93%]
Twitter	−3.51%	−18.54%	[−29.91%, −7.17%]	−3.91%	−11.91%	[−18.74%, −5.08%]	−3.78%	−24.75%	[−45.26%, −4.25%]	−1.90%	−12.50%	[−25.14%, 0.13%]
YouTube	−1.72%	−18.82%	[−30.19%, −7.45%]	−1.59%	−4.51%	[−11.34%, 2.32%]	−2.06%	−32.09%	[−52.59%, −11.58%]	−0.57%	−4.63%	[−17.27%, 8.01%]

Note: $N_{democrats} = 672$. $n_{sessions} = 220{,}445$. Values represent the implied percentage reduction of visits to misinformation websites if a given platform were removed. Est. = estimate. CI = Confidence Interval (95% confidence interval).

Table 4. Direct versus embedded referrer paradigm results among Independents.

	All Misinformation			Conservative Misinformation			Liberal Misinformation			Non-Ideological Misinformation		
Platform	Direct Referrers Est.	Embedded Referrers Est.	CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI	Direct Referrers Est.	Embedded Referrers Est.	Embedded Referrers CI
Facebook	−4.73%	−23.47%	[−31.33%, −15.61%]	−5.29%	−22.00%	[−27.11%, −16.90%]	−7.50%	−31.45%	[−46.36%, −16.54%]	−1.98%	−26.20%	[−57.81%, 5.41%]
Instagram	−0.04%	−0.46%	[−8.32%, 7.40%]	−0.06%	−0.97%	[−6.08%, 4.14%]	0.00%	3.07%	[−11.84%, 17.98%]	0.00%	0.23%	[−31.38%, 31.83%]
Reddit	−0.08%	−12.53%	[−20.39%, −4.67%]	0.00%	−0.79%	[−5.90%, 4.32%]	0.00%	−5.36%	[−20.27%, 9.55%]	−0.40%	−55.14%	[−86.75%, −23.53%]
Snapchat	0.00%	0.02%	[−7.84%, 7.88%]	0.00%	0.00%	[−5.11%, 5.11%]	0.00%	0.01%	[−14.90%, 14.92%]	0.00%	0.01%	[−31.60%, 31.62%]
TikTok	0.00%	0.20%	[−7.66%, 8.07%]	0.00%	0.30%	[−4.81%, 5.41%]	0.00%	−0.76%	[−15.67%, 14.15%]	0.00%	0.02%	[−31.59%, 31.62%]
Twitter	−1.61%	−17.27%	[−25.13%, −9.41%]	−1.88%	−7.74%	[−12.85%, −2.63%]	−1.87%	−23.01%	[−37.92%, −8.10%]	−0.59%	−48.64%	[−80.25%, −17.03%]
YouTube	−1.61%	0.33%	[−7.54%, 8.19%]	−1.94%	−0.73%	[−5.84%, 4.38%]	0.00%	−13.04%	[−27.95%, 1.87%]	−0.99%	7.56%	[−24.05%, 39.17%]

Note: $N_{independents} = 146$. $n_{sessions} = 44{,}172$. Values represent the implied percentage reduction of visits to misinformation websites if a given platform were removed. Est. = estimate. CI = Confidence Interval (95% confidence interval).

(95% CI = 23.53, 86.75) more visits to non-ideological misinformation sites than the scenario without Reddit.

In comparing the two paradigms, we generally find that analyses under the embedded paradigm yields larger "effects" compared to the direct referrer paradigm for platforms with more visits in our dataset (i.e., Facebook, Twitter, YouTube). In contrast, platforms with fewer web visits in our dataset (i.e., Instagram, TikTok, Snapchat) demonstrate statistically similar results between the two paradigms. One platform (Reddit) produces mixed results. However, when examining differences by device use (see Supplemental Materials K, L, and M; Tables S29, S32, S35), we find distinct patterns in which platform removals drive these effects. For desktop users, removing Facebook and Twitter shows the strongest effects on misinformation exposure, while for smartphone users, removing Facebook and YouTube shows the largest effects. These device-specific differences are particularly pronounced for liberal and non-ideological misinformation, where significant platform removal effects are observed primarily among desktop users.

While the above results are informative of platforms' role in misinformation exposure, we also examined the ratio of high-quality news to misinformation visits to quantify the relative impact of platform removal on different types of content (see Table 5). This ratio, derived from the point estimates of the embedded referer models (for high-quality news websites model results, see Supplemental Materials G; for low-quality news websites model results, see Supplemental Materials H), represents how many high-quality news visits would be reduced for each reduced misinformation visit in a world without a given platform. Across platforms, we find that removing access would disproportionately affect high-quality news exposure compared to misinformation exposure. For Facebook, the ratio is 73.3, indicating that for every reduced misinformation visit, there would be 73.3 fewer visits to high-quality news sites. Other platforms show even larger ratios, such as Reddit (317.0). These ratios vary substantially by content ideology. For conservative content, Facebook shows a ratio of 0.5, indicating that removal would reduce conservative misinformation visits more than conservative high-quality news visits. In contrast, platforms show larger positive ratios for liberal content (e.g., Reddit: 328.0) and especially non-ideological content (Facebook: 917.9, Instagram: 10,354.0), indicating that removal would have a much stronger effect on high-quality news exposure than misinformation exposure for these content types. For more details, see Supplemental Materials I.

Table 5. Embedded Referrers paradigm results for ratio of high-quality news to misinformation websites.

Platform	All High: Misinformation	Conservative High: Misinformation	Liberal High: Misinformation	Non-Ideological High: Misinformation
NoFacebook	73.3	0.5	4.3	917.9
NoInstagram	160.2	−18.3	−Inf	10,354.0
NoReddit	317.0	13.7	328.0	539.0
NoSnapchat	−12.1	25.0	−Inf	2099.0
NoTiktok	−290.3	121.3	−Inf	1114.7
NoTwitter	78.0	5.2	11.6	645.5
NoYoutube	101.2	4.4	−0.8	3751.6

Note: $N_{participants}$ = 1,240. $N_{sessions}$= 395,936. Values are the ratio of the coefficients of high-quality news to misinformation. The scenario variables represent the change in the number of exposures per path in the scenario without the given social media platform. "-Inf" indicates the ratio is undefined: the estimated change in misinformation exposure (the denominator) was zero for these platforms, so dividing the (negative) high-quality-news estimate by zero yields negative infinity.

Proportion of Individuals Exposed to Misinformation

The embedded referrer paradigm enables the calculation of alternative metrics that are central in the extant misinformation literature; for example, the proportion of individuals exposed to at least one misinformation website (A. Guess et al., 2020; R. C. Moore et al., 2023) (Table 6). Our model estimates indicate that 43.5% of participants were exposed to at least one misinformation website in the observed world, though this varies substantially by device type (45.8% for desktop users vs 31.3% for mobile users; see Supplemental Materials K, L, and M; Tables S30, S33, S36). In the counterfactual scenario without Facebook, the overall percentage exposed drops by 6.4% (95% CI = 5.5, 7.3) to 37.1%, with this effect driven primarily by desktop users, where removing Facebook reduces exposure by 8.0% compared to minimal effects on mobile. Similarly, while removing Twitter and YouTube shows significant effects on exposure in desktop browsing, the impact of their removal is minimal in mobile-only paths. These effects also vary by partisan identity (see Supplemental Materials E, Tables S9-S11), with Republicans showing the highest baseline exposure (55.4%) compared to Democrats (35.9%) and Independents (45.2%). While the effects of removing Facebook are relatively consistent across partisan groups (reducing exposure by 6.7% for Republicans, 6.0% for Democrats, and 6.8% for Independents), removing other platforms shows asymmetric effects. Democrats show significant reductions in exposure from removing both Twitter (−2.7%) and YouTube (−3.0%), while removing these platforms shows no significant effects on overall exposure among Republicans or Independents.

Exposures to Misinformation Among the Exposed

Similarly, we can calculate counterfactual changes among another metric of interest to the extant misinformation literature, the average number of exposures to misinformation among those exposed (A. Guess et al., 2020; R. C. Moore et al., 2023) (Table 7). For example, individuals who visited at least one misinformation website were exposed to 44.2 misinformation websites (95% CI = 29.7, 58.8). Then, in the scenario without Facebook, those exposed had 6.2 (95% CI = 3.6, 8.9) fewer visits. These effects vary substantially by device type (see Supplemental Materials K, L, and M; Tables S31, S34, S37). Desktop users show both higher baseline exposure (49.5 visits vs 22.9 visits for mobile) and stronger effects of platform removal, with removing Facebook reducing visits by 7.4 visits on desktop compared to 1.7 visits on mobile. While removing Twitter significantly reduces exposure on desktop (−6.7 visits for all misinformation), the effects of its removal are not significant on mobile. Conversely, removing YouTube shows significant effects on mobile (−1.3 visits) but not desktop. The intensity of exposure also varies markedly by partisan identity (see Supplemental Materials F, Tables S12-S14), with Republicans showing substantially higher baseline exposure (79.1 visits) compared to Democrats (13.7 visits) and Independents (35.8 visits). The effects of removing Facebook are strongest among Republicans (−7.6 visits) and Independents (−8.0 visits), while the effects of removing Twitter are primarily concentrated among Republicans (−9.0 visits for all misinformation).

Table 6. Proportion of individuals exposed to at least one misinformation website (embedded referrer paradigm).

	All Misinformation		Conservative Misinformation		Liberal Misinformation		Non-Ideological Misinformation	
Term	Est.	CI	Est.	CI	Est.	CI	Est.	CI
Fixed Effects								
β_0 (Constant)	0.435***	[0.408, 0.463]	0.094***	[0.078, 0.109]	0.094***	[0.078, 0.109]	0.148***	[0.128, 0.167]
β_1 NoFacebook	−0.064***	[−0.073, −0.055]	−0.023***	[−0.028, −0.017]	−0.023***	[−0.028, −0.017]	−0.030***	[−0.037, −0.023]
β_2 NoInstagram	−0.004	[−0.013, 0.005]	−0.002	[−0.007, 0.004]	−0.002	[−0.007, 0.004]	−0.003	[−0.010, 0.004]
β_3 NoReddit	−0.009	[−0.018, 0.000]	−0.003	[−0.009, 0.002]	−0.003	[−0.009, 0.002]	−0.005	[−0.012, 0.002]
β_4 NoSnapchat	0.000	[−0.009, 0.009]	−0.000	[−0.006, 0.006]	0.000	[−0.006, 0.006]	0.000	[−0.007, 0.007]
β_5 NoTikTok	−0.001	[−0.010, 0.008]	−0.001	[−0.007, 0.005]	−0.001	[−0.007, 0.005]	−0.001	[−0.008, 0.006]
β_6 NoTwitter	−0.018***	[−0.027, −0.009]	−0.015***	[−0.020, −0.009]	−0.015***	[−0.020, −0.009]	−0.009*	[−0.016, −0.002]
β_7 NoYouTube	−0.019***	[−0.028, −0.010]	−0.005	[−0.011, 0.001]	−0.005	[−0.011, 0.001]	−0.023***	[−0.030, −0.016]
Random Effects								
sd(u_0)	0.480	[0.461, 0.500]	0.273	[0.262, 0.284]	0.273	[0.262, 0.284]	0.334	[0.321, 0.348]
Model Fit								
LogLikelihood	4439		9001		9001		7034	
AIC	−8858		−17982		−17982		−14048	
BIC	−8786		−17910		−17910		−13976	

Note: $N_{participants}$ = 1,240. The Constant represents the proportion of participants exposed to at least one misinformation website in the observed world. The scenario variables represent the change in the proportion of individuals exposed in the scenario without the given social media platform. Est. = estimate. CI = Confidence Interval (95% confidence interval). ***$p < .001$. **$p < .01$. *$p < .05$. p-values are two-sided.

Table 7. Average number of exposures to misinformation websites among those exposed to at least one misinformation website (embedded referrer paradigm).

	All Misinformation		Conservative Misinformation		Liberal Misinformation		Non-Ideological Misinformation	
Term	Est.	CI	Est.	CI	Est.	CI	Est.	CI
Fixed Effects								
β_0 (Constant)	44.244***	[29.725, 58.763]	19.681**	[5.608, 33.754]	19.681**	[5.608, 33.754]	8.285***	[4.223, 12.348]
β_1 NoFacebook	−6.217***	[−8.855, −3.579]	−7.353**	[−12.547, −2.160]	−7.354**	[−12.547, −2.160]	−1.422	[−3.191, 0.347]
β_2 NoInstagram	−0.433	[−3.071, 2.205]	−0.001	[−5.195, 5.193]	−0.000	[−5.193, 5.193]	−0.012	[−1.781, 1.757]
β_3 NoReddit	−0.846	[−3.484, 1.792]	−0.027	[−5.220, 5.167]	−0.027	[−5.220, 5.167]	−1.379	[−3.147, 0.390]
β_4 NoSnapchat	−0.004	[−2.642, 2.634]	0.008	[−5.186, 5.201]	0.008	[−5.186, 5.201]	0.004	[−1.765, 1.773]
β_5 NoTikTok	0.011	[−2.627, 2.649]	−0.001	[−5.194, 5.193]	−0.001	[−5.195, 5.193]	−0.034	[−1.803, 1.734]
β_6 NoTwitter	−5.428***	[−8.066, −2.790]	−3.043	[−8.237, 2.151]	−3.044	[−8.237, 2.150]	−1.805*	[−3.573, −0.036]
β_7 NoYouTube	−2.206	[−4.843, 0.432]	−3.681	[−8.875, 1.513]	−3.680	[−8.874, 1.514]	−0.160	[−1.929, 1.609]
Random Effects								
sd(u_0)	170.660	[160.759, 181.170]	74.542	[65.464, 84.878]	74.542	[65.464, 84.878]	26.651	[24.023, 29.566]
Model Fit								
LogLikelihood	−21170		−4377		−4377		−5629	
AIC	42,360		8773		8773		11,279	
BIC	42,423		8822		8822		11,332	

Note: $N_{participants\ All} = 540$. $N_{participants\ Conservative} = 116$. $N_{participantsLiberal} = 116$. $N_{participantsNonIdeological} = 183$. The Constant represents the average number of exposures among those who were exposed to at least one misinformation website. The scenario variables represent the change in the average number of exposures in the scenario without the given social media platform. Est. = estimate. CI = Confidence Interval (95% confidence interval). ***$p < .001$. **$p < .01$. *$p < .05$. *p*-values are two-sided.

Discussion

Summary of Findings

The effects of removing social media platforms on citizens' exposure to political misinformation differ dramatically between the direct and embedded referrer paradigms. While previous political communication research has focused on direct exposure pathways, our analysis reveals a more complex picture of how the removal of platforms influences political information environments. The direct referrer paradigm suggests a modest role for social media platforms in driving visits to misinformation websites (e.g., Twitter directly preceded 2.17% of misinformation visits in our data). However, our analysis under the more comprehensive embedded referrer paradigm reveals that the effects of removing social media on political information exposure are substantially larger than previously understood. By comparing simulated counterfactual worlds with and without seven different social media platforms, we find that the total effects of platform removal on political misinformation exposure are consistently larger than direct referral measures would suggest. For example, in the observed world, there were 12.27% more visits to misinformation websites than in the simulated world without Twitter, an effect over five times larger than Twitter's estimated role under the direct referrer paradigm. Although the general pattern of greater effects from platform removal held across platforms and misinformation types, the magnitude of effects varied meaningfully across partisan groups, suggesting the distinct effects of removing different platforms on exposure to ideologically aligned versus cross-cutting misinformation.

Beyond these misinformation exposure patterns, our embedded referrer paradigm reveals differential effects on information quality that have implications for understanding platforms' role in the broader information ecosystem. Our analysis of the ratio of high-quality news to misinformation visits demonstrates that platform removal would disproportionately affect high-quality news exposure compared to misinformation exposure. For Facebook, removing the platform would reduce high-quality news visits by 73.3 times for every reduced misinformation visit, while other platforms show even larger ratios (e.g., Reddit: 317.0). These ratios vary substantially by content ideology. For conservative content, Facebook shows a ratio of 0.5, indicating that removal would reduce visits to conservative misinformation more than visits to conservative high-quality news. In contrast, platforms show much larger ratios for liberal content (Reddit: 328.0) and especially non-ideological content (Facebook: 917.9, Instagram: 10,354.0). These findings reveal that social media platforms may function as conduits for high-quality news consumption and that their role in the information ecosystem extends well beyond their role in misinformation exposure. This pattern holds important implications for both the theoretical understanding of how platforms shape political information environments and for evaluating potential interventions aimed at reducing misinformation exposure.

Contributions

Our study makes four core contributions to political communication theory and methodology. First, we introduce the embedded referrer paradigm as a new framework for understanding how social media shapes citizens' exposure to political information. This paradigm advances political communication research by accounting for both direct and indirect pathways through which platforms influence information exposure, including (1) exposure through intermediary websites that direct to political misinformation (e.g., Facebook → blog post → misinformation website) and (2) activation of political information-seeking behaviors that persist after leaving the platform. The embedded referrer paradigm provides a more comprehensive theoretical account of how social media shapes political information environments than the direct referrer paradigm used in existing literature. The direct referrer paradigm, which focuses only on instances in which social media directly precedes visits to political information, does not account for the two possibilities for social media's additional influence. We argue that this conceptualization presents a limited view of platforms' roles and effects. While studies operating under the direct referrer paradigm often make conclusions that suggest more comprehensive documentation of social media's effects than the limited set of scenarios they ultimately analyze, our embedded referrer paradigm encapsulates a more comprehensive account. Furthermore, while our analysis is limited to social media's role in exposure to misinformation, the embedded referrer paradigm can be used to conceive of social media platforms' role in exposing people to any political information source.

Second, we use a simulation-based approach to operationalize the embedded referrer paradigm and reveal the extent to which conclusions about social media's role in leading to misinformation may differ depending on which paradigm is used. While studies under the direct referrer paradigm simply count instances of social media use that precede misinformation visits to attribute responsibility to social media, our analysis under the embedded referrer paradigm accounts for more of the ways social media can influence people's online behaviors. This more comprehensive approach compares misinformation visits in counterfactual worlds without social media.

Third, our empirical findings reveal how the direct referrer paradigm systematically underestimates the effects of social media removal on political misinformation exposure, with important variations across both platforms and partisan groups. While this underestimation is consistent across all seven platforms we study, the magnitude varies in theoretically meaningful ways. For certain combinations of platforms and misinformation types, the embedded referrer paradigm reveals effects 15+ times larger than those estimated under the direct referrer paradigm, suggesting that removing some platforms has particularly strong indirect effects on political information exposure. These platform-specific effects are further shaped by partisan dynamics. For instance, removing Twitter shows especially strong effects among Republicans compared to Democrats, while removing other platforms like Facebook shows more symmetric partisan effects. In terms of selective exposure dynamics, in which social media platforms systematically expose partisans to congenial misinformation, there are both important differences in effects across platforms (Tables 2 and 3, direct referrers estimates), important as many of the previous studies only focus on one platform (e.g., Bakshy et al., 2015; Shin, 2020), but perhaps more important are the drastic differences between the direct and embedded referrers paradigms for certain platforms. The embedded referrers estimate for the effect of removing Twitter on conservatives' exposure to congenial misinformation is over 5x the direct referrers estimate, and for the effect of removing YouTube on liberals' exposure to congenial misinformation is nearly 16x the direct referrers estimate. If one conceptualizes selective exposure as being consequential for misinformation exposure not only in exposing people directly to like-minded dubious content but also for creating a personalized information ecosystem that is likely to beget exposure to and belief in misinformation (Rhodes, 2022), the effects we estimate are an important contribution. Together, these findings advance our understanding of how different platforms may reinforce or challenge partisan information diets, while demonstrating that previous research has underestimated social media's overall influence on political misinformation exposure.

Finally, our fourth contribution is the potentially important evidence it offers for regulators and policy makers considering the role of platforms in the media ecosystem. There have been a number of policy proposals worldwide to ban some large social media platforms, such as a proposed ban of TikTok at the federal level in the United States (e.g., Shepardson, 2024). Our simulation approach offers some counterfactual evidence for policy makers to consider regarding the effect of such a ban on news consumption patterns. Both to illuminate policy implications, and as a robustness check, we calculate results comparable to our main misinformation findings for high-quality (see Supplemental Materials G) and low-quality news websites (see Supplemental Materials H), as well as the ratio of high-quality news to misinformation (see Supplemental Materials I) and low-quality news websites (see Supplemental Materials J). These additional results allow for comparisons for policy implications. Consider, for example, a scenario in which Facebook was banned. Our simulation data suggest that such a ban would reduce exposures to misinformation by

14.1% (see Table 1). However, for every one reduced exposure to misinformation, the ban would reduce visits to high-quality news by 73.3 times (see Supplemental Materials I, Table S23). Taken together, these data suggest that banning a platform like Facebook could potentially do more harm than good, from a news quality perspective. While the evidence from these simulations are obviously subject to important limitations, they offer otherwise unavailable counterfactual data for various regulatory options regarding restricting access to social media platforms.

Future Research

The present study inspires three broad directions for future research. First, future research should explore the mechanisms behind the embedded referrer paradigm. While our findings demonstrate that this more comprehensive paradigm attributes greater responsibility to social media in leading people to misinformation compared to the direct referrer paradigm, it is unclear how these influences occur. Platforms may expose people to intermediary websites (e.g., blog posts, forums) that then expose them to misinformation. They may also direct people to content that prompts psychological responses (e.g., priming a specific topic) that lead to searches or other online behaviors and ultimately to misinformation. Future work should also seek to illuminate the mechanisms behind these results, for example, if there are specific intermediate websites that may be pathways between social media and misinformation websites under the embedded referrer paradigm.

In addition, in seeking to explicate and provide an operational example of the direct versus embedded paradigms, we analyze social media platforms because of the emphasis extant literature has placed on social media platforms and their role in referring individuals to misinformation websites (e.g., Allcott & Gentzkow, 2017; Allen et al., 2020; Fox, 2018; Frenkel, 2021; A. Guess et al., 2020; Jones-Jang et al., 2021; Lee et al., 2023; Luk et al., 2021; R. C. Moore et al., 2023, 2024; Myers, 2022; Neyazi et al., 2022; Warner et al., 2021). However, our methodological framework can also be applied to any type of website. While the present research treats social media platforms as the counterfactual point of interest, future research can use the framework to treat social media as an intermediate state and conduct analyses to understand how other counterfactual scenarios may impact social media visits. In other words, social media and misinformation may be an exemplary case for our framework, but it is far from the only potential use of the framework.

Furthermore, future research should carefully analyze people's online journeys to understand the social media to misinformation pipeline more precisely; for example, via qualitative analysis of web browsing histories (e.g., Moore et al., 2024). Doing so not only refines our theoretical understanding of how social media shapes exposure to misinformation but may also suggest viable opportunities for interventions or policies to lessen social media's influence on misinformation consumption. Another avenue for future work is to combine the URL-level visit data with web scraping to capture the content on the pages people visited (Wojcieszak et al., 2024). This technique could provide further insight into the mechanisms of influence that alter the relationship between social media platforms and misinformation website visits. However, this web scraping task requires nuance due to systematic bias in the underlying technical infrastructure of different types of websites, like misinformation and news websites (Dahlke et al., 2025).

Second, future research should build upon our work by conducting experiments to infer causality more directly. To measure social media's total effects, researchers can manipulate whether individuals have access to particular platforms, similar to Allcott et al. (2022) paying individuals not to use Facebook, Aridor (2023) paying people not to use YouTube or Instagram, and Walsh et al. (2021) paying participants not to use Facebook, Twitter, Snapchat, or Instagram. This experimental elimination is notably similar to our in silico scenarios, where we remove platforms and measure the system's outcomes. These past studies examined other outcomes, including well-being (Allcott et al., 2022; Walsh et al., 2021) and social media use (Aridor, 2023). Still, future research could measure misinformation exposure as the primary outcome of interest. One advantage of our study is that, by using simulation, we could examine many different combinations of platforms and misinformation types, and future research can more closely examine specific platforms and misinformation.

Third, neither social media (Fan et al., 2024; Fu, 2023; Lee et al., 2022; Munger et al., 2024; Papathanassopoulos et al., 2013; Rossini, 2023; Scherman et al., 2022; Tong et al., 2021; Vaccari, 2021) nor misinformation is a U.S.-only phenomenon (Lukito, 2024). Thus, our study only covers a portion of the global relationship between social media and misinformation. Other research uses similar web-browsing data as the present study in contexts outside of the United States, including Brazil (e.g., Mont'alverne et al., 2024), France (e.g., Altay et al., 2022), Germany (e.g., Altay et al., 2022), India (e.g., Mukerjee, 2024), Netherlands (e.g., Wojcieszak et al., 2024), Poland (e.g., Wojcieszak et al., 2022, 2024), and the United Kingdom (e.g., Altay et al., 2022; Fletcher et al., 2023; Nielsen & Fletcher, 2022). These non-U.S. data can be used for analyses under the embedded referrer paradigm. However, a challenge to expanding the method to outside of the U.S. is that, to our knowledge, there are few equivalent experimental deactivation studies from outside the U.S. that can provide an estimates of diversion ratios, a key input to the present method (i.e., how we use the results from Aridor, 2023). Therefore, future experimental studies should expand on deactivation studies outside the U.S. to provide researchers with the inputs needed to conduct embedded referral analyses in contexts outside of the United States, particularly the Global South, to provide a more global, comprehensive view on the connection between social media platforms and misinformation websites. Moreover, to meet calls for more work that combines experimental and simulation-based studies (e.g., Hammond & Barkin, 2024; Schwabl et al., 2024), both inside and outside the United States, more experimental work needs to be conducted to estimate the substitution of social media in scenarios in which a given platform cannot be visited. More experimental studies can provide additional experimental inputs to the embedded referrer paradigm (beyond Aridor, 2023) and thus provide more robust estimates of the effect of social media on visits to different types of websites.

Limitations

First, the multinomial logit model makes IID (independently and identically distributed [error terms]), IIA (independence of irrelevant alternatives), and Markov assumptions. Similar work operates under these assumptions (e.g., Vermeer & Trilling, 2020). In addition, we do not explicitly model for substitution to nonsocial media use, something that recent work has found occurs when use of social media platforms is restricted (e.g., Allcott et al., 2024). For example, in a scenario where one could not use Facebook, one could possibly increase face-to-face interactions instead of fully substituting time to alternative social media. This assumption, while strong, likely means we underestimate the magnitude of our effect sizes. For example, spending time on a substituted social media platform likely comes with some probability of ultimately visiting a misinformation website whereas offline activities would have a lower probability of leading to such websites. Of course, these offline interactions could also expose people to more misinformation (e.g., misinformation from a family member), but it is unclear how these offline interactions could lead to more misinformation website visits. Future work can alleviate these concerns by explicitly modeling temporal interdependence between observations within individuals, implementing flexible choice models allowing for varying degrees of substitution (including to offline and nonsocial media), and including higher-order Markov models to consider more than a user's immediate past state.

Our data from YouGov's Pulse browser plugin has limitations. Our data's representativeness, gathered from U.S. residents via YouGov Pulse, is limited due to self-selection bias, impacting generalizability (Bosch et al., 2023; Gil-López et al., 2023). This issue is common in web-browsing studies requiring software installation or data donation (Casas et al., 2022; Guess, 2021; Stier et al., 2022; Weeks et al., 2021; Wojcieszak et al., 2021, 2021). The YouGov Pulse panel underrepresents individuals without a high school diploma and leans Democratic but broadly resembles the American adult population (A. Guess et al., 2020). Concerns about altered browsing behavior due to observation (Hawthorne Effect) seem minimal (Bail et al., 2018; Yeykelis et al., 2018), as evidenced by sensitive website visits, such as pornography, in our data. The software also records website visits in web browsers but not activities in standalone apps, a notable gap given the prevalence of app usage on smartphones (Wurmser, 2019).

Additionally, the URL data only captures instances leading to new URLs. For example, misinformation exposure while scrolling through a Facebook feed would not generate a new URL (Ellison et al., 2020) and would be unobserved. Relatedly, the device composition of our sample has important implications for understanding different platforms' roles in misinformation spread and the generalizability of our estimates. Our desktop-heavy sample shows substantially different patterns across devices (see Supplementary Materials K and L). Desktop users exhibit both higher baseline exposure rates (45.8% vs 31.3% for mobile) and more frequent visits when exposed (49.5 vs 22.9 visits on average). Moreover, platform effects are consistently stronger on desktop. For instance, removing Facebook reduces misinformation exposure by 7.4 visits on desktop but only 1.7 visits on mobile. However, this desktop-centric view may not fully capture the impact of mobile-first platforms. Platforms like TikTok, Instagram, and Snapchat show minimal effects in our analysis, but this may reflect our data collection limitations, as their primary mode of content discovery and sharing occurs through mobile app interactions that our browser-based tracking cannot capture. The differences between desktop and mobile results, particularly the weaker and fewer significant platform effects on mobile, suggest our findings may be most representative of desktop-based information consumption patterns. Given the prevalence of mobile social media use in the general population, our results should be interpreted as conservative estimates, particularly for mobile-centric platforms, and the total impact of social media on misinformation exposure may be larger than our analysis indicates.

In the present study, we discuss social media exposing partisans to congenial misinformation as evidence of selective exposure. While we note that this operationalization follows others who have explored dynamics of partisan exposure to news via web browsing data (Cardenal et al., 2019; A. Guess et al., 2020; Rao et al., 2022), important assumptions are being made that are necessary to consider when evaluating our findings. Specifically, participants in our data may have clicked on links to sites out of habit or due to being lured by certain features of content (e.g., clickbait titles, social endorsement cues), different from canonical selective exposure behavior of deliberately seeking congenial information to support existing beliefs. Without knowing participants' intentions as they browsed the web, it is hard to know which processes and motivations specifically drove the website visits we observed. One approach that may bolster digital trace data, such as web browsing data, is to pair it with targeted survey questions prompting deeper reflection and reporting from participants about their selective and incidental exposure activities (e.g., Shin, 2020). Regarding incidental exposure, we also do not make claims about incidental exposure dynamics in our data. We cannot know the exact processes of deliberate seeking vs. accidental or habitual exposures that precipitated news and misinformation exposures in our data. Following past work, we frame some of our findings in terms of selective exposure because of the concordance between user and misinformation outlet partisanship. Still, we again reiterate caution in interpreting those findings as a specific selective exposure behavior or process, as our behavioral data leave questions of user intention unanswered.

Furthermore, our data were collected during a politically consequential period: the 2020 U.S. Presidential Election, which coincided with the COVID-19 pandemic. This unique temporal context likely shaped both political information seeking and platform usage patterns. During this period, Americans possibly spent more time online and on social media, while simultaneously navigating two major challenges to democratic discourse: unprecedented electoral misinformation about voter fraud and widespread health-related falsehoods (Dahlke & Pan, 2024; Kasimov et al., 2023; Prochaska et al., 2023; Starbird et al., 2023). The convergence of these factors created a distinctive political communication environment where citizens faced competing narratives about both democratic institutions and public health. While studying this period provides valuable insights into how social media shapes exposure to political misinformation during moments of democratic stress, future work should examine whether these patterns hold during non-election periods and times of relative political stability (Munger, 2023). Such comparative analyses would help distinguish between platform effects that are specific to high-stakes political moments versus those that persist across different political contexts.

Finally, we operationalize misinformation exposure as visits to domains known for false and misleading news content, a common approach in research, in particular in work using the direct referrer paradigm (Allcott & Gentzkow, 2017; Allen et al., 2020; Grinberg et al., 2019; A. Guess et al., 2020; R. C. Moore et al., 2023). However, this approach is not without important limitations. Classifying misinformation websites at the domain-level means that we do not capture all misinformation exposures, such as misinformation found on generally credible sites (e.g., on published pages, in comments sections). While identifying misinformation at the webpage-level rather than the domain-level would be ideal, several persistent issues complicate identifying misinformation at the webpage-level. For one, it is notoriously difficult to use automated approaches to identify misinformation on the web at scale – for example Conti et al. (2017) found that classification F-scores never exceeded 0.7, a minimum standard for accurate classification. Moreover, manual annotation is not feasible considering the 21 million website visits in our data. Wisdom of Crowds (e.g., Allen et al., 2021) or large language model-based approaches may help make progress in this area in the future, but creating and validating such a technique would require a methodology paper on its own. Dahlke et al. (2025) also documented how content on misinformation websites was more difficult to scrape retrospectively compared to other website types, further complicating the identification of specific misinformation articles regardless of the specific approach taken. While these challenges, in addition to facilitating direct comparability with existing research using the direct referrer paradigm, are reasons why we took a domain-level approach to identifying misinformation in this paper, we acknowledge the important limitations of the approach. An important priority for research in this area will be to develop methods and tools to facilitate accurate, reliable, and replicable classification of misinformation at the webpage-level. Finally, and relatedly, our approach may overlook misinformation prevalent in emerging media such as in-app videos on platforms like TikTok. Future work should develop tools and databases for identifying misinformation across various formats, including images and videos (Peng et al., 2023).

Notes

1. Some studies have examined scenarios in which individuals cannot visit a given social media site to study various outcomes. These studies include experimentally restricting access to a given platform (e.g., Allcott et al., 2020, 2022, 2024; Aridor, 2023) or examining periods of time when a given platform was unavailable because of a company-side outage (e.g., Andersson & Danielsson, 2023; Liao & Sundar, 2022; Motta et al., 2023; Sekścińska & Jaworska, 2022; Shousha & Abdelgawad, 2021) or government-created social media blackout (e.g., Dunn, 2011; Rydzak, 2019; Shah et al., 2021). Other studies have manipulated on-platform content, such as newsfeeds, in various ways (e.g., Bond et al., 2012; Guess et al., 2023a, 2023b; Nyhan et al., 2023).
2. https://www.newsguardtech.com/

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This project was funded by a Multidisciplinary University Research Initiative (MURI) grant number [W911NF-20-1-0252] and support from Stanford's Freeman Spogli Institute for International Studies (no grant number). R.D. was supported by graduate fellowship awards from Knight-Hennessy Scholars and Stanford Data Science at Stanford University. R.C.M. was supported by a Stanford Interdisciplinary Graduate Fellowship. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Notes on contributors

Ross Dahlke (PhD, Stanford University) is an Assistant Professor in the University of Wisconsin–Madison School of Journalism and Mass Communication.

Ryan C. Moore (PhD, Stanford University) is an Assistant Professor in the School of Information at The University of Texas at Austin.

Jeffrey T. Hancock is the Harry and Norman Chandler Professor of Communication at Stanford University

References

Ahmad, W., Sen, A., Eesley, C., & Brynjolfsson, E. (2024). Companies inadvertently fund online misinformation despite consumer backlash. *Nature*, *630*(8015), 123–131. https://doi.org/10.1038/s41586-024-07404-1

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *The American Economic Review*, *110*(3), 629–676. https://doi.org/10.1257/aer.20190658

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Allcott, H., Gentzkow, M., Mason, W., Wilkins, A., Barberá, P., Brown, T., Cisneros, J. C., Crespo-Tenorio, A., Dimmery, D., Freelon, D., González-Bailón, S., Guess, A. M., Kim, Y. M., Lazer, D., Malhotra, N., Moehler, D., Nair-Desai, S., Nait El Barj, H., Nyhan, B., . . . Tucker, J. A. (2024). The effects of Facebook and Instagram on the 2020 election: A deactivation experiment. *Proceedings of the National Academy of Sciences*, *121*(21), e2321584121. https://doi.org/10.1073/pnas.2321584121

Allcott, H., Gentzkow, M., & Song, L. (2022). Digital addiction. *The American Economic Review*, *112*(7), 2424–2463. https://doi.org/10.1257/aer.20210867

Allen, J., Arechar, A. A., Pennycook, G., & Rand, D. G. (2021). Scaling up fact-checking using the wisdom of crowds. *Science Advances*, *7*(36), eabf4393. https://doi.org/10.1126/sciadv.abf4393

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, *6*(14), eaay3539. https://doi.org/10.1126/sciadv.aay3539

Altay, S., Nielsen, R. K., & Fletcher, R. (2022). Quantifying the "infodemic": People turned to trustworthy news outlets during the 2020 coronavirus pandemic. *Journal of Quantitative Description: Digital Media*, *2*. https://doi.org/10.51685/jqd.2022.020

Anderl, E., Becker, I., von Wangenheim, F., & Schumann, J. H. (2016). Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *International Journal of Research in Marketing*, *33*(3), 457–474. https://doi.org/10.1016/j.ijresmar.2016.03.001

Andersson, L., & Danielsson, M. (2023). Where were you when Facebook went out? Experiences of involuntary disconnection from social media. *Convergence: The International Journal of Research into New Media Technologies*, *30*(5), 13548565231223488. https://doi.org/10.1177/13548565231223488

Aridor, G. (2023). Market definition in the attention economy: An experimental approach. *Ssrn scholarly paper no. 4069567*. https://doi.org/10.2139/ssrn.4069567

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Basch, C. H., Meleo-Erwin, Z., Fera, J., Jaime, C., & Basch, C. E. (2021). A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. *Human Vaccines and Immunotherapeutics*, *17*(8), 2373–2377. https://doi.org/10.1080/21645515.2021.1894896

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. https://doi.org/10.1038/nature11421

Bosch, O. J., Sturgis, P., Kuha, J., & Revilla, M. (2023). *Uncovering digital trace data biases: Tracking undercoverage in web tracking data*. OSF. https://doi.org/10.31235/osf.io/t2dbj

Cardenal, A. S., Aguilar-Paredes, C., Galais, C., & Pérez-Montoro, M. (2019). Digital Technologies and selective exposure: How choice and Filter bubbles shape news media exposure. *International Journal of Press/Politics*, *24*(4), 465–486. https://doi.org/10.1177/1940161219862988

Casas, A., Menchen-Trevino, E., & Wojcieszak, M. (2022). Exposure to extremely partisan news from the other political side shows scarce boomerang effects. *Political Behavior*, *45*(4), 1491–1530. https://doi.org/10.1007/s11109-021-09769-9

Conti, M., Lain, D., Lazzeretti, R., Lovisotto, G., & Quattrociocchi, W. (2017). It's always April fools' day!: On the difficulty of social network misinformation classification via propagation features. In *2017 IEEE Workshop on Information Forensics and Security (WIFS)* (pp. 1–6). https://doi.org/10.1109/WIFS.2017.8267653

Dahlke, R., & Hancock, J. (2025). Untrustworthy website exposure and election beliefs: Selective exposure and ideological asymmetry. *Journal of Online Trust and Safety*, *3*(1), 1–37. https://doi.org/10.54501/jots.v3i1.250

Dahlke, R., Kumar, D., Durumeric, Z., & Hancock, J. T. (2025). Quantifying the systematic bias in the accessibility and inaccessibility of web scraping content from url-logged web-browsing digital trace data. *Social Science Computer Review*, *43*(5), 1071–1086. https://doi.org/10.1177/08944393231218214

Dahlke, R., Moore, R. C., Bengani, P., & Hancock, J. T. (2026). The consumption of pink slime journalism: Who, what, when, where, and why?. *Digital Journalism*, 1–23.

Dahlke, R., Moore, R., & Hancock, J. (2026). *Exposure to (AI-generated) untrustworthy websites in the 2024 U.S. presidential election*. PsyArXiv [Preprint]. https://doi.org/10.31234/osf.io/qtdmg_v1

Dahlke, R., & Pan, J. (2024). January 6 arrests and media coverage do not remobilize conservatives on social media. *Proceedings of the National Academy of Sciences*, *121*(23), e2401239121.

de Haan, E., Wiesel, T., & Pauwels, K. (2016). The effectiveness of different forms of online advertising for purchase conversion in a multiple-channel attribution framework. *International Journal of Research in Marketing*, *33*(3), 491–507. https://doi.org/10.1016/j.ijresmar.2015.12.001

Dunn, A. (2011). Unplugging a nation: State media strategy during Egypt's January 25 uprising. *The Fletcher Forum of World Affairs*, *35*(2), 15 https://www.jstor.org/stable/45289529.

Ellison, N. B., Triệu, P., Schoenebeck, S., Brewer, R., & Israni, A. (2020). Why we Don't Click: Interrogating the relationship between viewing and clicking in social media contexts by Exploring the "non-Click". *Journal of Computer-Mediated Communication*, *25*(6), 402–426. https://doi.org/10.1093/jcmc/zmaa013

Ernala, S. K., Burke, M., Leavitt, A., & Ellison, N. B. (2020). 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*. Association for Computing Machinery. (pp. 1–14). https://doi.org/10.1145/3313831.3376435

Fan, Y., Pan, J., & Sheng, J. (2024). Strategies of Chinese state media on Twitter. *Political Communication*, *41*(1), 4–25. https://doi.org/10.1080/10584609.2023.2233911

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*(S1), 298–320. https://doi.org/10.1093/poq/nfw006

Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2023). More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the United Kingdom. *New Media and Society*, *25*(8), 2118–2139. https://doi.org/10.1177/14614448211027393

Fox, M. (2018, March 9). Want something to go viral? Make it fake news. *NBC news*. https://www.nbcnews.com/health/health-news/fake-news-lies-spread-faster-social-media-truth-does-n854896

Frenkel, S. (2021, July 19). White House dispute exposes Facebook blind spot on misinformation. *The New York Times*. https://www.nytimes.com/2021/07/19/technology/facebook-misinformation-blind-spot.html

Fu, K.-W. (2023). Propagandization of relative gratification: How Chinese state media portray the international pandemic. *Political Communication*, *40*(6), 788–809. https://doi.org/10.1080/10584609.2023.2207492

Gil-López, T., Christner, C., de León, E., Makhortykh, M., Urman, A., Maier, M., & Adam, S. (2023). Do (not!) track me: Relationship between willingness to participate and sample composition in online information behavior tracking research. *Social Science Computer Review*, *41*(6), 08944393231156634. https://doi.org/10.1177/08944393231156634
Google. (2022). *How a web session is defined in Universal Analytics-Analytics Help*. https://support.google.com/analytics/answer/2731565?hl=en#zippy=%2Cin-this-article
Gordon, B. R., Jerath, K., Katona, Z., Narayanan, S., Shin, J., & Wilbur, K. C. (2021). Inefficiencies in digital advertising markets. *Journal of Marketing*, *85*(1), 7–25. https://doi.org/10.1177/0022242920913236
Green, J., Hobbs, W., McCabe, S., & Lazer, D. (2022). Online engagement with 2020 election misinformation and turnout in the 2021 Georgia runoff election. *Proceedings of the National Academy of Sciences*, *119*(34), e2115900119. https://doi.org/10.1073/pnas.2115900119
Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378. https://doi.org/10.1126/science.aau2706
Guess, A. (2021). (Almost) everything in moderation: New evidence on Americans' online media diets. *American Journal of Political Science*, *65*(4), 1007–1022. https://doi.org/10.1111/ajps.12589
Guess, A., Lyons, B., Montgomery, J. M., Nyhan, B., & Reifler, J. (2019). *Fake news, Facebook ads, and misperceptions: Assessing information quality in the 2018 U.S. midterm election campaign*. 36 https://www.researchgate.net/publication/340004341_Fake_news_Facebook_ads_and_misperceptions_Assessing_information_quality_in_the_2018_US_midterm_election_campaign.
Guess, A. M., & Lyons, B. A. (2020). Misinformation, disinformation, and online propaganda. In N. Persily & J. A. Tucker (Eds.), *Social media and democracy: The state of the field, prospects for reform* (pp. 10–33). Cambridge University Press.
Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, D. R., . . . Tucker, J. A. (2023a). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, *381*(6656), 398–404. https://doi.org/10.1126/science.abp9364
Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, D. R., . . . Tucker, J. A. (2023b). Reshares on social media amplify political news but do not detectably affect beliefs or opinions. *Science*, *381*(6656), 404–408. https://doi.org/10.1126/science.add8424
Guess, A., Munger, K., Nagler, J., & Tucker, J. (2019). How accurate are survey responses on social media and politics? *Political Communication*, *36*(2), 241–258. https://doi.org/10.1080/10584609.2018.1504840
Guess, A., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, *4*(5), Article 5. https://doi.org/10.1038/s41562-020-0833-x
Haenschen, K. (2020). Self-reported versus digitally recorded: Measuring political activity on Facebook. *Social Science Computer Review*, *38*(5), 567–583. https://doi.org/10.1177/0894439318813586
Hall, B., Limaye, A., & Kulkarni, A. B. (2009). Overview: Generation of gene knockout mice. *Current Protocols in Cell Biology*, *44*(1).19.12.1–.19.12.17. https://doi.org/10.1002/0471143030.cb1912s44
Hammond, R. A., & Barkin, S. (2024). Making evidence go further: Advancing synergy between agent-based modeling and randomized control trials. *Proceedings of the National Academy of Sciences*, *121*(21), e2314993121. https://doi.org/10.1073/pnas.2314993121
Houidi, Z. B., Scavo, G., Traverso, S., Teixeira, R., Mellia, M., & Ganguly, S. (2019). The news we like are not the news we visit: News categories popularity in usage data. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 13., pp. 91–102). https://doi.org/10.1609/icwsm.v13i01.3212
Hwang, Y., & Jeong, S.-H. (2023). Misinformation exposure and acceptance: The role of information seeking and processing. *Health Communication*, *38*(3), 585–593. https://doi.org/10.1080/10410236.2021.1964187

Jones-Jang, S. M., Kim, D. H., & Kenski, K. (2021). Perceptions of mis- or disinformation exposure predict political cynicism: Evidence from a two-wave survey during the 2018 US midterm elections. *New Media and Society*, *23*(10), 3105–3125. https://doi.org/10.1177/1461444820943878

Kasimov, A., Johnston, R., & Heer, T. (2025). "Pepe the frog, the greedy merchant and #stopthesteal": A comparative study of discursive and memetic communication on twitter and 4chan/pol during the insurrection on the US Capitol. *New Media and Society*, *27*(1), 127–150.

Kieffer, B. L., & Gavériaux-Ruff, C. (2002). Exploring the opioid system by gene knockout. *Progress in Neurobiology*, *66*(5), 285–306. https://doi.org/10.1016/S0301-0082(02)00008-4

Kim, H. K., & Tandoc, E. C. (2022). Consequences of online misinformation on COVID-19: Two potential pathways and disparity by eHealth literacy. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.783909

Knobloch-Westerwick, S. (2014). *Choice and preference in media use: Advances in selective exposure theory and research*. Routledge. https://doi.org/10.4324/9781315771359

Knoll, J., Matthes, J., & Heiss, R. (2020). The social media political participation model: A goal systems theory perspective. *Convergence: The International Journal of Research into New Media Technologies*, *26*(1), 135–156. https://doi.org/10.1177/1354856517750366

Kwon, K. H., Moon, S.-I., & Stefanone, M. A. (2015). Unspeaking on Facebook? Testing network effects on self-censorship of political expressions in social network sites. *Quality and Quantity*, *49*(4), 1417–1435. https://doi.org/10.1007/s11135-014-0078-8

Lasser, J., Aroyehun, S. T., Simchon, A., Carrella, F., Garcia, D., & Lewandowsky, S. (2022). Social media sharing by political elites: An asymmetric American exceptionalism (No. arXiv: 2207.06313). [Preprint].arXiv. https://doi.org/10.48550/arXiv.2207.06313

Lee, G. (2010). Death of 'last click wins': Media attribution and the expanding use of media data. *Journal of Direct, Data and Digital Marketing Practice*, *12*(1), 16–26. https://doi.org/10.1057/dddmp.2010.14

Lee, J. J., Kang, K.-A., Wang, M. P., Zhao, S. Z., Wong, J. Y. H., O'Connor, S., Yang, S. C., & Shin, S. (2020). Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: Cross-sectional online study. *Journal of Medical Internet Research*, *22*(11), e22205. https://doi.org/10.2196/22205

Lee, S., Rojas, H., & Yamamoto, M. (2022). Social media, messaging apps, and affective polarization in the United States and Japan. *Mass Communication & Society*, *25*(5), 673–697. https://doi.org/10.1080/15205436.2021.1953534

Lee, T., Johnson, T., Jia, C., & Lacasa-Mas, I. (2023). How social media users become misinformed: The roles of news-finds-me perception and misinformation exposure in COVID-19 misperception. *New Media and Society*, *27*(3), 14614448231202480. https://doi.org/10.1177/14614448231202480

Li, H. (., Kannan, P. K., Viswanathan, S., & Pani, A. (2016). Attribution strategies and return on keyword investment in paid search advertising. *Marketing Science*, *35*(6), 831–848. https://doi.org/10.1287/mksc.2016.0987

Liao, M., & Sundar, S. S. (2022). #facebookdown: Time to panic or detox? Understanding users' reactions to social media outage. Extended Abstracts of the *2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–8). https://doi.org/10.1145/3491101.3519674

Lin, H., Lasser, J., Lewandowsky, S., Cole, R., Gully, A., Rand, D. G., & Pennycook, G. (2023). High level of correspondence across different news domain quality rating sets. *PNAS Nexus*, *2*(9), pgad286. https://doi.org/10.1093/pnasnexus/pgad286

Luce, R. D. (1959). *Individual choice behavior* (pp. xii, 153). John Wiley.

Luk, T. T., Zhao, S., Weng, X., Wong, J. Y.-H., Wu, Y. S., Ho, S. Y., Lam, T. H., & Wang, M. P. (2021). Exposure to health misinformation about COVID-19 and increased tobacco and alcohol use: A population-based survey in Hong Kong. *Tobacco Control*, *30*(6), 696–699. https://doi.org/10.1136/tobaccocontrol-2020-055960

Lukito, J. (2024). Global misinformation & disinformation special issue introduction. *International Journal of Public Opinion Research*, *36*(3), edae030. https://doi.org/10.1093/ijpor/edae030

Lukito, J., Brown, M. A., Dahlke, R., Suk, J., Yang, Y., Zhang, Y., Chen, B., Kim, S. J., & Soorholtz, K. (2023). *The state of digital media data research*, 2023. Social Science Research Councilx. https://doi.org/10.26153/tsw/46177

Malinen, S., Koivula, A., Keipi, T., & Koiranen, I. (2018). Exploring selective exposure and selective avoidance behavior in social media. In *Proceedings of the 9th International Conference on Social Media and Society* (pp. 350–354). https://doi.org/10.1145/3217804.3217943

Matassi, M., & Boczkowski, P. J. (2021). An agenda for comparative social media studies: The value of understanding practices from cross-national, cross-media, and cross-platform perspectives. *International Journal of Communication*, *15*, 22–22.

Möller, J., van de Velde, R. N., Merten, L., & Puschmann, C. (2020). Explaining online news engagement based on browsing behavior: Creatures of habit? *Social Science Computer Review*, *38* (5), 616–632. https://doi.org/10.1177/0894439319828012

Mont'alverne, C., Arguedas, A. R., Banerjee, S., Toff, B., Fletcher, R., & Nielsen, R. K. (2024). The electoral misinformation nexus: How news consumption, platform use, and trust in news influence belief in electoral misinformation. *Public Opinion Quarterly*, *88*(SI), 681–707. https://doi.org/10.1093/poq/nfae019

Moore, R. C., Dahlke, R., Forberg, P. L., & Hancock, J. T. (2024). The private life of QAnon: A mixed Methods investigation of Americans' exposure to QAnon content on the web. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW2), 1–34. https://doi.org/10.1145/3687057

Moore, R. C., Dahlke, R., & Hancock, J. T. (2023). Exposure to untrustworthy websites in the 2020 US election. *Nature Human Behaviour*, *7*(7), Article 7. https://doi.org/10.1038/s41562-023-01564-2

Motta, M., Hwang, J., & Stecula, D. (2023). What goes down must come up? Pandemic-related misinformation search behavior during an unplanned Facebook outage. *Health Communication*, 1–12. https://doi.org/10.1080/10410236.2023.2254583

Mukerjee, S. (2024). Online news in India: A quantitative appraisal of the digital news consumption landscape in the world's largest democracy (2014-2018). *Information Communication & Society*, *27* (8), 1650–1670. https://doi.org/10.1080/1369118X.2024.2317898

Mukerjee, S., & Yang, T. (2021). Choosing to avoid? A conjoint experimental study to understand selective exposure and avoidance on social media. *Political Communication*, *38*(3), 222–240. https://doi.org/10.1080/10584609.2020.1763531

Munger, K. (2023). Temporal validity as meta-science. *Research & Politics*, *10*(3), https://doi.org/10.1177/20531680231187271.

Munger, K., Villegas-Cruz, A., Gallego, J., & Vasquez-Cortes, M. (2024). Reenviado muchas veces": How platform warnings affect WhatsApp users in Mexico and Colombia. *Political Communication*, *41*(5), 719–742. https://doi.org/10.1080/10584609.2024.2326130

Myers, S. L. (2022, October 14). How social media amplifies misinformation more than information. *The New York Times*. https://www.nytimes.com/2022/10/13/technology/misinformation-integrity-institute-report.html

Naab, T. K., Karnowski, V., & Schlütz, D. (2019). Reporting mobile social media use: How survey and experience sampling measures differ. *Communication Methods and Measures*, *13*(2), 126–147. https://doi.org/10.1080/19312458.2018.1555799

Neely, S. R., Eldredge, C., Ersing, R., & Remington, C. (2022). Vaccine hesitancy and exposure to misinformation: A survey analysis. *Journal of General Internal Medicine*, *37*(1), 179–187. https://doi.org/10.1007/s11606-021-07171-z

Neyazi, T. A., Yi Kai Ng, A., Kuru, O., & Muhtadi, B. (2022). Who gets exposed to political misinformation in a hybrid media environment? The case of the 2019 Indonesian election. *Social Media + Society*, *8*(3), 20563051221122792. https://doi.org/10.1177/20563051221122792

Nielsen, R. K., & Fletcher, R. (2022). Concentration of online news traffic and publishers' reliance on platform referrals: Evidence from passive tracking data in the UK. *Journal of Quantitative Description: Digital Media*, *2*. https://doi.org/10.51685/jqd.2022.015

Nyhan, B., Settle, J., Thorson, E., Wojcieszak, M., Barberá, P., Chen, A. Y., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Guess, A. M., Kennedy, E., Kim, Y. M., Lazer, D., Malhotra, N., Moehler, D., … Tucker, J. A. (2023). Like-minded sources on Facebook are prevalent but not polarizing. *Nature*, *620*(7972), 137–144. https://doi.org/10.1038/s41586-023-06297-w

Oeldorf-Hirsch, A. (2018). The role of engagement in learning from active and incidental news exposure on social media. *Mass Communication & Society*, *21*(2), 225–247. https://doi.org/10.1080/15205436.2017.1384022

Papathanassopoulos, S., Coen, S., Curran, J., Aalberg, T., Rowe, D., Jones, P., Rojas, H., & Tiffen, R. (2013). Online threat, but television is still dominant: A comparative study of 11 nations' news consumption. *Journalism Practice*, *7*(6), 690–704. https://doi.org/10.1080/17512786.2012.761324

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, *5*(11), Article 11. https://doi.org/10.1038/s41562-021-01117-5

Peng, Y., Lu, Y., & Shen, C. (2023). An agenda for studying credibility perceptions of visual misinformation. *Political Communication*, *40*(2), 225–237. https://doi.org/10.1080/10584609.2023.2175398

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., & Starbird, K. (2023). Mobilizing manufactured reality: How participatory disinformation shaped deep stories to catalyze action during the 2020 US presidential election. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW1), 1–39.

Rao, A., Morstatter, F., & Lerman, K. (2022). Partisan asymmetries in exposure to misinformation. *Scientific Reports*, *12*(1), 15671. https://doi.org/10.1038/s41598-022-19837-7

Revilla, M., Ochoa, C., & Loewe, G. (2017). Using passive data from a meter to complement survey data in order to study online behavior. *Social Science Computer Review*, *35*(4), 521–536. https://doi.org/10.1177/0894439316638457

Rhodes, S. C. (2022). Filter bubbles, echo chambers, and fake news: How social media conditions individuals to be less critical of political misinformation. *Political Communication*, *39*(1), 1–22. https://doi.org/10.1080/10584609.2021.1910887

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing partisan audience bias within Google search. *Proceedings of the ACM on Human-Computer Interaction*, *2* (CSCW), 148: 1–148. 22. https://doi.org/10.1145/3274417

Rossini, P. (2023). Full article: Reassessing the role of inclusion in political communication research. *Political Communication*, *40*(5), 676–680. https://doi.org/10.1080/10584609.2023.2220666

Rydzak, J. (2019). Of blackouts and bandhs: The strategy and structure of disconnected protest in India (*SSRN Scholarly Paper No. 3330413*). https://papers.ssrn.com/abstract=3330413

Schäfer, S. (2023). Incidental news exposure in a digital media environment: A scoping review of recent research. *Annals of the International Communication Association* 47(2) , 242–260. https://doi.org/10.1080/23808985.2023.2169953

Scharkow, M., Mangold, F., Stier, S., & Breuer, J. (2020). How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences*, *117*(6), 2761–2763. https://doi.org/10.1073/pnas.1918279117

Scherman, A., Etchegaray, N., Browne, M., Mazorra, D., & Rojas, H. (2022). WhatsApp, polarization, and non-conventional political participation: Chile and Colombia before the social outbursts of 2019. *Media and Communication*, *10*(4), 77–93. https://doi.org/10.17645/mac.v10i4.5817

Schwabl, P., Haim, M., & Unkel, J. (2024). Aligning agent-based testing (ABT) with the experimental research paradigm: A literature review and best practices. *Journal of Computational Social Science*, *7*(2), 1625–1644. https://doi.org/10.1007/s42001-024-00283-6

Sekścińska, K., & Jaworska, D. (2022). Who felt blue when Facebook went down? - The role of self-esteem and FoMO in explaining people's mood in reaction to social media outage. *Personality & Individual Differences*, *188*, 111460. https://doi.org/10.1016/j.paid.2021.111460

Shah, N., Wang, D. I., Huang, P., & Ling, P. (2021). Digital infrastructure, liminality, and world-making via Asia| (Dis)information blackouts: Politics and practices of internet shutdowns. *International Journal of Communication*, *15*, Article 0. https://doi.org/10.1016/j.amepre.2012.10.018

Shepardson, D. (2024, June 20). *Tiktok says US ban is inevitable without a court order blocking law*. *Reuters*. https://www.reuters.com/legal/tiktok-says-us-ban-is-inevitable-without-court-order-blocking-law-2024-06-20/

Shin, J. (2020). How do partisans consume news on social media? A comparison of self-reports with digital trace measures among Twitter users. *Social Media + Society*, *6*(4), 2056305120981039. https://doi.org/10.1177/2056305120981039
Shousha, N. M., & Abdelgawad, L. R. (2021). Down in minutes, out for six hours: A brief report on feelings during the outage of WhatsApp, Instagram, and Facebook. *British Journal of Psychology Research*, *9*(2), Article 2.
Southwick, L., Guntuku, S. C., Klinger, E. V., Seltzer, E., McCalpin, H. J., & Merchant, R. M. (2021). Characterizing COVID-19 content posted to TikTok: Public sentiment and response during the first phase of the COVID-19 pandemic. *Journal of Adolescent Health*, *69*(2), 234–241. https://doi.org/10.1016/j.jadohealth.2021.05.010
Staddon, J., Acquisti, A., & LeFevre, K. (2013). Self-reported social network behavior: Accuracy predictors and implications for the privacy paradox. In *2013 International Conference on Social Computing* (pp. 295–302). https://doi.org/10.1109/SocialCom.2013.48
Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media + Society*, *9*(2). https://doi.org/10.1177/20563051231177943), 20563051231177943.
Stier, S., Mangold, F., Scharkow, M., & Breuer, J. (2022). Post post-broadcast democracy? News exposure in the age of online intermediaries. *The American Political Science Review*, *116*(2), 768–774. https://doi.org/10.1017/S0003055421001222
Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, *30*(3), 341–366. https://doi.org/10.1007/s11109-007-9050-9
Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of Communication*, *60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x
Sultan, M., Tump, A. N., Ehmann, N., Lorenz-Spreen, P., Hertwig, R., Gollwitzer, A., & Kurvers, R. H. J. M. (2024). Susceptibility to online misinformation: A systematic meta-analysis of demographic and psychological factors. *Proceedings of the National Academy of Sciences*, *121*(47), e2409329121. https://doi.org/10.1073/pnas.2409329121
Tong, C., Winckler, H., & Rojas, H. (2021). La conexión entre las percepciones de sesgos mediáticos y su influencia en la polarización afectiva: Un examen en Brasil, México y Estados Unidos. *Revista de Comunicación Política*, *3*(1), Article 1. https://doi.org/10.29105/rcp3-1
Vaccari, C. (2021). Full article: The international and post-disciplinary journey of political communication: Reflections on "Media-centric and politics-centric views of media and democracy: A longitudinal analysis of political communication and the International Journal of Press/Politics". *Political Communication*, *39*(2), 286–290. https://doi.org/10.1080/10584609.2021.1966599
Vegetti, F., & Mancosu, M. (2022). Perceived exposure and concern for misinformation in different political contexts: Evidence from 27 European countries. *The American Behavioral Scientist*, *69*(2), 00027642221118255. https://doi.org/10.1177/00027642221118255
Vermeer, S., & Trilling, D. (2020). Toward a better understanding of news user journeys: A Markov chain approach. *Journalism Studies*, *21*(7), 879–894. https://doi.org/10.1080/1461670X.2020.1722958
Vermeer, S., Trilling, D., Kruikemeier, S., & de Vreese, C. (2020). Online news user journeys: The role of social media, news websites, and topics. *Digital Journalism*, *8*(9), 1114–1141. https://doi.org/10.1080/21670811.2020.1767509
Vraga, E., Bode, L., & Troller-Renfree, S. (2016). Beyond self-reports: Using eye tracking to measure topic and style differences in attention to social media content. *Communication Methods and Measures*, *10*(2–3), 149–164. https://doi.org/10.1080/19312458.2016.1150443
Walsh, L. C., Regan, A., Okabe-Miyamoto, K., & Lyubomirsky, S. (2021). *Does putting away your smartphone make you happier? The effects of restricting digital media and social media on well-being*. PsyArXiv. https://doi.org/10.31234/osf.io/c3phw
Warner, E. L., Waters, A. R., Cloyes, K. G., Ellington, L., & Kirchhoff, A. C. (2021). Young adult cancer caregivers' exposure to cancer misinformation on social media. *Cancer*, *127*(8), 1318–1324. https://doi.org/10.1002/cncr.33380

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., & Kwak, N. (2017). Incidental exposure, selective exposure, and political information sharing: Integrating online exposure patterns and expression on social media. *Journal of Computer-Mediated Communication*, *22*(6), 363–379. https://doi.org/10.1111/jcc4.12199

Weeks, B. E., Menchen-Trevino, E., Calabrese, C., Casas, A., & Wojcieszak, M. (2021). Partisan media, untrustworthy news sites, and political misperceptions. *New Media and Society*, *25*(10), 14614448211033300. https://doi.org/10.1177/14614448211033300

Wojcieszak, M., Clemm von Hohenberg, B., Casas, A., Menchen-Trevino, E., de Leeuw, S., Gonçalves, A., & Boon, M. (2022). Null effects of news exposure: A test of the (un)desirable effects of a 'news vacation' and 'news binging'. *Humanities and Social Sciences Communications*, *9*(1), 1–10. https://doi.org/10.1057/s41599-022-01423-x

Wojcieszak, M., Menchen-Trevino, E., Clemm von Hohenberg, B., de Leeuw, S., Gonçalves, J., Davidson, S., & Gonçalves, A. (2024). Non-news websites expose people to more political content than news websites: Evidence from browsing data in three countries. *Political Communication*, *41*(1), 129–151. https://doi.org/10.1080/10584609.2023.2238641

Wojcieszak, M., Menchen-Trevino, E., Goncalves, J. F. F., & Weeks, B. (2021). Avenues to news and diverse news exposure online: Comparing direct navigation, social media, news aggregators, search queries, and article hyperlinks. *International Journal of Press/Politics*, *27*(4), 19401612211009160. https://doi.org/10.1177/19401612211009160

Wurmser, Y. (2019, May 30). *US time spent with mobile 2019*. Insider Intelligence. https://www.emarketer.com/content/us-time-spent-with-mobile-2019

Xiao, X. (2022). Not doomed: Examining the path from misinformation exposure to verification and correction in the context of COVID-19 pandemic. *Telematics and Informatics*, *74*, 101890. https://doi.org/10.1016/j.tele.2022.101890

Yang, Y., Davis, T., & Hindman, M. (2023). Visual misinformation on Facebook. *Journal of Communication*, *73*(4), 316–328. https://doi.org/10.1093/joc/jqac051

Yeykelis, L., Cummings, J. J., & Reeves, B. (2018). The fragmentation of work, entertainment, e-Mail, and news on a personal computer: Motivational predictors of switching between media content. *Media Psychology*, *21*(3), 377–402. https://doi.org/10.1080/15213269.2017.1406805

Zhou, A., Yang, T., & González-Bailón, S. (2023). The puzzle of misinformation: Exposure to unreliable content in the United States is higher among the better informed. *New Media and Society*, *27*(3), 14614448231196863. https://doi.org/10.1177/14614448231196863