Understanding Misinformation Engagement Through Digital Trace Data: From Easy-Access to Donated Data

By

Júlia Számely

Submitted to
Central European University
Department of Network and Data Science

In partial fulfillment of the requirements for the degree of
PhD in Network Science

Supervisor: Prof. Elisa Omodei
Associated Supervisor: Prof. Júlia Koltai

Vienna, Austria

2025

Author's Declaration

I, the undersigned, **Júlia Számely**, candidate for the PhD degree in Network Science declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 31 October 2025

Számely Júlia

Signature

Ideas, results, and figures appearing in this thesis are based on the publications listed below:

[I] Júlia Számely, Alessandro Galeazzi, Júlia Koltai & Elisa Omodei (2025) Easy-access online social media metrics are associated with misinformation sharing activity, Scientific Reports (in press).

[II] Zoltán Kmetty*, Ádám Stefkovics*, Júlia Számely*, Dongning Deng, Anikó Kellner, Edit Pauló, Elisa Omodei & Júlia Koltai (2025) Determinants of willingness to donate data from social media platforms, Information, Communication & Society, 28:7, 1324-1349. DOI: 10.1080/1369118X.2024.2340995.
**These authors contributed equally.*

[III] Júlia Számely, Júlia Koltai & Elisa Omodei (2025) Socio-demographic and Online Behavioural Drivers of Misinformation Engagement in Hungary: Evidence from Linked Survey and Social Media Data (manuscript submitted for review).

Other works not covered in this thesis include:

[I] Júlia Számely, Thomas Costello, David Rand & Elisa Omodei (2025) LLM-Based Persuasion for Democracy in Hungary (in preparation).

Copyright Notice

[1]Icon by Font Awesome.

Abstract

With the growing prevalence of misinformation online, which undermines trust, distorts democratic discourse, and hinders responses to global challenges, understanding the characteristics of individuals who are susceptible to misinformation is essential for mitigating its far-reaching consequences.

While personal and socio-demographic characteristics associated with online misinformation engagement have received extensive research attention in the past years, much of this work has relied on self-reported accounts of misinformation engagement or on observations in experimental settings. This thesis argues that to better understand who engages with misinformation online, researchers should complement traditional methods with the use of digital trace data. On the one hand, self-reports provide valuable insights into attitudes and motivations as well as a straightforward approach to obtaining socio-demographic information on individuals, and on the other hand, digital trace data captures actual instances of online behaviour—including misinformation engagement—within the environments in which they originally occur. Integrating these sources allows for a more comprehensive and accurate understanding of misinformation engagement.

The three empirical studies included in this thesis address various aspects of this approach. The first study, based on a sample of users from the U.S. and Western Europe, demonstrates that even simple, easily accessible behavioural metrics, proxying for personal characteristics, can serve as initial indicators of misinformation-sharing behaviour. We find that user characteristics on X (previously known as Twitter), such as activity levels and account age, were significantly associated with misinformation sharing. The second study examines the conditions under which richer and more detailed digital behavioural data can be ethically and practically collected. The study investigates the determinants of willingness to donate digital trace data from social media platforms in an experimental setting, in two distinct political and cultural contexts—the United States and Hungary. Across the two samples, the monetary incentive offered for participation emerged as the strongest predictor, demonstrating that, with appropriate incentives, data donation can be effectively achieved. The final study combines survey data with donation based digital trace data from a representative sample of Hungarian internet users to examine the extent to which personal and socio-demographic characteristics, and online behavioural patterns of individuals are associated with misinformation engagement in an underexplored, but socio-politically important context. The findings show that age, and online behaviours, such as activity levels and engagement with certain communities, were significantly associated with misinformation engagement.

Together, these studies illustrate a progression of approaches to studying misinformation engagement—from large-scale, low-granularity data to rich, survey-linked digital trace data—and contribute to the understanding of the characteristics and behaviours of who engages with misinformation online.

Dedication

I would like to dedicate this PhD thesis to myself. I did it! I started it. I worked through it. I finished it. *I did it!*

Acknowledgements

This PhD meant so many things to me. It was a period in which I was hoping I could stay a child for just a little while longer, and in which I eventually first felt like an adult. It was a place where I expected to continue feeling comfortable and free – to continue being a student – but still move forward and show the world and myself what I can do. And as a bonus, become a *doctor* at the end. This, too, is finally (close to) complete! It was a place where I was expecting to meet people similar to me, who could become my best friends, and the place where this in fact came true. If I am honest, I was also hoping to find a really great boyfriend, which I also did! Thanks PhD!

First I want to thank my supervisors, Elisa and Juli, who walked this whole journey with me. Thank you Elisa for being there for me at all times, for caring so much about how I do in every aspect of being a PhD student, for always paying attention to and remembering the details, for encouraging me, for standing up for me, and generally for leading me through the longest and closest work relationship I have had in my life. Thank you Juli for always showing up and being invested in our common work and my development, for your additional insight, for making me feel included in your group, and mostly for your enthusiasm! I think we were a good team!

I also want to thank each and every friend for making their contribution to shaping my life in these past 4 years. To me, maybe the most important thing in this period was the truly exceptional community we have in DNDS. This community is the reason why I am absolutely certain that wherever my career might take me, this was the best way I could have spent my late twenties. They provided me with a family away from my family in this period of my life, which I will forever cherish.

Timur, thank you for being my gossip bestie as well as my biggest cheerleader – without you who knows if I had gone to MIT! Onkar, thank you for being the reliable and constantly positive presence in my life, who is there to talk to about everything, from relationships to admin! Clara, thank you for being so attentive and caring, and for standing our women ground when there were just sooo many boys around. Jan, thank you for our psychotherapy sessions, your simultaneous involvement and outsider perspective, as well as your own sharing made you the best therapist/client. Sina, thank you for being my *best friend* from the beginning, and for always listening to my drama with empathy and without judgement. Berné, thank you for being the sunshine you are, your presence here always energises the mood in the best way! Bojan, thank you for initiating the wildest conversations I've ever been a part of – please keep my secrets. Helcio, thank you for being the entertainer you are, while also being such a great and genuine person who I can relate to deeply (*and generally hot*)! Onur, thank you for making me feel special by calling me your bestie all the time! Jasper, thank you for being someone

who is so fun to do sports as well as to have deep conversations with, someone I aspire to be like! Martí, thank you allowing me to think that we are alike – if I am even a little bit like you, I am very proud of myself! Martina, thank you for being such a fun girl, who is simultaneously being grounded and reasonable, and by whom I feel liked and completely accepted. Max, thank you for always being up for activities and providing constant entertainment! Piero, thank you for inspiring me with philosophical conversations and by having a curious and chill attitude at the same time, these conversations actually make me want to be/stay a scientist! Basti, thank you for always being so enthusiastic to see me, your encouragement always meant a lot! Adri, thank you for being the main cultivator of the radically inclusive nature of our community, and for showing by example how to unconditionally accept others and myself. Elsa, thank you for always being kind to me, that meant a lot. Lorenzo, thank you for always being so chill and easy to talk to, as well as the cookies and freshly pressed olive oil! Jun, thank you for sharing your stressful moments with us and laughing about our misery together. Yijing, thank you for the shared academic journey, and mostly the trips and deep conversations that happened on the way. Aranka, thank you for the intense conversations, and for making me feel like my advice is valuable. Sjoerd, thank you for being such a great partner in life, for that I never have to doubt your love, but that you also give me space to explore my own journey. Italian boys, Leo Rizzo, Leo Di Gaetano, Filippo, Thomas, Ludo, Luca, you are not a monolith, but let me thank you together for the charm, the laughter, and the occasional confidence boost! Visitors and new first years, thank you for the very fun months together!

I also want to thank my closest friends outside of DNDS, Dia, Emma, Marie, Julcsi, Luca, Orsi, Zsombor, Dani, Juli, and Zuza, who, while we were in long-distance friendships, I could always return to and pick up exactly where we left off.

Köszönöm az elfogadást és szeretetet a nagy családtól, Kinga, Geri, Lilla, Zóra, Buda, István, Kriszta, Isti, Ákos, Zsombor, Péter, Andi, Kata, Livi, Éva, Titi, Pista, Nagymama, és Nagypapa. És a köszönöm ugyanezt a Szelekovszky családnak, és Tamásnak is, akik sokszor a saját családomnak érződnek.

És akik nélkül ez mind nem történhetett volna meg, Mama, Papa, Anna, Dani, és Boti baba! Akármi van, ti ott vagytok nekem ahogy csak tudtok, és ennek az értékét újra és újra realizálom.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The rapid spread of misinformation has become one of the most pressing socio-political challenges of the digital age, and understanding who engages with misinformation remains critical in addressing this challenge. This thesis explores how digital trace data—used both independently and in combination with survey data—can help us understand the characteristics and behaviours that underpin this phenomenon.

Through three interrelated studies, the proposed research demonstrates a progression of approaches to effectively studying online misinformation engagement: from large-scale, publicly available behavioural traces that enable broad analyses, to detailed, donation-based datasets that link online behaviour with survey measures.

Together, these studies demonstrate how data donation approaches can help overcome barriers to data access and generate insights that would otherwise remain out of reach, and how increasing levels of data granularity and participant involvement enable progressively deeper understanding of who engages with misinformation online.

The thesis places particular emphasis on socio-politically significant but academically underexplored contexts. It culminates in an in-depth analysis of Hungary, where rising authoritarian tendencies and a centralised media ecosystem create a distinctive information environment. In doing so, it contributes to an empirically grounded and context-aware understanding of misinformation engagement, with implications for both research design and policy responses.

1.1 The broader context and motivation

1.1.1 Misinformation as a Continuing Societal Concern

Misinformation is widely recognised as a serious challenge with consequences that extend across multiple domains of society as well as to individuals' cognition and behaviour. The World Economic Forum's Global Risks Report 2025 identifies misinformation as the foremost threat to political cohesion and societal trust in the short- to medium-term, emphasising its potential to erode democratic institutions (1). Scholars have likewise documented its effects in areas such as media, politics, science, economics, and public health, underscoring that misinformation constitutes a pressing social problem that warrants sustained scholarly and policy attention (2).

At the societal level, misinformation has been linked to four broad domains of concern: media, politics, science and the economy (2). In the media sphere, false and misleading information erodes trust in journalism and blurs the line between true and false news. Research shows that misinformation can draw disproportionate attention online, but in doing so it fuels scepticism toward both true and false reporting (3). This contributes to disengagement with news altogether, declining trust in news institutions, and negative perceptions of journalistic quality (4; 5; 6). In politics, misinformation undermines democratic processes by distorting public opinion, reinforcing polarization, and shaping support for policies or political figures based on false premises (7; 8; 9). High-profile examples include the role of false claims during the 2016 U.S. presidential election and the Brexit referendum, both of which raised concerns that misinformation can sway outcomes counter to voters' own interests (10; 11). Beyond elections, misinformation has historically functioned as state propaganda and rumour, used as a tool of social control and sometimes connected to violence and political repression (12; 13). The consequences of misinformation are also evident in science and public health. Falsehoods about health and science can directly endanger human well-being. The COVID-19 pandemic offered vivid examples, such as chloroquine overdoses, hoarding behaviours, and resistance to protective measures linked to misinformation (14; 15). Vaccine hesitancy, partly rooted in long-debunked claims about links between the MMR vaccine and autism, demonstrates the enduring power of false beliefs (16; 17). Misinformation also hampers efforts to address climate change by fostering doubt and delaying political action (18; 19). Economically, misinformation can destabilize markets, harm brand reputations, and impose significant costs for fact-checking and countermeasures (20; 21). A striking example is the 2013 false tweet about an explosion at the White House, which briefly caused the Dow Jones to drop by 140 points (22). Estimates suggest that misinformation may be a multibillion-dollar problem globally (23).

At the individual level, misinformation affects both cognition and behaviour. One of the processes from the cognitive side that enhance the effect of misinformation is that false claims can persist in memory even after being retracted, a phenomenon known as the "continued influence effect" (16; 24). Because misinformation is often designed to be emotionally appealing and sensational, it gains an advantage in the attention economy, shaping information-seeking and sharing behaviours (25; 26), and repeated exposure can also make it harder for people to distinguish between truth and falsehood (27). Additionally, people tend to overestimate their own objectivity and underestimate their susceptibility to bias, a cognitive blind spot that further entrenches false beliefs (28; 29). In terms of behaviour, although establishing direct causal links is challenging, research suggests that misinformation influences health behaviours (e.g., vaccine avoidance, rejection of protective measures during pandemics), environmental behaviours (e.g., resistance to sustainable practices), and political participation (e.g., disengagement or polarization) (30; 31; 32). Misinformation also spreads through interpersonal sharing, often driven by motivations beyond accuracy, such as signalling group identity, entertaining others, or gaining social validation (33; 34). Such dynamics help explain how misinformation circulates widely even without malicious intent.

Taken together, these findings show that misinformation is an issue with far-reaching consequences. It undermines trust in media, disrupts democratic deliberation, jeopardizes public health measures, slows responses to urgent challenges like climate change, destabilizes economies, and it does so by exploiting well-documented cognitive biases at the individual level. Even where causal pathways remain difficult to pin down, the breadth of evidence underscores that misinformation represents a multifaceted threat that modern societies should not ignore.

1.1.2 Individual Characteristics Associated with Misinformation Engagement

Because social media lacks traditional gatekeepers and the applied tools against sharing misinformation are limited, large volumes of misleading or false information can spread rapidly (35), and users often struggle to detect it (3; 36; 37). As discussed above, this poses risks across media, politics, science, and the economy, where misinformation can distort beliefs and lead to harmful consequences (2). While only a minority of users share misinformation, estimates of 10–40% (38; 39) still translate into substantial reach given the above mentioned processes and the massive user base of platforms such as Facebook and Twitter (now X) (40). Addressing the problem of misinformation—for example, through the development and implementation of effective interventions—requires understanding of who is most likely to share misinformation, including their socio-demographic traits, motivations, behaviours, and social media habits, as such knowledge can inform tailored measures to reduce misinformation engagement and enhance the effectiveness of interventions (35; 41). This thesis extends prior work on individual differences in misinformation engagement by combining survey and digital trace data to move beyond self-reported measures.

A substantial body of research has explored how individual traits contribute to misinformation engagement. To better understand the influence of these traits, previous work (35) has grouped them into five dimensions: (1) media literacy; (2) psychological factors, including various forms of trust and motivational drivers; (3) social media engagement, such as social media use and prior exposure to misinformation; (4) personality traits; and (5) demographic factors.

Leveraging the potential of combined survey and digital trace data, the present thesis examines behavioural factors, socio-demographic and personal characteristics associated with misinformation sharing. The empirical analysis integrates measurable behavioural correlates of misinformation engagement derived from trace data with individual characteristics derived from survey data.

Media literacy

Media literacy has been conceptualised as a multifaceted construct encompassing the ability to access, evaluate, and create media across diverse formats (42; 43). Related forms of literacy—information, digital, and new media literacy—highlight distinct yet overlapping competencies (44). Information literacy refers to the capacity to recognise when information is needed and to locate, evaluate, and use that information effectively (45; 46). It is typically measured through objective assessments of verification and information-seeking ability rather than self-report (44). Digital literacy, by contrast, reflects one's capability to navigate and evaluate information on digital platforms (47), often operationalised as knowledge of internet-related terms and functions such as "PDF", "spyware", "wiki" or "phishing", which predict success in identifying accurate online information (48; 49). New media literacy extends these notions further, integrating classical, audiovisual, digital, and information literacy (50). Individuals high in new media literacy can access and interpret media meanings, critically evaluate content from social and cultural perspectives, and create media responsibly (51; 52).

Empirical evidence regarding the relationship between media literacy and misinformation susceptibility remains mixed. Some studies find that greater literacy reduces the spread of misinformation—such as when new media literacy negatively predicts fake or unverified information sharing (53) or when low information literacy predicts greater willingness to share misinformation (54). However, other findings suggest a more complex pattern: lower digital literacy has been linked to less misinformation sharing on Twitter (34), and new media literacy has at times positively predicted both intentional and accidental sharing (40). These results indicate that while literacy skills may aid in recognising unreliable content, higher awareness can also increase self-reported recognition of past misinformation engagement. This suggests that literacy-related competencies may influence perceived rather than behavioural susceptibility, highlighting the gap between awareness and action. Relatedly, a growing body of research highlights the role of cognitive ability and cognitive reflection in misinformation susceptibility, showing that individuals with stronger analytical reasoning skills are more likely to distinguish true from false information and less likely to believe or share misinformation (55; 56; 57) The thesis engages with this literature by incorporating indicators of digital experience and platform tenure as potential proxies for media literacy-related competencies in real-world settings, particularly on social media platforms where explicit media literacy measures are unavailable.

Trust

Trust represents another key dimension of misinformation susceptibility, encompassing confidence in mainstream news media, social media platforms, and the credibility of specific information. Declining trust in professional journalism has been identified as a driver of fake news proliferation (58). Individuals with lower trust in mainstream outlets are more likely to turn to alternative information sources (59; 60; 61), leaving them potentially less informed and less capable of distinguishing false from factual information (62). Consistent with this, higher trust in mainstream media is negatively associated with misinformation sharing on social platforms such as Facebook (63) and with misinformation susceptibility more broadly (64). Conversely, trust in social media correlates positively with misinformation engagement: individuals who regard social platforms as credible sources are more likely to share unverified content (65; 66; 67). Finally, perceived trust in individual pieces of information—often operationalised as perceived credibility or believability—strongly predicts sharing behaviour, even when content is false (40; 64; 68; 69; 70; 71; 72). Although trust is not explicitly measured in the studies in this thesis, it may be indirectly reflected in the measures used to capture users' participation across different online communities. Patterns of engagement within communities—whether centred on mainstream, alternative, or politically aligned content—may capture some aspects of shared trust profiles, as users with similar orientations of trust could exhibit comparable media engagement tendencies.

Motivation

Motivational factors also shape misinformation sharing, as suggested by uses and gratifications theory (73; 74; 75). This framework posits that people use media to fulfil psychological and social needs. Four motivations have been found particularly relevant to misinformation engagement: altruism, socialisation, entertainment, and passing time (76). Altruistic motives—sharing to help others—can unintentionally promote misinformation if individuals believe that doing so informs or protects others (77; 78; 79; 80). Similarly, socialisation motives, or the desire to connect with others, have been linked to greater sharing of both news and misinformation (76; 79; 81; 82). Entertainment motives drive people to share amusing or emotionally engaging content without verifying accuracy (62; 82; 83; 84; 85), while those who share to pass time may do so impulsively and without scrutiny (86; 87; 88). Similarly to trust, while individual motivations are not directly measured in the thesis, they are approached through users' community affiliations and activity patterns. Distinct types of online communities often cater to different motivational needs—such as information seeking, social belonging, or entertainment. Analysing engagement across these clusters therefore allows for indirect observation of motivational dynamics shaping misinformation engagement. Additionally, some behavioural metrics were also considered as proxies for underlying motivations when formulating hypotheses. For instance, it was hypothesised that individuals using social media primarily for social connection—reflected, for example, in a higher number of followers—might be more inclined to share misinformation, as their focus may lie more on the social consequences of sharing than on content accuracy.

Personality

Personality traits, most often studied through the Big Five framework—extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience (89; 90)—are also linked to misinformation susceptibility. Individuals high in agreeableness and conscientiousness and low in extraversion tend to better distinguish true from false information and to verify content before sharing (85; 91). Conversely, extraversion and neuroticism have been positively associated with misinformation sharing (40), while openness shows no consistent relationship. Our final study included measures of personality traits, to examine their relationship with misinformation engagement.

Social media use

Among average users, social media platforms are used primarily for hedonic purposes—such as entertainment and social interaction—rather than utilitarian purposes, which involve task completion or goal-oriented use (69). A hedonic mindset is typically associated with reduced critical thinking compared to a utilitarian mindset (92). Consequently, individuals who approach social media with hedonic intentions may be less inclined to scrutinize the accuracy of the information they encounter. Given that social media environments are rife with misleading information and fake news (58), frequent users may face higher exposure to misinformation and a greater likelihood of sharing it. Empirical evidence supports this notion: higher intensity of social media use has been linked to both unintentional and intentional sharing of false information (40), and frequent Facebook use has been positively associated with the intention to share fake news (69). However, counter-evidence also exists. For example, (93) found that frequent WeChat use among young people in China reduced misinformation sharing by lowering anxiety, suggesting that the relationship between social media use and misinformation engagement may depend on contextual and emotional factors. Building on these findings, the thesis examines social media activity as both a determinant and a behavioural outcome. The analyses assess whether measures of activity intensity—such as posting and interaction frequency—are systematically related to engagement with misinformation.

Prior exposure to misinformation

Repeated exposure to information increases its perceived fluency—that is, the ease with which it is processed—which in turn enhances its perceived truthfulness (94; 95). This phenomenon, known as the *illusory truth effect* (96), implies that prior exposure can increase belief in misinformation regardless of its veracity. Empirical studies have confirmed this effect in online misinformation contexts: participants rated previously viewed fake news headlines as more accurate, even when these were politically inconsistent or explicitly labelled as false (97). Similarly, (98) found that individuals were more likely to share fake news headlines they had seen before, as repeated exposure made the content feel intuitively true. Across four studies, (40) demonstrated that prior exposure to a story on social media significantly increased the likelihood of sharing it. These findings suggest that familiarity alone can enhance the perceived credibility and shareability of misinformation. While repeated exposure is not directly examined in the studies in this thesis, it is approximated indirectly by analysing the online communities and content environments that users engage with.

Demographics

Age-related differences in misinformation sharing have been widely studied, though results remain mixed. In the United States, individuals aged 65 and above shared significantly more fake news on Facebook than younger users, even after controlling for education, ideology, and posting frequency (39). Similar findings were reported in South Korea, where older adults demonstrated higher intentions to share misinformation (99). These results have been attributed to factors such as declines in cognitive ability, lower digital literacy, motivated reasoning, or cohort effects (39). Conversely, other studies have found negative or non-significant relationships between age and misinformation susceptibility. For instance, (59) reported that age was negatively correlated with unintentional misinformation sharing in the U.S., and (40) found that older adults in the U.K. were less likely to intentionally, unintentionally, or innocently share false content. This may reflect older individuals' stronger trust in traditional information sources such as newspapers, television, or official channels (100). Still, other studies have observed no significant association between age and misinformation sharing (101).

Gender differences in misinformation sharing appear modest but noteworthy. Men have been found to verify news content less frequently than women before sharing it (65), which may explain their higher likelihood of sharing false political information on social media in both the U.K. and U.S. (38; 40; 69; 102). Similar trends have also been observed in China (93).

Education level also influences misinformation engagement, though the direction of effects varies. In the U.K., lower educational attainment predicted higher likelihood of sharing false political content (40). In the U.S., individuals with graduate degrees were less likely to share fake news compared to those with undergraduate degrees (102). However, other findings indicate that higher education may correlate with unintentional misinformation sharing (40; 59). One explanation is that more educated individuals are better able to recognise and report instances of having shared misinformation (40). Meanwhile, no significant relationship between education and intentional misinformation sharing was found in Spain (101).

Income levels show similarly mixed results. In the U.S., income was positively correlated with unintentional misinformation sharing (59), and in China, higher income predicted greater intention to share misinformation (93). When controlling for education, income was not significantly related to misinformation susceptibility in Chile (64), suggesting that economic and socio-political context may moderate this relationship.

Political orientation has emerged as a consistent correlate of misinformation engagement.

Studies in multiple countries have found that conservatism predicts greater likelihood of misinformation sharing. In the U.K., conservatism was a positive predictor of sharing disinformation on Facebook and Twitter (40), and similar results were reported in the U.S. (103) and Spain (104). This may reflect the ideological asymmetry of misinformation, as much online fake news content tends to lean conservative in tone or source (39), and individuals are more likely to believe (36) and share (40) information consistent with their political identity.

The last empirical chapter also engages with demographic explanations, particularly to test whether these characteristics retain explanatory power once behavioural indicators are included.

In sum, the literature identifies a broad range of individual characteristics associated with misinformation engagement. Although many studies have examined overlapping personal characteristics, findings on factors such as age, education, income, and media literacy remain inconsistent. For example, research on media literacy yields mixed results (e.g, (34; 40; 53; 54)), which may reflect that individuals with higher literacy are more aware of having shared misinformation and therefore more likely to report it. A meta-analysis of 60 studies by (35) sought to clarify these inconsistencies and found that, based on these mainly survey based studies, misinformation sharing is primarily driven by psychological motivations—particularly passing time, trust in information, and socialization—while demographic and personality traits play a weaker role. Yet, as the authors note, much of the variation across studies may stem from the reliance on self-reported misinformation engagement only, as individuals' awareness and willingness to acknowledge such engagement can significantly bias results. Therefore, in this thesis, we combine survey and digital trace data to examine how all of these factors play a role in actual rather than self-reported misinformation engagement behaviour.

1.1.3 Studying Misinformation through Digital Trace Data

Research on misinformation faces a fundamental measurement challenge: it is difficult to determine who actually encounters and engages with misleading content online. As mentioned, many influential studies on misinformation, like most of those introduced in the previous section, rely on self-reported survey measures of misinformation encounter and sharing or are conducted in experimental settings. Self-report based studies are inherently subject to limitations such as recall error (105; 106) and social desirability bias (107), suffer from declining response rates (108), and weak correspondence with observed behaviour (109; 110; 111); while experimental studies carry the risk of potential behavioural changes due to artificial experimental settings (112).

This is particularly problematic for measuring misinformation engagement, where self-reports cannot reliably capture the extent of individual exposure to and engagement with misleading content (113; 114). While survey responses about misinformation sharing sometimes align with observed behaviour, the correspondence is far from perfect. For instance, (115) compared MTurk participants' stated sharing intentions with Twitter sharing data, and (39) linked self-reports of past sharing to actual Facebook and Twitter behaviour. Both studies found meaningful correlations, but the inconsistencies observed point to the potential for further refinement of the measures.

Furthermore, while survey and experiment-based studies provide important insights into individual-level predictors of misinformation engagement, they often fail to account for the environments in which misinformation engagement occurs—such as other platform-related habits, like activity levels and engagement with specific online communities—that shape this behaviour. Recent work highlights that misinformation sharing is not merely the result of cognitive or partisan biases, captured well in survey-based studies, but can emerge from habitual sharing patterns reinforced by social media design itself. (116) demonstrate that users develop reward-based habits of sharing attention-grabbing information, leading to automatic sharing of both true and false content, independent of accuracy considerations. These findings underscore that misinformation engagement is partly a by-product of users' social media habits rather than purely individual traits. Other recent research indicates that the online communities individuals participate in can influence their engagement with misinformation. For instance, studies have shown that the structure and dynamics of online networks, such as echo chambers and filter bubbles, can amplify exposure to misinformation and reinforce existing beliefs (117; 118; 119). Additionally, within these communities, individuals may face social costs if they refrain from sharing certain content, which can pressure individuals into disseminating misinformation (120).

Integrating behavioural data with survey responses can help researchers link individual characteristics to observed online behaviours—such as engagement with misinformation or other forms of social media activity—, providing a more nuanced understanding of the interplay between the two.

Digital trace data offers a promising complement to traditional approaches to studying misinformation. Defined as "records of activity undertaken through an online information system" (121, p.769), digital trace data are unobtrusively collected as users interact with platforms, generating continuous and fine-grained logs of online activity (122; 123). The rise of social media has created vast, naturally occurring datasets capturing behaviours, interactions, and preferences of millions of individuals in near real time. These traces reflect not only user-generated content (e.g., posts, comments), but also user interest (e.g., searches), sometimes also algorithmic curation (e.g., recommendations, filtering), and passive exposure while scrolling (124). Such data make it possible to observe the circulation of (mis)information at scale and analyse engagement patterns down to the individual level. Unlike surveys or experiments, digital trace data capture actual behaviour rather than self-reported intentions, reducing certain types of measurement bias (125). Compared to traditional methods, digital trace data thus allows researchers to move beyond coarse estimations and self-reported recall, and instead to observe actual content engagement in situ. This granularity makes it possible to directly link observed misinformation engagement to potential causes or downstream outcomes such as beliefs and opinions, captured through surveys, or behaviours (126; 127), captured through digital trace data. While digital trace methods bring their own ethical, theoretical, and methodological challenges (124; 128), they remain a crucial addition to traditional methods for building a more accurate understanding of misinformation engagement in contemporary media environments. By combining digital trace data with survey-based measures of socio-demographic and individual characteristics, this thesis demonstrates how linking online behaviour to offline attributes can provide a more comprehensive picture of misinformation engagement as compared to survey or digital trace data based understanding alone.

Ethical, Legal, and Practical Considerations for Collecting Digital Trace Data

Studying digital behaviour necessarily raises ethical and legal questions. As an increasing share of human interaction takes place online, digital traces have become deeply intertwined with individuals' private lives and social identities. In 2025, the global average time spent on social media was approximately 2 hours and 21 minutes per day—with an estimated 5.4 billion users worldwide (129). Such ubiquity makes behavioural data from social platforms a powerful source of insight into social and political phenomena, yet it also poses significant risks for privacy, autonomy, and consent—especially when sensitive behaviours are under study. Researchers must balance the need for sufficiently detailed data with the imperative to protect participants. Informed consent, anonymisation, and secure data storage remain essential safeguards, but they also constrain what data can be collected and how it can be analysed.

The introduction of the European Union's General Data Protection Regulation (GDPR) in 2018 codified many of these principles into law. Designed to strengthen individuals' control over personal data and promote transparency and accountability in data processing, the GDPR has fundamentally reshaped the landscape of online behavioural research (130). By restricting passive data collection without explicit consent, it has compelled researchers to explore new, participant-centred approaches that preserve individual agency.

At the same time, platform policies have further tightened access to digital data. In the aftermath of the Cambridge Analytica scandal, major social media companies curtailed data-sharing practices—restricting previously common data collection methods such as large-scale scraping as well as limiting API-based access—, citing privacy and commercial concerns (131; 132; 133). Twitter introduced steep API fees and rate limits that rendered comprehensive data collection nearly unfeasible without commercial partnerships (134). Meta and Reddit likewise reduced or curated research access (135). These policy changes not only limit overall access, but also exacerbate inequalities between well-resourced and less-resourced researchers (136).

These regulatory and technical shifts have left researchers with a narrower range of options for ethically and legally collecting digital trace data to study misinformation engagement. Publicly available indicators—such as posting frequency or follower counts—remain accessible but capture only coarse behavioural signals, offering limited insight into the motivations and contexts underlying misinformation engagement. To obtain richer, individual-level data while maintaining ethical integrity, scholars have increasingly turned to consent-based, user-centric methods such as web tracking, browser plug-ins, and, most notably, data donation (137; 138; 139). In data donation studies, participants are typically recruited using standard survey sampling techniques and are then asked to download their own social media data and share it with researchers for analysis under informed consent (132; 139).

Among these, the donation of Data Download Packages (DDPs) has emerged as a particularly valuable approach. DDPs contain detailed archives of users' social-media activity that can be directly linked to survey measures of the same participants, offering both depth and transparency (132; 139). This GDPR-compliant, consent-based model enhances participant agency and enables researchers to directly link behavioural data with survey responses, but also introduces practical challenges: willingness to donate such data is far from universal, and low participation rates can generate selection bias and limit generalisability (139; 140). Existing evidence suggests that willingness depends on factors such as monetary incentive, perceived legitimacy, and data sensitivity (138; 141; 142), yet systematic empirical studies remain scarce. Understanding these determinants is therefore crucial for developing feasible, representative, and ethically robust research designs that link observed behavioural and self-reported data.

Working within existing ethical, legal, and practical constraints, this thesis demonstrates how data donation–based approaches that enable the integration of different data sources can expand what is possible in the study of misinformation. It begins by illustrating the value of publicly accessible digital trace data for identifying patterns of misinformation engagement at scale, then motivated to move towards richer datasets, it examines how data donation enables more detailed, representative and ethically grounded research designs, and finally advances to exploring how participant-donated digital trace data can enable a deeper, more nuanced understanding of the behaviours and personal characteristics underlying this phenomenon—all while maintaining full compliance with regulatory standards and respect for participants' autonomy and privacy.

1.1.4 Understudied Socio-Political Contexts

Research on misinformation engagement has predominantly focused on Western democracies, particularly the United States, United Kingdom, and Western Europe (143; 144; 145). These studies typically examine how factors such as social media dynamics, political polarization, and individual cognitive traits shape susceptibility to misinformation in contexts where media systems are decentralized, digital penetration is high, and democratic institutions are well established (64; 143). This dominant focus has provided important insights into how misinformation circulates in pluralistic information environments, but it has also generated a strong

Western and Global North bias in the literature (144).

By contrast, Central and Eastern Europe and several non-Western democracies have received far less attention (146; 147; 148). Contexts such as Hungary, Romania, Poland, Serbia, or Indonesia, feature distinct political and media structures: centralized or state-controlled media, low trust in institutions, and populist or authoritarian governance (147; 149; 150). These are environments in which the flow of information is heavily shaped by political control, yet where digitalization and social media are also expanding. As several studies and reviews emphasize, this combination produces unique vulnerabilities to misinformation, distinct from those observed in Western democracies (144; 147).

Studying these under-represented socio-political contexts is therefore important for two reasons. First, it allows researchers to identify both universal and context-specific mechanisms of misinformation engagement—understanding not only how misinformation spreads in general, but how structural features of media and governance shape this process differently across societies. Second, these contexts often represent frontlines of harmful media and political transformation, including democratic erosion, populist mobilization, and media capture. Understanding misinformation within such settings provides critical insights into how misinformation can become institutionalized and how resilience interventions can be tailored to local cultural and political realities (149; 150).

A Case for Hungary

Among understudied contexts, Hungary has emerged as a particularly salient case study. Over the past decade, the country has undergone a pronounced shift toward authoritarian governance, marked by the consolidation of media ownership and the systematic weakening of institutional checks and balances (151; 152). This transformation has produced a highly centralised and politically aligned media environment in which pro-government narratives dominate, while independent journalism faces persistent structural and financial constraints (153; 154). Recent empirical research describes Hungary as an informational autocracy—a regime that preserves the formal institutions of democracy while using media control and state-aligned communication to sustain political dominance (149; 150; 155). Within this system, pro-government outlets shape a one-sided flow of information that hinders citizens' ability to distinguish between real and false news; pro-government voters, in particular, struggle with news discernment, reflecting the cognitive and informational effects of sustained exposure to politically controlled media (149; 150).

The Hungarian context also challenges common assumptions from Western-democracy-focused misinformation research about trust. (155) identified two distinct profiles of disinformation believers in Hungary: one characterized by excessive trust (sceptical of all official institutions), and another by deep distrust (characterized by extreme trust in the media and politicians, especially those aligned with the government). Both profiles were associated with heightened susceptibility to misinformation, revealing that vulnerability can emerge not only from overconfidence in information sources but also from cynicism—a pattern rarely observed in more pluralistic contexts. Similarly, (147) found in Central and Eastern Europe that low trust in experts and mainstream media led individuals to seek information through personal networks and social media, further amplifying misinformation exposure. These findings illustrate how mechanisms underlying misinformation engagement can differ across political and cultural contexts, underscoring the importance of studying diverse socio-political environments rather than assuming that patterns identified in Western democracies generalize universally.

Beyond describing misinformation susceptibility, research from Hungary also demonstrates context-sensitive intervention success. (150) developed a family-based prosocial intervention leveraging Hungary's strong family orientation, which significantly improved participants' ability to distinguish fake from real news. This approach outperformed generic interventions and yielded lasting effects, underscoring the importance of cultural tailoring in misinformation mitigation. Such findings indicate that interventions grounded in local value systems may be particularly effective in environments where trust in media and institutions is weak.

The Hungarian case illustrates the global significance of studying diverse socio-political contexts. It shows how political and media systems shape misinformation differently from Western democracies, where pluralism, while imperfect, still constrains information monopolies. In contrast, Hungary's centralized, pro-government media system enables systematic disinformation campaigns that blur the line between propaganda and misinformation, making the latter a tool of statecraft rather than a side effect of digital media (150). This dynamic exemplifies the risks of democratic backsliding and media capture, phenomena increasingly relevant across the globe (149; 150).

Comparative research across illiberal or hybrid regimes—such as those in Central and Eastern Europe—can therefore deepen our understanding of how misinformation interacts with media structures which, like in Hungary, are products of these political regimes (155). Insights from Hungary are likely to extend to other post-communist and semi-authoritarian contexts, where similar media monopolies and value orientations prevail (147; 150). By diversifying research settings, scholars can construct more nuanced theoretical models of misinformation engagement and design more effective, culturally grounded interventions (149; 150).

1.2 Research gap and contribution

This thesis contributes to the study of misinformation by showing how digital behavioural data and survey-based individual characteristics can be combined to identify who engages with misinformation online. It addresses two interrelated gaps in the literature: the need for approaches that link individuals' psychological and socio-demographic profiles with their observed online behaviours, and the need for evidence from political contexts beyond Western democracies. The thesis expands both the methodological and geographic scope of research in this domain, while navigating the technical challenges of, and adhering to ethical and legal standards of detailed personal data collection.

Methodologically, the thesis demonstrates progression toward this integration. The first empirical study (Study 1) shows that even simple, publicly obtainable behavioural metrics can offer insight for identifying users likely to engage with misinformation. The second empirical study (Study 2) tests the feasibility of obtaining richer and ethically sourced digital behavioural data through data donation, identifying factors that shape individuals' willingness to share such data and thereby establishing the conditions for combining survey and trace data. The third empirical study (Study 3) then builds on this approach in studying misinformation engagement in a large-scale, representative Hungarian sample, showing that combining survey measures with rich donated trace data enables a more nuanced understanding of the socio-demographic and behavioural drivers of misinformation engagement.

Substantively, the thesis contributes to a more globally inclusive understanding of misinformation dynamics. By analysing behavioural data from both Western and Central-Eastern European contexts, it reveals how predictors of misinformation engagement—such as age, online activity intensity and community membership—manifest in a sociopolitical environment that diverges from the extensively studied Western democracies. The Hungarian case, in particular, highlights the value of studying misinformation engagement in settings where democratic norms are under strain and where misinformation has been politically salient yet empirically understudied.

Taken together, the studies demonstrate a progressive research trajectory: from scalable but limited digital indicators toward richer, individual-level data that combine self-reported and observed behaviours. This progression across levels of data richness and contextual specificity not only offers new empirical insights into misinformation engagement but also illustrates how researchers can deepen their understanding of misinformation-related behaviour by linking what people do with who they are, across diverse contexts and under data constraints.

1.3 Conceptual grounding: Defining misinformation

1.3.1 What is Misinformation?

It is first necessary to clarify what is meant by misinformation. The term and its close synonyms—such as "fake news"—have become heavily politicised in recent years; for example, both former U.S. President Donald Trump (156) and Hungarian Prime Minister Viktor Orbán (157) have used it to characterise established institutions or their political opponents. In this thesis, however, the term is employed in line with academic conventions, while acknowledging that conceptual clarity remains elusive across the literature.

A prominent example of a definition of misinformation in the relevant academic literature is by (158), who define the term as information contradicting the best available expert evidence, adopting a broad perspective that focuses on accuracy regardless of intent (158). By contrast, some definitions emphasize the role of intent to mislead as the defining feature (159). For example, (160) distinguish between three types of "information disorder": *misinformation*, *disinformation*, and *mal-information*. In practice, discerning intent is rarely feasible at scale. Following (161) and (162), this thesis therefore adopts a broad definition of *misinformation* as any information that is false, regardless of intent.

1.3.2 What is Misinformation Engagement?

In this thesis, misinformation engagement is conceptualised as the observable ways in which individuals interact with content originating from low-factuality or misinformation sources. Interactions could include amplifying misinformation or low-quality content in various ways, such as sharing, liking, commenting on such content. This definition aligns with the thesis's focus on integrating observed digital behaviour into the study of misinformation engagement, in that alongside considering self-reported characteristics, it focuses on observable user activity rather than on inferred engagement or exposure. This approach is also consistent with prior large-scale studies that classify misinformation at the outlet level rather than on a post-by-post basis, thereby providing a scalable and reliable measure of misinformation engagement (39; 114).

1.4 Overview of the empirical chapters

This thesis develops a progressive empirical strategy for studying misinformation engagement through online behavioural data, moving from simple, large-scale digital indicators toward integrated survey–trace data that enable a more comprehensive understanding of who engages with misinformation and why. The three studies build sequentially on one another: Study 1 uses easily accessible digital metrics to identify basic behavioural correlates of misinformation sharing; Study 2 examines the feasibility and determinants of data donation as a means to collect richer behavioural data; and Study 3 brings these elements together by combining survey measures with donated digital traces in a representative Hungarian sample. Together, they demonstrate how the integration of behavioural and self-reported data can deepen our understanding of misinformation engagement across different socio-political contexts.

Study 1 — Easy-access social media metrics as predictors of misinformation sharing. This first study tests whether widely accessible metrics from Twitter — such as tweet frequency, account age, and number of followers — can help identify users who are more likely to share misinformation. Using a sample largely from the US and UK, it finds robust relationships: higher tweet frequency is associated with greater sharing of low-factuality content, while older accounts tend to share less, and these effects differ depending on the number of accounts a user follows. These results show that even minimal behavioural data can yield useful signals for identifying potential spreaders of misinformation.

These findings demonstrate the value of large-scale, low-granularity behavioural indicators in understanding who engages with misinformation, yet they also underscore the limitations of what can be inferred without richer information about individual users, motivating a search for data access that enables richer behavioural analyses and a closer examination of how such access can be achieved.

Study 2 — Determinants of willingness to donate social media data. To enable richer behavioural analyses, this study investigates the factors that influence individuals' willingness to donate their social media data via Data Download Packages (DDPs). Two vignette experiments, embedded in surveys conducted in Hungary and the United States, manipulated request parameters—including monetary and non-monetary incentives, number of requested platforms, data type, and estimated upload time—to assess their effects on willingness to participate. The findings show that monetary incentives consistently increase participation, although other effects vary by country. Hungarian respondents, for instance, showed stronger sensitivity to the platforms and types of data requested, suggesting that attitudes toward research participation can shape data-donation willingness in context-dependent ways. These results highlight the importance of request framing and participant characteristics in securing high-quality donated datasets.

Building on these insights, Study 3 implements a data-donation design with a representative sample in the politically sensitive context of Hungary.

Study 3 — Socio-demographic and online behavioural factors associated with misinformation engagement in Hungary. Drawing on a representative sample of Hungarian internet users who donated their Facebook data, this study links survey measures with observed online behaviour, including engagement with misinformation sources. Respondent age was consistently and positively associated with misinformation sharing. While some other socio-demographic factors showed associations when considered independently, these effects disappeared once online behavioural variables were included as explanatory variables, suggesting that differences in digital behaviour account for much of the apparent demographic variation. In particular, membership in certain online communities and higher overall Facebook activity emerged as robust factors associated with misinformation engagement, mirroring some of the behavioural patterns identified in Study 1. By demonstrating the analytical value of combining survey and digital trace data, the study shows how integrating self-reported characteristics and observed behaviour can yield more nuanced insights into the social and behavioural drivers of misinformation engagement, as illustrated through the Hungarian case.

The final chapter synthesises the findings across studies, providing a comprehensive summary of results and examining how the combination of digital data with self-reported data sources can enhance the study of misinformation. It further addresses the thesis's key limitations and proposes avenues for future research, highlighting both methodological refinements and broader implications for policy and practice.

Chapter 2

Study 1: Examining the Association between Misinformation Engagement and Easy-Access Online Social Media Metrics

This chapter is based on the following published article: Júlia Számely, Alessandro Galeazzi, Júlia Koltai & Elisa Omodei (2025) Easy-access online social media metrics are associated with misinformation sharing activity, Scientific Reports (in press).

2.1 Introduction

Building on the thesis's overarching aim to understand misinformation engagement through digital behavioural data, this first empirical chapter examines the extent to which easily obtainable social media metrics are associated with users' propensity to share misinformation. As discussed in the Introduction, misinformation poses persistent risks to democratic discourse and social trust, yet much of the evidence about who engages with it remains limited by reliance on self-reports. Publicly available digital traces—such as account activity indicators—offer a practical alternative for large-scale behavioural analysis, though they necessarily capture only a narrow slice of online behaviour.

This study therefore serves as an initial test of what can be inferred from minimal, openly accessible digital footprints. Using data from Twitter users in largely Western contexts, it assesses whether simple account characteristics—tweet frequency, account age, follower count, and the number of accounts followed by the user~~—can identify~~are more typical characteristics of individuals who are more likely to share low-factuality content. In doing so, it evaluates both the capacity of such coarse-grained behavioural indicators to capture relevant associations and the conceptual limitations of such coarse-grained behavioural indicators.

2.2 Motivation

Researchers have focused on a variety of areas to understand misinformation, examining the phenomenon from the perspectives of the misinformation producers, consumers, the messages themselves, and the contexts in which the messages are embedded (162). While bots may play a significant role in spreading misinformation, the role of human users in sharing it remains a key element in the diffusion process (118). Furthermore, as the spreading of misinformation carries the risk of subjecting people to manipulation and might influence behaviour in undesirable ways (163; 164), understanding the factors that are associated with individuals' likelihood of sharing misinformation is essential.

The reasons behind people's tendency to share misinformation have been widely studied, and findings have shown that in many cases the decision of users to share misinformation does not stem from malicious intent, but rather reflects a failure to consider the accuracy of the piece of news shared (165). Two pathways have been proposed for this failure: limited attention to accuracy in the moment of sharing, or difficulty in distinguishing between true and false content (165). In particular, even individuals who can distinguish between true and false information more easily often fail to apply this discernment at the point of sharing, due to contextual or cognitive factors. First, when sharing content on social media, individuals are prone to prioritising social considerations, such as how sharing certain content affects their social status. Users might also use the content they share as a way to connect to others (84; 87; 166). Second, limited attention caused by high information volume on social media platforms reduces users' ability to reflect on the accuracy of each piece of content they encounter (167). These processes have been linked to various individual-level traits, including limited cognitive skills (168), lack of analytical reasoning (165), and lower levels of digital literacy (169; 170).

Interventions on social media have been shown to be successful across various contexts. Examples include prompting users to think about accuracy (168), and 'inoculating' users with small doses of misinformation and then refuting that piece of misinformation to increase awareness of how one can be manipulated into trusting content (171). Given their success, it is compelling to explore ways in which social media users who are more likely to engage with misinformation sharing could be targeted with such interventions. While previous research presented above provides ample evidence on where to look for identifying individuals prone to misinformation sharing, obtaining data on these predictor characteristics can be complicated, time-consuming, and/or expensive (172).

In this paper, we set out to provide a low-barrier, ~~explanatory~~ exploratory approach to assess whether simple, easily accessible account-level characteristics are systematically associated with users' likelihood of sharing misinformation, drawing on insights from previous literature and using easy-access data from Twitter. Our aim in utilising Twitter metrics as accessible proxies for the above outlined phenomena is to explore whether they can serve as early-warning signals of lower factuality tendencies among users. We aim to show that basic information available on Twitter accounts ~~can help surface broad patterns associated with lower factuality, offering a potential foundation for scalable and low-cost interventions~~ is systematically associated with variation in content factuality, thereby surfacing broad descriptive patterns that may inform future research and exploratory, low-cost approaches to addressing misinformation.

For this purpose, we selected four metrics to test: (1) follower count, (2) average daily tweet count, (3) followed account count, and (4) account age. In our criteria for selecting these metrics we prioritised metrics that are widely available across platforms; not reliant on user self-report, thereby reducing biases associated with self-disclosure; and retrievable in a structured and standardized way, facilitating potential cross-platform generalizability; ~~as well as the metrics' potential ability to serve as proxies for the significant determinants of misinformation sharing found in the literature~~ as well as the extent to which these metrics align with determinants of misinformation sharing discussed in the literature, without implying that they constitute direct or exhaustive operationalisations of those constructs.

While we focus on these metrics due to their accessibility, we recognize that other account-level characteristics may also be useful for understanding users' likelihood of sharing low-factuality content. For example, features such as whether a user provides a location or the complexity of their bio text could add additional explanatory power. Additionally, we acknowledge that alternative methods—such as timeline content or network-based analysis—may yield greater accuracy in detecting misinformation (173). However, these approaches typically require substantially more data collection and processing effort, or more advanced computation. In contrast, the metrics we rely on—follower and following counts, tweet frequency, and account age—are embedded in the basic profile metadata and can be obtained and analysed with minimal technical and infrastructural requirements. This makes them especially valuable in contexts where data access is restricted, resources are limited, or scalability is a priority. In other words, our metric selection was informed by the literature to ensure efficiency, however we prioritised data accessibility instead of the selection of perfect proxies.

Drawing on prior literature, we use three broad strands of research—social considerations, information overload, and digital experience—to motivate the selection of a small set of easily observable account-level metrics. Importantly, these strands are used as conceptual guides for variable choice rather than as claims that the selected metrics directly capture, or exhaustively represent, the underlying psychological or behavioural processes discussed in this literature.

Based on these criteria, we chose the following metrics to represent three lines of misinformation sharing determinants: follower count for social considerations, average daily tweet count and number of users followed for cognitive overload, and account age for digital literacy. We treat these metrics as descriptive indicators that plausibly relate to these lines of prior research, rather than as validated proxies for specific psychological constructs.

Our expectation for each metric is not that it captures a specific mechanism, but that it may be empirically informative in distinguishing between users who tend to share lower- versus higher-factuality content. We expect to find users with more followers and the ones who tweet frequently and follow more accounts to be less factual, while users with older accounts are expected to be more factual.

The expectation regarding follower count is based on research suggesting that social motivations—such as maintaining reputation or appealing to an audience—can divert attention from content accuracy (165). A higher number of followers may therefore coincide with differences in sharing patterns, without implying a specific motivational pathway.

Similarly, we hypothesize that users who produce large volumes of tweets or follow many accounts experience a form of cognitive strain or information overload, reducing the attention they can dedicate to assessing accuracy (167). More cautiously, higher posting activity and larger followed networks may be associated with differences in information exposure and engagement intensity, which prior work has linked to challenges in accuracy-related judgement (174).

We acknowledge that platform algorithms shape users' content exposure, and that a user's feed may include messages from both followed and non-followed accounts. Nonetheless, following more accounts likely increases the potential content diversity and volume a user may encounter, even within algorithmically filtered feeds.

Prior work in the misinformation literature suggests that individual differences in cognitive reflection are systematically related to selectivity in both content sharing and network formation. For example, individuals lower in cognitive reflection tend to be less selective in the accounts they follow and the content they choose to share (175; 176). The behavioural patterns observed in our dataset—particularly the higher posting frequency and substantially larger following networks among users who predominantly share lower-factuality content—are consistent with these established findings. Lower selectivity may coincide with broader but less curated information environments, increasing exposure to low-quality or unreliable content. Conversely, users with higher factuality scores exhibit more constrained network choices and lower posting activity, which aligns with more selective and deliberative engagement. While we do not measure cognitive traits directly, these behavioural regularities parallel theorised attentional and reasoning differences documented in prior work on misinformation sharing.

Finally, we interpret account longevity as a rough proxy for digital experience or literacy. More precisely, we treat account age as a descriptive marker of platform tenure rather than as a direct measure of digital literacy. We also note that account age may correlate with user age or other demographic traits, which complicates its interpretation. Account longevity could reflect a mix of factors — including but not limited to digital literacy — such as habitual platform use, early adopter behaviour (which has been linked to higher media savvy (177)), and greater exposure to platform norms and misinformation corrections. This interpretation aligns with prior work on digital behaviour and literacy (48), while also acknowledging the conceptual limitations of using account age as a stand-in for digital skills.

Hence, although several techniques exist to infer users' propensity to consume unreliable sources, focusing on easily accessible account-level metrics allows us to characterise broad and robust patterns of misinformation-related behaviour. These patterns are intended to be descriptive rather than predictive or causal, and to inform future theory-building and more targeted data collection efforts.

2.3 Methods

2.3.1 Data

Our study builds upon the data used in the studies by (178), which was approved by the University of Cambridge Psychology Research Ethics Committee (PRE.2020.144). All participants in the study provided informed consent before answering surveys and providing their Twitter usernames (178). This dataset was selected in part due to its unique ability to capture regular, human user behaviour at scale. Most public misinformation datasets rely on keyword-based collection methods that tend to over-represent highly active or partisan users. Our focus on followers of original survey participants allowed us to build a larger, more behaviourally representative dataset for this purpose.

The study collected 463 survey responses through the online recruitment platform Prolific. Starting from those usernames, the publicly accessible followers of those accounts were collected, resulting in a dataset of n = 1,670,127 users. Such data includes information about the number of followers and following, as well as the number of content posted and the subscription date. For each user, the latest 500 tweets containing a URL to external sources, posted within one year of the most recent published content, were collected. In this study, we utilize users' data, including the number of followers of the account, the number of people followed by the account, the total number of tweets produced by the account, registration dates, and their tweets containing a URL, all directly extracted using the official Twitter API. All methods in this study were performed in accordance with the relevant guidelines and regulations.

2.3.2 User factuality

The reliability of users' diets was measured using a metric called Factuality, based on the domains of the content they published. The quantification of this metric relies on data from Media Bias Fact Check (MBFC), an independent fact-checking organisation that classifies news outlets based on both their reliability and political bias (179). Each news outlet on MBFC is assigned labels indicating its reliability (Very Low, Low, Mixed, Mostly Factual, High, Very High). News outlets' factuality corresponds to MBFC's reliability label. These metrics have been applied individually to each Twitter user (denoted as u), generating a vector, $F_u \in \mathbb{R}^{6\times 1}$, representing the percentage of links shared by the user for Factuality categories. We assign each user a unique factuality score, calculated as the average of the factuality scores associated with their posted tweets. To assign factuality scores, we initially collected up to the latest 500 tweets containing URLs, posted within one year of the most recent published content, for the approximately 1.6 million users. The restriction to URLs shared within one year of each user's most recent activity was chosen to focus on contemporaneous information consumption. Because domain-level factuality ratings and users' sharing behaviour may change over time, older URLs may reflect outdated exposure patterns or domain reputations. Limiting the observation window helps ensure that the outcome captures recent information-sharing habits rather than long-run historical behaviour. Similarly, the cap of 500 URLs per account was implemented to prioritize recent consumption among highly active users. For accounts that share very large numbers of links, a large but finite number of recent URLs is sufficient to obtain a stable estimate of information-sharing tendencies, while avoiding disproportionate influence from older content. Prior work inferring users' information diets or political orientation from URLs has relied on substantially fewer observations per user (180; 181), suggesting that this threshold is conservative. Finally, these choices also reflect practical considerations related to data collection and computational feasibility when gathering content from over one million users. Out of these, only 240,915 users had ever shared a URL, and among them, only 59,610 had shared at least one domain we were able to classify using MBFC ratings. While domain-level factuality ratings provide a scalable way to characterise information quality, they do not capture article-level variation: high-quality outlets occasionally publish misleading content, and low-rated outlets may publish accurate information. Consequently, the outcome in this chapter should be interpreted as a propensity to share content from low-credibility sources, rather than misinformation sharing per se.

We segmented users into low, middle, and high factuality groups based on their respective factuality scores. This involved dividing users into the three factuality categories such that the bottom and top 25%, 30%, or 35% would be the low and high factuality users, respectively, and the remaining middle would be the middle factuality group, resulting in three distinct factuality groups. Our primary goal with this approach is to identify the most prominent misinformation spreaders and characterize their behaviours, rather than provide a continuous score for all users. While continuous scores may offer more granularity, prior work has shown that misinformation sharing on social media follows a highly skewed distribution, with a small fraction of users responsible for the majority of spread (114). Therefore, segmenting users allows us to better focus on actionable insights relevant to platform moderation and intervention strategies. Although our segmentation approach may appear to contrast with literature highlighting power-law distributions in misinformation sharing, it serves a different purpose: to provide a broad behavioural overview rather than pinpoint only extreme cases. Grouping users in the bottom 30% (or 25%, 35%) allows us to capture general patterns among users with lower factuality. While the main focus of our paper lies on the results derived from the factuality categorisation taking the top and bottom 30% as low and high factuality users, respectively, we also present findings obtained applying the other two thresholds in the Supplementary Information to validate the robustness of our conclusions.

2.3.3 Filtering the dataset for regular users

As the purpose of our study is understanding the behaviour of “human users” (182), we initially curated our datasets to focus on users who are more representative of organic human social media users. I.e., we wanted to exclude accounts from the study that appeared to be either business, political or celebrity accounts, or bots. To ensure the selection of regular individuals, we implemented specific criteria as outlined below. These criteria were derived from various sources and were applied iteratively to curate our dataset.

Verified status: Verified accounts, typically belonging to recognized individuals or brands, were identified and removed from the dataset as they do not represent typical user behaviour. This was based on X’s verification policy—a rigorous review process of the authenticity, notability, and activity of the account at the time of the data collection (183). Such accounts comprise 2.9% of all accounts in the dataset.

Average number of tweets per day: Accounts with an exceptionally high average number of tweets per day were considered unlikely to belong to regular users, as this behaviour is more characteristic of commercial or political accounts, sometimes bots (184). We set thresholds to remove accounts with tweet/day ratios exceeding certain values expected to be limits of regular user behaviour.

Number of followers: Accounts with a very high number of followers were considered atypical for regular users, as such levels of fame are usually attained outside of Twitter or through exceptionally high activity on the platform. We removed accounts with follower counts surpassing a certain threshold.

Follower/followed account ratio: A high proportion of followers compared to followed accounts is characteristic of celebrity-like accounts rather than typical users. Therefore, we removed accounts with follower/followed account ratios exceeding certain values.

Number of followed accounts: Accounts with an unusually high number of followed accounts may indicate bot-like behaviour, particularly if the followings are not reciprocal (185). We set thresholds to remove accounts with followed account counts exceeding specific values.

Followed account/follower Ratio: A high proportion of non-reciprocal followings may also indicate bot-like behaviour (185). Therefore, accounts with followed account/follower ratios exceeding certain values were removed. Finally, to filter for bots, we applied an additional filtering step using Botometer (186), specifically removing users with a Botometer score above 0.4, following the “human threshold” suggested by (187).

Based on the stringency of the above introduced criteria we created three distinct groups of datasets: relaxed, middle, and strict. By applying these criteria, we aimed to ensure the selection of regular individuals whose behaviour is more representative of typical social media users for our analysis. The specific values for each strictness category are shown in Table 2.1.

To characterize the groups resulting from our filtering criteria, we present the distribution of Factuality scores in Figure A.1 in the Supplementary Information, illustrating the differences between the selected groups.

In addition to providing the distributions, we briefly characterize the groups based on their factuality scores. As shown in Supplementary Figure A.1, low factuality users are predominantly characterized by scores below 60, middle factuality users span scores between 60 and 70, and high factuality users are characterized by scores above 70, with a strong concentration around 80. We note that the observed peaks in the distributions arise from the underlying labelling procedure: factuality scores are assigned based on categorical source credibility ratings, leading to discrete jumps at 0, 20, 60, 80, and 100. As many users' sharing histories contain sources with identical credibility ratings, sharp spikes are observed in the plotted distributions at scores 40, 60, and 80.

Criteria	Relaxed	Middle	Strict
Verified Status	-	Not Verified	Not Verified
Average Tweets Per Day	≤ 56	≤ 32	≤ 16
Follower Count	$\leq 10,000$	$\leq 10,000$	$\leq 5,000$
Follower/Followed Account Ratio	≤ 10	≤ 5	≤ 3
Followed Account Count	$\leq 10,000$	$\leq 10,000$	$\leq 5,000$
Followed Account/Follower Ratio	≤ 10	≤ 10	≤ 5
Botometer score	≤ 0.4	≤ 0.4	≤ 0.4

Table 2.1: Criteria for dataset variation for robustness check.

As in the case of factuality grouping, the main focus of our paper lies on the results derived from the middle category, but we also present findings obtained applying the two other criteria in the Supplementary Information to validate the robustness of our conclusions. The number of users in each dataset are shown in Table 2.2.

Criteria	Relaxed	Middle	Strict
Top/bottom 25%	4492	4218	2878
Top/bottom 30%	5427	5113	3643
Top/bottom 35%	6166	5806	4072
Total	17616	16586	11503

Table 2.2: Number of observations in each of the datasets used for validation of results.

2.3.4 Randomised dataset

To ensure the conclusions drawn from our results are not due to chance, we constructed a counterfactual dataset consisting of 1,000 datasets mirroring the structure of our original dataset. In these counterfactual datasets, factuality scores were randomly shuffled across individuals while keeping all other characteristics constant.

2.3.5 Analysis of combined effects

We employed multinomial regression analysis to investigate the combined effects of our four social network metrics—follower count, followed account count, tweets per day, and days since registration—on factuality, which is categorised into three levels: low, middle, and high. Multinomial regression is particularly suitable for this analysis as it allows for the modelling of outcomes with multiple discrete categories, enabling us to examine the association of multiple predictors with the probability of an observation belonging to each factuality level simultaneously. To interpret the results, we utilised Average Marginal Effects (AMEs). AMEs offer a clear and intuitive measure of the impact of each social network metric on the probability of an account being classified into each factuality category. By translating regression coefficients into average changes in probabilities, AMEs enhance the interpretability of our findings, making it easier to understand and compare the effects of different metrics. This approach aligns with the recommendations by (188), who highlights the advantages of using marginal effects in complex regression models.

2.4 Results

We analysed data collected from Twitter in 2022, consisting of 1,670,127 users and 14,688,374 tweets. Each user in the dataset is assigned a factuality score based on the credibility of the news sources they share, determined from their last 500 tweets containing a URL. Using these scores, we categorised users into high, middle, and low factuality groups. Additionally, we retrieved various social network metrics on the same users, including follower count, number of followed accounts, tweet count, and registration date. Finally, we created 1,000 randomised versions of the dataset, where we shuffled users' factuality scores while keeping all other features fixed, to serve as a basis for testing the statistical significance of our results.

In the following, we present our findings on easy-access social network metrics' effectiveness in the initial identification of individuals who tend to share a large amount of misinformation. We draw on the literature introduced above to guide the selection of the social network metrics tested. ~~Specifically, drawing upon three strands of relevant literature – on social motivations, limited attention, and digital literacy – we select four metrics: follower count, daily average tweet count, followed account count, and the number of days since registration.~~ Specifically, drawing upon prior literature on social visibility, engagement intensity, and platform tenure, we select four easily observable account-level metrics that have been previously linked to online information sharing behaviour. Importantly, these constructs serve only to motivate the choice of variables rather than to imply that the selected metrics validly operationalize specific psychological mechanisms.

2.4.1 Popularity is associated with lower factuality

Building on the literature discussed above, highlighting the dominance of social considerations in social media sharing over accuracy concerns (165), we selected follower count on Twitter (popularity) as our first metric. We hypothesise that follower count on Twitter will serve as an indicator to identify users more inclined to share misinformation. To this end, we employed one-sided Mann-Whitney U (MWU) tests to compare the follower count distributions of the low and high factuality users, and verified their statistical significance by comparing them with the MWU statistics obtained on the 1,000 randomised datasets. Low factuality users were on average found to have a higher number of followers compared to high factuality ones, confirming our initial expectations (Figure 2.1a). The results were found to be statistically significant by comparing the MWU test statistic with the MWU test statistics obtained using the 1000 shuffled datasets (Figure 2.1b). This finding thus suggests that follower count can serve as an initial indicator in distinguishing between low and high factuality users, confirming our hypothesis.

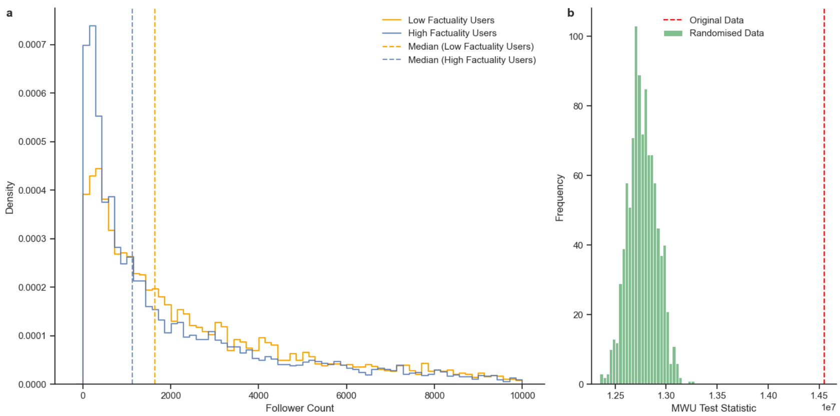


Figure 2.1: **Low factuality users tend to have higher follower count than high factuality users.** Panel a: Comparing the distributions of follower count between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly higher number of followers compared to high factuality users (p-value < 0.0001). The median value for low factuality users (orange dotted line) is 1642 followers, whereas for high factuality users (blue dotted line) it is 1125. Panel b: The MWU test score obtained from the empirical data (red dotted line) is higher than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars).

2.4.2 High tweeting rate and followed account count are also associated with lower factuality

Based on the related literature on information overload (167), we selected the number of tweets per day and the number of accounts followed as two additional metrics for distinguishing between low and high factuality users, and hypothesise that these two metrics are useful in distinguishing between low and high factuality users. Similarly to the previous metric, we utilised

MWU scores to compare the distributions of these two metrics between low and high factuality users, and validated their significance against the 1,000 shuffled datasets. Low factuality users were observed both to post a higher number of tweets per day (Figure 2.2a) and to follow a greater number of users compared to their high factuality counterparts (Figure 2.3a). Both results were found to be statistically significant by comparing the MWU test statistics obtained from the empirical data to the MWU test statistics obtained from the 1000 shuffled datasets (see Figure 2.2b and Figure 2.3b, respectively). These results suggest that the two metrics tested can be useful in distinguishing between low and high factuality users.

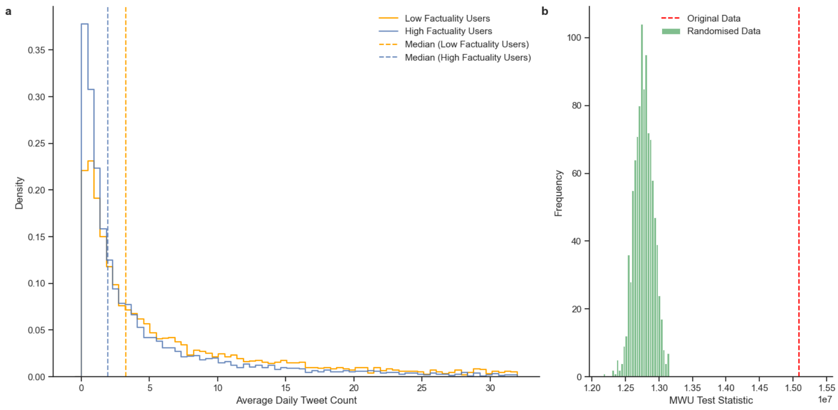


Figure 2.2: **Low factuality users tend to have higher tweet count than high factuality users.** Panel a: Comparing the distributions of tweet count between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly higher number of tweets compared to high factuality users (p-value < 0.0001). The median value for low factuality users (orange dotted line) is 3.26 tweets per day, whereas for high factuality users (blue dotted line) it is 1.94. Panel b: The MWU test score obtained from the empirical data (red dotted line) is higher than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars).

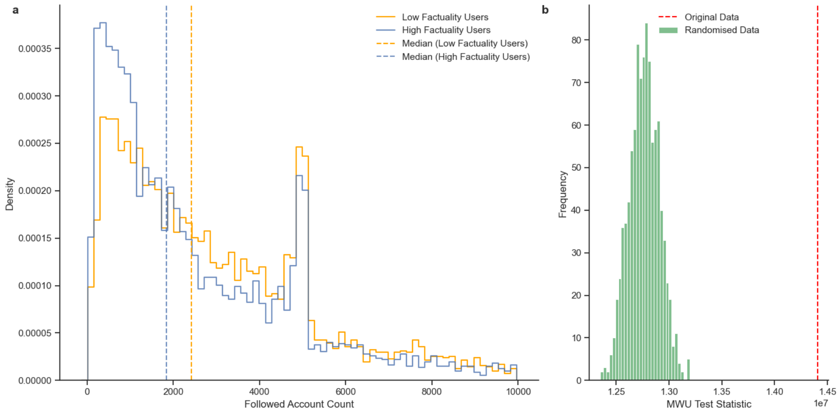


Figure 2.3: **Low factuality users tend to have higher followed account count than high factuality users.** Panel a: Comparing the distributions of followed account count between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly higher number of followed accounts compared to high factuality users (p-value < 0.0001). The median value for low factuality users (orange dotted line) is 2415 followed accounts, whereas for high factuality users (blue dotted line) it is 1845. The peak around 5000 is due to an Twitter policy that limits the number of new followed accounts until the user obtains more followers. Panel b: The MWU test score obtained from the empirical data (red dotted line) is higher than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars).

2.4.3 Longer social media presence is associated with higher factuality

Based on the literature on the relationship between digital literacy and misinformation sharing behaviour (169), we selected the number of days since registration as our last key metric for distinguishing between low and high factuality users. We employed MWU scores to compare distributions and ensured their statistical significance through comparisons with the 1,000 shuffled datasets. Low factuality users displayed a lower number of days since registration compared to high factuality users, i.e. older accounts tend to be more factual, aligning with our expectation (Figure 2.4a). The results were found to be statistically significant by comparing the MWU test statistics obtained from the empirical data to the MWU test statistics obtained from the 1000 shuffled datasets (Figure 2.4b). This finding supports the relevance of this metric in distinguishing between low and high factuality users.

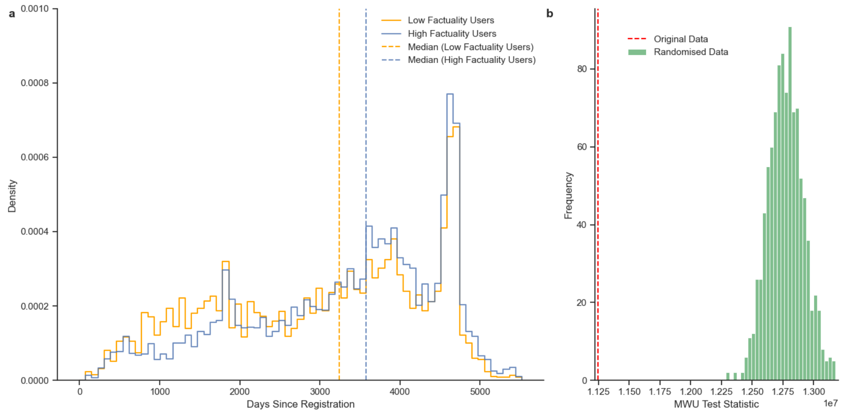


Figure 2.4: **Low factuality users tend to have lower number of days since registration than high factuality users.** Panel a: Comparing the distributions of the number of days since registration between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly lower number of days compared to high factuality users (p-value < 0.0001). The median value for low factuality users (orange dotted line) is 3242 days since registration, whereas for high factuality users (blue dotted line) it is 3578. Panel b: The MWU test score obtained from the empirical data (red dotted line) is lower than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars).

2.4.4 Combined effect of social network characteristics on factuality

Next, we employed multinomial regression to move from analysing the associations of factuality with individual social network features separately to considering the features all together. We do this in two stages. First, we perform multinomial regression with no interactions between the independent variables, second, we add interactions between explanatory variables. With conducting regression analysis, we are able to examine the effect of each key metric while simultaneously controlling for the effects of all other metrics. Thus, these models help us analyse the interplay of the metrics on factuality, and evaluate the extent to which these social network characteristics are associated with factuality scores.

We considered factuality as the outcome variable with three levels—low, middle, and high factuality. To capture the factors associated with notably low or high factuality, in each analysis we chose middle factuality to be the reference group that the other two are compared to. Explanatory variables included the four Twitter metrics used above: follower count, followed account count, tweets per day, and days since registration. The explanatory variables were standardised for comparability. To measure the effects, we assessed the average marginal effects (AMEs) of the four social network metrics, as well as interactions between them. According to the results, each of the four social network metrics was found to have a significant relationship with both the low and high factuality groups—compared to the middle factuality group—with a number of significant interactions between variables.

First, we consider and report the AMEs of the four social network metrics on factuality (Figure 2.5). Examining the AMEs of follower count on factuality group membership, we find that the more followers users have, the more likely they are to be in the low factuality group and less likely to be in the high factuality group, compared to the middle factuality group. Specifically, a one standard deviation increase in follower count is associated with a 1.3 percentage point increase on average in the probability of belonging to the low factuality group, and on average a 2.3 percentage point decrease in the probability of belonging to the high factuality group, compared to the middle factuality group. This suggests that, even when accounting for other variables, follower count remains a useful indicator for distinguishing between high and low factuality users.

The number of users followed is negatively associated with both the high and low factuality groups, meaning that holding all other variables constant, following a high number of accounts makes a user more likely to be in the middle factuality category, while it makes them less likely to be in the low or high factuality categories. Quantitatively, a one standard deviation increase in the number of followed accounts is associated with a 1.0 percentage point decrease on average in the probability of being in the low factuality group and on average a 1.9 percentage point decrease in the probability of being in the high factuality group, compared to the middle factuality group. These findings are on the one hand consistent with our expectations, as we see a negative association between the number of users followed and high factuality, compared to middle factuality. Opposite to our expectation, we also see a negative association between the number of users followed and belonging to the low factuality category compared to belonging to the middle factuality category, however this effect is significantly weaker than the one belonging to high factuality. These effects show that there is a reverse U-shaped relationship between followed account count and the three factuality groups, both with low and high factuality users typically following fewer users than middle factuality users. These results suggest that the power of followed account count as a measure to classify low and high factuality users is therefore limited.

The average number of tweets produced by the user per day is negatively associated with being in the high factuality group compared to the middle factuality group, i.e. holding all other variables constant, high factuality users tweet less on average. For the low factuality group, we find a positive association compared to the middle factuality group, i.e. holding all other variables constant, low factuality users tweet more on average. More specifically, a one standard deviation increase in average daily tweet count is associated with a 0.23 percentage point increase on average in the probability of being in the low factuality group, and on average a 1.2 percentage point decrease in the probability of being in the high factuality group, compared to the middle factuality group. These results are in line with our hypothesis and previous results, confirming that average daily tweet count is a useful indicator in distinguishing between low and high factuality users.

For days since registration, we see a positive effect for the high factuality, and the opposite for the low factuality group compared to the middle one. Numerically, a one standard deviation increase in days since registration is associated with a 2.2 percentage point decrease on average in the probability of being in the low factuality group, and on average a 1.5 percentage point increase in the probability of being in the high factuality group, compared to the middle factuality group. The results are therefore consistent with the outlined expectations and hypothesis, indicating that, holding all other variables constant, longer presence on Twitter is associated with higher factuality, and that the metric is useful in distinguishing between low and high factuality users.

To assess the robustness of our results, we tested the consistency of the observed effects across eight alternative dataset constructions, varying both the user filtering criteria and the factuality thresholding. We found that tweet frequency and days since registration consistently distinguished between high and low factuality users across all variations, with stable effect directions and magnitudes. Follower count also showed a qualitatively consistent pattern, with a stable difference between high and low factuality groups, although the effect size varied somewhat across conditions. The number of followed accounts showed more variation, but the distinction between high and low factuality users—particularly the stronger negative association for high factuality users—was preserved across nearly all dataset versions. These results reinforce the general robustness of our findings while also highlighting that some metrics (e.g., follower count and followed account count) are more sensitive to dataset construction choices than others (see Figure A.10 in the Supplementary Information).

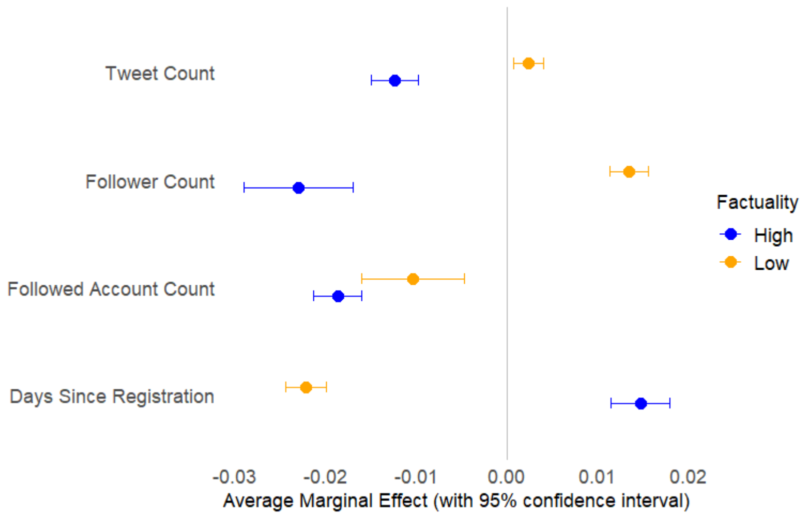


Figure 2.5: **Average marginal effects of social network characteristics on factuality.** All examined social network metrics are significantly associated with factuality, and the differences in the effects between low factuality (marked with orange) and high factuality (marked with blue) users is significant in all four cases. Zero effect would mean users are equally likely to be low or high factuality as middle factuality. Tweet count: higher tweet count means the user is more likely to be low factuality, and less likely to be high factuality than middle factual. Followed account count: higher followed account count means the user is less likely to be low factuality, as well as to be high factuality, but the latter effect is stronger. Follower count: higher follower count means the user is more likely to be low factuality, and less likely to be high factuality. Days Since Registration: higher number of days means the user is more likely to be high factuality, and less likely to be low factuality.

Finally, we examined the interactions between the effects of the four social network metrics on factuality to unfold potential interplay of the independent variables. We have found that some of the social network metrics interact significantly with one another on factuality. To illustrate the interactions we divided the users into two equally sized groups based on the given social network metrics, e.g. followed account count, and plotted the effects at the median of each group.

Our analysis reveals differences regarding the effect of the average daily tweet count on factuality (Figure 2.6, panel a). In examining average daily tweet count across different levels of followed account count and its association with factuality levels we find that among those with few accounts followed the average tweet count is not associated with the likelihood of belonging to the high factuality group. However, an increase in tweet activity is associated with an increased likelihood of belonging to the high factuality group among those who follow more accounts. For the low factuality group, tweet count is negatively associated with low factuality and this effect is stronger among those who follow more accounts. There is no significant difference in the likelihood of belonging to the low factuality group among those who tweet infrequently, however, among those who tweet frequently, users following few accounts are way more likely to have low factuality than users who follow more accounts.

In relation to days since registration, our analysis (Figure 2.6, panel b) shows that among users who follow fewer accounts, those who have been registered on the platform for a longer duration are more likely to be in the high factuality category, while among those who follow more accounts, the trend is the opposite: the longer they are on the platform, the less likely that they belong to the high factuality group. The relationship between days since registration and low factuality does not vary based on followed account count. The relationship is consistently negative between days since registration and low factuality, indicating that longer registration periods are associated with a decreased likelihood of low factuality, however, this effect is somewhat more pronounced among users who follow fewer accounts.

In examining the interactions of the effects among the rest of the explanatory variables we found only slight differences. In only one of the other cases was there a difference in the direction of the effect based on another variable (effects of followed account on high factuality by follower count). The rest of the observed differences are small, regarding either the magnitude or the strength of the effects (see the full set of explanatory variable interactions in Supplementary Figure A.14).

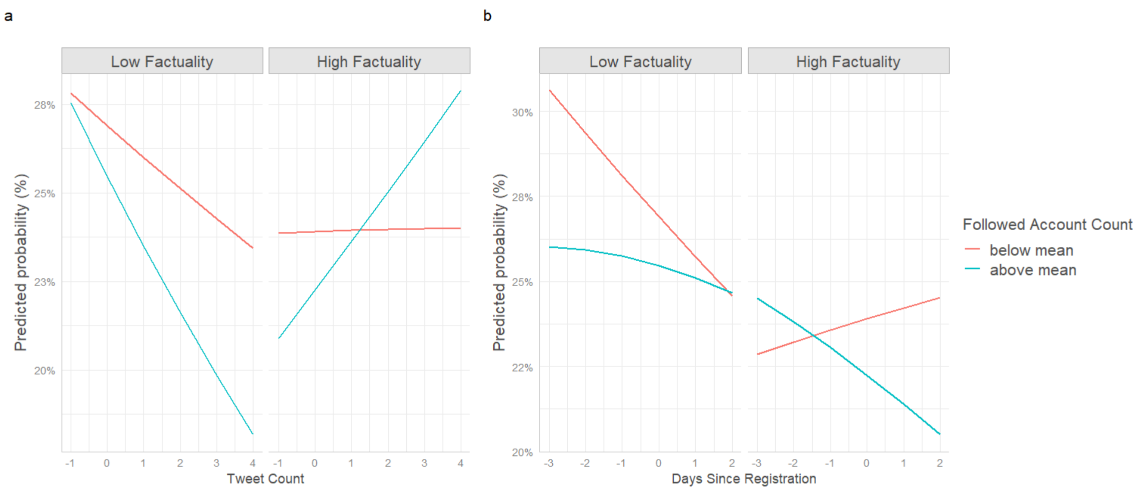


Figure 2.6: **Average marginal effect interactions between followed account count, average daily tweet count, and days since registration.** Panel a: The effect of tweet activity on factuality differs by followed account count, with positive effects on high factuality and negative effects on low factuality primarily among those who follow many accounts. Panel b: The effect of days since registration on factuality also varies by followed account count, with opposite trends for high factuality and consistent negative association with low factuality.

2.5 Discussion

In this study, we aimed to examine whether simple, low-barrier metrics are systematically associated with differences in users' likelihood of sharing misinformation on Twitter. Drawing upon existing literature and leveraging easily accessible data from X, we tested four social network metrics as possible correlates of user factuality. Our analysis focused on follower count, average daily tweet count, number of users followed, and account age. The results offer several insights into the association of these metrics as group-level indicators with misinformation sharing tendencies.

Our findings indicate that when considered individually, each of the four metrics—follower count, average daily tweet count, number of users followed, and account age—has a significant relationship with user factuality scores. However, when all metrics were analysed together using multinomial regression, the distinctive power of follower count and followed account count diminished for the separation of low and high factuality, while average daily tweet count and number of users followed remained robust correlates of both low and high factuality scores. In contrast, the marginal effects associated with follower count and following count were more sensitive to model specification, likely reflecting their high correlation. Nevertheless, robustness analyses excluding each of these variables in turn yielded substantively unchanged results for the remaining predictors (see Figures A.12 and A.13).

Higher average daily tweet counts were linked to lower factuality scores, consistent with our outlined expectation grounded in previous research which suggests that the sheer volume of content users engage with can detract from their ability to discern true from false information. This may be because the higher the number of tweets one shares, the less attention each individual piece of content would receive, making the user more likely to share non-accurate content (167).

Older accounts were associated with higher factuality scores, in line with our expectations that longer presence on the platform correlates with higher content accuracy, which may be moderated by better digital literacy. The finding might be explained by the fact that users with older accounts might have developed stronger digital literacy (specifically on X) as they have had more time to do so, whereas this is not yet true for users with newer accounts, in line with findings from (169). Furthermore, according to diffusion of innovations theory (189), those who adopt new technologies earlier are often more digitally literate, which could explain why older accounts are more factual. We might also find fewer old accounts with low factuality because of certain policies of Twitter removing strongly non-factual accounts.

Several of the activity metrics analysed here—most notably follower count and following count—are strongly correlated, reflecting overlapping aspects of network embeddedness on the platform. While this collinearity limits the precision with which their unique marginal contributions can be disentangled when included simultaneously, additional analyses excluding each metric in turn demonstrate that the core patterns reported in this chapter are robust. Accordingly, these features should be interpreted as jointly characterising broader styles of platform engagement rather than as independent causal predictors. This perspective aligns with the chapter's emphasis on descriptive and interpretive insight rather than individual-level prediction.

The finding that lower-factuality users tend to have higher follower counts should be interpreted in light of the sample construction. Accounts with very large follower counts—typically associated with public figures, organizations, or professional media actors—were excluded by design to focus on regular users. As a result, the follower count variation analysed here reflects differences within the range of non-elite accounts, rather than comparisons between ordinary users and highly prominent public figures. Within this restricted range, higher follower counts may reflect greater engagement intensity or embedding in particular online communities, rather than elite status or institutional visibility.

Finally, the multinomial regression analysis revealed certain significant interaction effects in the case of follower and followed account count. Among users who follow few accounts, tweeting more is less strongly associated with low factuality than among those who follow many accounts, while the effects on high factuality are positive for users who follow many accounts. Among users who follow many accounts, account age is less strongly associated with low factuality than among those who follow few accounts, while among users who follow few accounts, longer time on the platform is linked to higher factuality, whereas among those who follow many accounts, longer platform use is associated with lower factuality. These cross-effects highlight the context-dependent nature of the relationship between these social network metrics and factuality.

These findings have several interpretive implications. For social media platforms, the results indicate that basic metrics such as tweet count and account age are systematically associated with differences in misinformation sharing at the aggregate level, and may therefore help characterize broad user groups that tend to differ in their average sharing behaviour. Rather than enabling individual-level identification, such patterns may inform the design of exploratory, population-level approaches—for example, the contextual deployment of accuracy prompts or digital literacy messaging. Similarly, policymakers may draw on these descriptive regularities when considering low-cost, scalable strategies for addressing misinformation, without relying on these metrics as stand-alone decision tools. It is important to note that the effect sizes identified in this study are modest, and the examined metrics have limited predictive value when used as classifiers of individual behaviour. Accordingly, the contribution of this chapter is interpretive rather than predictive. The findings demonstrate that even highly accessible, low-cost account-level features are systematically associated with misinformation sharing, thereby revealing behavioural regularities in platform use rather than enabling accurate individual-level identification. These regularities may nonetheless be informative for theory-building, descriptive measurement, and the design of future studies or interventions that incorporate richer behavioural or cognitive information. Future research should continue to explore additional accessible data points and refine existing metrics to better understand the behavioural tendencies associated with misinformation sharing.

While our approach provides valuable insights into the identification of users likely to share misinformation on X, several limitations must be acknowledged. First, it is important to note that the analysis presented here is not aimed at individual-level prediction. Instead, we aim to identify broad trends in user behaviour that may be associated with misinformation sharing, rather than develop a robust predictive model. As such, the overlap in distributions among high- and low-factuality users limits the ability to draw definitive conclusions about individual users' likelihood of spreading misinformation.

Moreover, the dataset we used originates from a specific population—followers of a sample of users recruited via Prolific who self-reported their Twitter handles in a survey on vaccine hesitancy. While this initial sample may introduce some bias, our analyses were not limited to the survey participants themselves but expanded to include their first-degree followers. This significantly larger group—over 11,000 users—likely offers a broader behavioural snapshot of the general Twitter population, though some residual similarities may remain. However, rather than aiming for full population representativity, the purpose of our dataset is to compare the behaviour of users who are more prone to share content from unreliable sources to those who are more likely to share reliable ones. However, findings should be interpreted as conditional on the sampling strategy.

A further limitation is that we cannot distinguish between genuine human followers and automated or coordinated accounts. Accordingly, follower count should be interpreted as a descriptive measure of platform-visible reach rather than as an indicator of social endorsement, audience quality, or interpersonal influence.

Importantly, our decision to use this dataset was guided by the goal of studying regular users. Many available public datasets that support factuality scoring rely on keyword-based data collection, which tends to over-represent highly active or politically engaged accounts. These are often not reflective of typical user behaviour. In contrast, our dataset—filtered to exclude bots, verified users, and extreme activity levels—better represents everyday users and their misinformation-related behaviours. As such, while not perfectly representative, it provides an appropriate test case for our research goals.

Beyond sampling considerations, another limitation of our method is the exclusion of users who share domains without Media Bias/Fact Check (MBFC) ratings. This decision restricts the scope of our analysis, as it may not be fully representative of the broader user base, particularly those sharing content from sources that are not evaluated by MBFC. We collected the latest 500 tweets containing URLs for approximately 1.6 million users. Out of these, 240,915 users had ever shared a URL, and among them, only 59,610 had shared at least one URL that could be classified using MBFC ratings. Future research could explore the implications of including a broader range of users to better understand the full spectrum of misinformation-sharing behaviour.

Additionally, our measure of information quality relies on domain-level factuality ratings rather than article-level classifications. As discussed in the Methods, this approach captures systematic patterns of source selection rather than isolated instances of false content, and should therefore be interpreted as reflecting propensity to share from low-credibility information environments rather than misinformation per se.

Beyond limitations related to sampling and content coverage, the generalizability of our findings is influenced by the platform-specific nature of user behaviour. Twitter, a highly public, text-based platform, is distinct from other social media platforms such as Facebook, which emphasizes private networks and long-form content. The social network metrics we use to identify misinformation spreaders on Twitter may not be as effective on platforms where user engagement patterns differ significantly. Therefore, while our findings offer insights specific to X, they may not fully translate to other platforms with different interaction styles.

Network structure variations between platforms also present challenges. Twitter operates on a directed follower-following model, where information flows asymmetrically, whereas platforms such as Facebook and LinkedIn use mutual connections. This structural difference may result in variations in the usefulness of network-based features like follower count and following count when applied to other platforms with more reciprocal interactions.

The impact of content moderation policies further complicates the application of our approach across platforms. Each platform has unique strategies for content distribution and moderation—such as visibility algorithms, labelling practices, and content removal—that shape the dynamics of misinformation spread. These differences may affect how observable and measurable misinformation sharing is, and thus influence the performance of detection methods like ours.

Despite these limitations, our methodological framework, which leverages social network metrics for early identification of misinformation spreaders, holds promise for adaptation to other platforms. However, careful consideration of platform-specific differences is necessary, and future work should explore these variations to determine whether similar patterns of behaviour emerge across different social media environments.

In conclusion, this study demonstrates that basic, easily accessible social network metrics—particularly tweet count and account age—are systematically associated with differences in the factuality of shared content. Rather than enabling individual-level identification or causal inference, the studied metrics provide a descriptive means of characterising broad behavioural patterns that distinguish groups of users who tend to differ in their misinformation-sharing behaviour. Accordingly, these metrics should be interpreted with caution and within the limits of their observational nature. Their use may be most valuable for informing high-level understanding of misinformation dynamics and for guiding future research or exploratory, population-level approaches, rather than serving as stand-alone tools for intervention. Any practical relevance is therefore contingent on contextual factors, and their usefulness may vary across platforms due to differences in user behaviour, content moderation policies, and network structures.

Chapter 3

Study 2: Determinants of Willingness to Donate Data from Social Media Platforms

This chapter is based on the following published article: Zoltán Kmetty, Ádám Stefkovics*, Júlia Számely*, Dongning Deng, Anikó Kellner, Edit Pauló, Elisa Omodei & Júlia Koltai (2025) Determinants of willingness to donate data from social media platforms, Information, Communication & Society, 28:7, 1324-1349,*
DOI: 10.1080/1369118X.2024.2340995.
**These authors contributed equally.*

3.1 Introduction

In the previous study, we explored how easily accessible social media metrics can be used to identify users who are more likely to share misinformation. While such metrics offer a low-barrier way to detect potentially problematic content-sharing patterns, they do not provide deeper insight into the personal characteristics or contextual factors that shape users' behaviour. To move toward a richer understanding, this chapter examines the feasibility of obtaining detailed, individual-level behavioural data through donation-based methods. Specifically, it investigates the determinants of willingness to donate data from social media platforms in two distinct political and cultural contexts: the United States and Hungary.

As outlined in the Introduction, collecting richer behavioural data presents significant ethical and legal challenges. Traditional approaches to digital data access—such as large-scale scraping or platform APIs—have become increasingly constrained by privacy regulations and corporate policies. Data donation, in which individuals voluntarily share their digital traces for research purposes, provides a viable and ethically sound alternative. Yet, the success of such an approach depends on people's readiness to participate, which remains insufficiently understood.

To address this gap, this study employs two vignette experiments embedded in national surveys to test how different request features—such as monetary incentives, number of requested platforms, type of data, and estimated upload time—influence individuals' willingness to donate their Data Download Packages (DDPs). Conducted during the premiership of Viktor Orbán in Hungary and the presidency of Joe Biden in the United States, the cross-national design allows for comparison between contexts that differ both culturally and politically, offering insight into how contextual factors may condition individuals' perceptions of data sharing.

By empirically identifying the factors that enhance or hinder participation in data-donation studies, this chapter contributes practical guidance for the ethical and effective collection of detailed behavioural data. It also provides a bridge between the scalable but limited analyses of Study 1 and the rich, consent-based dataset analysed in Study 3. In doing so, it lays the foundation for combining survey and digital-trace measures to achieve a more comprehensive understanding of misinformation engagement.

3.2 Motivation

Digital traces on social media platforms can be promising sources of information for researchers in various fields (190). In contrast to self-reports from surveys, which may imply measurement errors due to recall or social desirability bias (110; 191; 192; 193), digital traces are assumed to provide reliable behavioural data. Surveys, on the other hand, allow researchers to gain access to self-reported beliefs and attitudes, as well as the socio-demographic information of the respondents. Linking individual digital behavioural data with survey responses holds the promise of giving context of behaviour on social media platforms, and therefore scientific attention in combining these data sources has substantially grown recently.

However, accessing digital social media data has recently become increasingly difficult. After the Cambridge Analytica scandal in 2018, many platforms, including Facebook, decided to strongly restrict the use of Application Programming Interfaces (APIs) (131; 132; 133), which pushed researchers to explore other approaches, such as web-scraping (194; 195), novel partnerships between industry and academia (196), passive data collection methods as web-tracking and browser plug-ins (137; 138) or data donation (132; 139; 140). The chosen data collection approach may be influenced by the research question, the resources available, and even the skills and capabilities of the research team (197). Out of these solutions, this article focuses on data donation involving data download packages (DDPs, (139)). DDPs are a collection of historical user data stored on social media platforms consisting of behavioural (e.g., likes), textual (messages), media (photos, videos), or location data. Social media users have been able to download their DDPs from major platforms because GDPR obliged platforms to provide access to their users to the data collected about them. Given the unique scientific value of this data, researchers have recently shown great interest in data donation. In data donation studies participants are typically recruited with standard survey sampling techniques, then asked to download their own social media data through DDPs and provide them to the researchers for analysis with informed consent (132; 139). One key advantage of the approach is that it helps overcome some of the privacy concerns of the API approach, such as consent giving. Whilst in the case of platform-centric data collection users' ability to monitor and understand who uses their data and how they use it is limited (198), user-centric approaches offer users greater transparency and the chance to provide consent under clear research terms (132). Users' data archives, moreover, provide a richer set of data compared to API or scraping approaches, which can support various analytical goals or can be easily linked to other (such as survey) data. Nevertheless, data donation is more burdensome compared to other, passive modes of data sharing such as web-browser plug-ins (137; 138), since participants are required to actively download their data and then share it with the researchers. While data donation holds great promise for studying online human behaviour and some recent studies reported promising results (132; 199; 200; 201), little is known about how to best optimize such approaches.

This study aimed to understand the underlying mechanisms of willingness to participate in academic research as a social media data donor. Earlier research has shown that digital data-sharing behaviour in a broader sense can be a function of, for instance, the offered incentives, the study's sponsor and various other factors (see e.g., 138; 141; 142; 202). Our research expands existing knowledge by examining willingness to share DDPs and addressing previously uncovered factors (such as the type of social media data). We further contribute to the literature by running the same vignette experiments in two countries (the United States and Hungary) allowing for a cross-national comparison. Our results can add to the understanding of the characteristics of non-response bias in the case of data donation, which can raise awareness of the dimensions, thereby limiting the generalization of the results (140). These results also provide grounds for non-response adjustments. Additionally, uncovering the determinants of willingness to donate such data can directly help researchers improve their recruitment strategies and research design.

3.3 Determinants of data sharing behaviour and hypotheses

3.3.1 Incentives

Social exchange theory (203) implies that research participants' willingness of engagement is affected by their assessment of the rewards and costs of the action (204; 205). Survey respondents are similarly expected to weigh the pros and cons of participation as proposed by both survey participation theory (206) and leverage-saliency theory (207). Perceived rewards include internal rewards (e.g., feeling satisfied from scientific contribution) and external rewards (e.g., getting incentives such as price rewards). Monetary incentives are one of the most common approaches to motivate participation (205). Many empirical findings have proven the effectiveness of monetary incentives in motivating people in research participation, such as responding to web surveys (see (208) for a meta-analysis or (205) for a review), and sharing digital trace data (137; 138). The results of (138) showed that providing monetary incentives can have a positive effect on the willingness to share social media data. Similar effects have been found regarding willingness to share active and passive mobile data (137; 209; 210). Additionally, earlier studies (137; 211) have found that incentives that are too low were among the two main reasons for not participating in digital data sharing[1]. By contrast, (212) and (213) found no effect of incentives on downloading an app or sharing different types of data. (138) also reported that incentives do not necessarily motivate all types of data-sharing (e.g., they do not work in the case of sharing highly sensitive health data).

In the case of data donation, little is known about the optimal amount and method of offering incentives. Too high incentives may make respondents suspicious and assume that their data are highly valuable, eventually decreasing their likelihood of participation (138). The challenge here is to estimate the optimal amount of incentive. In the context of data donation, relatively high effort is required from participants. To match with this high cost, we assume that a higher monetary incentive will make respondents more willing to donate their data. As we employed incentive levels only somewhat higher than incentives of regular survey requests, we did not expect any non-linear or counterproductive effects associated with the highest incentive amounts.

H1a. Higher monetary incentives increase the willingness to donate data.

While various types of non-monetary incentives could be considered, in this paper, we study the effect of a summary report as a proxy for non-monetary incentives for comparability with the mentioned use of summary reports by other studies (205; 214; 215). Summary reports may contain interesting findings such as contrasts between participant-specific data and the overall sample. Offering summary reports has been successfully used to motivate participation in business-to-business (205) and medical web surveys (214; 215). However, for other web surveys, such non-monetary incentives had no or even a negative impact (see (205) for a review). Despite the mixed evidence in the literature (which may be due to the diverse range of outcomes examined in the literature), theoretically it is more plausible to expect a positive association. Thus, we hypothesize that offering summary reports will have a positive impact on willingness to donate data.

H1b. Offering summary reports increases willingness to donate data.

3.3.2 The number of platforms and the time required to download and upload data

When deciding on participation, respondents also take into account the costs that derive from the perceived *difficulty* and *burden* of the data-sharing process. Both difficulty and burden can be higher in the case of data donation compared to other passive digital data collection methods (138; 205). (216) distinguished four elements of respondents' burden regarding surveys:

[1] The other main reason was privacy concern.

length, frequency, required effort, and caused stress. Following Bradburn's dimensions, the length of the task increases with the number of platforms involved in the request (clicks, download and upload time), as well as the required effort. Therefore, the more platforms involved in the request, the greater burden the participants may have. Each platform requires a somewhat different task from that of the downloader because the download procedure varies by platform. Moreover, the more platforms involved, the more likely respondents' privacy concerns emerge.

Time can also be a type of burden. In the case of data donation, the downloading, processing, and uploading time can last for hours or days, which is longer than usual survey participation.

According to these results, we expect that willingness to donate will decrease with the increase in the number of requested platforms, and that a declared longer download/ upload time will also hinder cooperation.

H2a. Asking for more platforms decreases the willingness to donate data.

H2b. A longer download/upload time of the data decreases the willingness to donate data.

3.3.3 Types of data

On social media platforms, users can share different types of user-generated content, such as textual content (tweets, posts, comments) or audio-visual materials (pictures and videos), which can be valuable data sources for research (132). When people are asked about their participation in a data donation study, the type of data requested may affect how they decide. The level of sensitivity of the data to be donated is a particularly relevant aspect of these decisions. One of (216)'s respondent burdens is the stress caused by the task, i.e., the discomfort the respondents feel while participating in the study. Such stress may derive from privacy issues: participants can be reluctant to be involved in the data collection because, for instance, the data to be shared is sensitive or confidential (132). (141) examined willingness to donate data in non-experimental hypothetical settings and found that higher perceived data sensitivity was associated with lower willingness to donate; moreover, the level of perceived sensitivity of data was the most influential factor in determining the willingness to donate.

Nevertheless, it is not completely clear which types of digital data are considered sensitive by the users. Photos and videos can be sensitive because people can be directly or indirectly identified, which can lead to users' (and others') exposure to several risks (217). Geolocation data can similarly uncover personal information, such as home, work, or school addresses, therefore likely evoking privacy concerns (218).

(138) reported the highest willingness to share Spotify data (59.1%, musical data), and the lowest for Facebook (31.2%, various types of data) with Twitter in the middle (41.4%, mostly textual content). In their other survey, 24% of the users shared their Twitter data, whereas less than 10% shared their health app data. (213) compared consent rates for seven types of domains and found that willingness to share Facebook data was relatively high (above 50%) compared

to other administrative domains such as bank account or health insurance data sharing. (141) found that compared to their activities on Google, people demonstrate a higher willingness to share what they encounter on social media. Additionally, individuals are least inclined to donate data that includes details about their personal social network, encompassing friends and followers.

Research on sensor-based data sharing also shows that willingness to share is partly a function of the requested data type. (218) found that respondents were less willing to share the GPS location of their smartphone compared to other tasks involving less confidential data (e.g., completing a questionnaire, or downloading an app). (211) reported that respondents' willingness to take and share pictures was higher compared to providing access to their Facebook account or sharing geolocation data. In contrast, in the study of (202), tasks that involved photographing or video-taking one's surroundings at home yielded far lower willingness-to-share rates compared to providing geolocation data with a sensor (see (142) for a similar approach with somewhat different results). Nevertheless, (219) and (210) did not find significant differences between willingness to share different, passively collected data.

To sum up, earlier findings suggest that the type of data matters in the decision on participation, and respondents perceive geolocation data as less private compared to photos and videos. Nevertheless, reluctance to share audio-visual materials strongly depends on the content that is being recorded. Additionally, most of these studies requested respondents to actively take pictures or videos for the current study, thus participants had more control over what they shared. In the case of a social media data donation, screening all content before the donation is more burdensome. For this very reason, in our hypotheses, we did not differentiate between willingness to share audio-visual materials and geolocation data but expected that willingness to share would be lower in both cases, so when the request involves pictures and videos, as well as geolocation data.

H3a. Asking for pictures/videos of the user decreases the willingness to donate data.

H3b. Asking for geolocation data of the user decreases the willingness to donate data.

3.3.4 Respondents' characteristics

Finally, we summarize respondent characteristics that might affect willingness to donate data. Given that the main focus of this paper was the impact of the details of a hypothetical request on participation willingness, no particular hypotheses were developed for the following factors.

Participating in a data donation study is an active task that requires a basic level of digital literacy (199), and therefore can place more burden on people with lower technical skills. Self-assessed smartphone skills in earlier studies, however, show contradictory results: (218) and (192) found that participants with higher phone skills were more willing to participate, but in the study of (137) and (142) these skills were not associated with the willingness to participate.

As we discussed earlier, privacy concerns are likely to evoke when sensitive and private data is involved in data donation. Previous studies show that privacy and security concerns lower the willingness to participate (141; 142; 202; 211; 212; 218), but see (220) and (138) for null results. Indeed, they are the most mentioned reasons for not being willing to participate in digital data sharing (137; 211; 212).

Psychological traits can also play a role. Earlier research examined the effect of the Big Five traits on item/unit non-response, or attrition in surveys (221; 222; 223; 224), and on downloading an app and sharing passive (220) or social media data (138). The results of these studies are inconsistent. Some studies found that conscientiousness, one of the Big Five traits, can increase participation while sharing GPS data is higher among introverts (220). However, (138) did not find such positive effects regarding these psychological traits on sharing social media or health data.

Willingness to share data can also be a function of usage of the platform or device. Generally, it can be assumed that heavy users may be more open to data sharing (132). (138), for instance, found that a higher platform usage increased the likelihood of sharing social media data, but not health app data. Heavy users of social media may be more likely to share their data because they often have a deeper engagement and familiarity with these platforms. Presumably, self-reports about their activity may be more motivating for them. Additionally, they are typically more tech-savvy thus they likely perceive data-sharing tasks as less difficult.

Finally, among socio-demographic factors, age is one of the major determinants of digital technology use and attitudes toward digital technologies as well (210). Age likely correlates with the willingness to participate in digital data collection, as participation willingness is typically higher in the younger generation (210; 220; 225), and it decreases after age 50 (226). Gender was also found to influence willingness to donate digital data in earlier studies (138). One possible explanation for the potential differences is that males and females use social media in different ways. For instance, (227) found that men predominantly utilize Facebook to initiate new relationships, while women primarily engage in sustaining existing connections. Privacy concerns were also found to differ between males and females with females showing a tendency to have higher privacy concerns and to consistently exhibit behaviours aimed at protecting privacy safeguards on Facebook (138; 228). Such differences can mediate willingness to donate social media data (see (138) who found that men were more likely to share Facebook data). Prior research also shows that those with higher levels of education may be less likely to share their Facebook or Twitter data (138). Educational differences are likely associated with differences in social media use, privacy concerns, etc. Lastly, financial situation and subjective wealth may play a role in the willingness to participate. People with different financial situations likely vary in how individuals perceive and react to incentives, in their altruistic behaviour or access to technology.

3.4 Data and Methods

3.4.1 Data

We collected two datasets for this study and conducted one survey experiment in Hungary and another one with a similar design in the U.S[2]. In Hungary, the survey was administered by an online polling company, NRC, on a non-probability access panel. The NRC panel consists of more than 140,000 people. Compared to the general population of Hungary, individuals with a high level of education and from bigger cities are overrepresented in the panel. We used a quota sampling method (with quotas for gender, age, and geographical region) to ensure equal representation. The company gives regular incentives for the respondents of their studies. Altogether 1,000 respondents participated in the Hungarian study. The fieldwork was carried out between 11-25, May 2022.

The U.S. dataset is a based on a panel of Harvard University called Harvard Digital Lab for Social Science (DLABSS). DLABSS is a pool of survey respondents primarily recruited using social media and other free sources. Respondents do not get incentives for filling out surveys in the panel. The size of DLABSS's pool is growing rapidly, currently counting nearly 30,000 volunteers. A study by (229) found that such volunteer panels can replicate classic and contemporary social science findings and produce high levels of overall response quality comparable to paid subjects. In the end, 844 respondents participated in the U.S. study. The fieldwork was carried out at the end of 2022, between October 14 and November 8. Both data collections and studies were pre-registered before the data collection[3].

3.4.2 Design of the survey experiment

We built up the survey experiments similarly in the two countries. We applied a mixed factorial vignette design. Respondents had to evaluate multiple situations (called vignettes), in which we

[2]Countries were chosen on a practical basis and opportunity to collect data. Nevertheless, we believe that collecting data from different countries allows us to test the robustness of our findings and assess whether cultural factors influence attitudes and behaviors towards data donation. There are several reasons why comparing these two countries is relevant. For instance, compared to Hungary, the U.S. is a highly developed economy with a strong focus on technological innovation, which may impact how citizens react to digital requests. Internet and social media usage (or platforms) and digital literacy are also different in the two countries. These factors can influence people's understanding of data donation and their readiness to participate in such initiatives. Trust in science, trust in institutions or interpersonal trust may all affect willingness to donate data. The levels of trust also show some, although not large differences between the two countries. Lastly, data privacy regulations are markedly different in Europe compared to the US which can lead to differing levels of public awareness and trust in how data is handled and protected. At the same time, identifying cultural differences was not the main focus of this study, thus we did not develop any comparative hypotheses. Conducting data collection across two countries enhances the robustness of our study by introducing a more diverse sample, thereby increasing the generalizability and reliability of our findings.

[3]The anonymised pre-registration for the Hungarian study is available here: https://osf.io/r4kxm?view_only=af4b7da2aba14495b2e5df280b68a37d ; and for the U.S. study is here: https://osf.io/tvefj/?view_only=728b0fa3b66a4c47abcebc30dd07b08e

manipulated various dimensions of a fictional research. At each vignette, they had to provide the likelihood of their willingness to donate their digital data in such research on a 0 to 10 scale. These manipulated dimensions of the fictional research were the following (See Table 3.1.). An example of the vignettes is available in the Supplementary Information.

Dimension	**Levels**	**Explanation**
Platform	- Facebook - Facebook and Google - Facebook and other social media sites you use (Instagram, Twitter, Spotify) - Facebook, Google and other social media sites you use (Instagram, Twitter, Spotify)	Our research was the first step in a more extensive data donation project. Facebook data plays a major role in our data donation research, so we wanted to keep this platform as a reference. Facebook was the most popular social media platform in both the Hungarian (92%) and U.S. (71%) samples.
Range of data	- all, except: private messages - all, except: private messages and location - all, except: private messages and photo, videos - all, except: private messages, and photo, videos, and location	Private messages include messages from the user and their conversation partners, so we did not consider this to be shareable data, despite the participant's consent.
Time to download/upload data	- Less than an hour - More than an hour	DDP data is never made immediately available by the platforms and would have to wait hours or even days to become available for download. With a download/upload time of more than one hour, we wanted to explore whether it makes a difference if the respondent cannot resolve the request within one session.
Incentives	- 3000 HUF/ $10 - 5000 HUF / $20 - 10 000 HUF /$30	3000 HUF was around 8 U.S. dollars during the data collection. In order to standardize the money incentives in the two surveys, we converted the Hungarian Forint to U.S. dollars and adjusted it with purchasing power parities (see: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm)
Additional report	- Yes - No	Additional reports mean feedback and summaries on social media usage, such as activity patterns, closest friends based on activity, or the number of friends over time.

Table 3.1: Manipulated dimensions and their levels in the survey experiment

Altogether we had 192 possible combinations of the dimensions in the following way: 4 [platform] * 4 [range of data] * 2 [time] * 3 [incentive] * 2 [report]. In the Hungarian study, we created 16 decks (packages of vignettes) and assigned 12 vignettes to each deck (16*12=192). Thus, one respondent had to evaluate 12 different situations. With the expected (and then realized) 1,000 respondents we had around 67 respondents per deck. The twelve vignettes per respondent is a relatively high number. To assess the validity of the results, we also conducted a robustness check (see the results section for more details).

The U.S. data comes from a volunteer panel, where we expected a higher dropout rate in a repetitive task than among paid panel members, like the Hungarian ones. To overcome the possible bias and high dropout rate caused by the large number of vignettes per respondent, we followed a slightly different strategy in the U.S. study. In this study, we created 16 decks as well, but only assigned five vignettes to each deck. We used the optBlock function of the AlgDesign package (230) in R to find the best design. This function uses the D criterium to optimize vignette allocation. With this design, we had around 52 respondents per deck.

According to our previous power calculation, with this design 700 respondents would have been enough to achieve a 0.95 power with a 0.1 (small) effect size in both countries. The final sample sizes were over this limit.

3.4.3 Variables

We used the following independent variables in the analysis: gender, highest level of education, age, subjective wealth, frequency of social media usage, number of social media platforms used by the respondent, Internet Users' Information Privacy Concerns (IUIPC) scale, privacy concerns, Affinity for Technology Interaction Scale big five (BF) inventory. A detailed description of the independent variables is available in the Supplementary Information (Section A, and Table A1).

There were no missing values in the dependent variable, only in the independent ones. In 26 and 28 percent of the cases in the Hungarian and U.S. study, there was at least one missing variable. To handle these missing values in the dataset in order not to lose too many cases, we applied multiple imputations. We included all the independent variables in the imputation process and used predictive-mean-matching (PMM) for the procedure. We created five imputed datasets and calculated the pool results in the regression models. We used the 'mice' package of R (231) for these calculations.

3.4.4 Analytical strategy

As the first step of our analysis, we applied a variance component model to understand how much of the variation in the response variable – willingness to donate digital footprint data – is explained by vignette level and respondent level characteristics. It is important to note that this variance component model partitions the total variance of the outcome into respondent-level and vignette-level components, rather than indicating variance explained by predictors. Thus, the reported proportions of vignette and individual level variance sum to 100% by definition and do not imply the absence of within-respondent (residual) variance. Substantial within-respondent variance remains, as indicated by the estimated residual variance in all models.

As a next step, a set of multilevel regressions were performed as we had two levels in the data: one for the vignettes and another one for the respondents (because one respondent evaluated multiple vignettes). In the regression models, we allowed for random intercepts by the respondents, first regressing only vignette-level variables on willingness to donate data, then we added respondent-level characteristics as well.

Some respondents did not use the platforms we mentioned in the situations. This is particularly important in the case of Facebook, which was used as a reference, or Google data, which was mentioned in many situations. To deal with the potential bias that could arise from this, we also ran models that filtered out those who actively used Facebook and those who had Google data in addition to Facebook.

We carried out the analysis using the 'lme4' (232) and related packages in R.

3.5 Results

3.5.1 Study 1. Hungary

Twelve vignettes were assigned to 1000 respondents in the Hungarian study, thus we had 12,000 cases on the vignette level. In thirty percent of the vignettes, respondents indicated that it is not likely at all that they would donate their data under the given circumstances, while maximal willingness was shown in 14 percent of the vignette cases. The mean value of the willingness questions was 4.2 on the 0 to 10 scale[4]. On the respondent level, 18 percent refused any kind of data sharing, regardless of the vignette content (answered zero to all twelve vignette situations they evaluated).

In the first step of the analysis, we applied a variance component model on the Hungarian dataset (Table 2, first column). The results of this analysis showed that variation at the deck level is not significant, while variation on the respondent level explains 79.7 percent of the variation in the willingness to donate data, and the remaining 20.3 percent of the variation is explained by the vignette level.

In the next step of the analysis, we added the vignette-level variables. Regressing the outcome variable on the vignette-level explanatory variables (Table 2, second column) showed that *incentive*, *platform*, and *data type* have significant effects on the outcome variable (with

[4]For a robustness check of the results, we calculated the standard deviation of the willingness probability for the first and second six vignettes. A smaller standard deviation might have been a sign of fatigue for the respondent. Based on Barlett's test, we did not find differences between the standard deviations (p =.39) of the two sets.

incentive having the strongest effect), while the effect of *report* provision and the *time* to download/upload data are not significant (see the relative strength of effects in Figure 3.1). The effect of the *incentive* variable on the willingness to provide data is positive, with a 0.25-point increase in the expected value of the outcome variable with each additional amount of HUF worth 10 USD at Purchasing Power Parity (PPP). The effect of the *platform* variable means that compared to donating digital footprint data from the respondent's Facebook account only, the more platforms the respondent is required to provide digital footprint data from, the less likely they are to do so. The effect of the *type* variable means that as compared to providing all data except private messages, respondents are on average less likely to provide their data if private messages and photos and videos are excluded, as well as if private messages, photos, videos, and location are excluded. As these results are quite counter-intuitive, we will get back to their explanation in the discussion[5]. The alternative models, in which we restricted our sample to Facebook users and Facebook and Google users, showed no difference compared to the results obtained on the total sample (see Table A3 in the Supplementary Information).

Next, we added a set of respondent-level explanatory variables to the regression, allowing for random intercept by respondents (Table 2, third column). The results show that with the inclusion of control variables, the same vignette-level variables remain significant as in the previous setup, i.e. H1a, H2a, H3a, and H3b remain confirmed, and H1b, H2b remain contradicted. The strength of the vignette-level variables does not change significantly either as compared to the regression with only vignette-level variables. Moreover, *gender* and *age*, as well as *education* are significant in explaining willingness to donate data. Female respondents are on average less likely to donate data, the older the respondent the less likely they are to donate, and the higher the level of the respondent's education the less likely they are to be willing to provide their digital footprint data. *Subjective wealth* showed no significant effect on the outcome variable.

Of the other control variables, only the number of platforms visited had a significant effect on the likelihood of sharing data. Those using multiple platforms were more likely to share their data. After including the control variables, the explanatory power of the model went up to 7.3 percent.

3.5.2 Study 2. USA

We had 844 respondents in the U.S. study with 5 vignette evaluations. resulting in 4,174 vignette evaluations. 59 percent of the cases on the vignette level respondents answered that it is not likely at all that they would donate their data under the given circumstances, and we

[5]For robustness check, we re-ran this multilevel model with the first six and second six vignettes separately (see Table A2 in the supplementary). There were some differences between the evaluations of the first and second six vignettes. Still, the incentive had the most pronounced positive effect in the regressions fitted to both vignette groups. For the second six vignettes, fewer variables have a significant impact which may indicate fatigue and less attentive evaluation.

	Null model			Model with vignette dimensions			+ controls		
	Estimate	SE	p	Estimate	SE	p	Estimate	SE	p
(Intercept)	4.25	0.11	<0.001	3.55	0.12	<0.001	4.93	1.17	<0.001
Incentive				0.25	0.01	<0.001	0.25	0.01	<0.001
Report				0.01	0.03	0.66	0.01	0.03	0.68
Platform: FB + Google				-0.19	0.04	<0.001	-0.19	0.04	<0.001
Platform: FB + other				-0.21	0.04	<0.001	-0.21	0.04	<0.001
Platform: FB + Google + Other				-0.18	0.04	<0.001	-0.18	0.04	<0.001
Time				-0.03	0.03	0.34	-0.03	0.03	0.34
Type of data: no PM/loc				-0.06	0.04	0.14	-0.06	0.04	0.13
Type of data: no PM/vid				-0.19	0.04	<0.001	-0.20	0.04	<0.001
Type of data: no PM/loc/vid				-0.22	0.04	<0.001	-0.22	0.04	<0.001
Controls									
Gender							-0.59	0.22	0.01
Age							-0.03	0.01	<0.001
Education							-0.25	0.10	0.01
Subjective wealth							-0.11	0.12	0.36
IUIPC_control							0.07	0.10	0.46
IUIPC_collect							-0.12	0.08	0.20
Privacy beliefs							0.12	0.09	0.23
Tech attitudes							0.12	0.11	0.27
BF: openness							0.09	0.16	0.57
BF: conscientiousness							-0.05	0.17	0.77
BF: extroversion							0.00	0.10	0.98
BF: agreeability							0.06	0.17	0.73
BF: neuroticism							-0.01	0.09	0.88
Social media usage frequency							0.19	0.14	0.18
No. of platforms							0.14	0.06	0.03
Variances of random effects									
Variance: constant	11.14			11.18			10.37		
Variance: residual	2.84			2.59			2.59		
Proportion of Level 1 variance	20.3%			18.8%			20.0%		
Proportion of Level 2 variance	79.7%			81.2%			80.0%		
Model fit									
Variance explained (Level 1)				8.8%			8.8%		
Variance explained (Level 2)				0.0%			6.9%		
Variance explained (overall)				1.5%			7.3%		

Table 3.2: **Results about willingness to donate data – Hungary** (multilevel mixed-effects linear regression)
Note: Values are standardized coefficients. SE = standard error.

observed the highest level of willingness in only 5 percent of the cases. The mean value of willingness was 2.1 on a 0 to 10 scale, where higher values mean higher willingness. On the respondent level, 52 percent mentioned that it is not likely at all that they would share their data regardless of the vignette content (answered zero to all 5 vignette situations). Overall, the willingness rate was much lower in the U.S. sample, than in the Hungarian one.

Similarly to the analysis of the Hungarian dataset, in the first step of the analysis, we applied a variance component model on the U.S. dataset (Table 3, first column). The results of this analysis show that in the U.S. dataset, the individual level explains 85.2 percent of the variation of the outcome variable, the deck level explains 0.6 percent of the variation, while the remaining 14.2 percent of the outcome variable's variation is explained by the vignette level.

When we added the vignette-level variables (Table 3, second column), we found that with the exception of *data type*, all vignette-level variables have significant effects on the willingness to donate data. Increasing the *incentive* has a positive effect on willingness to donate, 0.35-point increase in willingness with every additional ten USD, the same effect as in the Hungarian case. Offering a *report* has a significant positive effect on the outcome variable. Asking for data from more *platforms* affects willingness positively – contrary to the effect found in the Hungarian data. The effect of the *time* of download/upload is negative, and while pointing in the same direction, it is an order of magnitude larger than in the Hungarian data. As the results of the regressions with standardized variables show (Figure 3.1), similarly to the Hungarian dataset, the effect of *incentive* is the strongest among the vignette-level variables. These results confirm hypotheses H1a, H1b, H2b, and H3b, and contradict hypotheses H2a and H3a in the U.S. dataset. H1a was therefore confirmed by both the Hungarian and U.S. datasets, while the evaluation of the rest of the hypotheses varied across the two datasets. When we narrowed the models down to Facebook users and Facebook and Google users, the platform effect disappeared (see A4 tables in the Supplementary Information). We will return to this result in the discussion section.

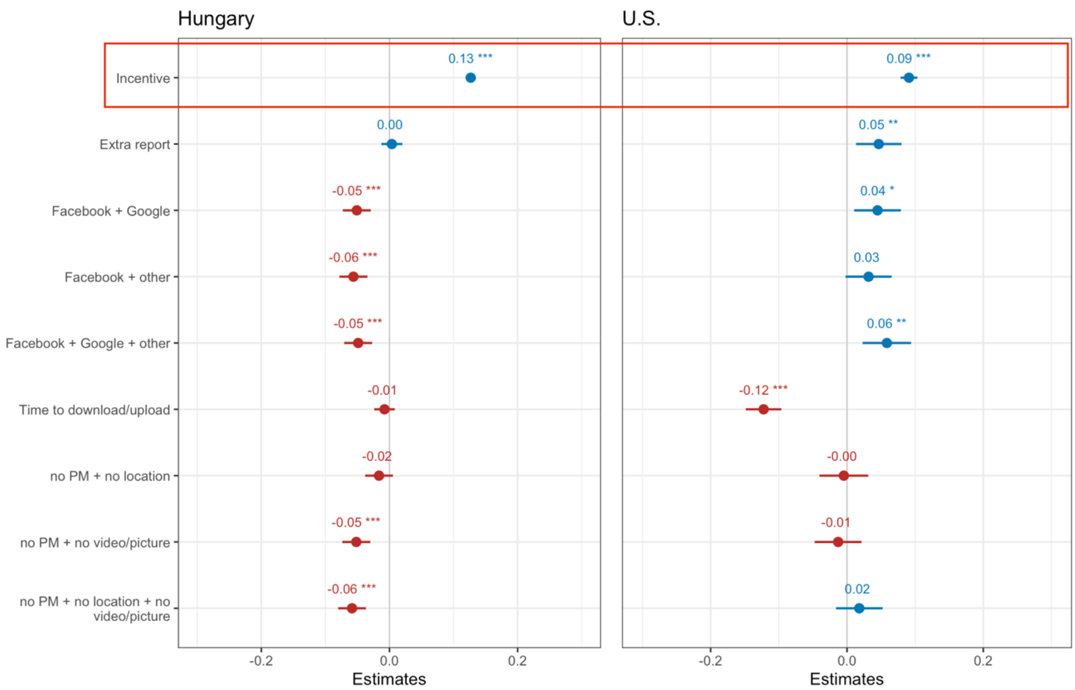


Figure 3.1: **Relative effects of vignette-level characteristics on willingness to donate data.** The figure displays standardized regression coefficients estimated from models including only vignette-level independent variables, allowing for random intercepts. Results are shown separately for Hungary and the United States.

Next, introducing respondent-level variables in the regression, while allowing for random intercepts by respondents (Table 3, third column), we found that the significance of the effects of vignette-level variables do not change, and the magnitudes of coefficients change only slightly compared to the previous model, which only included vignette-level independent variables. Gender has a significant effect, such that female respondents are more likely to donate their data (opposite as in Hungary)[6]. The effect of one of the *IUIPC* indicators is significant, specifically, having a more positive opinion about how one's personal data is generally collected affects willingness to donate data positively. The number of used *platforms* is positively associated with the outcome variable, and also the frequency of social media usage. After including the control variables, the model's explanatory power went up to 14,3, which is higher than the Hungarian case.

3.6 Discussion

This study aimed at understanding the mechanisms underlying the respondents' willingness to participate in an academic study as a social media data donor. To this end, we designed two

[6] In the US sample, 7 percent of respondents said they belonged to Other gender category or did not answer the gender question. We imputed data of these respondents in the main models, but we also run an alternative model where the Other and DK/NA categories are combined and included as a third category in the analysis. The results of these modelling runs are presented in Table A5. The three-category variable performed very similarly to the two-category variable, with female having higher willingness to participate.

	Null model			Model with vignette dimensions			+ controls		
	Estimate	SE	p	Estimate	SE	p	Estimate	SE	p
(Intercept)	2.07	0.12	<0.001	1.39	0.12	<0.001	0.47	0.96	0.62
Incentive				0.35	0.02	<0.001	0.35	0.02	<0.001
Report				0.15	0.05	0.01	0.15	0.05	<0.001
Platform: FB + Google				0.14	0.05	0.01	0.14	0.06	0.01
Platform: FB + other				0.10	0.05	0.07	0.10	0.05	0.06
Platform: FB + Google + Other				0.18	0.06	<0.001	0.19	0.06	<0.001
Time				-0.38	0.04	<0.001	-0.38	0.04	<0.001
Type of data: no PM/loc				-0.01	0.06	0.80	-0.01	0.06	0.82
Type of data: no PM/vid				-0.04	0.05	0.46	-0.03	0.06	0.53
Type of data: no PM/loc/vid				0.06	0.05	0.30	0.06	0.06	0.30
Controls									
Gender							0.35	0.16	0.03
Age							-0.01	0.01	0.08
Education							-0.01	0.06	0.84
Subjective wealth							-0.14	0.10	0.17
IUIPC_control							0.18	0.17	0.32
IUIPC_collect							-0.54	0.13	<0.001
Privacy beliefs							0.12	0.13	0.39
Tech attitudes							0.06	0.12	0.62
BF: openness							0.07	0.18	0.70
BF: conscientiousness							-0.36	0.18	0.05
BF: extroversion							-0.02	0.10	0.80
BF: agreeability							0.24	0.19	0.23
BF: neuroticism							0.03	0.05	0.57
Social media usage frequency							0.27	0.12	0.03
No. of platforms							0.24	0.09	0.01
Variances of random effects									
Variance: constant	8.27			8.28			6.99		
Variance: residual	1.44			1.32			1.33		
Proportion of Level 1 variance	14.8%			13.8%			16.0%		
Proportion of Level 2 variance	85.2%			86.3%			84.0%		
Model fit									
Variance explained (Level 1)				8.3%			7.6%		
Variance explained (Level 2)				0.0%			15.5%		
Variance explained (overall)				1.1%			14.3%		

Table 3.3: **Results about willingness to donate data – U.S.** (multilevel mixed-effects linear regression)
Note: Values are standardized coefficients. SE = standard error.

vignette experiments embedded in two online surveys conducted in Hungary and in the U.S. In hypothetical requests for donating social media via DDPs, we manipulated the amount of monetary incentives (1), the presence or lack of non-monetary incentives (2), the number of platforms to which one is requested to donate (3), the estimated upload/download time (4), and the type of data to be donated (5). The results revealed that data-sharing attitudes are subject to the parameters of the actual study, and some respondent characteristics.

Monetary incentives were the strongest motivators of willingness to donate data in both countries, although the effect was stronger in the Hungarian sample. This finding is consistent with earlier results (138; 210; 233; 234).

Non-monetary incentives had a positive effect on willingness in the U.S. sample but not in the Hungarian one. This difference can be linked to the differences between the two online panels. While the Hungarian panel is a standard access panel where panel members regularly receive points and even monetary incentives for completion, the DLABSS panel is fully volunteer based. Receiving a summary report of the participant's digital behavior compared to others can be more motivating for volunteer panel members than for members who normally answer surveys for monetary incentives. The impact of the summary report was examined in terms of respondent participation, but there may also be a positive effect of such a report in that if participants share their completed reports with their friends, it may even bring new participants into the research, as a kind of snowballing method.

The perceived cognitive burden or difficulty of the task inconsistently influenced the willingness to donate in the two countries. In line with our hypotheses, the more platforms were included in the request the less likely Hungarian respondents would have participated in the study, while we found an opposite effect in the U.S. sample. ~~The unexpected platform effect in the U.S. results might be because those who did not register for some platforms were not motivated to consider the role of the platforms in the data-sharing process. This hypothesis is supported by the models where we filter our sample to those with platform access. In this sub-sample, the platform effect was no longer significant, but the effect of the other variables remained unchanged.~~ Several complementary mechanisms may help explain this divergence. First, differences in digital literacy may shape how respondents evaluate the marginal burden associated with additional platforms. Individuals with lower levels of digital literacy may perceive each additional platform request as imposing a substantially higher technical and cognitive cost, as data donation typically requires navigating platform-specific interfaces and completing multiple procedural steps. If respondents in the Hungarian sample—recruited from a paid access panel—have, on average, lower digital proficiency than respondents in the U.S. volunteer sample, each additional platform may disproportionately reduce willingness. This interpretation is consistent with evidence that realized data donation depends not only on stated willingness but also on capacity-related factors such as age and education, which may proxy digital literacy (235). Second, platform salience and familiarity may differ across contexts. If respondents in the U.S. sample are more likely to actively use or be familiar with multiple platforms beyond Facebook, the incremental burden of including additional platforms may be relatively small. In contrast, unfamiliarity with certain platforms in the Hungarian sample may increase perceived complexity, leading to a negative marginal effect of requesting data from additional sources. In other words, adding platforms may feel more burdensome when those platforms are unfamiliar, even if the number of required steps increases only modestly. Third, differences in perceived legitimacy and participation norms may further moderate responses to multi-platform requests. Respondents in a volunteer-based research panel may approach data donation with a "contribution" mindset, interpreting broader data requests as enhancing the scientific value of the study. In contrast, respondents recruited through paid panels may approach participation more instrumentally, evaluating additional platforms primarily in terms of added burden relative to compensation. These differing frames may help explain why requesting more platforms increases willingness in the U.S. but decreases it in Hungary. Finally, as noted in the manuscript, the observed positive platform effect in the U.S. sample appears to be driven by respondents who do not have accounts on all requested platforms. When restricting the analysis to respondents with access to the relevant platforms, the platform effect becomes statistically insignificant, while the effects of other experimental features remain stable. This pattern suggests that the initial positive association may reflect that respondents without accounts on these platforms did not carefully consider the platform-related aspects of the task, rather than a genuine increase in willingness among active multi-platform users. Taken together, these findings suggest that the effect of platform breadth on willingness to donate data is highly context-dependent and may be shaped by digital literacy, platform familiarity, and recruitment norms. As such, cross-country differences should not be interpreted as reflecting stable preferences toward platform scope per se, but rather as the interaction of donation requests with distinct participation contexts and respondent capabilities.

In line with earlier studies (210; 234) longer download and upload time was strongly associated with lower willingness in the U.S. sample, but not in the Hungarian one. Future research could also consider investigating how access to technology (such as PCs or laptops) influences willingness to share digital data.

Earlier research suggested that participation can be a function of the requested data type, and especially depends on the sensitivity of the data type. Our results are not clearly in line with these results. The type of data had no effect on the responses in the U.S. sample, while in the Hungarian sample, respondents were somewhat more likely to share their data when more data types (including sensitive data) were asked from them. A possible explanation of this result is in the phrasing of the situation: when respondents saw the vignettes where the excluded data types were explicitly listed after each other, it decreased their willingness to share data compared to the condition that asked for the most data, so did not specify, and list the types of excluded data. This suggests that detailed information about the different types of data included in the DDPs may decrease willingness.

Lastly, some respondent characteristics influenced willingness to donate in a significant way. For instance, older and highly educated respondents were less likely to share their data in the Hungarian sample. The results reinforced that privacy and security concerns lower the willingness to participate (142; 202; 211; 218; 236), at least among U.S. respondents. Consistently in the two samples, participants with multiple platform usage were more likely to donate their data, but for instance, the affinity for technology or personality traits (e.g., openness) of respondents did not influence willingness significantly.Nevertheless, to the extent that these self-reports overlap with actual sharing behavior, our findings altogether suggest that data donation studies should expect strong and systematic selection bias.

The differences in the observed mechanisms found between the two countries may have several reasons. The relatively large cultural differences between American and Hungarian society might cause varying levels of trust in technology companies, cultural norms surrounding data sharing, and awareness and understanding of the benefits and implications of data donation. Moreover, even if these differences are small, the two online panels were somewhat different in their nature. Cross-national differences observed in this study should be interpreted with caution, as they are inseparable from differences in sampling strategies. The Hungarian sample was drawn from a paid access panel in which respondents routinely receive incentives for participation, whereas the U.S. sample consisted of volunteers recruited without monetary compensation. These recruitment contexts likely shape expectations about compensation, norms of participation, and trust in research requests. For example, the difference in the effect size of monetary incentive might be a consequence of the financial situations and attitudes of the respondents in each country. Respondents recruited through paid access panels—such as the Hungarian sample in this study—may be more strongly motivated by monetary incentives, as this group has an already demonstrated sensitivity to monetary compensation in exchange for providing their data. As individuals in the U.S. sample did not receive monetary compensation for their study participation, they necessarily demonstrated that they have other reasons for why they are donating their data. The positive effect of a study report, and the negative effect of the time to download/upload data may be examples of such other motivators. Neither of these effects were significant in the Hungarian sample. Therefore, comparisons between the countries should be done cautiously, together with considering differences in the samples too. To enhance our comprehension of cross-country differences, further research is needed. However, our results indicate that extrapolating findings obtained from one country to another may be limited when investigating participation in data donation.

Choosing study dimensions is always a crucial point in a similar study. An essential purpose of our data collection was to select a framework for the data donation research that was later (and has since been) launched. We adapted the dimensions primarily to this goal. In our project, it was a given that we were thinking in an academic research framework, so we did not vary, for example, the organization requesting the data or the purpose for which the research was being done. However, other factors could be also crucial for such a study. Research has also shown that when participants are given more control over the content of the data they share, they are more willing to share their data (141). Regarding platforms, we treated the different Google products as one, but we are unsure if they all meant similar data sets. YouTube, for example, is shared more willingly than other Google data (141). Before a real data-donation study, researchers must consider all the possible aspects of such a project and develop their own design that fits their research goals and the target sample.

Our study has multiple practical implications as well. First, the use of monetary incentives for social media data donation requests is advised given the perceived high cost of the study from the perspective of the respondents. We did not find that a too high incentive would backfire (138), although our level of incentives was moderate[7]. Future research could explore the effect of higher incentives to identify the point at which incentives begin to have a counterproductive impact. Relatedly, the strong effect of monetary incentives—especially in the Hungarian sample—also raises ethical considerations related to undue inducement. When incentives represent a meaningful financial benefit, they may disproportionately motivate individuals with fewer economic resources to consent to sharing highly sensitive data. Although participation in the present study was hypothetical and clearly framed as voluntary, real-world data donation designs must carefully balance adequate compensation with the risk of exerting undue influence. This concern underscores the importance of transparency, clear communication about risks, and robust data protection safeguards. Researchers should also be attentive to the demographic composition of donor samples and assess whether incentive structures systematically bias participation toward particular socioeconomic groups.

Second, non-monetary incentives such as personal reports on the results may also be worth considering, although they may not boost willingness in every context or culture.

Third, the burden, and particularly the difficulty of sharing data should be kept as low as possible. DDP requests are not routine tasks for most of the platform users. (141) reported that only 7.75% of their study participants previously undertook such a request. High download and upload times can deter people from participating. Choosing easy-to-do tasks and providing helping materials can help researchers to reduce participants' burden. To better understand our inconsistent findings about the number of platforms included in the request, further research is needed.

Fourth, although our results do not suggest strong evidence against asking for specific data types, data donation requests should be carefully designed in this regard. Several earlier studies suggested that the sensitivity of the data can lower the willingness to donate (e.g., (141)).

Fifth, reinforcing earlier findings, privacy concerns were strongly associated with willingness in our U.S. sample, therefore, addressing privacy concerns is key in social media data donation projects. Future studies should be conducted to explore how privacy concerns can be mitigated, for instance, by using different framing, providing more information about data

[7] Our models were also tested by including the variable measuring the incentive as a factor in the model rather than continuously. The explanatory power of the models was not higher in the alternative runs, and the B values of the incentive variable indicated that the effect of the variable of interest was linear.

protection, etc. In this study, we did not have enough room to capture the complexity of privacy concerns. We advise future research to address this gap, for instance by designing similar experiments that manipulate different privacy aspects. Data obtained from these experiments may aid in crafting more targeted mitigation strategies.

Lastly, our study highlighted that nonresponse is expected to be high in such data donation requests, and as we can assume that this nonresponse is not random, this likely translates to strong selection bias. Future data donation studies need to develop strategies to handle different types of nonresponses.

One limitation of this study is the use of non-probability-based samples, which limits the generalizability of our findings. Generally, non-panel member internet users are expected to be less likely to participate in data donation studies (138), although the extent to which the underlying mechanisms differ between panel- and non-panel members is unclear. Nevertheless, an advantage of using these panels can be that data donation research tends to be based on similar platforms and sampling frames. This population, compared to a random population sample, is more open to a data donation study and, due to its higher digital capabilities, is more likely to be able to retrieve and deliver its data. While the used samples may not completely reflect the overall populations of these two countries, they are likely to closely represent the types of people that researchers commonly utilize in studies involving DDP. Another potential limitation is that we relied on self-reports and do not know to what extent willingness transfers to actual donation. A possible direction for further research is to develop similar experimental designs in which self-reports and real participation can be contrasted (see e.g., (202; 237)). Also, while our data collection could ”only” examine a hypothetical situation and not a ”real” situation, it can still provide important information on how to design ”real” research. What platforms to ask for and what data to ask for within those platforms is essential when designing a survey, as is the ”minimum” amount of money to ask for. In the context of willingness to participate, it is clear that the textual framing of the research is necessary because participants do not have a clear understanding of what data is available on these platforms. There are platforms such as Instagram and TikTok, where users have no option to select which specific data they want to download, so users with lower digital literacy may unknowingly provide data for research purposes they would not otherwise want to. Thus, framing research on these platforms significantly impacts the actual donation. Although the detailed description of the data requested will reduce the willingness to participate, ethical considerations should override data collection efficiency, and it is vital to be as clear and precise as possible in telling participants what data we ask them for and what this data will be used for. Data donation research must be on very firm ground from an ethical and data protection point of view. In different countries, especially continents, there can be significant differences in the data protection and ethical principles that should be applied to data donation research. In recent years, several European infrastructure projects (datadonation.eu, datadonation.uzh.ch/en/infrastructure) have been launched to support data donation projects at the platform level (139). These platforms support data pre-processing, selection, and aggregation before data transfer, making such research more transparent, ethically clear, and privacy-friendly. These platforms can increase trust in data donors and support higher participation rates and lower sample biases.

When evaluating the results, it is also worth considering that we asked people's opinions on a complex task that they may not have been able to see through the situations they were given. In general, it is thought that people have very little information about how such a process works and precisely what data is shared. As a clue to the latter, they may have been helped by the fact that we listed data types for them in the vignettes and even manipulated the latter on the cards. As for the Hungarian sample, we know that no panel members have participated in a similar study, but we assume that the U.S. sample would not have included people who would have provided DDP data similarly. The problem has two sides: one is about data protection, and the other is the technical part of data sharing. We controlled for these with independent variables, but it is certainly possible that people underestimated the difficulty of sharing and the sensitivity of the data. This may mean that privacy sensitivity and online technical skills may play a more significant role in a real data-sharing situation.

Finally, another important limitation of this study is that it measures stated willingness to donate data in hypothetical scenarios rather than actual donation behaviour. A well-known challenge in studying data donation is the gap between stated intentions and realized behaviour (235; 238; 239), particularly in contexts where privacy preferences are moralized. Recent work shows that this gap can arise for multiple, partly opposing reasons. (235), comparing the vignette-based willingness to share digital trace data studied in this chapter with actual donation behaviour measured nine months later, find that hypothetical willingness is only moderately correlated with realized participation. Importantly, the determinants of willingness and behaviour only partially overlap: willingness is primarily shaped by attitudinal and normative considerations, whereas actual data donation depends more strongly on capacity-related factors such as age and education, reflecting the technical demands of the donation task. As a result, realized behaviour may fall short of stated willingness because individuals who are normatively willing lack the practical capacity to follow through. These findings caution against interpreting vignette-based willingness measures as direct proxies for actual data donation behaviour, instead highlighting their role in capturing stated preferences under hypothetical conditions. At the same time, evidence from (239) suggests that stated unwillingness may itself be downward biased. In their meta-analysis drawing on a sample of studies conducted in a variety of countries they show that respondents may under-report their willingness to donate digital trace data due to moralized privacy concerns, even though these concerns do not always translate into behaviour when concrete situational factors are present. In real-life contexts, incentives or other practical considerations—such as monetary compensation in the present study—may outweigh privacy concerns, leading individuals who initially report low willingness to donate to nonetheless do so in practice. Taken together, these findings indicate that willingness to donate data should not be interpreted as a direct behavioural forecast. Rather, willingness primarily reflects normative orientations and abstract evaluations of privacy trade-offs, whereas actual behaviour emerges from the interaction of these orientations with practical constraints, incentives, and situational factors that are not fully captured in vignette designs. Understanding willingness is therefore substantively important in its own right, as it helps distinguish between reluctance and constraint and sheds light on how individuals reason about data sharing under hypothetical conditions.

Collecting individual-level social media data through DDPs with informed consent and linking this data with survey data is a promising area of research (139; 240). Our study contributes to the understanding of the circumstances under which individuals are more likely to share their data, and to the assessment of the self-selection bias in data donation studies.

Chapter 4

Study 3: Socio-demographic and Online Behavioural Drivers of Misinformation Engagement in Hungary: Evidence from Linked Survey and Social Media Data

This chapter is based on the following manuscript (submitted for review): Júlia Számely, Júlia Koltai & Elisa Omodei (2025) Socio-demographic and Online Behavioural Drivers of Misinformation Engagement in Hungary: Evidence from Linked Survey and Social Media Data.

4.1 Introduction

Having shown that with the right incentive it is possible to obtain high resolution donated digital trace data in Hungary, this final empirical chapter presents an in-depth case study of Hungary, a context where a highly centralised media system and growing political polarisation make misinformation both socially significant and theoretically important. As discussed in the Introduction, Hungary represents an under-studied but valuable setting for understanding how structural features of the information environment interact with individual-level characteristics to shape engagement with misinformation.

The study combines survey data with donated Facebook traces from a nationally representative sample of Hungarian internet users. This approach enables a detailed examination of how socio-demographic factors, online activity patterns, and participation in online communities relate to interactions with misinformation sources. Unlike Study 1, which demonstrated what can be inferred from publicly accessible but low-granularity behavioural data, the present study relies on high-resolution, participant-donated data to explore these dynamics in greater depth. It also complements Study 1 by integrating higher-resolution online behavioural data with detailed personal and socio-demographic information obtained from the linked survey data—dimensions that were not available in Study 1. In this sense, it complements the earlier large-scale analyses by demonstrating the analytical value of richer behavioural data, even within today's constrained research environment.

4.2 Motivation

While misinformation research has expanded rapidly, important gaps remain in our understanding of how both socio-demographic characteristics and online behavioural patterns jointly shape individual engagement with misleading content. Despite increasing attention to the psychological and political antecedents of misinformation engagement (39; 241), comparatively less is known about the role of socio-demographic characteristics (172) and online behavioural patterns (242) in shaping individuals' engagement with misinformation.

Additionally, much of this research has focused on the United States and a few Western European countries, and less attention has been paid to Central and Eastern Europe, where some countries—such as Hungary—have experienced notable democratic backsliding (243) and an increase in the use of propaganda (244) in recent years. In Hungary, government communication campaigns and state-controlled media play a central role in shaping the information environment, making it a critical case for studying how citizens engage with misinformation and how democratic institutions are eroded (244). Examining these settings in more detail may contribute to a fuller understanding of the diversity of factors associated with misinformation engagement.

Approaching misinformation from multiple angles—both personal characteristics (such as socio-demographic) and online behavioural patterns—offers a compelling basis for unpacking the dynamics of misinformation sharing. Research in social psychology suggests that thinking and behavioural patterns are highly contingent on socioeconomic status (245), and more broadly, that personal background shapes how individuals engage with and interpret information.

In this chapter, we set out to explore what socio-demographic characteristics and digital behavioural patterns are associated with misinformation engagement and redistribution in Hungary today. To this end, we leverage a novel data donation framework to examine how a combination of socio-demographic characteristics and digital behaviour relate to misinformation engagement and sharing in Hungary. By combining these perspectives, we aim to provide a richer understanding of what drives misinformation engagement in Hungary, a setting that offers a particularly relevant context for studying the interplay between socio-demographic characteristics, online behaviour, and exposure to misleading content.

By integrating observed and self-reported data, this chapter offers a more nuanced account of the behavioural and social mechanisms underlying misinformation engagement in a politically centralised information environment. It shows how linking survey responses with ethically obtained digital trace data can enhance the understanding of engagement with misinformation, yielding findings that are both empirically grounded and sensitive to local context. In doing so, the chapter extends the empirical component of the thesis, demonstrating how analyses across varying levels of data richness can collectively advance a more comprehensive understanding of misinformation engagement. It is important to note that the sheer volume of misinformation encountered in everyday online environments is often relatively small compared to legitimate information. However, low prevalence does not necessarily imply low impact. Even limited engagement with false or misleading content can have disproportionate effects on beliefs, attitudes, and evaluations of evidence, particularly in politically polarized contexts (2). Accordingly, the present study focuses not on the volume of misinformation per se, but on individual differences in engagement with such content, which may signal broader vulnerabilities to online manipulation.

4.3 Related Work

In recent years, there has been a considerable body of research dedicated to exploring the potential sources of the increasing prevalence of misinformation worldwide. While the literature covers a range of topics—such as health-related misinformation (246) or science-related misinformation (247)—and geographical regions—such as France, Germany, Spain, Italy, and the United Kingdom (248; 249; 250) or Saudi Arabia (251)—most research on misinformation so far has been concentrated on politics-related fake news in the United States (172). At the same time, recent comparative studies show that misinformation exposure and vulnerability vary substantially across countries, suggesting the importance of cross-national analyses beyond the U.S. context (250). In a study on Central and Eastern Europe, for instance, (149) demonstrate in the Hungarian case that analytic thinking plays an important role in predicting susceptibility to political misinformation, highlighting how cognitive style interacts with specific political and media environments.

Most research studying the human aspect of misinformation engagement and spreading has focused on cognitive factors underlying individual level misinformation engagement, such as confirmation bias or motivated reasoning (246; 252; 253). Some studies have included demographic or socioeconomic factors as well, and yet a smaller subset has focused explicitly on demographic and socioeconomic status variables as determinants (248; 254). Having identified this research gap, Tucker et al. (172)—among others ((242))—call for further research efforts on who shares misinformation, as gaining insight into this is a necessary step towards better understanding, and potentially finding solutions to the problem.

Arin and coauthors (248) address this in their study of misconceptions and fake news during the COVID-19 pandemic in the United States using survey data. The survey asks respondents about their experience and past behaviour regarding encountering and sharing fake news and relates these answers to a host of socio-demographic characteristics. The study finds that individuals with high or medium income levels encounter more fake news but are less likely to share it compared to individuals with low income. The authors compare their results to previous literature concerned with the socio-demographic determinants of fake news consumption and sharing. They find that while their results show many similarities, there also appear to be a number of inconsistencies regarding the identified demographic groups that are the most vulnerable to fake news. For example, using social media data, Grinberg et al. (114) and Guess et al. (39) find that engagement with fake news is more likely among conservative-leaning, older, and politically engaged subpopulations; based on a survey experiment Buchanan (40) finds that younger, less educated males are the most likely to spread fake news; finally, based on survey data, Barthel et al. (255) find that fake news induces confusion across all income and education levels as well as across most other demographic characteristics and both Democratic and Republican party affiliations. More recent survey data based work shows that these effects are often contingent and inconsistent across settings, with some studies finding strong demographic predictors of misinformation engagement, while others do not (254; 256).

Most of the studies discussed above rely on self-reported survey measures of misinformation encounter and sharing. While these approaches provide useful insights into individuals' perceptions and reported behaviours, they are inherently subject to limitations such as recall error (106) or social desirability bias (107). This phenomenon underscores the value of approaches, including the present study, that complement survey data with observed digital behaviour.

These studies and comparisons point to two important facts. First, individual socio-demographic characteristics do seem to play a significant role in the engagement with and spreading of misinformation. Second, consensus has not yet been reached regarding which groups are the most vulnerable to misinformation.

While the need for further exploration of socio-demographic characteristics as determinants of misinformation engagement and sharing has been recognised in the literature, the reason for the scarcity of relevant work appears to be the lack of appropriate data. Indeed, Tucker and co-authors (172) identify a key data gap: survey data paired with social media data of the same survey respondents would be needed for a more nuanced and accurate characterisation of the identities of those susceptible to misinformation.

While survey-based studies have provided insights into individual-level predictors of misinformation engagement, they often overlook the social and informational contexts between users and their online environments. Recent research indicates that the online communities individuals participate in can influence their engagement with misinformation. For instance, studies have shown that the structure and dynamics of online networks, such as echo chambers and filter bubbles, can amplify exposure to misinformation and reinforce existing beliefs (119). Furthermore, within these communities, individuals may face social costs if they refrain from sharing certain content, which can pressure individuals into disseminating misinformation (120). Integrating behavioural data with survey responses can help researchers examine how membership in specific online communities correlates with patterns of misinformation engagement, providing a more nuanced understanding of these dynamics.

In their review article (172) also emphasise the need for the increased use of Facebook data, arguing that Facebook's relative dominance in terms of number of users over other social media—notably Twitter/X, a data source providing the basis for an outstanding number of studies—indicates that Facebook might provide better insight into issues affecting the general population.

The work of Guess et al. (39) fits precisely into these data and research gap by studying the prevalence and predictors of fake news dissemination in the context of the 2016 US elections. In their study, the authors link a representative online survey to behavioural data on respondents' Facebook sharing history during the 2016 US election campaign. They measure fake news consumption by counting the number of times respondents shared a Facebook post containing links to a list of external websites identified as fake news publishers by academics and journalists. The authors test the robustness of results by using several alternative lists of fake news publishers. They find that apart from age—over-65 users shared more fake news—no other demographic characteristics had a significant effect on sharing fake news. Further investigation with similar data but in different contexts might have the potential to reveal novel insights about engagement with misinformation consumption and sharing.

4.4 Data and Methods

4.4.1 Data Collection and Sample

This study draws on a unique dataset combining detailed survey responses and behavioural digital trace data collected via a structured data donation process. Data were collected in Hungary between February and June 2023. Participants were recruited from online panels managed by a digital market research company, and invited via email to take part in the study. The procedure involved (1) completing an eligibility screening and providing informed consent, (2) downloading their data archives from major platforms (Facebook, Google, and in some cases, Instagram, TikTok, Twitter) following step-by-step instructions, and (3) securely uploading the unaltered archives to the project server, where files were validated through automated scripts. Respondents then completed a final survey. Although multiple platforms were included, this study focuses on Facebook data, as it is the most widely used social media platform in Hungary (257), and is equally popular among all gender and age groups (258). Moreover, Facebook was the only platform for which complete data were available from all respondents. The final analytic sample consists of 758 individuals who successfully provided valid Facebook data and completed all survey components. Smaller deviations of the sample from the 16+ years old Hungarian internet user population were adjusted with iterative proportional fitting. The final sample is representative of the above-mentioned population by gender, age, education, settlement type, and geographic regions. The procedure was fully complying with legal regulations of Hungary and the European Union and was approved by the Ethical Board of the HUN-REN Centre for Social Sciences (resolution number 1-FOIG/130-37/2022). For more information on the data collection, see (259).

4.4.2 Survey Measures as Explanatory Factors

The survey captured a wide range of personal characteristics, including socio-demographic characteristics (e.g., age, gender, education, subjective and objective indicators of income), political attitudes (e.g., political interest), psychological traits, and self-reported digital behaviour. These variables serve as explanatory factors in our analyses (see the section on Models). Some of these characteristics are directly included in our final models, while for sets of highly correlated variables, we use their principal components to reduce dimensionality and mitigate multicollinearity in the final regression models.

4.4.3 Behavioural Data as Explanatory Factors

Participants' donated Facebook archives provided detailed records of their activity on the platform, including for example pages followed, posts created, and comments or reactions left on others' posts. We structured and analysed this behavioural data to quantify individual engagement patterns, including activity related to misinformation, which we describe in detail in the next section.

From this, we extract two categories of behavioural characteristics to serve as explanatory variables in our models. First, we measure general Facebook activity through metrics such as the number of posts shared, comments written, and likes or love reactions given. Second, we examine the types of content users engage with by analysing their interactions with Facebook pages—such as following, reacting to, or commenting on page content, detailed in the following subsections.

To focus on contemporary patterns of platform engagement and reduce potential biases introduced by historical changes in platform dynamics, we restrict the dataset to platform activities occurring after January 2019 only.

Facebook Activity and Engagement Types

To capture participants' behavioural activity on Facebook, we constructed two explanatory variables.

We combined participants' original posts and comments into a single measure of activity. Specifically, we counted, for each participant, the total number of posts they created and the total number of comments they wrote, and then summed these values. We refer to this measure as *Number of Posts + Comments*, which was treated as a continuous variable. This provides a baseline measure of how much content each participant contributed to the platform during the study period.

Participants' Facebook archives also contain records of their broader engagement with the platform, including interactions such as pages followed, reactions left on posts, comments made on page or group content, and group memberships. For each participant, we aggregated the total number of such interactions across different engagement types (e.g., total page followings, total page reactions, total group memberships). To reduce these multiple correlated measures into a smaller set of summary dimensions, we applied principal component analysis (PCA). When including all interaction variables–pages liked, page reactions, page comments, groups liked, group reactions, group comments–the first principal component explained 39.1% of the total variance. The communalities of these variables with respect to the first principal component indicate that page-related activities (e.g., pages liked, page reactions, page comments) are relatively well represented (0.54–0.62), whereas group-related activities moderately captured (0.16–0.26). To address the low communalities, we iteratively removed variables with the weakest scores (those with communalities below 0.25) and reran the PCA. The final analysis retained pages liked (communality = 0.72), page reactions (0.51), page comments (0.64), and groups liked (0.30). This revised model explained 54.21% of the total variance and captured the dominant axis of variation in users' activity counts, serving as an overall index of engagement intensity across different parts of the platform. The first principal component was then included in our models as *Facebook Interactions (PC1)*.

These two measures—Number of Posts + Comments and Facebook Interactions (PC1)—were included as control variables in the regression models described under *Models*. Together, these variables allow us to distinguish between basic content production (total posts and comments) and a broader latent measure of engagement across the diverse interaction features available on Facebook.

Finally, because our dataset covers Facebook interactions over the past five years, but some respondents joined the platform more recently, we included a control variable–*Months since Registration*–to account for differences in users' opportunity to engage.

Content Engagement Analysis on the Facebook Page Network

Facebook pages represent a useful unit of analysis for examining patterns of online content engagement. Pages are maintained by organizations, media outlets, public figures, and other actors that share information with broad audiences. Interactions with such pages—whether by following, reacting, or commenting—leave observable traces of how users connect with different types of content. Studying these traces provides a structured way to describe the kinds of information environments users encounter and engage with on the platform.

Utilizing this, we analyse patterns of content engagement by quantifying the similarity of page audiences. Specifically, for each page i, we define a user vector $\mathbf{u}_i$ of length U, where

U denotes the total number of unique users who interact with pages in the dataset. The n-th element of $\mathbf{u}_i$ corresponds to the number of times user n has engaged with page i. For any pair of pages i and j, we then compute the cosine similarity between $\mathbf{u}_i$ and $\mathbf{u}_j$ and use this cosine similarity to create a network of pages, where an edge between two pages is weighted with the value of their cosine similarity. This approach captures not only the extent of audience overlap (i.e., the number of shared users) but also the relative intensity of their engagement with each page. In doing so, it preserves information that is typically lost when projecting page–user interactions into a lower-dimensional representation, thereby offering a more nuanced view of how audiences cluster around different content sources. To ensure that edges reflect substantive overlaps rather than incidental ties, we apply a cosine similarity threshold of 0.5, yielding a weighted, undirected page–page network. The procedure for selecting this threshold is detailed in the Supplementary Information (Figure C.1).

We then apply community detection to the page network to identify clusters of pages that are co-engaged by similar sets of users, using the Louvain algorithm (260). Since the Louvain algorithm can yield slightly different partitions across runs, we assess its stability by repeating the algorithm 100 times and computing normalized mutual information (NMI) (261) across results. To mitigate instability, we implement a consensus clustering approach (262) that aggregates information across runs via a co-association matrix, yielding a final, robust set of communities.

To incorporate community information as respondent-level characteristics, we assigned each page to its detected community and then merged this information with the respondent–page interaction data. In other words, we translated the page-level communities into user-level features by counting how much each individual engaged with them. For each respondent, we summed the number of interactions with pages belonging to each community, producing a set of community-level engagement scores. These scores quantify the extent to which each individual engaged with different communities of pages, and were subsequently used as explanatory variables in the models.

These page communities reveal the latent structure of the Facebook information environment in which Hungarian internet users are embedded. They serve as proxies for interest clusters and are subsequently included as features in our models to capture the role of content engagement in shaping misinformation engagement patterns.

4.4.4 Constructing an Individual-Level Misinformation Engagement Measure

For the dependent variables of our models, our objective is to capture individual-level engagement with misinformation from digital trace data. Accordingly, we focus on observable behaviours that may indirectly indicate misinformation engagement. These behaviours differ in the degree to which they imply endorsement, attention, or exposure. We therefore construct a measure that combines multiple types of engagement, recognising that misinformation engagement can manifest both through deliberate dissemination and through more routine, everyday interaction with low-credibility content (116).

We build on two complementary components. The first captures direct sharing of misinformation, by identifying posts and comments authored by respondents that contain hyperlinks to domains previously classified as misinformation outlets (39). Sharing behaviour typically requires conscious effort and carries reputational costs, which makes it a strong signal of engagement (62). While some shares may be intended as critique, satire, or humour, our manual review of posts and comments confirmed that explicitly disapproving shares were negligible in our data. Only 41.5% of posts (34 in total) containing a link to a misinformation website included any additional text, none of which were disapproving. For comments, 42.5% (17 in total) contained extra text, of which only two adopted a sarcastic tone, suggesting likely disapproval of the shared content. We therefore treat sharing as a meaningful indicator of active engagement with misinformation.

The second component extends beyond sharing to capture more subtle forms of interaction with misinformation sources. We identify cases where respondents followed, reacted to, or commented on Facebook pages and groups associated with misinformation domains. We identified these Facebook pages and groups by matching their names to known misinformation websites and manually verified them, following the procedure used by (263). We found 127 pages in the Facebook dataset whose names corresponded to an identified misinformation website. Recognizing that these pages also contribute to the construction of the interest cluster explanatory variables, we carefully verified that the overlap between pages used in the explanatory and outcome variables was limited. Specifically, the proportion of misinformation pages and groups in the dataset was 0.08%, indicating a negligible overlap. Facebook interactions demand less effort than creating a post but can nonetheless reflect a sustained interest or openness to content from these sources. Following a page, for example, increases the likelihood of future exposure, while liking or reacting to content (including “like” and “love”, but excluding “haha”, “surprise”, and “angry” to avoid ambiguity) can indicate agreement or approval. These interactions thus represent weaker but still relevant signals of misinformation engagement (263).

By combining high-effort behaviours such as sharing with lower-effort interactions such as liking or following, our approach captures a spectrum of misinformation engagement behaviours. We therefore treat these two components—direct sharing and indirect interaction—as the raw behavioural building blocks of our misinformation engagement indices, described in detail in the next subsection.

In both components we rely on a domain-level classification of misinformation. Following established practice (39; 113; 264; 265; 266; 267), we assign the “misinformation” label at the publisher level rather than at the level of individual articles. A domain-based approach allows for consistent classification across content and is practical for large datasets where article-level fact-checking is rarely feasible.

Our reference list of misinformation websites was assembled by aggregating several Hungarian-language fact-checking outlets, journalist-curated sources, and relevant Wikipedia entries. We checked that all entries appeared in at least two independent misinformation lists to ensure broad coverage and to minimize reliance on potentially biased sources (see Table 4.1).

Table 4.1: Sources Used to Compile Misinformation Website List

Source	**Description**
HVG	List of fake news sites curated by journalists
Urbanlegends.hu	Hungarian fact-checking site
Neten a Videóm	Journalist-compiled guide to deceptive sites
Kelecsényi.info	Resource by a Hungarian computer scientist
Wikipedia	Community-curated list of unreliable sources

Outcome Variable Construction: Misinformation Engagement Indices

Based on the components described above, we constructed four indices to capture different assumptions about the relative importance of various engagement behaviours. Each index aggregates five types of interactions in a different way: user-authored posts containing links to known misinformation websites; user-authored comments containing such links; followings of Facebook pages and groups identified as misinformation sources; reactions (including "like" and "love", but excluding "haha", "surprise", and "angry" to avoid ambiguity) to posts by these pages and groups; and comments on such posts.

We assigned multipliers to these components to reflect their likely indicative strength, guided by the effort and cost required and the degree of intentionality they appear to involve. For example, creating a post that shares a misinformation link requires more effort than liking a page, and thus was given the highest multiplier. Comments containing links to misinformation sources were generally assigned relatively high multipliers, although in some specifications these multipliers were reduced to account for the possibility that such sharing may not always reflect intentional endorsement. Follows were assigned moderate multipliers, as they can indicate interest or repeated exposure but involve less effort. Reactions were given the lowest multiplier, since they represent quick, low-effort signals of engagement. Finally, one specification excluded comments on misinformation content altogether, in order to test the sensitivity of results to the possibility that some comments were critical rather than endorsing.

The constructed misinformation engagement indices are not intended to capture precise quantitative magnitudes of different types of interaction. Rather, they are designed to reflect ordinal differences between qualitatively distinct actions.

The assumed ordering of engagement types is informed by prior research. Posts and comments containing links to misinformation sources are widely treated as the strongest observable indicators of misinformation engagement. For example, (62) highlight the high reputational costs associated with sharing misinformation online; (39) treat the sharing of links as the core behavioural indicator of misinformation engagement; and (118) focus on misinformation sharing cascades rather than reactions or comments, underscoring both the centrality and the greater downstream consequences of sharing. Reactions and comments, while still meaningful indicators of interaction with misinformation content, are generally understood to be lower-cost actions and less consequential for information diffusion.

To our knowledge, no established weighting scheme exists that jointly incorporates posts, comments, and reactions into a unified behavioural index of misinformation engagement. Accordingly, the multipliers used here should be interpreted as pragmatic modelling choices grounded in converging evidence from the literature rather than as empirically validated parameters. The index is therefore best understood as a behaviourally meaningful composite of misinformation engagement, even in the absence of an established weighting convention.

As different choices of multipliers reflect certain assumptions, we constructed four indices to ensure that our results are not driven by any single specification and to enable robustness checks and sensitivity analyses. The "main" index serves as our primary outcome measure and is reported in the main text. Results based on the alternative indices—emphasising higher intent, greater exposure, or excluding potentially sceptical engagement—are presented in the Supplementary Information.

Each index was computed by multiplying each component by its assigned multiplier and summing the results. For example, the main index was calculated as:

$$
\begin{aligned}
\texttt{misinfo\_count\_main} = {} & 1.0 \times \text{misinfo\_posts} + 1.0 \times \text{misinfo\_comments} + \\
& 0.6 \times \left(\text{misinfo\_pages\_liked\_count} + \text{misinfo\_groups\_liked\_count}\right) + \\
& 0.3 \times \left(\text{misinfo\_page\_reactions\_count} + \text{misinfo\_group\_reactions\_count}\right) + \\
& 0.7 \times \left(\text{misinfo\_page\_comments\_count} + \text{misinfo\_group\_comments\_count}\right).
\end{aligned}
$$

Analogous formulas were applied to construct the other indices using the relevant multipliers shown in Table C.1 in the Supplementary Information.

4.4.5 Models

Given the large number of zeros in the dependent variable (see Figure C.2a), to examine the predictors of misinformation engagement, we estimated hurdle models (268) that combine (1) logistic regressions predicting whether individuals engaged with any misinformation content ("zero models") and (2) Gamma regressions with a log link predicting the volume of engagement among those who engaged with misinformation ("positive models"). These models allow us to separately estimate, first, the likelihood of engaging with misinformation at all, and second, conditional on engagement, the effect of the explanatory factors on the extent of that engagement.

For each stage, we estimated two nested model specifications. Model 1 included socio-demographic characteristics from the survey. Model 2 incorporated measures of individuals' online behaviour with interests clusters derived from individuals' Facebook page engagements as well as indicators of their general Facebook activity.

Model fit was evaluated separately for the zero and positive components. For both components, we report adjusted McFadden's pseudo-R^2 as the indicator of explanatory power relative to a null model (269).

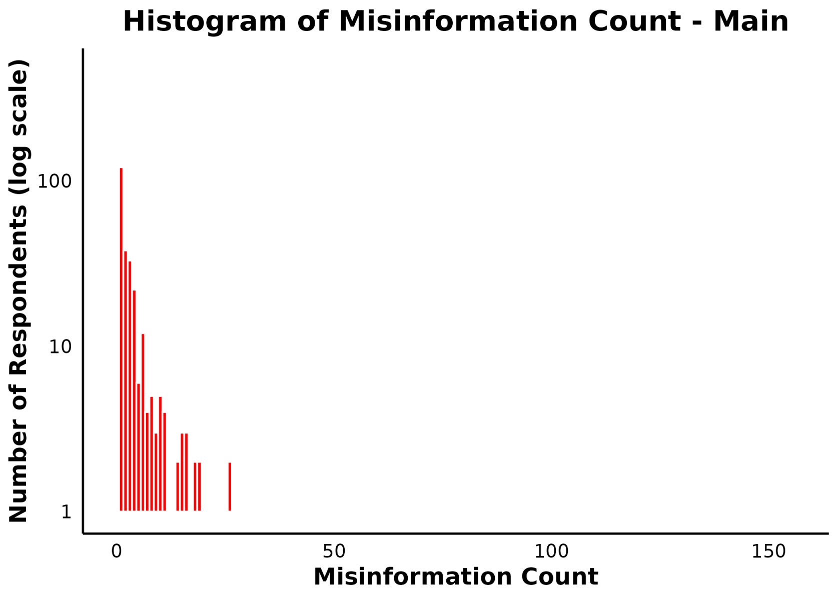


Figure 4.1: Distribution of the dependent variable measuring misinformation engagement. A large share of observations take the value zero (no engagement), while the remainder form a long right-skewed tail.

4.5 Results

4.5.1 Identified Interest Clusters in the Facebook Page Network

Applying the community detection procedure described above, we identified 46 distinct clusters of Facebook pages. These clusters represent coherent spheres of interest that vary in both size and thematic focus. Their size distribution is skewed, such that a few large clusters account for most pages, while many smaller clusters capture more specialized domains (see Figure C.3 in the Supplementary Information).

Taken together, the communities reflect a diverse information landscape, spanning both everyday lifestyle and politically charged content. Some clusters are explicitly connected to partisan or civic engagement—such as pro-government politics, opposition-aligned outlets, or civic NGOs—whereas others center on non-political domains, including food, retail, culture, or health and wellness. This underscores how online engagement is structured by clusters that encompass a diverse range of topics.

Figure 4.2 shows a network visualisation of the Facebook pages, with nodes coloured according to their assigned cluster, providing a visual overview of the cluster structure.

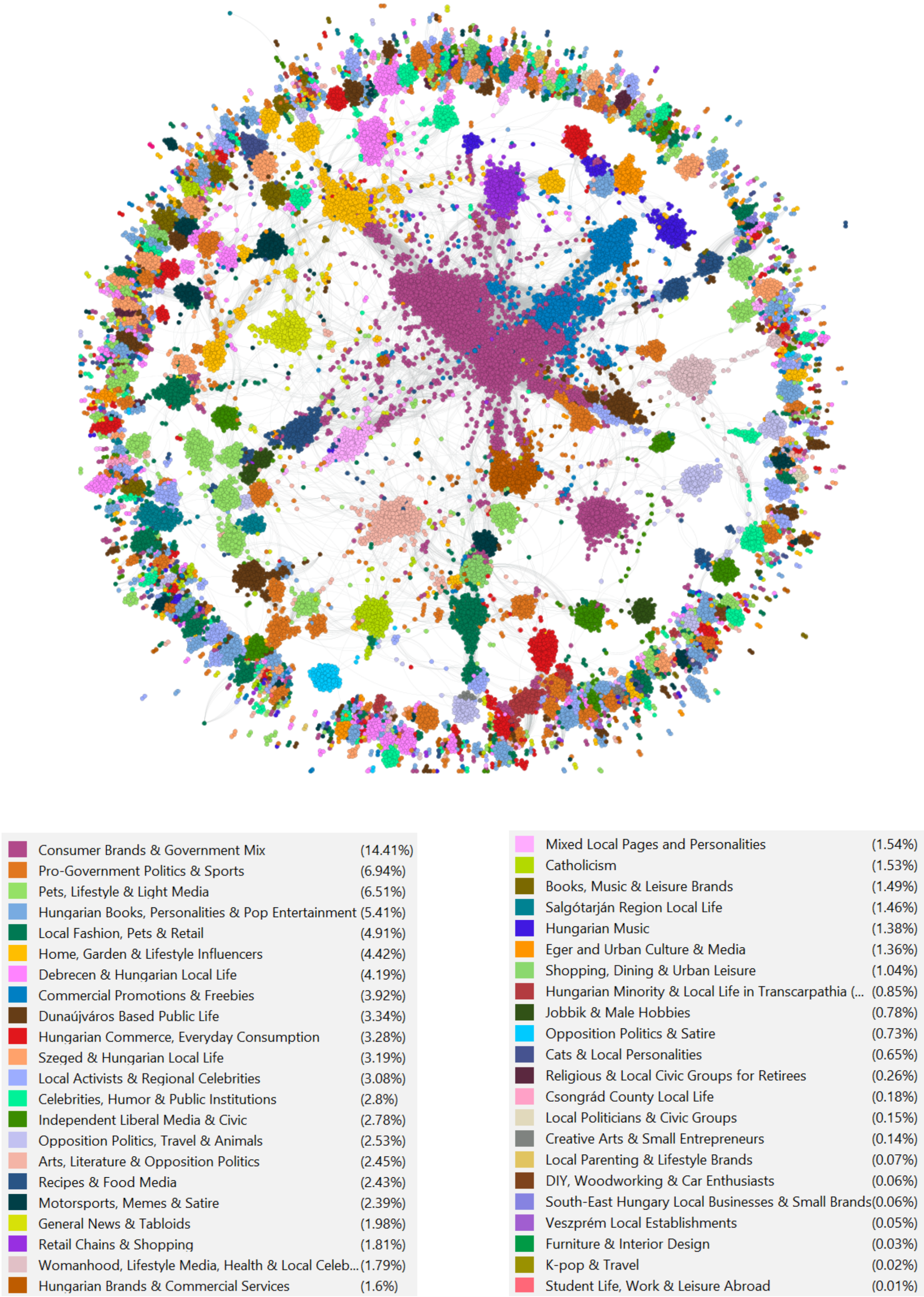


Figure 4.2: Facebook page network, with nodes coloured by interest cluster. The legend below displays the names and corresponding colours of each interest cluster, with the numbers in parentheses indicating their relative sizes.

Below, we describe the interest clusters significantly associated with our outcome variables.

Pro-Government Politics & Sports combines government-supportive political pages with sports content, including Orbán Viktor, Novák Katalin, and M4 Sport. The audience is politically engaged and supportive of pro-government narratives while actively following sports content.

Hungarian Books, Personalities & Pop Entertainment focuses on Hungarian culture, literature, celebrity news, and pop entertainment, with pages such as Hadházy Ákos, Sláger FM (radio), and Barátok Közt (tv show). It attracts users interested in contemporary Hungarian media, entertainment, and literary culture.

General News & Tabloids includes a variety of Hungarian news outlets, such as 24.hu, Blikk, Ripost, Portfolio, Híradó.hu, Euronews Magyarul, Metropol, and Propeller.hu. While all these pages are technically news sources, many are sensationalized, tabloid-style, or of lower quality, with some featuring propagandistic content. The cluster also contains pages focused on lifestyle, cars, weather, and local communities (e.g., Totalcar.hu, Időkép, Abony.hu, Bpiautósok.hu), reflecting a mix of political, cultural, and practical everyday interests.

Independent Liberal Media & Civic consists of independent news outlets, civic organizations, and advocacy pages, including Telex.hu, 444.hu, and Amnesty International Hungary. It features educational and cultural content, highlighting liberal, civic-minded, and socially engaged audiences.

Motorsports, Memes & Satire combines motorsport fan pages, meme communities, and satirical content, such as 9GAG, Formula.hu, and Michael Schumacher Forever Rajongók Klubja (fan club). The cluster is characterized by high engagement with entertainment, sporting, and humor-focused content.

Recipes & Food Media focuses on cooking, gastronomy, and food-related lifestyle content, with pages such as Nosalty, Cookrate.hu, and Street Kitchen. The audience is primarily interested in home cooking, food trends, and related consumer products, often sharing and engaging with recipes.

Retail Chains & Shopping centers on retail brands, supermarkets, and consumer goods, including Lidl Magyarország, Aldi Magyarország, and Media Markt Magyarország. The audience is oriented toward consumer products, discounts, and everyday shopping needs.

Books, Music & Leisure Brands covers literature, music, leisure activities, and hobbyist content, including pages such as Libri Könyvesboltok (bookshop), Halott Pénz (music band), and Vadász-Szó-Tár (hunting dictionnary). The audience is culturally oriented, with an interest in arts, literature, and hobbyist pursuits.

Hungarian Brands & Commercial Services encompasses domestic brands, service providers, and commercial products, featuring pages such as Nescafé Dolce Gusto Magyarország, Mezőkövesd Zsóry Sportegyesület (sports club), and Optic World. It reflects users interested in Hungarian products and commercial offerings.

Eger and Urban Culture & Media is centered around the city of Eger and combines local news, cultural heritage content, and urban lifestyle media, including pages such as Eger hírek (Eger news) and Gárdonyi Géza Színház (theater). It attracts users interested in local culture, history, arts, and broader urban media trends.

Jobbik & Male Hobbies combines politically conservative content linked to the Jobbik party with male-oriented hobbies such as off-roading, fishing, and tactical/outdoor pages, including Jobbik-Konzervatívok (Jobbik conservatives), Russian Extreme Offroad Trucks, and Hobbikert.hu (hobby garden). Users engage with both political and recreational interests.

DIY, Woodworking & Car Enthusiasts is characterized by do-it-yourself projects, woodworking, and automotive interests, featuring pages such as DIY & Crafts, The Dusty Lumber Co., and BMW E36 Best Car. Audiences are hands-on, hobby-focused, and highly engaged in practical creative or mechanical pursuits.

A complete overview of all clusters, including descriptive labels, summaries, and representative example pages is provided in Table C.3 in the Supplementary Information.

The association of the interest clusters with misinformation engagement is detailed below, under the discussion of model results.

4.5.2 Zero Models: Predicting Any Misinformation Engagement

Figure 4.3 presents two nested zero models predicting whether participants engaged with misinformation-related content at all. Model 1 includes socio-demographic variables from the survey, and Model 2 incorporates online behavioural variables, such as Facebook interest cluster variables, and Facebook activity measures.

In Model 1, age was positively associated with misinformation engagement, while education and income-related variables showed non-significant associations. All these effects were consistent across the two models. Being a woman was positively associated with engagement in Model 1, but this effect disappeared in Model 2, where we included interest clusters and controlled for online activity. Model fit of Model 1 was modest, with adjusted McFadden's $R^2 = 0.048$.

Model 2 added Facebook interest cluster variables and Facebook activity controls, improving model fit significantly (adjusted McFadden's $R^2 = 0.277$).

The two interest clusters associated with a higher likelihood of engagement with misinformation were *General News & Tabloids* and *Eger and Urban Culture & Media*. In contrast, several clusters were linked to a lower likelihood of engagement, including *Independent Liberal Media & Civic*, *Motorsports, Memes & Satire*, *Recipes & Food Media*, *Books, Music & Leisure Brands*, and *Jobbik & Male Hobbies*–where Jobbik is a formerly far-right opposition party in Hungary. The *DIY, Woodworking & Car Enthusiasts* cluster showed the strongest negative association with misinformation, displaying particularly pronounced effects. The total number of posts and comments and the first principal component capturing interaction intensity were positively associated with misinformation engagement, while months since registration showed a negative association. Overall, these results suggest that membership in general and local media clusters substantially increased the odds of any misinformation engagement, whereas cultural, lifestyle, and opposition-oriented clusters generally decreased these odds, sometimes sharply so.

All results are confirmed by robustness checks using the three alternative outcome variable specifications. Only a few additional weakly significant coefficients emerged in the specification excluding comments (to avoid potential ambiguity from sceptical or critical comments); full details are provided in Figures C.8, C.10, C.12 in the Supplementary Information.

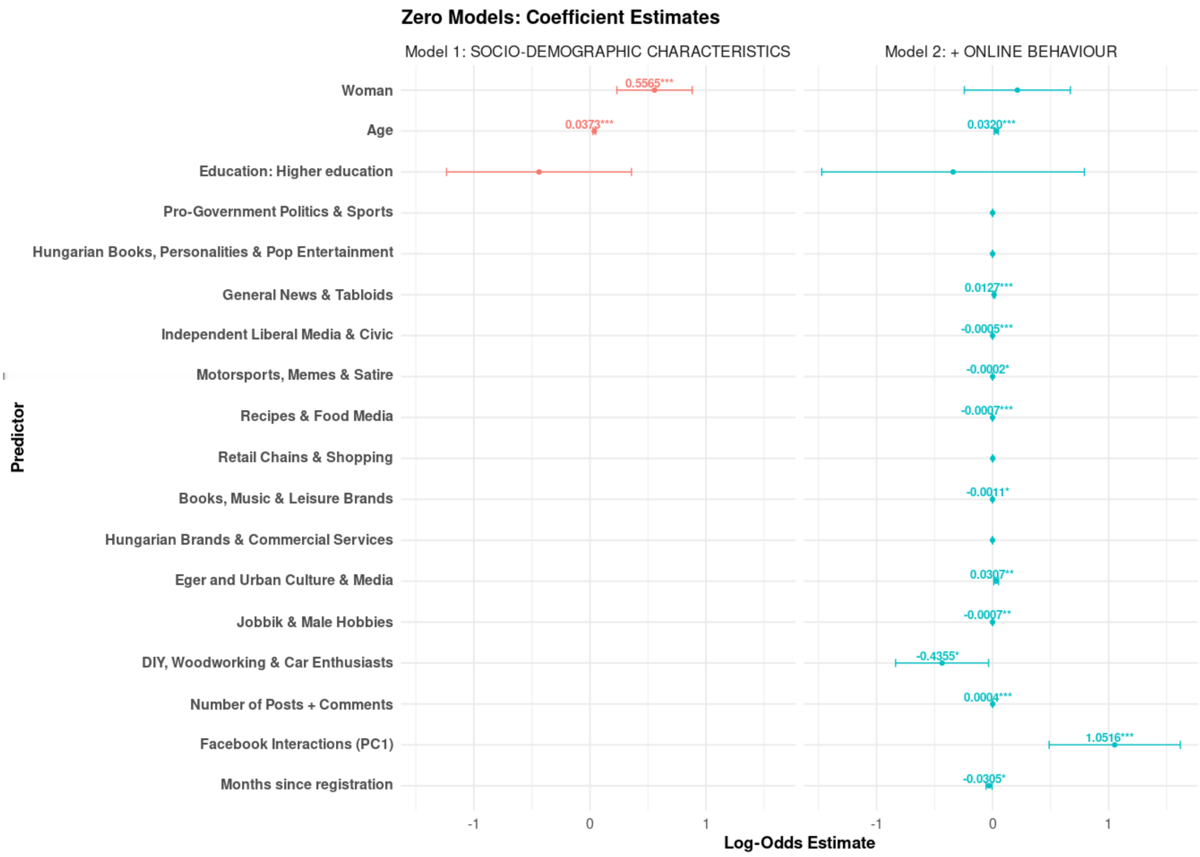


Figure 4.3: Coefficient estimates for the Zero models. For demonstration purposes, this figure includes coefficients which were significant in at least one model specification (including all zero and positive nested models) only. The figure with the full set of explanatory variables is available in the Supplementary Information (Figure C.4). *Note*: Significance levels are indicated such that: * $p < .05$, ** $p < .01$, *** $p < .001$.

4.5.3 Positive Models: Predicting the Volume of Misinformation Engagement

Figure 4.4 shows two nested positive models predicting the amount of misinformation engagement among participants who engaged at least once. As for the zero models, Model 1 includes socio-demographic variables, Model 2 adds Facebook activity measures and Facebook interest cluster variables.

In Model 1, among those who had any engagement with misinformation, age was positively associated with the volume of engagement. Higher education was negatively associated with the volume of engagement, however this effect disappeared once online behaviour variables were included. Other socio-demographic variables, including gender and income, were not significant in this models. Model fit of Model 1 was moderate, with adjusted McFadden's $R^2 = 0.048$.

Model 2 added Facebook interest cluster variables and Facebook activity measures as controls, significantly increasing model fit (adjusted McFadden's $R^2 = 0.155$)–but to a lesser extent compared to the zero models. Three clusters were associated with higher levels of misinformation engagement once participants had engaged at least once, including *Pro-Government Politics & Sports*, *Hungarian Books, Personalities & Pop Entertainment*, and *DIY, Woodworking & Car Enthusiasts*, the latter showing the largest increase in misinformation engagement volume. The *Retail Chains & Shopping*, and *Hungarian Brands & Commercial Services* interest clusters show small, but significant negative coefficients. Among Facebook activity measures, only the number of posts and comments was positively associated with the volume of engagement among those who had any engagement with misinformation–although the effect size of the number of posts and comments is small. Taken together, these findings indicate that once engaged with misinformation, members of pro-government, lifestyle, and DIY-oriented clusters engage more intensely, whereas consumer and retail clusters dampen misinformation engagement.

As in the case of the zero models, all results are confirmed by robustness checks using the three alternative outcome variable specifications. Similarly to the zero models, a few additional weakly significant coefficients emerged in the specification excluding comments; full details are provided in Figures C.9, C.11, C.13 in the Supplementary Information.

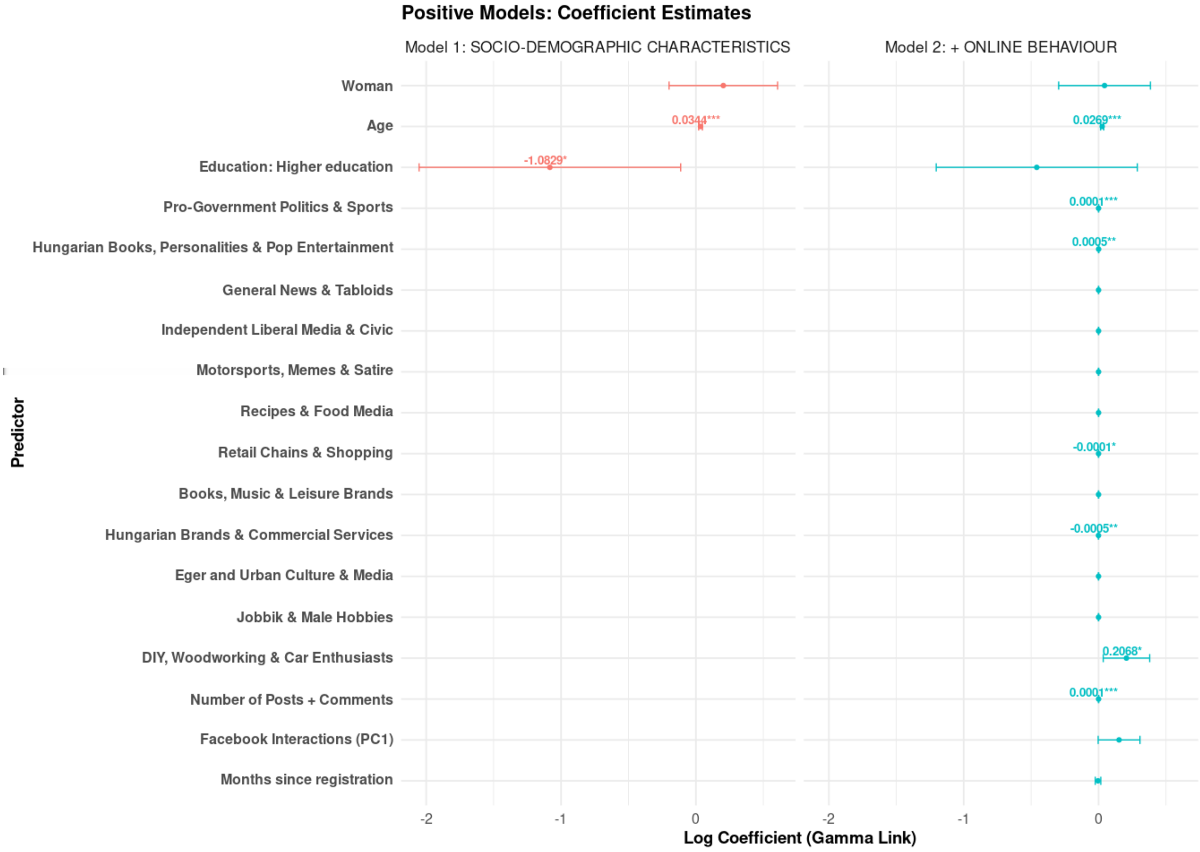


Figure 4.4: Coefficient estimates for the Positive models. For demonstration purposes, this figure includes coefficients which were significant in at least one model specification (including all zero and positive nested models) only. The figure with the full set of explanatory variables is available in the Supplementary Information (Figure C.6). *Note*: Significance levels are indicated such that: * $p < .05$, ** $p < .01$, *** $p < .001$.

Overall, these nested models indicate that age, certain interest cluster affiliations, and Facebook activity are associated with misinformation engagement. The notable increases in adjusted McFadden's R^2 suggest that interest clusters and Facebook activity variables account for additional variation beyond socio-demographic factors alone.

At the same time, the zero model performed better than the positive model, indicating that our independent variables are more effective in explaining whether someone engages with misinformation at all, rather than the intensity of such engagement once it occurs. Across both models, observed digital behaviour explains misinformation engagement to a much greater extent than socio-demographic characteristics.

Beyond these differences in overall fit, the factors associated with the likelihood of any misinformation engagement are not identical to those linked to the amount of engagement once it occurs. In the zero models, clusters centred on general news, tabloids, and local media substantially increased the likelihood of engagement, while lifestyle- and leisure-focused clusters, as well as opposition-aligned clusters, reduced it. In the positive models, however, engagement intensity was more strongly tied to pro-government political, sports, entertainment, and

DIY-oriented clusters, while consumption-related clusters dampened it. These patterns suggest that the environments facilitating entry into misinformation engagement are not necessarily the same as those driving its escalation.

These results highlight that considering socio-demographic characteristics, online interest affiliations, and behavioural measures together can provide a more complete picture of both whether users engage with misinformation and the extent of their engagement.

4.6 Discussion

This study contributes to our understanding of who engages with misinformation, by leveraging a unique data donation framework that combines digital trace data with self-reported characteristics, thereby enabling the assessment of misinformation engagement through a combination of self-reported socio-demographic characteristics and observed behaviour rather than exclusively relying on self-reports.

This study examined the factors associated with misinformation engagement on Facebook using a two-part hurdle modelling approach. By separately analysing the likelihood of any misinformation engagement (zero models) and the volume of misinformation engagement among participants who engaged at least once (positive models), we were able to disentangle the distinct processes shaping whether individuals interact with misinformation and how extensively they do so. This design allowed us to move beyond a binary view of misinformation engagement and instead capture both the entry point into misinformation interactions and the intensity of participation once engaged.

The results provide three main insights. First, among socio-demographic predictors age was consistently positively associated with both the likelihood of engaging with misinformation and the volume of misinformation engagement. Gender was positively associated with the likelihood of misinformation engagement when no online behavioural patterns were accounted for. The change in its coefficient across models may reflect that engagement with specific online topics and communities, as well as Facebook activity are correlated with gender, which absorbs its predictive power. Higher education was negatively associated with the intensity of misinformation engagement in the model that included only socio-demographic variables. However, this effect disappeared once online behavioural variables were added, suggesting that differences in online activity patterns may account for the initial association. These findings suggest that structural socio-demographic factors should be studied together with online behaviour to provide a more complete understanding of the drivers of misinformation engagement, rather than capturing certain traits alone.

Second, and most novel, membership in specific interest clusters on Facebook strongly shaped misinformation engagement. Interest clusters centred on tabloids, low-quality news outlets, regional and local media were positively associated with misinformation engagement, indicating that these information environments can amplify exposure and interaction. While these relationships cannot be interpreted causally, they invite several hypotheses about the mechanisms that might underlie them, which we discuss below. Specifically, the General News & Tabloids cluster is characterised by sensationalized, tabloid-style, or lower quality news outlets, with some featuring propagandistic content. This blend of entertainment and news may contribute to a less critical consumption pattern among users. (270) found that when news is framed as entertainment, users engage more readily with content, sometimes reducing scrutiny of factual accuracy. This suggests that entertainment framing can lower the cognitive defences of users, leading to increased susceptibility to misinformation. Furthermore, (175) demonstrated that analytic thinking and cognitive reflection are associated with reduced susceptibility to misinformation. However, the entertainment framing prevalent in this cluster may discourage users from engaging in the deeper cognitive processing necessary to critically evaluate information. This lack of engagement with reflective thinking could make users more vulnerable to engaging with misinformation. Therefore, the combination of sensationalized content, entertainment framing, and reduced cognitive reflection within the General News & Tabloids cluster may create an environment where users are more likely to engage with and propagate misinformation. Additionally, users who begin to engage with tabloid-like content are also more likely to be exposed to such content in the future because of the way social media recommendation algorithms reinforce revealed preferences. Therefore, the implications for user preferences should be expanded on by also recognising that social media algorithms may also contribute to the continued misinformation engagement of users associated with the General News & Tabloids cluster. The Eger and Urban Culture & Media cluster, centred around the city of Eger and combining local news, cultural heritage content, and urban lifestyle media attracts users interested in local culture, history, arts, and broader urban media trends. As (271) suggest, users embedded in such tightly-knit, identity-driven communities may be more likely to engage with misinformation due to heightened trust in in-group sources and reinforcement of shared narratives. By contrast, clusters focused on independent liberal media, opposition politics, civic issues, and lifestyle topics such as books, food, hobbies, and leisure were associated with decreased likelihood and lower intensity of misinformation engagement. On the one hand, both opposition-oriented interest clusters—Independent Liberal Media & Civic, representing the political left, and Jobbik & Male Hobbies, representing the political right—were negatively associated with misinformation engagement. This pattern suggests that divisions in misinformation susceptibility among Hungarian Facebook users may not primarily follow traditional left–right ideological lines, but rather align with support for or opposition to Fidesz, the governing party led by Viktor Orbán. One plausible explanation is that individuals engaging with independent media and opposition politics may display higher levels of critical thinking—manifested in their tendency to question the status quo—,a cognitive style that have been found to reduce susceptibility to misinformation (150; 175). On the other hand, several interest clusters centred on leisure and lifestyle—such as Motorsports, Memes & Satire, Recipes & Food Media, and Books, Music & Leisure Brands—were associated with a lower likelihood of misinformation engagement. One explanation is that these communities are oriented toward entertainment and everyday interests rather than political or controversial topics, which substantially reduces exposure to misinformation in the first place. Research shows that misinformation circulates most intensely in politicized contexts (114). Moreover, the mode of interaction typical within these clusters—centred on sharing recipes, humour, or aesthetic content—is predominantly light, prosocial, and community-oriented rather than argumentative or identity-driven. Such interaction styles are associated with lower levels of misinformation diffusion, as emotionally less contentious content tends to be less conducive to the virality of false information (118). In short, users embedded in hobby- and leisure-oriented environments tend to operate in informational spaces that are both less exposed to and less conducive to the spread of misinformation.

Regarding misinformation engagement volume, government-friendly, pop-culture–focused, and male-oriented handy work–related interest clusters were positively associated with misinformation engagement, indicating that these information environments can amplify exposure and interaction. Potential explanations for these patterns may resemble those observed for the clusters with significant associations in the zero model. For instance, the Hungarian Books, Personalities & Pop Entertainment cluster likely overlaps with lower-quality news media ecosystems. Similar to the dynamics seen in the General News & Tabloids cluster, the prevalence of entertainment framing and sensationalized content within such environments may foster more passive, less critical forms of engagement. (270) show that when news or quasi-informational content is presented as entertainment, audiences engage more readily but scrutinize factual accuracy less carefully. This lowered cognitive vigilance can make users more vulnerable to misinformation, especially when combined with the reduced analytic processing identified by (175). However, in this case, it is plausible that much of the misinformation circulating within these clusters may concern lifestyle, celebrity, or soft-news domains rather than explicitly political topics. The Pro-Government Politics & Sports cluster's positive association with misinformation engagement further supports the interpretation that the key divide in misinformation susceptibility among Hungarian Facebook users follows support for versus opposition to Fidesz, rather than traditional left–right ideological lines. Pro-government and sport-oriented communities often mix political content with entertainment and group identity cues, which may create information environments where partisan narratives circulate more easily and users are less likely to engage in reflective evaluation. Prior research shows that such identity-driven and emotionally charged contexts increase the spread of misinformation by encouraging intuitive rather than analytic processing (34; 272). In short, the presence of a distinct pro-government cluster that amplifies interaction supports the interpretation that the key dimension structuring misinformation engagement in our data is support for versus opposition to Fidesz, rather than conventional left–right ideology. However, these patterns should be interpreted in light of Hungary's distinctive media environment. Over the past decade, the Hungarian government has consolidated extensive control over traditional broadcast and print media, leading to widespread exposure to pro-government narratives that is not solely the result of active ideological self-selection (273). Importantly, several outlets that are now government-aligned were previously independent, meaning that audience composition has often been inherited through institutional transitions rather than exclusively formed through deliberate partisan sorting. As a result, higher engagement volumes observed within pro-government online clusters may partly reflect structural exposure and path-dependent media consumption, rather than uniquely strong psychological susceptibility. This context suggests that the political divide observed here is best understood as support for versus opposition to the governing party within a captured media system, rather than a generalizable left–right ideological cleavage. Finally, the DIY, Woodworking & Car Enthusiasts cluster may reflect a sub-community characterized by strong in-group trust and practical problem-solving discourse, where information is circulated in closed, self-validating networks. While such communities are not inherently political, their strong identity orientation and reliance on peer validation could occasionally facilitate uncritical sharing of misleading content, particularly when related to consumer products, technology, or everyday risks. This interpretation remains tentative but suggests that even ostensibly apolitical clusters can be associated with misinformation engagement under certain social and informational conditions. Interest clusters centred on retail and commercial services–Retail Chains & Shopping and Brand & Commercial Services–were negatively associated with misinformation engagement, similar to hobby- and leisure-oriented clusters. These communities focus on transactional and everyday interests rather than political or controversial topics, which reduces exposure to misinformation (114). Furthermore, interactions within these clusters—such as sharing product reviews, shopping tips, and practical advice—are functional and cooperative rather than argumentative or identity-driven, making them less conducive to the spread of false information (118).

These results suggest that online communities serve as powerful filters of information exposure: some ecosystems amplify engagement with misinformation, while others are associated with lower levels of such engagement. This pattern could indicate a protective role of certain communities, but it might also simply reflect that members of lifestyle- or leisure-focused clusters (e.g., cooking or arts and crafts) are less exposed to political and news-related content in the first place.

Overall, these differences are consistent with broader evidence that the structure and focus of online communities can influence patterns of misinformation exposure and interaction, echoing findings discussed in the Related Work section (119; 120). Regardless of the underlying mechanism, these differences highlight the critical role that the informational and cultural environments within online communities may play in shaping the spread of false information.

Finally, Facebook activity was included as a set of control variables, since participants who are generally more active on the platform—posting, commenting, and reacting more frequently—are statistically more likely to interact with or share misinformation at some point. These measures were strongly associated with the likelihood misinformation engagement, and partially with the volume of misinformation engagement. Importantly, the inclusion of these activity controls did not explain away the effects of other predictors, many of which remained significant even when accounting for overall intensity of Facebook use (see details in the Supplementary Information Figures C.14 and C.15). This underscores that, while overall Facebook activity partly explains engagement with misinformation, the patterns associated with age and online interest community affiliations provide additional, meaningful insight into misinformation engagement.

An additional contribution of this study lies in showing that our models better explain whether individuals engage with misinformation at all, than the volume of engagement among those who do. This suggests that the determinants of the likelihood of any misinformation engagement may be clearer and more systematic, while the intensity of subsequent activity is likely shaped by additional unobserved factors. Importantly, across both types of models, observed digital behaviour proved to be a far stronger predictor than socio-demographic variables, underscoring the importance of incorporating behavioural data into future research.

Taken together, these findings contribute to ongoing debates on misinformation by emphasizing that engagement with misinformation is not randomly distributed, but structured by online communities and behavioural patterns. This has important implications for interventions: individual-level solutions, such as fact-checking and digital literacy campaigns, may be necessary but insufficient if the broader community context remains unaddressed. More effective strategies may involve targeting specific communities or network clusters and supporting alternative online spaces that provide credible content and foster civic engagement.

Several limitations of this study should however be acknowledged. First, our analysis identifies associations but was not designed for causal inference, meaning that both directions of influence remain possible—for example, interest cluster membership may shape misinformation engagement, but misinformation engagement could also drive cluster membership. As in the case of any modelling, unobserved covariates could still influence our results. To reduce the number of potential confounders, we conducted robustness checks with psychological characteristics and political attitudes (see Figures C.14 and C.15 in the Supplementary Information) and found no significant effects, but other unmeasured factors, such as for example network structure, offline participation, or media literacy may matter.

Our sample may be affected by voluntary data donation, introducing potential selection bias as participants who contribute their data may differ systematically from the broader population even if their composition is similar to the population by main demographic variables. While the sample was collected to be representative of the Hungarian internet-using population with respect to gender, age, education, region, and settlement type, participation in a data donation study requires not only willingness but also the ability to complete a technically demanding task. As a result, individuals who are willing and able to donate digital trace data may differ from the broader population in ways not fully captured by these demographic variables—particularly in terms of digital literacy. A recent study (235) shows that realized data donation in this project was shaped not only by stated willingness but also by capacity-related factors such as education and age, likely reflecting underlying digital proficiency. Therefore, even conditional on observed demographic characteristics, data-donating individuals may possess stronger digital literacy skills than their demographic counterparts who do not participate. Individuals with higher digital literacy may engage differently with online information environments, for example by being more aware of the prevalence and forms of misinformation or by exhibiting greater selectivity in content consumption. Accordingly, the results should be interpreted with caution, as they may not fully generalize to populations with lower levels of digital literacy.

A further limitation concerns the relatively low prevalence of misinformation in the observed data. While this reflects real-world base rates—misinformation remains a minority of overall content—it nonetheless has implications for interpretation. When the outcome is rare, estimates may become more sensitive to modelling choices and harder to interpret substantively. At the same time, the fact that several associations remain robust despite the low prevalence of misinformation suggests that the observed relationships are unlikely to be mere artifacts of noise.

Additionally, the low prevalence of misinformation observed in the present dataset also highlights the importance of definitional choices in misinformation research. While explicit falsehoods originating from low-credibility domains appear to be relatively rare, prior work suggests that misleading, propagandistic, or selectively framed content may be substantially more prevalent and potentially more influential (113; 274). The present study intentionally adopts a narrow definition to provide a conservative benchmark. Future research could extend this approach by incorporating broader, graded, or topic-based classifications of information quality, allowing for a more comprehensive assessment of susceptibility to online manipulation.

Furthermore, the study focuses exclusively on Facebook, so the findings may not generalize to platforms like YouTube, TikTok, or X/Twitter, which have different user bases and engagement dynamics. Additionally, misinformation engagement was measured through observable actions—likes, comments, and group membership—excluding private sharing (e.g., Messenger) and passive exposure. As a result, our estimates capture visible misinformation engagement rather than total exposure or consumption.

Finally, the context-specific nature of Hungary's polarized media system limits generalizability. Extensive government influence over traditional broadcast and print media shapes patterns of political exposure in ways that differ from those in more decentralized and competitive media environments, and these structures may indirectly affect observable online engagement through path-dependent audience composition and displacement effects. Nonetheless, Hungary represents a critical case where vulnerabilities are amplified, given the country's well-documented democratic backsliding, increasing concentration of media ownership, and persistent political polarization. These factors make Hungary an especially salient setting for examining how individuals engage with misinformation, as dynamics that are particularly visible here may foreshadow or mirror developments in other countries facing similar pressures.

Despite these limitations, the study demonstrates the value of integrating both socio-demographic and online behavioural predictors into the study of misinformation. Our findings show that while socio-demographic factors alone may not strongly predict misinformation engagement, combining survey-based characteristics with digital behavioural indicators allows for a richer understanding than either source could provide in isolation. Whereas survey data have long been central to the study of misinformation, their integration with behavioural traces adds important dimensions to the existing approaches. Future work should continue to incorporate both sources of data, as their complementarity offers a more comprehensive picture of the factors shaping misinformation engagement. Misinformation engagement is shaped by both individual characteristics and digital behaviour, and taking these together provides a more comprehensive basis for understanding and addressing the phenomenon.

To conclude, this study used a data donation framework to link self-reported characteristics with Facebook trace data, offering insight into how socio-demographic factors and online behaviour relate to engagement with misinformation. The findings point to the relevance of age, to some extent education and gender, overall activity on the platform (e.g., sharing, commenting and reactions), and the online communities individuals participate in. The results also point out that patterns of digital behaviour appear to play a larger role than socio-demographic characteristics in explaining engagement with misinformation. Conducted in Hungary, the study contributes evidence from a context that has received limited attention, while also addressing the gap in research that connects socio-demographic characteristics measured via survey to observed misinformation-related behaviour. More broadly, it illustrates how survey and digital trace data can be combined through data donation to examine patterns of individual-level engagement with misinformation.

Chapter 5

Conclusion and Outlook

This thesis set out to examine how digital trace data can be used to deepen our understanding of the individual, socio-demographic and behavioural foundations of misinformation engagement. Across three empirical studies, it demonstrated a progression from using accessible but limited online indicators toward integrating rich, participant-donated digital traces with survey measures, each addressing different levels and challenges of data accessibility and granularity, and together illustrating the opportunities and limitations of studying online behaviour in an increasingly regulated digital environment.

The first study explored the potential of publicly accessible social media metrics for identifying users who are more likely to share misinformation. Drawing on openly available indicators of social media activity, a large dataset of Twitter users from mostly Western contexts, it demonstrated that even coarse-grained, easily retrievable data can reveal meaningful patterns in misinformation sharing behaviour. The study showed that certain engagement characteristics are systematically associated with misinformation propagation. Specifically, the study found that higher activity (more tweets per day) and greater numbers of followers were robust correlates of lower factuality, while older accounts tended to be associated with higher factuality. Several interaction effects revealed that these relationships depend on users' network characteristics: for example, the link between high activity and low factuality was weaker among users following fewer accounts. These findings highlight that meaningful behavioural insights can still be derived from limited data, offering an accessible way to study misinformation engagement in the absence of detailed individual-level data.

The second study turned to the question of how more detailed, individual-level data might be obtained in a transparent and ethically sustainable way in order to study misinformation in greater depth. Recognising the constraints imposed by data protection regulations and platform restrictions, it examined the determinants of individuals' willingness to donate their social media data for research purposes. Using two vignette experiments conducted in Hungary and the United States, the study identified both contextual and attitudinal factors that influence willingness to participate, such as monetary incentive, perceived legitimacy of the research, and the type of data requested. Specifically, across both countries, monetary incentives emerged as the strongest motivator—particularly in Hungary—while non-monetary incentives (such as feedback reports) were effective only in the U.S. sample. Task complexity (time or number of platforms involved) reduced willingness in Hungary but had limited effect in the U.S., where respondents were generally more digitally active. Among individual-level factors, older and more educated respondents in Hungary were less open to donation, whereas in the U.S., privacy and security concerns reduced willingness. Personality traits and tech affinity showed no consistent effects. Together, these findings underscore that data donation is both feasible and context-dependent,and also highlight that systematic selection biases may arise in donation-based datasets, particularly across cultural and institutional contexts.

The third study complements these earlier analyses by focusing on a case study of Hungary, combining survey data with participant-donated Facebook traces. This chapter offered a detailed examination of how socio-demographic characteristics, online activity patterns, and community structures relate to interactions with misinformation sources. Two modelling approaches distinguished between the likelihood of any engagement ("zero models") and the volume of engagement among engagers ("positive models"). The results revealed that behavioural variables—such as posting frequency, comment activity, and interaction intensity—were powerful predictors in explaining misinformation engagement. Among socio-demographic variables the study found that older users were more likely to engage, while education effects disappeared once online behaviour was included. The analysis of interest clusters further clarified the social environments of misinformation. Users interested in general news, tabloid, and urban culture pages had higher likelihoods of engaging with misinformation, while those oriented toward independent media, civic activism, satire, and leisure were less prone. Among active misinformation engagers, higher engagement volumes were concentrated in clusters related to pro-government politics, entertainment, and DIY culture. Model fits improved substantially when behavioural and interest-based variables were included, demonstrating that what users do and where they interact contributes significantly to explaining misinformation engagement as compared to just considering socio-demographic factors obtained from self-report based surveys.

Unlike the first study, which relied on widely accessible but low-granularity behavioural data, the last study leveraged high-resolution, participant-donated data to explore these dynamics in greater depth. In this sense, the first and last studies together illustrate different pathways for investigating misinformation engagement under current data-access constraints, from large-scale but coarse analyses to smaller-scale, high-granularity research built on informed consent.

Taken together, the three studies form a coherent methodological and empirical arc. They show how meaningful insights into misinformation engagement can be derived from data that vary in accessibility and depth, while maintaining ethical integrity throughout. Methodologically, the thesis demonstrates how different types of digital and survey data can be integrated to address research questions about misinformation engagement. It illustrates the value of large-scale and publicly accessible traces with limited variable set, as well as the added value of smaller, but richer and representative datasets to balance breadth, depth, and ethical rigour in digital behavioural research. Substantively, it advances understanding of how misinformation engagement is shaped by individual traits, such as socio-demographic characteristics, and online activity patterns.

Beyond its empirical contributions, this thesis engages with a central tension of contemporary social research: how to study human behaviour in digital spaces responsibly, at a time when some of the most informative data are also some of the most sensitive. By focusing on approaches that respect participant autonomy and data protection, the work aligns with a broader movement toward more participatory and transparent forms of digital research. The data donation framework, while still emerging, exemplifies this shift, transforming individuals from passive data sources into active collaborators whose informed choices shape what kinds of knowledge can be produced. Together, the studies suggest that ethical and methodological rigour are not opposing goals but interdependent ones, each essential for sustaining both the credibility and the social value of research in the digital age. In this way, the thesis demonstrates a pathway for conducting socially responsible, methodologically robust research that can guide the study of misinformation and broader patterns of online human behaviour in increasingly complex digital environments.

5.1 Limitations and Future Directions

While the studies collectively advance knowledge about misinformation engagement, they also highlight several limitations and point to directions for future research. First, differences in platform data accessibility constrain the comparability of digital behavioural research across contexts. The datasets analysed here, though diverse, represent only a subset of online environments and behaviours. Future research could expand to other platforms and media formats, integrating multimodal data (e.g., text, images, and network structures) to capture more complex interaction patterns. Additionally, theoretical work is needed to link behavioural indicators with psychological mechanisms such as motivation, identity, and trust. Digital trace data can reveal what people do online, but understanding why they do it requires integrating these behavioural patterns with attitudinal and cognitive measures.

Furthermore, behavioural data must always be interpreted within the social and technological contexts in which they are produced. Even high-granularity data do not automatically yield clear explanations; careful theoretical framing and transparency in analytical choices, as well as the knowledge of the local context remain essential to ensure that digital evidence supports meaningful social inference. While this thesis examined Hungary as a politically salient case, future studies could adopt a comparative design to explore how misinformation engagement operates across different political and media systems. Such cross-national comparisons would clarify whether the observed dynamics—for example the role of online community participation—generalise beyond this context.

Related to data donation, although the combination of survey and trace data enriches behavioural analysis, it also introduces challenges of representativeness and sample bias. Data donation remains a selective process, as detailed in Study 2, and individuals willing to share their data may differ systematically from those who do not. Improving participation rates—through clearer communication, ethical incentive structures, and transparent data handling—remains a critical avenue for methodological innovation.

Finally, another challenge concerns the unequal distribution of resources required for ethically robust research that includes detailed digital trace data on individuals. Such projects demand secure data management infrastructure, technical expertise, and adequate participant incentives—capacities that are not equally available across research institutions or geographical regions. These disparities shape who can feasibly undertake high-quality digital behavioural research and risk reinforcing existing inequalities in access to rich data sources.

Overall, this thesis contributes to the growing field of digital behavioural research by demonstrating that the integration of survey and digital trace data offers a powerful way to study misinformation engagement effectively and responsibly. It shows that even under restrictive data-access conditions, researchers can produce meaningful insights by combining scalable, rich, and ethically collected data sources. As data environments and regulations continue to evolve, this integrative approach provides a foundation for future research that is not only methodologically rigorous and ethically sound, but also sensitive to the social and political contexts in which misinformation spreads.

Bibliography

[1] World Economic Forum, "Global risks report 2025," January 2025. Accessed: 2025-10-27.

[2] Z. Adams, M. Osman, C. Bechlivanidis, and B. Meder, "(why) is misinformation a problem?," *Perspectives on Psychological Science*, vol. 18, no. 6, pp. 1436–1463, 2023.

[3] M. Luo and J. T. Hancock, "Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues.," *Communication Research*, vol. 49, no. 2, pp. 171–195, 2022.

[4] S. Altay, A.-S. Hacquin, C. Chevallier, and H. Mercier, "Why do so few people share fake news? it hurts their reputation.," *New Media & Society*, vol. 22, no. 5, pp. 900–923, 2020.

[5] J. R. Axt and J. F. Landy, "Political news and trust: The role of misinformation in undermining credibility," *Political Communication*, vol. 37, no. 6, pp. 725–740, 2020.

[6] C. Fisher and S. Park, "Perceived journalistic quality and misinformation exposure," *Journalism Studies*, vol. 22, no. 14, pp. 2005–2022, 2021.

[7] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, vol. 31, p. 211–36, May 2017.

[8] S. Lewandowsky and S. van der Linden, "Misinformation and its correction: Continued influence and successful debiasing," *Psychological Science in the Public Interest*, vol. 22, no. 3, pp. 1–79, 2021.

[9] C. R. Sunstein, *Republic: Divided democracy in the age of social media*. Princeton, NJ: Princeton University Press, 2017.

[10] M. T. Bastos and D. Mercea, "Brexit: The role of fake news and misinformation," *Social Science Computer Review*, vol. 35, no. 3, pp. 379–399, 2017.

[11] N. A. Cooke, "Post-truth, fake news, and the media," *Library Technology Reports*, vol. 53, no. 8, pp. 1–33, 2017.

[12] P. Pomerantsev, *Nothing is true and everything is possible: The surreal heart of the new Russia*. New York: PublicAffairs, 2015.

[13] T. Snyder, *Our malady: Lessons in liberty from a hospital diary*. New York: Crown, 2021.

[14] S. Tasnim, M. M. Hossain, and H. A. Mazumder, "Impact of rumors and misinformation on covid-19 in social media," *Journal of Preventive Medicine and Public Health*, vol. 53, no. 3, pp. 171–174, 2020.

[15] A. Shiina, T. Niitsu, O. Kobori, K. Idemoto, T. Hashimoto, T. Sasaki, M. Nakazato, and M. Iyo, "Relationship between perception and anxiety about covid-19 infection and risk behaviors for spreading infection: A national survey in japan," *Brain, Behavior, and Immunity*, vol. 87, pp. 174–187, 2020.

[16] S. Lewandowsky, U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook, "Misinformation and its correction: The continued influence effect and successful debiasing," *Psychological Science in the Public Interest*, vol. 13, no. 3, pp. 106–131, 2012.

[17] S. Pluviano, C. Watt, and S. Della Sala, "Persistence of vaccine misinformation: Evidence from experimental studies," *Health Psychology*, vol. 41, no. 2, pp. 109–120, 2022.

[18] T. Bolsen, J. N. Druckman, and F. L. Cook, "The influence of partisan motivated reasoning on public opinion," *Political Behavior*, vol. 38, no. 3, pp. 761–789, 2016.

[19] J. Cook, S. Lewandowsky, and E. Maibach, *Climate change communication and the challenge of misinformation*. Oxford University Press, 2018.

[20] M. Kogan and A. Mosseri, "The business of misinformation: Social and economic costs," *Business Horizons*, vol. 64, no. 5, pp. 589–601, 2021.

[21] B. G. Southwell, E. A. Thorson, and L. Sheble, *Misinformation and mass audiences*. Austin: University of Texas Press, 2015.

[22] B. F. Liu, S. Kim, and H. Song, "Fake news on twitter during crisis events: A case of the white house bombing hoax," *Public Relations Review*, vol. 46, no. 2, pp. 101–117, 2020.

[23] J. Cavazos, T. Rutherford, and U. Akcigit, "The global costs of misinformation: Economic impacts and challenges," *World Development*, vol. 123, p. 104623, 2019.

[24] U. K. Ecker, S. Lewandowsky, M. Apai, and J. Freidrich, "Explicit warnings reduce but do not eliminate the continued influence of misinformation," *Memory & Cognition*, vol. 39, no. 3, pp. 486–503, 2011.

[25] A. Acerbi, "Cultural evolution of misinformation: An analytical framework," *Royal Society Open Science*, vol. 6, no. 9, p. 190382, 2019.

[26] T. T. Hills, "The dark side of information proliferation," *Perspectives on Psychological Science*, vol. 14, no. 3, pp. 323–330, 2019.

[27] N. Newman, R. Fletcher, A. Schulz, S. Andı, and R. K. Nielsen, "The hostile media phenomenon revisited: Revisiting effects of repeated exposure to misinformation," *International Journal of Public Opinion Research*, vol. 31, no. 1, pp. 1–22, 2019.

[28] M. D. Robinson and A. K. Fetterman, "Biases in the perception of self versus others," *Personality and Social Psychology Bulletin*, vol. 21, no. 12, pp. 1278–1289, 1995.

[29] E. Pronin, D. Y. Lin, and L. Ross, "The bias blind spot: Perceptions of bias in self versus others," *Personality and Social Psychology Bulletin*, vol. 28, no. 3, pp. 369–381, 2002.

[30] M. V. Bronstein, G. Pennycook, A. Bear, D. G. Rand, and T. D. Cannon, "Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking," *Journal of Applied Research in Memory and Cognition*, vol. 10, no. 1, pp. 108–120, 2021.

[31] A. R. B. Soutter, T. C. Bates, and R. Mottus, "Misinformation and individual differences in sharing behavior on social media," *PLoS One*, vol. 15, no. 1, p. e0227823, 2020.

[32] S. J. Drucker, "Rumors and collective behavior: A review," *Sociological Inquiry*, vol. 75, no. 2, pp. 189–215, 2005.

[33] G. Pennycook and D. G. Rand, "The psychology of fake news," *Trends in Cognitive Sciences*, vol. 25, no. 5, pp. 388–402, 2021.

[34] M. Osmundsen, A. Bor, P. B. Vahlstrup, A. Bechmann, and M. B. Petersen, "Partisan polarization is the primary psychological motivation behind political fake news sharing on twitter," *American Political Science Review*, vol. 115, no. 3, pp. 999–1015, 2021.

[35] C. Sun and Y. Xie, "Individual traits and misinformation sharing: A meta-analysis," *Journal of Communication*, 2024. Advance online publication.

[36] P. L. Moravec, R. Minas, and A. R. Dennis, "Fake news on social media: People believe what they want to believe when it makes no sense at all," *MIS Quarterly*, vol. 43, pp. 13430–11360, 2019.

[37] W. Yan and Z. Pan, "Believing and sharing false news on social media: The role of news presentation, epistemic motives, and deliberative thinking," *Media Psychology*, 2023.

[38] T. Buchanan and J. Kempley, “Individual differences in sharing false political information on social media: Direct and indirect effects of cognitive-perceptual schizotypy and psychopathy,” *Personality and Individual Differences*, vol. 182, p. 111071, 2021.

[39] A. M. Guess, J. Nagler, and J. A. Tucker, “Less than you think: Prevalence and predictors of fake news dissemination on facebook,” *Science Advances*, vol. 5, no. 1, p. eaau4586, 2019.

[40] T. Buchanan, “Why do people spread false information online? the effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation,” *PLoS One*, vol. 15, no. 10, pp. 1–33, 2020.

[41] L. D. Scherer and G. Pennycook, “Who is susceptible to online health misinformation?,” *American Journal of Public Health*, vol. 111, pp. S276–S277, 2020.

[42] T. Koltay, “The media and the literacies: Media literacy, information literacy, digital literacy,” *Media, Culture & Society*, vol. 33, no. 2, pp. 211–221, 2011.

[43] E. Commission, “A european approach to media literacy in the digital environment.” https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM%3A2007%3A0833%3AFIN%3AEN%3APDF, 2007.

[44] S. M. Jones-Jang, T. Mortensen, and J. Liu, “Does media literacy help identification of fake news? information literacy helps, but other literacies don’t,” *American Behavioral Scientist*, vol. 65, no. 2, pp. 371–388, 2021.

[45] A. L. Association, “Presidential committee on information literacy: Final report.” https://www.ala.org/acrl/publications/whitepapers/presidential, 1989.

[46] R. Hobbs, “Multiple visions of multimedia literacy: Emerging areas of synthesis,” in *International handbook of literacy and technology* (M. McKenna, L. D. Labbo, R. D. Kieffer, and D. Reinking, eds.), pp. 15–26, Lawrence Erlbaum, 2006.

[47] P. Gilster, *Digital literacy*. John Wiley, 1997.

[48] E. Hargittai, “Survey measures of web-oriented digital literacy,” *Social Science Computer Review*, vol. 23, no. 3, pp. 371–379, 2005.

[49] E. Hargittai and Y.-P. Hsieh, “Succinct survey measures of web-use skills,” *Social Science Computer Review*, vol. 30, pp. 95–107, 2011.

[50] U. A. de Barcelona, “Studies on the current trends and approaches to media literacy in europe.” http://ec.europa.eu/avpolicy/media_literacy/docs/studies/study.pdf, 2007.

[51] D.-T. Chen, J. Victor, Wu, and Y. Wang, "Unpacking new media literacy," *Journal of Systemics, Cybernetics and Informatics*, vol. 9, no. 2, pp. 84–88, 2011.

[52] M. Koc and E. Barut, "Development and validation of new media literacy scale (nmls) for university students," *Computers in Human Behavior*, vol. 63, pp. 834–843, 2016.

[53] C. González-Cabrera, C. Ugalde, C. Figueroa, and J. Pesántez, "The impact of media literacy on the intention to share fake information in social networks," in *11th international conference on education and new learning technologies*, 2019.

[54] A. Keselman, C. A. Smith, G. Leroy, and D. R. Kaufman, "Factors influencing willingness to share health misinformation videos on the internet: Web-based survey," *Journal of Medical Internet Research*, vol. 23, no. 12, 2021.

[55] J. De Keersmaecker and A. Roets, ""fake news": Incorrect, but hard to correct. the role of cognitive ability on the impact of false information on social impressions," *Intelligence*, vol. 65, pp. 107–110, 2017.

[56] M. Mosleh, G. Pennycook, A. A. Arechar, and D. G. Rand, "Cognitive reflection correlates with behavior on twitter," *Nature Communications*, vol. 12, no. 1, p. 921, 2021.

[57] M. Mosleh, Q. Yang, T. Zaman, G. Pennycook, and D. G. Rand, "Differences in misinformation sharing can lead to politically asymmetric sanctions," *Nature*, 2024.

[58] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news: Addressing fake news requires a multidisciplinary effort," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.

[59] S. Ahmed, "Navigating the maze: Deepfakes, cognitive ability, and social media news skepticism," *New Media and Society*, pp. 1–22, 2021.

[60] A. Ardèvol-Abreu and H. Gil de Zúñiga, "Effects of editorial media bias perception and media trust on the use of traditional, citizen, and social media news," *Journalism & Mass Communication Quarterly*, vol. 94, no. 3, pp. 703–724, 2017.

[61] A. Kalogeropoulos and J. Suiter, "News media trust and news consumption: Factors related to trust in news in 35 countries," *International Journal of Communication*, vol. 13, pp. 3672–3693, 2019.

[62] S. Altay, A. V. Hacquin, C. Chevallier, and H. Mercier, "Why do so few people share fake news? it hurts their reputation," *New Media & Society*, vol. 24, no. 6, pp. 1303–1324, 2022.

[63] T. Hopp, P. Ferrucci, and C. J. Vargo, "Why do people share ideologically extreme, false, and misleading content on social media? a self-report and trace data-based analysis of countermedia content dissemination on facebook and twitter," *Human Communication Research*, vol. 46, no. 4, pp. 357–384, 2020.

[64] S. Valenzuela, D. Halpern, J. E. Katz, and J. P. Miranda, "The paradox of participation versus misinformation: Social media, political engagement, and the spread of misinformation," *Digital Journalism*, vol. 7, no. 6, pp. 802–823, 2019.

[65] S. Laato, A. K. M. N. Islam, M. N. Islam, and E. Whelan, "What drives unverified information sharing and cyberchondria during the covid-19 pandemic?," *European Journal of Information Systems*, vol. 29, no. 3, pp. 288–305, 2020.

[66] O. D. Apuke and B. Omar, "Social media affordances and information abundance: Enabling fake news sharing during the covid-19 health crisis," *Health Informatics Journal*, vol. 27, no. 3, 2021.

[67] D. R. Obada and D. C. Dabija, """in flow"! why do users share fake news about environmentally friendly brands on social media?," *International Journal of Environmental Research and Public Health*, vol. 19, no. 8, pp. 1–26, 2022.

[68] Y. Tsfati and J. N. Cappella, "Do people watch what they do not trust?: Exploring the association between news media skepticism and exposure," *Communication Research*, vol. 30, no. 5, pp. 504–529, 2003.

[69] A. Kim, P. L. Moravec, and A. R. Dennis, "Combating fake news on social media with source ratings: The effects of user and expert reputation ratings," *Journal of Management Information Systems*, vol. 36, no. 3, pp. 931–968, 2019.

[70] S. J. Kim, *How do people process and share fake news on social media? In the context of dual-process of credibility with partisanship, cognitive appraisal to threat, and corrective action*. PhD thesis, Syracuse University, 2020. Unpublished doctoral dissertation.

[71] S. Ahmed, "Fooled by the fakes: Cognitive differences in perceived claim accuracy and sharing intention of non-political deepfakes," *Personality and Individual Differences*, vol. 182, pp. 1–4, 2021.

[72] T. Buchanan, "Trust, personality, and belief as determinants of the organic reach of political disinformation on social media," *The Social Science Journal*, pp. 1–12, 2021.

[73] E. Katz, J. G. Blumler, and M. Gurevitch, "Uses and gratifications research," *Public Opinion Quarterly*, vol. 37, no. 4, pp. 509–523, 1974.

[74] T. E. Ruggiero, “Uses and gratifications theory in the 21st century,” *Mass Communication & Society*, vol. 3, no. 1, pp. 3–37, 2000.

[75] A. Whiting and D. Williams, “Why people use social media: A uses and gratifications approach,” *Qualitative Market Research: An International Journal*, vol. 16, no. 4, pp. 362–369, 2013.

[76] O. D. Apuke and B. Omar, “Fake news and covid-19: Modelling the predictors of fake news sharing among social media users,” *Telematics and Informatics*, vol. 56, pp. 1–26, 2021.

[77] A. Bykov, “Altruism: New perspectives of research on a classical theme in sociology of morality,” *Current Sociology*, vol. 65, no. 6, pp. 797–813, 2017.

[78] V. Y. Chen, “Examining news engagement on facebook: Effects of news content and social networks on news engagement,” *Mass Communication & Society*, vol. 23, no. 6, pp. 833–857, 2020.

[79] A. Duffy, E. Tandoc, and R. Ling, “Too good to be true, too good not to share: The social utility of fake news,” *Information, Communication & Society*, vol. 23, no. 13, pp. 1965–1979, 2020.

[80] M. J. Metzger, A. J. Flanagin, P. Mena, S. Jiang, and C. Wilson, “From dark to light: The many shades of sharing misinformation online,” *Media and Communication*, vol. 9, no. 1, pp. 134–143, 2021.

[81] J. An, D. Quercia, M. Cha, K. Gummadi, and J. Crowcroft, “Sharing political news: The balancing act of intimacy and socialization in selective exposure,” *EPJ Data Science*, vol. 3, pp. 1–21, 2014.

[82] C. S. Lee and L. Ma, “News sharing in social media: The effect of gratifications and prior experience,” *Computers in Human Behavior*, vol. 28, no. 2, pp. 331–339, 2012.

[83] K. Hur, T. T. Kim, O. M. Karatepe, and G. Lee, “An exploration of the factors influencing social media continuance usage and information sharing intentions among korean travellers,” *Tourism Management*, vol. 63, pp. 170–178, 2017.

[84] A. N. Islam, S. Laato, S. Talukder, and E. Sutinen, “Misinformation sharing and social media fatigue during covid-19: An affordance and cognitive load perspective,” *Technological Forecasting and Social Change*, vol. 159, p. 120201, 2020.

[85] B. Sampat and S. Raj, “Fake or real news? understanding the gratifications and personality traits of individuals sharing fake news on social media platforms,” *Aslib Journal of Information Management*, vol. 74, no. 5, pp. 840–876, 2022.

[86] K. Baek, A. Holton, D. Harp, and C. Yaschur, “The links that bind: Uncovering novel motivations for linking on facebook,” *Computers in Human Behavior*, vol. 27, no. 6, pp. 2243–2248, 2011.

[87] V. Balakrishnan, K. S. Ng, and H. A. Rahim, “To share or not to share – the underlying motives of sharing fake news amidst the covid-19 pandemic in malaysia,” *Technology in Society*, vol. 66, 2021.

[88] N. Thompson, X. Wang, and P. Daya, “Determinants of news sharing behavior on social media,” *Journal of Computer Information Systems*, vol. 60, no. 6, pp. 593–601, 2020.

[89] R. R. McCrae and O. P. John, “An introduction to the five-factor model and its applications,” *Journal of Personality*, vol. 60, no. 2, pp. 175–215, 1992.

[90] R. R. McCrae and P. T. Costa, “The five-factor theory of personality,” in *Handbook of personality theory and research* (O. P. John, ed.), pp. 139–153, Guilford Press, 1999.

[91] D. P. Calvillo, R. J. B. Garcia, K. Bertrand, and T. A. Mayers, “Personality factors and self-reported political news consumption predict susceptibility to political fake news,” *Personality and Individual Differences*, vol. 174, pp. 1–6, 2021.

[92] E. C. Hirschman and M. B. Holbrook, “Hedonic consumption: Emerging concepts, methods,” *Journal of Marketing*, vol. 46, no. 3, pp. 92–101, 1982.

[93] L. He, H. Yang, X. Xiong, and K. Lai, “Online rumor transmission among younger and older adults,” *Sage Open*, vol. 9, no. 3, 2019.

[94] A. L. Alter and D. M. Oppenheimer, “Uniting the tribes of fluency to form a metacognitive nation,” *Personality and Social Psychology Review*, vol. 13, no. 3, pp. 219–235, 2009.

[95] R. Reber and N. Schwarz, “Effects of perceptual fluency on judgments of truth,” *Consciousness and Cognition*, vol. 8, pp. 338–342, 1999.

[96] L. Hasher, D. Goldstein, and T. Toppino, “Frequency and the conference of referential validity,” *Journal of Verbal Learning and Verbal Behavior*, vol. 16, no. 1, pp. 107–112, 1977.

[97] G. Pennycook, T. D. Cannon, and D. G. Rand, “Prior exposure increases perceived accuracy of fake news,” *Journal of Experimental Psychology: General*, vol. 147, no. 12, pp. 1865–1880, 2018.

[98] D. A. Effron and M. Raj, “Misinformation and morality: Encountering fake-news headlines makes them seem less unethical to publish and share,” *Psychological Science*, vol. 31, no. 1, pp. 75–87, 2020.

[99] H. Lee and J. Kim, “Norm perceptions about rumor sharing on genetically modified foods: The interaction between facebook likes and a refuting comment,” *Health Communication*, pp. 1–11, 2021.

[100] J. Choudrie, S. Banerjee, K. Kotecha, R. Walambe, H. Karende, and J. Ameta, “Machine learning techniques and older adults’ processing of online information and misinformation: A covid-19 study,” *Computers in Human Behavior*, vol. 119, pp. 1–11, 2021.

[101] A. Ardévol-Abreu, P. Delponti, and C. Rodríguez-Wangüemert, “Intentional or inadvertent fake news sharing? fact-checking warnings and users’ interaction with social media content,” *Profesional de La Información*, vol. 29, no. 5, pp. 1–13, 2020.

[102] I. K. Ali, *The fake news phenomenon: Impact of heuristic cues on perceived credibility and sharing on social media*. PhD thesis, University of Miami, 2019. Unpublished doctoral dissertation.

[103] X. Xiao and Y. Su, “Wired to seek, comment and share? examining the relationship between personality, news consumption and misinformation engagement,” *Online Information Review*, vol. 46, no. 6, pp. 1152–1166, 2022.

[104] C. Pretus, C. Servin-Barthet, E. A. Harris, W. J. Brady, O. Vilarroya, and J. J. Van Bavel, “The role of political devotion in sharing partisan misinformation and resistance to fact-checking,” *Journal of Experimental Psychology: General*, vol. 152, no. 11, pp. 3116–3134, 2023.

[105] M. Stadel and G. Stulp, “Balancing bias and burden in personal network studies,” *Social Networks*, vol. 70, pp. 16–24, 2022.

[106] S. E. Kreps and D. L. Kriner, “Assessing misinformation recall and accuracy perceptions: Evidence from the covid-19 pandemic,” *Harvard Kennedy School (HKS) Misinformation Review*, 2023.

[107] S. M. Smallpage, A. M. Enders, H. Drochon, and J. E. Uscinski, “The impact of social desirability bias on conspiracy belief measurement across cultures,” *Political Science Research and Methods*, vol. 11, no. 3, pp. 555–569, 2023.

[108] A. Luiten, J. J. C. M. Hox, and E. D. De Leeuw, “Survey nonresponse trends and fieldwork effort in the 21st century: Results of an international study across countries and surveys,” *Journal of Official Statistics*, vol. 36, no. 3, pp. 469–487, 2020.

[109] T. Araujo, A. Wonneberger, P. Neijens, and C. D. Vreese, “How much time do you spend online? understanding and improving the accuracy of self-reported measures of internet use,” *Communication Methods and Measures*, vol. 11, no. 3, pp. 173–190, 2017.

[110] D. A. Parry, B. I. Davidson, C. J. R. Sewall, J. T. Fisher, H. Mieczkowski, and D. S. Quintana, "A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use," *Nature Human Behaviour*, vol. 5, no. 11, 2021.

[111] M. Scharkow, "The accuracy of self-reported internet use—a validation study using client log data," *Communication Methods and Measures*, vol. 10, no. 1, pp. 13–27, 2016.

[112] R. L. Winkler and A. H. Murphy, "Experiments in the laboratory and the real world," *Organizational Behavior and Human Performance*, vol. 10, no. 2, pp. 252–270, 1973.

[113] J. Allen, B. Howland, M. Mobius, D. Rothschild, and D. J. Watts, "Evaluating the fake news problem at the scale of the information ecosystem," *Science Advances*, vol. 6, Apr. 2020.

[114] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer, "Fake news on twitter during the 2016 u.s. presidential election," *Science*, vol. 363, no. 6425, pp. 374–378, 2019.

[115] M. Mosleh, G. Pennycook, and D. G. Rand, "Self-reported willingness to share political news articles in online surveys correlates with actual sharing on twitter," *PLoS ONE*, vol. 15, no. 2, p. e0228882, 2020.

[116] G. Ceylan, I. A. Anderson, and W. Wood, "Sharing of misinformation is habitual, not just lazy or biased," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 120, no. 4, p. e2216614120, 2023.

[117] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "The spreading of misinformation online," *Proceedings of the National Academy of Sciences*, vol. 113, no. 3, pp. 554–559, 2016.

[118] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[119] Y. Sun and J. Xie, "Who shares misinformation on social media? a meta-analysis of individual traits related to misinformation sharing," *Computers in Human Behavior*, vol. 158, p. 108271, 2024.

[120] M. A. Lawson, S. Anand, and H. Kakkar, "Tribalism and tribulations: The social costs of not sharing fake news," *Journal of Experimental Psychology: General*, vol. 152, no. 3, pp. 611–631, 2023.

[121] J. Howison, A. Wiggins, and K. Crowston, "Validity issues in the use of social network analysis with digital trace data," *Journal of the Association for Information Systems*, vol. 12, no. 12, pp. 767–797, 2011.

[122] B. Reeves, N. Ram, T. N. Robinson, J. J. Cummings, C. L. Giles, J. Pan, A. Chiatti, M. Cho, K. Roehrick, X. Yang, A. Gagneja, M. Brinberg, D. Muise, Y. Lu, M. Luo, A. Fitzgerald, and L. Yeykelis, "Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them," *Human–Computer Interaction*, vol. 36, no. 2, pp. 150–201, 2021.

[123] S. Stier, J. Breuer, P. Siegers, and K. Thorson, "Post-api research: On the ethics, methodology, and politics of social media data scraping," in *Computational Communication Research*, vol. 2, pp. 1–24, 2020.

[124] D. M. J. Lazer, A. Pentland, D. J. Watts, S. Aral, S. Athey, N. Contractor, D. Freelon, S. Gonzalez-Bailon, G. King, H. Margetts, A. Nelson, M. J. Salganik, M. Strohmaier, A. Vespignani, and C. Wagner, "Meaningful measures of human society in the twenty-first century," *Nature*, vol. 595, pp. 189–196, 2021.

[125] M. J. Salganik, *Bit by Bit: Social Research in the Digital Age*. Princeton University Press, 2017.

[126] C. H. De Vreese, M. Boukes, A. Schuck, R. Vliegenthart, L. Bos, and Y. Lelkes, "Linking survey and media content data: Opportunities, considerations, and pitfalls," *Communication Methods and Measures*, vol. 11, no. 4, pp. 221–244, 2017.

[127] H. Allcott, L. Braghieri, S. Eichmeyer, and M. Gentzkow, "The welfare effects of social media," *The American Economic Review*, vol. 110, no. 3, pp. 629–676, 2020.

[128] C. Wagner, M. Strohmaier, A. Olteanu, E. Kıcıman, N. Contractor, and T. Eliassi-Rad, "Measuring algorithmically infused societies," *Nature*, vol. 595, no. 7866, pp. 197–204, 2021.

[129] Statista, "Daily social media usage worldwide." https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/, 2025. Accessed: 2025-10-14.

[130] E. Union, "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)," 2016. Accessed: 2025-09-22.

[131] A. Bruns, "After the 'apicalypse': Social media platforms and their fight against critical scholarly research," *Information, Communication & Society*, vol. 22, no. 11, pp. 1544–1566, 2019.

[132] J. Breuer, Z. Kmetty, M. Haim, and S. Stier, "User-centric approaches for collecting facebook data in the 'post-api age': Experiences from two studies and recommendations

for future research," *Information, Communication & Society*, vol. 26, no. 14, pp. 2649–2668, 2023.

[133] D. Freelon, "Computational research in the post-api age," *Political Communication*, vol. 35, no. 4, pp. 665–668, 2018.

[134] E. Blakey, "The day data transparency died: How twitter/x cut off access for social research," *Contexts*, vol. 23, no. 2, pp. 30–35, 2024. Original work published 2024.

[135] D. Parry, "Restrictions on data access impede crucial societal research." University World News, May 2024. https://www.universityworldnews.com/post.php?story=20240509000643443.

[136] A. Nagaraj, E. Shears, and M. de Vaan, "Improving data access democratizes and diversifies science," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 38, pp. 23490–23498, 2020.

[137] F. Keusch, B. Struminskaya, C. Antoun, M. P. Couper, and F. Kreuter, "Willingness to participate in passive mobile data collection," *Public Opinion Quarterly*, vol. 83, no. S1, pp. 210–235, 2019.

[138] H. Silber, J. Breuer, C. Beuthner, T. Gummer, F. Keusch, P. Siegers, S. Stier, and B. Weiss, "Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behaviour," *Journal of the Royal Statistical Society Series A: Statistics in Society*, vol. 185, no. 2, pp. 387–407, 2022.

[139] L. Boeschoten, J. Ausloos, J. E. Möller, T. Araujo, and D. L. Oberski, "A framework for privacy preserving digital trace data collection through data donation," *Computational Communication Research*, vol. 4, no. 2, pp. 388–423, 2022. Published online: 01 Oct 2022.

[140] J. Breuer, L. Bishop, and K. Kinder-Kurlanda, "The practical and ethical challenges in acquiring and sharing digital trace data: Negotiating public-private partnerships," *New Media & Society*, vol. 22, no. 11, pp. 2058–2080, 2020.

[141] N. Pfiffner and T. N. Friemel, "Leveraging data donations for communication research: Exploring drivers behind the willingness to donate," *Communication Methods and Measures*, vol. 17, no. 3, pp. 227–249, 2023.

[142] B. Struminskaya, V. Toepoel, P. Lugtig, M. Haan, A. Luiten, and B. Schouten, "Understanding willingness to share smartphone-sensor data," *Public Opinion Quarterly*, vol. 84, no. 3, pp. 725–759, 2020.

[143] S. Morosoli, P. van Aelst, E. Humprecht, A. Staender, and F. Esser, "Identifying the drivers behind the dissemination of online misinformation: A study on political attitudes and individual characteristics," *American Behavioral Scientist*, 2022.

[144] M. Pérez-Escolar, D. Lilleker, and A. Tapia-Frade, "A systematic literature review of the phenomenon of disinformation and misinformation," *Media and Communication*, 2023.

[145] M. Hameleers and S. Minihold, "Constructing discourses on (un)truthfulness: Attributions of reality, misinformation, and disinformation by politicians in a comparative social media setting," *Communication Research*, 2020.

[146] J. Schulte-Cloos and V. Anghel, "Right-wing authoritarian attitudes, fast-paced decision-making, and the spread of misinformation about covid-19 vaccines," *Political Communication*, 2023.

[147] S. Mihelj, K. Kondor, and V. Štětka, "Establishing trust in experts during a crisis: Expert trustworthiness and media use during the covid-19 pandemic," *Science Communication*, 2022.

[148] T. A. Neyazi, A. Y. K. Ng, O. Kuru, and B. Muhtadi, "Who gets exposed to political misinformation in a hybrid media environment? the case of the 2019 indonesian election," *Social Media + Society*, 2022.

[149] L. Faragó, P. Krekó, and G. Orosz, "Hungarian, lazy, and biased: The role of analytic thinking and partisanship in fake news discernment on a hungarian representative sample," *Scientific Reports*, 2023.

[150] G. Orosz, L. Faragó, B. Paskuj, Z. Rakovics, D. Sam-Mine, G. Audemard, M. S. Modeliar, and P. Krekó, "Softly empowering a prosocial expert in the family: Lasting effects of a counter-misinformation intervention in an informational autocracy," *Scientific Reports*, 2024.

[151] Freedom House, "Nations in transit 2023: Hungary," 2023. URL: https://freedomhouse.org.

[152] Varieties of Democracy (V-Dem) Project, "V-dem democracy report 2023," 2023. URL: https://v-dem.net.

[153] J. Bayer, B. Holznagel, K. Lubianiec, A. Pintea, J. B. Schmitt, J. Szakács, and E. Uszkiewicz, "Disinformation and propaganda: Impact on the functioning of the rule of law and democratic processes in the eu and its member states," April 2021. ISBN (PDF): 978-92-846-7990-4, ISBN (Print): 978-92-846-7989-8, Catalogue number (PDF): QA-05-21-091-EN-N, Catalogue number (Print): QA-05-21-091-EN-C.

[154] M. Demeter, "Propaganda against the west in the heart of europe: A masked official state campaign in hungary," *Central European Journal of Communication*, vol. 11, no. 21, pp. 177–197, 2018.

[155] Z. Szebeni, I. Jasinskaja-Lahti, J.-E. Lönnqvist, and Z. Szabó, "The price of (dis)trust – profiling believers of (dis)information in the hungarian context," *Social Influence*, 2023.

[156] A. Woodward, "'fake news': A guide to trump's favourite phrase – and the dangers it obscures," *The Independent*, October 2020. Accessed: 2025-10-12.

[157] Vlad Makszimov, "Hungarian pm orbán accuses epp of spreading 'fake news'," 2020. Accessed: 2025-10-12.

[158] E. K. Vraga and L. Bode, "Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation," *Political Communication*, vol. 37, no. 1, pp. 136–144, 2020.

[159] C. Wardle, "The need for smarter definitions and practical, timely debates about information disorder," *Digital Journalism*, vol. 6, no. 8, pp. 951–963, 2018.

[160] C. Wardle and H. Derakhshan, "Information disorder: Toward an interdisciplinary framework for research and policymaking," 2017.

[161] N. Walter and S. T. Murphy, "How to counter misinformation: A taxonomy of anti-misinformation interventions," *Journal of Communication*, vol. 68, no. 2, pp. 233–255, 2018.

[162] S. Chen, L. Xiao, and A. Kumar, "Spread of misinformation on social media: What contributes to it and how to combat it," *Computers in Human Behavior*, vol. 141, p. 107643, 2023.

[163] E. Ferrara, S. Cresci, and L. Luceri, "Misinformation, manipulation, and abuse on social media in the era of covid-19," *Journal of Computational Social Science*, vol. 3, no. 2, pp. 271–277, 2020.

[164] S. van der Linden and Y. Kyrychenko, "A broader view of misinformation reveals potential for intervention," *Science*, vol. 384, pp. 959–960, 2024.

[165] G. Pennycook, Z. Epstein, M. Mosleh, A. A. Arechar, D. Eckles, and D. G. Rand, "Shifting attention to accuracy can reduce misinformation online," *Nature*, vol. 592, pp. 590–595, 2021.

[166] O. D. Apuke and B. Omar, "Fake news and covid-19: modelling the predictors of fake news sharing among social media users," *Telematics and Informatics*, vol. 56, p. 101475, 2021.

[167] O. D. Apuke, B. Omar, E. A. Tunca, and C. V. Gever, “Information overload and misinformation sharing behaviour of social media users: Testing the moderating role of cognitive ability,” *Journal of Information Science*, vol. 0, no. 0, pp. 1–17, 2022.

[168] A. A. Arechar, J. Allen, A. J. Berinsky, R. Cole, Z. Epstein, K. Garimella, A. Gully, J. G. Lu, R. M. Ross, M. N. Stagnaro, Y. Zhang, G. Pennycook, and D. G. Rand, “Understanding and combatting misinformation across 16 countries on six continents,” *Nature Human Behaviour*, vol. 7, pp. 1502–1513, 2023.

[169] A. M. Guess, M. Lerner, B. Lyons, J. M. Montgomery, B. Nyhan, J. Reifler, and N. Sircar, “A digital media literacy intervention increases discernment between mainstream and false news in the united states and india,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 27, pp. 15536–15545, 2020.

[170] I. P. Fainmesser and A. Galeotti, “The market for online influence,” *American Economic Journal: Microeconomics*, vol. 13, no. 4, pp. 332–372, 2021.

[171] J. Rozenbeek, S. Van Der Linden, B. Goldberg, and et al., “Inoculating against misinformation,” *Science Advances*, vol. 8, no. 34, 2022.

[172] J. A. Tucker, A. Guess, P. Barberá, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, and B. Nyhan, “Social media, political polarization, and political disinformation: A review of the scientific literature,” *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*, 2018.

[173] B. T. Truong, O. M. Allen, and F. Menczer, “Account credibility inference based on news-sharing networks,” *EPJ Data Science*, vol. 13, no. 10, pp. 1–19, 2024.

[174] A. M. Enders, J. E. Uscinski, M. I. Seelig, C. A. Klofstad, S. Wuchty, J. R. Funchion, M. N. Murthi, K. Premaratne, and J. Stoler, “The relationship between social media use and beliefs in conspiracy theories and misinformation,” *Political Behavior*, vol. 45, no. 2, pp. 781–804, 2023.

[175] G. Pennycook and D. G. Rand, “Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning,” *Cognition*, vol. 188, pp. 39–50, 2019.

[176] M. Mosleh, C. Martel, D. Eckles, and D. G. Rand, “Shared partisanship dramatically increases social tie formation in a twitter field experiment,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 7, p. e2022761118, 2021.

[177] T. Kelleher and K. D. Sweetser, “Social media adoption among university communicators,” *Journal of Public Relations Research*, vol. 24, no. 2, pp. 105–122, 2012.

[178] S. Rathje, J. He, J. Roozenbeek, J. J. Van Bavel, and S. van der Linden, "Social media behavior is associated with vaccine hesitancy," *PNAS Nexus*, vol. 1, no. 4, pp. 1–11, 2022.

[179] Media Bias Fact Check. https://mediabiasfactcheck.com/, 2023.

[180] M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media," *Proceedings of the National Academy of Sciences*, vol. 118, no. 9, p. e2023301118, 2021.

[181] J. Flamino, A. Galeazzi, S. Feldman, *et al.*, "Political polarization of news media and influencers on twitter in the 2016 and 2020 us presidential elections," *Nature Human Behaviour*, vol. 7, pp. 904–916, 2023.

[182] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini, "Online human-bot interactions: Detection, estimation, and characterization," *arXiv preprint arXiv:1703.03107*, 2017.

[183] Twitter Help Center, "Legacy verification policy." https://help.twitter.com/en/managing-your-account/legacy-verification-policy. Accessed: 2024-07-08.

[184] S. Wojcik and A. Hughes, "Sizing up twitter users." https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/, April 2019. Accessed: 2024-07-08.

[185] Z. Gilani, R. Farahbakhsh, G. Tyson, L. Wang, and J. Crowcroft, "An in-depth characterisation of bots and humans on twitter," *arXiv preprint arXiv:1704.01508*, 2017.

[186] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, "Botornot: A system to evaluate social bots," in *Proceedings of the 25th international conference companion on World Wide Web*, pp. 273–274, ACM, 2016.

[187] I. Pozzana and E. Ferrara, "Measuring bot and human behavioral dynamics," *Frontiers in Physics*, vol. 8, p. 125, 2020. Section: Social Physics.

[188] C. Mood, "Logistic regression: Why we cannot do what we think we can do, and what we can do about it," *European Sociological Review*, vol. 26, no. 1, pp. 67–82, 2010.

[189] E. M. Rogers, *Diffusion of Innovations*. New York: Free Press, 5th ed., 2003.

[190] S. Stier, J. Breuer, P. Siegers, and K. Thorson, "Integrating survey data and digital trace data: Key issues in developing an emerging field," *Social Science Computer Review*, vol. 38, no. 5, pp. 503–516, 2020.

[191] T. Araujo, A. Wonneberger, P. Neijens, and C. de Vreese, "How much time do you spend online? understanding and improving the accuracy of self-reported measures of internet use," *Communication Methods and Measures*, vol. 11, no. 3, pp. 173–190, 2017.

[192] J. Ohme, T. Araujo, C. H. de Vreese, and J. T. Piotrowski, "Mobile data donations: Assessing self-report accuracy and sample biases with the ios screen time function," *Mobile Media & Communication*, vol. 9, no. 2, pp. 293–313, 2021.

[193] M. Scharkow, "The accuracy of self-reported internet use—a validation study using client log data," *Communication Methods and Measures*, vol. 10, no. 1, pp. 13–27, 2016.

[194] C. Christner, A. Urman, S. Adam, and M. Maier, "Automated tracking approaches for studying online media use: A critical review and recommendations," *Communication Methods and Measures*, vol. 16, no. 2, pp. 79–95, 2022.

[195] M. Mancosu and F. Vegetti, "What you can scrape and what is right to scrape: A proposal for a tool to collect public facebook data," *Social Media + Society*, vol. 6, no. 3, p. 2056305120940703, 2020.

[196] G. King and N. Persily, "A new model for industry–academic partnerships," *PS: Political Science & Politics*, vol. 53, no. 4, pp. 703–709, 2020.

[197] J. Ohme, T. Araujo, L. Boeschoten, D. Freelon, N. Ram, B. B. Reeves, and T. N. Robinson, "Digital trace data collection for social media effects research: Apis, data donation, and (screen) tracking," *Communication Methods and Measures*, pp. 1–18, 2023. Online first.

[198] A. Halavais, "Overcoming terms of service: A proposal for ethical distributed research," *Information, Communication & Society*, vol. 22, no. 11, pp. 1567–1581, 2019.

[199] S. E. Baumgartner, S. R. Sumter, V. Petkevič, and W. Wiradhany, "A novel ios data donation approach: Automatic processing, compliance, and reactivity in a longitudinal study," *Social Science Computer Review*, vol. 41, no. 4, pp. 1456–1472, 2023.

[200] Z. Kmetty and K. Bozsonyi, "Identifying depression-related behavior on facebook—an experimental study," *Social Sciences*, vol. 11, no. 3, pp. 1–19, 2022.

[201] Z. Kmetty and R. Németh, "Which is your favorite music genre? a validity comparison of facebook data and survey data," *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, vol. 154, no. 1, pp. 82–104, 2022.

[202] B. Struminskaya, P. Lugtig, V. Toepoel, B. Schouten, D. Giesen, and R. Dolmans, "Sharing data collected with smartphone sensors: Willingness, participation, and nonparticipation bias," *Public Opinion Quarterly*, vol. 85, no. S1, pp. 423–462, 2021.

[203] J. W. Thibaut and H. H. Kelley, *The social psychology of groups*. Routledge, 2017.

[204] D. A. Dillman, *Mail and web-based survey: The tailored design method*. Wiley, 2000.

[205] F. Keusch, "Why do people participate in web surveys? applying survey participation theory to internet survey data collection," *Management Review Quarterly*, vol. 65, no. 3, pp. 183–216, 2015.

[206] R. M. Groves and M. P. Couper, "A conceptual framework for survey participation," in *Nonresponse in Household Interview Surveys*, pp. 25–46, Wiley, 1998.

[207] R. M. Groves, E. Singer, and A. Corning, "Leverage-saliency theory of survey participation: Description and an illustration," *The Public Opinion Quarterly*, vol. 64, no. 3, pp. 299–308, 2000.

[208] A. S. Göritz, "Incentives in web studies: Methodological issues and a review," *International Journal of Internet Science*, vol. 1, no. 1, pp. 58–70, 2006.

[209] G.-C. Haas, F. Kreuter, F. Keusch, M. Trappmann, and S. Bähr, "Effects of incentives in smartphone data collection," in *Big data meets survey science*, pp. 387–414, Wiley, 2020.

[210] B. Ságvári, A. Gulyás, and J. Koltai, "Attitudes towards participation in a passive data collection experiment," *Sensors*, vol. 21, no. 18, p. 6085, 2021.

[211] M. Revilla, M. P. Couper, and C. Ochoa, "Willingness of online panelists to perform additional tasks," *Methods, Data, Analyses*, vol. 13, no. 2, pp. 223–252, 2019.

[212] A. Jäckle, J. Burton, M. P. Couper, and C. Lessof, "Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases," *Survey Research Methods*, vol. 13, no. 1, pp. 1–22, 2019.

[213] C. Beuthner, B. Weiß, H. Silber, F. Keusch, and J. Schröder, "Consent to data linkage for different data domains – the role of question order, question wording, and incentives," *International Journal of Social Research Methodology*, pp. 1–14, 2023. Online first.

[214] M. Bietz, K. Patrick, and C. Bloss, "Data donation as a model for citizen science health research," *Citizen Science: Theory and Practice*, vol. 4, no. 1, p. 1, 2019.

[215] P. J. Edwards, I. Roberts, M. J. Clarke, C. Diguiseppi, R. Wentz, I. Kwan, R. Cooper, L. M. Felix, and S. Pratap, "Methods to increase response to postal and electronic questionnaires," *The Cochrane Database of Systematic Reviews*, vol. 2009, no. 3, pp. 1465–1858, 2009.

[216] N. M. Bradburn, "Respondent burden," in *Proceedings of the American Statistical Association, survey research methods section*, pp. 35–40, 1978.

[217] L. Bioglio and R. G. Pensa, "Analysis and classification of privacy-sensitive content in social media posts," *EPJ Data Science*, vol. 11, no. 1, p. 522, 2022.

[218] A. Wenz, A. Jäckle, and M. P. Couper, "Willingness to use mobile technologies for data collection in a probability household panel," *Survey Research Methods*, vol. 13, no. 1, pp. 1–22, 2019.

[219] F. Kreuter, G.-C. Haas, F. Keusch, S. Bähr, and M. Trappmann, "Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent," *Social Science Computer Review*, vol. 38, no. 5, pp. 533–549, 2020.

[220] A. Elevelt, P. Lugtig, and V. Toepoel, "Doing a time use survey on smartphones only: What factors predict nonresponse at different stages of the survey process?," *Survey Research Methods*, vol. 13, no. 2, pp. 195–213, 2019.

[221] E. Brüggen and U. M. Dholakia, "Determinants of participation and response effort in web panel surveys," *Journal of Interactive Marketing*, vol. 24, no. 3, pp. 239–250, 2010.

[222] A. Cheng, G. Zamarro, and B. Orriens, "Personality as a predictor of unit nonresponse in an internet panel," *Sociological Methods & Research*, vol. 49, no. 3, p. 3, 2020.

[223] Z. Kmetty and Á. Stefkovics, "Assessing the effect of questionnaire design on unit and item-nonresponse: Evidence from an online experiment," *International Journal of Social Research Methodology*, vol. 25, no. 5, pp. 659–672, 2021.

[224] P. Lugtig, "Panel attrition: Separating stayers, fast attriters, gradual attriters, and lurkers," *Sociological Methods & Research*, vol. 43, no. 4, pp. 699–723, 2014.

[225] R. Pinter, "Willingness of online access panel members to participate in smartphone application-based research," in *Mobile Research Methods* (R. Pinter, D. Toninelli, and P. de Pedraza, eds.), pp. 141–156, Ubiquity Press, 2015.

[226] J. Mulder and M. de Bruijne, "Willingness of online respondents to participate in alternative modes of data collection," *Survey Practice*, vol. 12, no. 1, pp. 1–11, 2019.

[227] N. L. Muscanell and R. E. Guadagno, "Make new friends or keep the old: Gender and personality differences in social networking use," *Computers in Human Behavior*, vol. 28, no. 1, pp. 107–112, 2012.

[228] M. G. Hoy and G. Milne, "Gender differences in privacy-related measures for young adult facebook users," *Journal of Interactive Advertising*, vol. 10, no. 2, pp. 28–45, 2010.

[229] A. M. Strange, R. D. Enos, M. Hill, and A. Lakeman, "Online volunteer laboratories for human subject's research," *PLOS ONE*, vol. 14, no. 8, p. e0221676, 2019.

[230] B. Wheller, "Algdesign: Algorithmic experimental design," 2022. R package version 1.2.1.

[231] S. van Buuren, K. Groothuis-Oudshoorn, A. Robitzsch, G. Vink, L. Doove, and S. Jolani, "Package 'mice'," 2015. Computer software.

[232] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[233] G.-C. Haas, F. Kreuter, F. Keusch, M. Trappmann, and S. Bähr, "Effects of incentives in smartphone data collection," in *Big Data Meets Survey Science*, pp. 387–414, John Wiley & Sons, Ltd., 2020.

[234] F. Keusch, B. Struminskaya, C. Antoun, M. P. Couper, and F. Kreuter, "Willingness to participate in passive mobile data collection," *Public Opinion Quarterly*, vol. 83, no. S1, pp. 210–235, 2019.

[235] Z. Kmetty and Á. Stefkovics, "Validating a willingness to share measure of a vignette experiment using real-world behavioral data," *Scientific Reports*, vol. 15, p. 9319, 2025.

[236] A. Jäckle, J. Burton, M. P. Couper, and C. Lessof, "Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases," *Survey Research Methods*, vol. 13, no. 1, 2019.

[237] F. Keusch, P. K. Pankowska, A. Cernat, and R. L. Bach, "Do you have two minutes to talk about your data? willingness to participate and nonparticipation bias in facebook data donation," *Field Methods*, pp. 1–19, 2023.

[238] P. Sheeran and T. L. Webb, "The intention–behavior gap," *Social and Personality Psychology Compass*, vol. 10, no. 9, pp. 503–518, 2016.

[239] Y. Kim, S. H. Kim, R. A. Peterson, and J. Choi, "Privacy concern and its consequences: A meta-analysis," *Technological Forecasting and Social Change*, vol. 196, p. 122789, 2023.

[240] J. Breuer, Z. Kmetty, M. Haim, and S. Stier, "User-centric approaches for collecting facebook data in the post-api age: Experiences from two studies and recommendations for future research," *Information, Communication & Society*, pp. 1–20, 2022.

[241] G. Pennycook and D. G. Rand, "The psychology of fake news," *Trends in Cognitive Sciences*, vol. 25, no. 5, pp. 388–402, 2021.

[242] E. Aïmeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: a review," *Social Network Analysis and Mining*, vol. 13, no. 1, p. 30, 2023.

[243] FreedomHouse, "Nations in transit 2024." https://freedomhouse.org/report/nations-transit/2024/region-reordered-autocracy-and-democracy, 2024. Accessed: 2025-07-11.

[244] G. Polyák, Á. Urbán, P. Szávai, and K. Horváth, "Disinformation under the guise of democracy: Lessons from hungary," in *State-Sponsored Disinformation Around the Globe: How Politicians Deceive Their Citizens* (M. Echeverría, S. G. Santamaría, and D. C. Hallin, eds.), Routledge Studies in Media, Communication, and Politics, New York and Abingdon, UK: Routledge, 2025.

[245] K. Eom, H. S. Kim, and D. K. Sherman, "Social class, control, and action: Socioeconomic status differences in antecedents of support for pro-environmental action," *Journal of Experimental Social Psychology*, vol. 77, pp. 60–75, 2018.

[246] Y. Wang, M. McKee, A. Torbica, and D. Stuckler, "Systematic literature review on the spread of health-related misinformation on social media," *Social Science Medicine*, vol. 240, p. 112552, 2019.

[247] D. A. Scheufele and N. M. Krause, "Science audiences, misinformation, and fake news," *Proceedings of the National Academy of Sciences*, vol. 116, no. 16, pp. 7662–7669, 2019.

[248] K. P. Arin, J. A. Lacomba, F. Lagos, D. Mazrekaj, and M. Thum, "Misperceptions and fake news during the covid-19 pandemic," Tech. Rep. 9066, CESifo Working Paper Series, 2021. Available at SSRN: https://ssrn.com/abstract=3842330.

[249] A. Chadwick and C. Vaccari, "The impact of online news consumption on misinformation and knowledge," *New Media & Society*, vol. 21, no. 7, pp. 1480–1497, 2019.

[250] A. Baqir, A. Galeazzi, and F. Zollo, "News and misinformation consumption: A temporal comparison across european countries," *PLoS ONE*, vol. 19, no. 5, p. e0302473, 2024.

[251] G. Rampersad and T. Althiyabi, "Fake news: Acceptance by demographics and culture on social media," *Journal of Information Technology Politics*, vol. 17, no. 1, pp. 1–11, 2020.

[252] U. K. H. Ecker, S. Lewandowsky, J. Cook, *et al.*, "The psychological drivers of misinformation belief and its resistance to correction," *Nature Reviews Psychology*, vol. 1, pp. 13–29, 2022.

[253] R. Rygula, M. Piksa, and A. Cieslik-Starkiewicz, "Cognitive, psychological, and pharmacological correlates of susceptibility to (mis)information: Why we believe and how we resist," *Neuroscience & Biobehavioral Reviews*, vol. 177, p. 106339, 2025.

[254] Y. Kyrychenko, H. J. Koo, R. Maertens, J. Roozenbeek, S. van der Linden, and F. M. Götz, "Profiling misinformation susceptibility," *Personality and Individual Differences*, vol. 241, p. 113177, 2025.

[255] M. Barthel, A. Mitchell, and J. Holcomb, "Many americans believe fake news is sowing confusion," 2016. https://www.pewresearch.org/journalism/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/.

[256] M. Sultan, A. N. Tump, N. Ehmann, P. Lorenz-Spreen, R. Hertwig, A. Gollwitzer, and R. H. J. M. Kurvers, "Susceptibility to online misinformation: A systematic meta-analysis of demographic and psychological factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 121, no. 47, p. e2409329121, 2024.

[257] Statista, "Leading mobile social media apps in hungary in january 2024, by reach." https://www.statista.com/statistics/1231772/hungary-leading-mobile-social-media/, 2024. Accessed: 2025-06-30.

[258] Statista, "Social media usage in Hungary – statistics & facts." https://www.statista.com/topics/6592/social-media-usage-in-hungary/#topicOverview, 2024. Accessed: 2025-09-04.

[259] J. Koltai, Z. Rakovics, Z. Kmetty, *et al.*, "Classifying social position with social media behavioral data," *EPJ Data Science*, vol. 14, no. 60, 2025.

[260] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.

[261] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge: Cambridge University Press, 2008.

[262] N. Nguyen and R. Caruana, "Consensus clusterings," *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, pp. 607–612, 2007.

[263] L. Edelson, M.-K. Nguyen, I. Goldstein, O. Goga, D. McCoy, and T. Lauinger, "Understanding engagement with u.s. (mis)information news sources on facebook," in *Proceedings of the 21st ACM Internet Measurement Conference*, IMC '21, (New York, NY, USA), p. 444–463, Association for Computing Machinery, 2021.

[264] M. Glenski, T. Weninger, and S. Volkova, "Identifying and understanding user reactions to deceptive and trusted social news sources," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (I. Gurevych and Y. Miyao, eds.), (Melbourne, Australia), pp. 176–181, Association for Computational Linguistics, July 2018.

[265] B. D. Horne, J. Nørregaard, and S. Adalı, "Different spirals of sameness: A study of content sharing in mainstream and alternative media," in *Proceedings of the Thirteenth International AAAI Conference on Web and Social Media*, pp. 257–266, Association for the Advancement of Artificial Intelligence, 2019.

[266] M. Samory, V. K. Abnousi, and T. Mitra, "Characterizing the social media news sphere through user co-sharing practices," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, pp. 602–613, 2020.

[267] G. Weld, M. Glenski, and T. Althoff, "Political bias and factualness in news sharing across more than 100,000 online communities," in *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, vol. 15, pp. 796–807, 2021.

[268] C.-X. Feng, "A comparison of zero-inflated and hurdle models for modeling zero-inflated count data," *Journal of Statistical Distributions and Applications*, vol. 8, no. 8, pp. 1–19, 2021.

[269] D. McFadden, "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics* (P. Zarembka, ed.), pp. 105–142, Academic Press, 1974.

[270] S. Bhalla, R. Ray, and H. Taneja, "When news is entertainment: explaining the persistence of misinformation through the information environment," *Information, Communication & Society*, vol. 28, no. 12, pp. 2081–2100, 2024.

[271] A. Bessi, M. Coletto, G. A. Davidescu, A. Scala, G. Caldarelli, and W. Quattrociocchi, "Science vs conspiracy: Collective narratives in the age of misinformation," *PLoS ONE*, vol. 10, no. 2, p. e0118093, 2015.

[272] W. J. Brady, M. J. Crockett, and J. J. Van Bavel, "The mad model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online," *Perspectives on Psychological Science*, pp. 1–33, 2020.

[273] N. Newman, A. Ross Arguedas, C. T. Robertson, R. K. Nielsen, R. Fletcher, and J. Szakács, "Digital news report 2025: Hungary section," tech. rep., Reuters Institute for the Study of Journalism, 2025.

[274] J. Allen, D. J. Watts, and D. G. Rand, "Quantifying the impact of misinformation and vaccine-skeptical content on facebook," *Science*, vol. 384, no. 6699, p. eadk3451, 2024.

Appendix A

Supplementary Information for Study 1

Robustness checks

Distributions of factuality scores

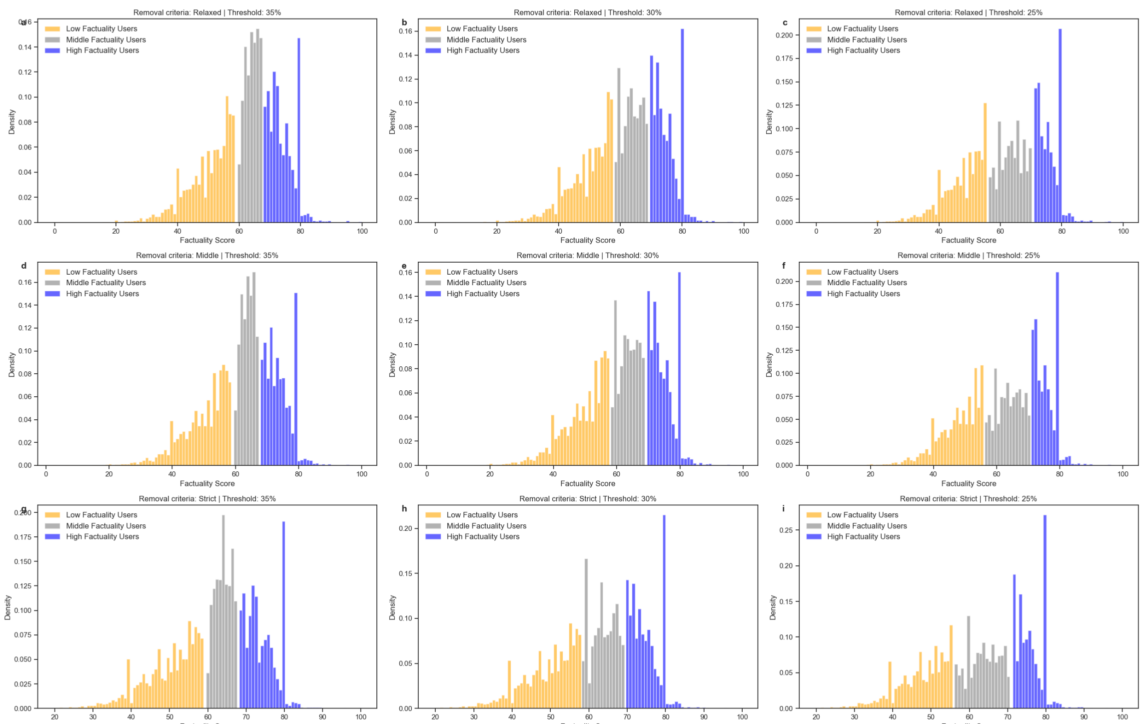


Figure A.1: **Distribution of factuality scores across the nine groups, defined by varying levels of regular user and bot filtering criteria (Relaxed, Middle, Strict) and low/high factuality group cutoffs (35/35, 30/30, 25/25).**

Distributions of follower count

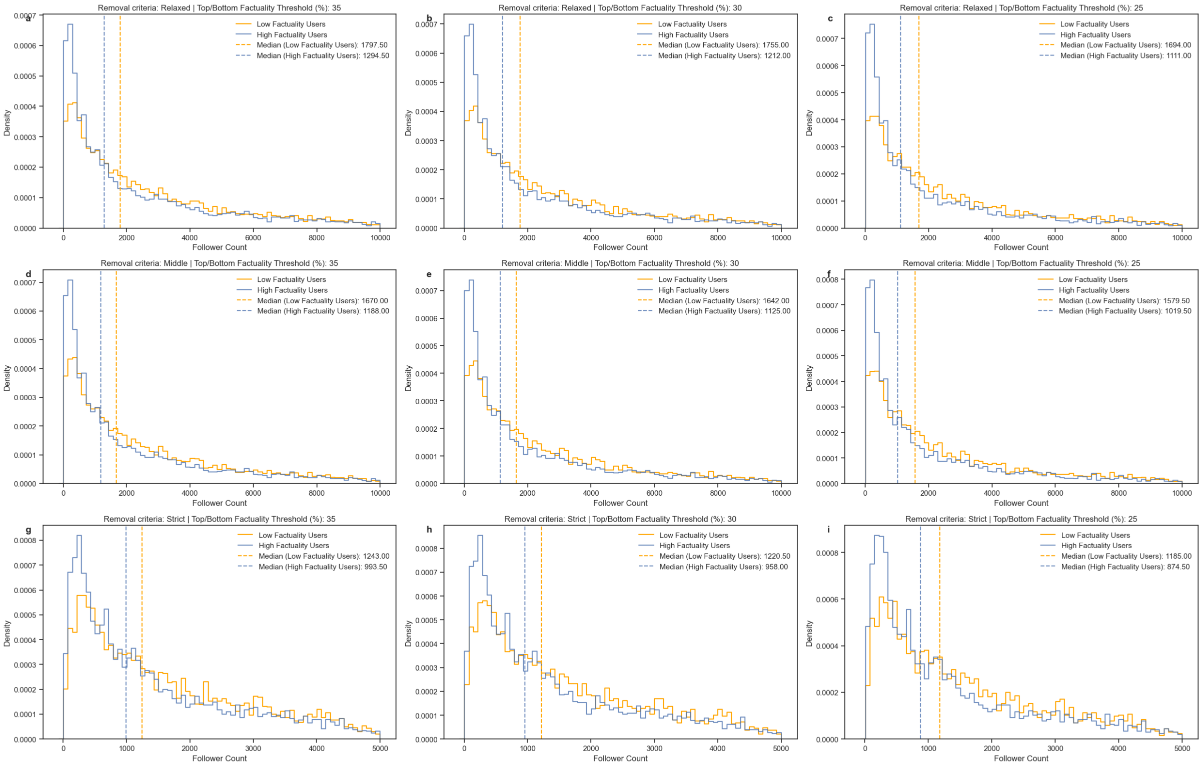


Figure A.2: **Low factuality users tend to have higher follower count than high factuality users.** Comparing the distributions of follower count between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly higher number of followers compared to high factuality users (p-value < 0.0001) in each of the tested datasets. The median value for low factuality users (orange dotted line) is higher than for high factuality users (blue dotted line) in each of the tested datasets.

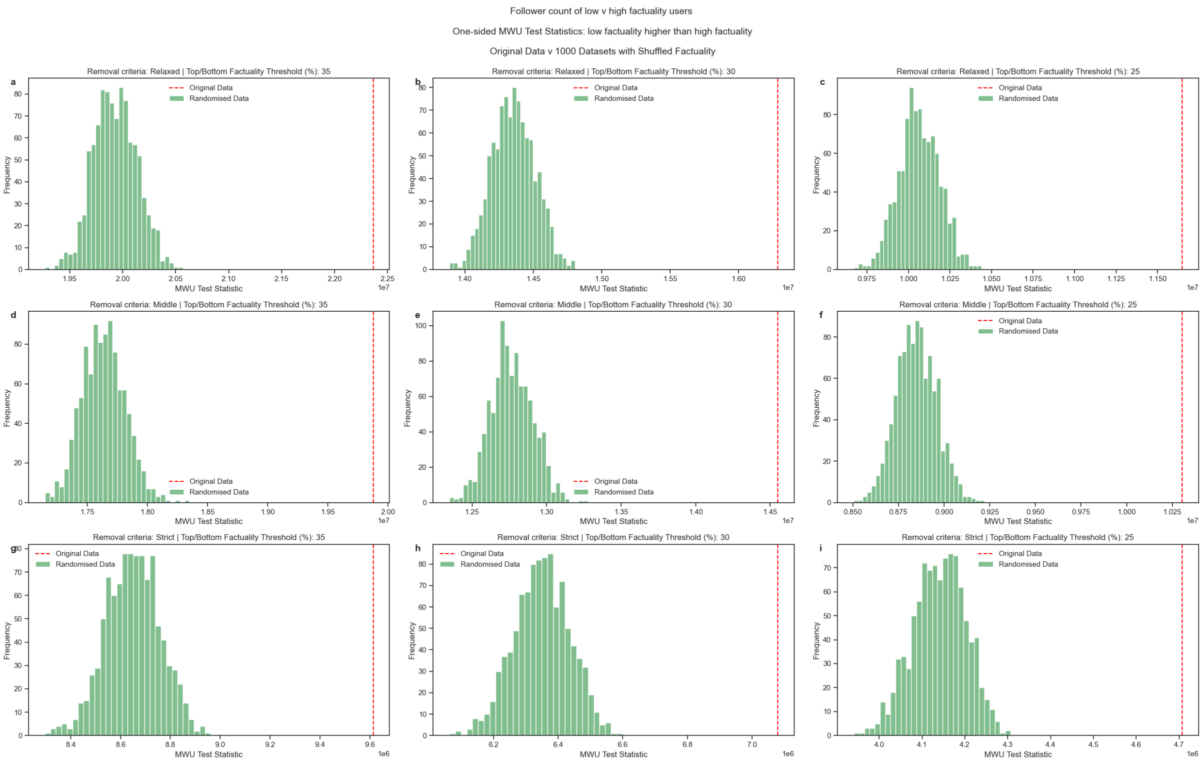


Figure A.3: **Low factuality users tend to have higher follower count than high factuality users.** The MWU test score obtained from the empirical data (red dotted line) is higher than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars) in each of the tested datasets.

Distributions of tweet count

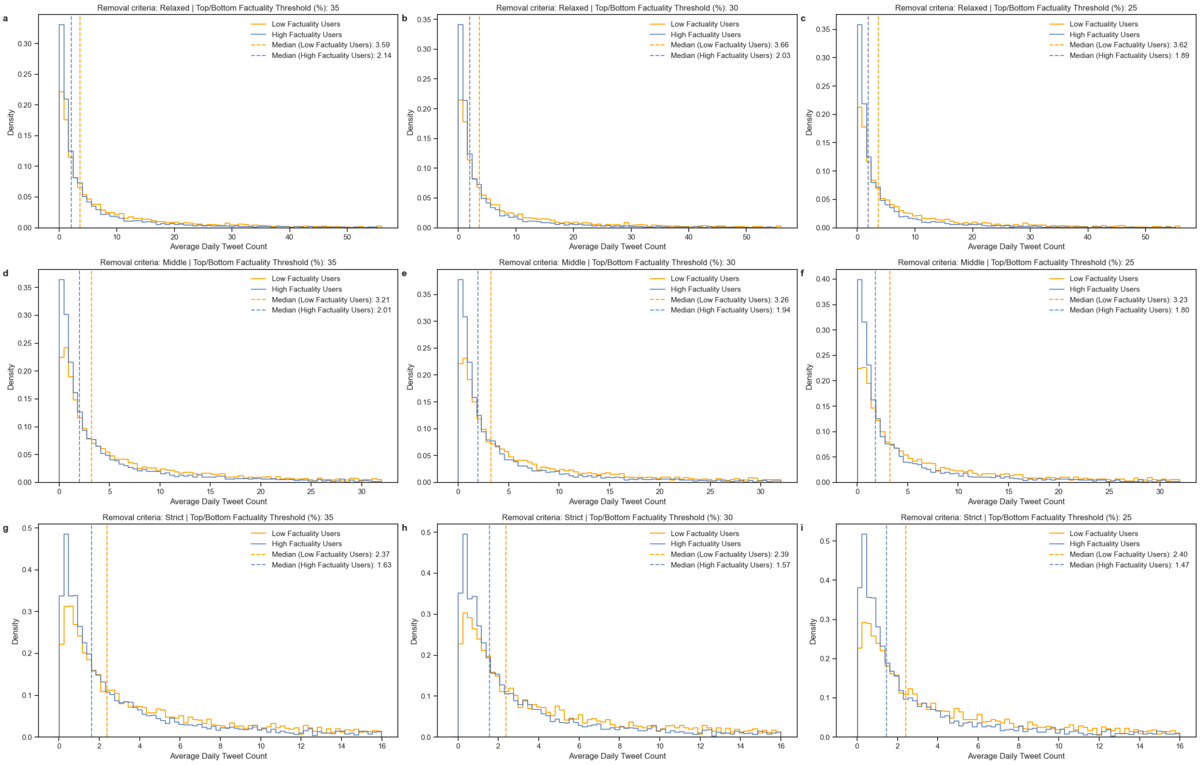


Figure A.4: **Low factuality users tend to have higher tweet count than high factuality users.** Comparing the distributions of tweet count between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly higher number of tweets compared to high factuality users (p-value < 0.0001) in each of the tested datasets. The median value for low factuality users (orange dotted line) is higher than for high factuality users (blue dotted line) in each of the tested datasets.

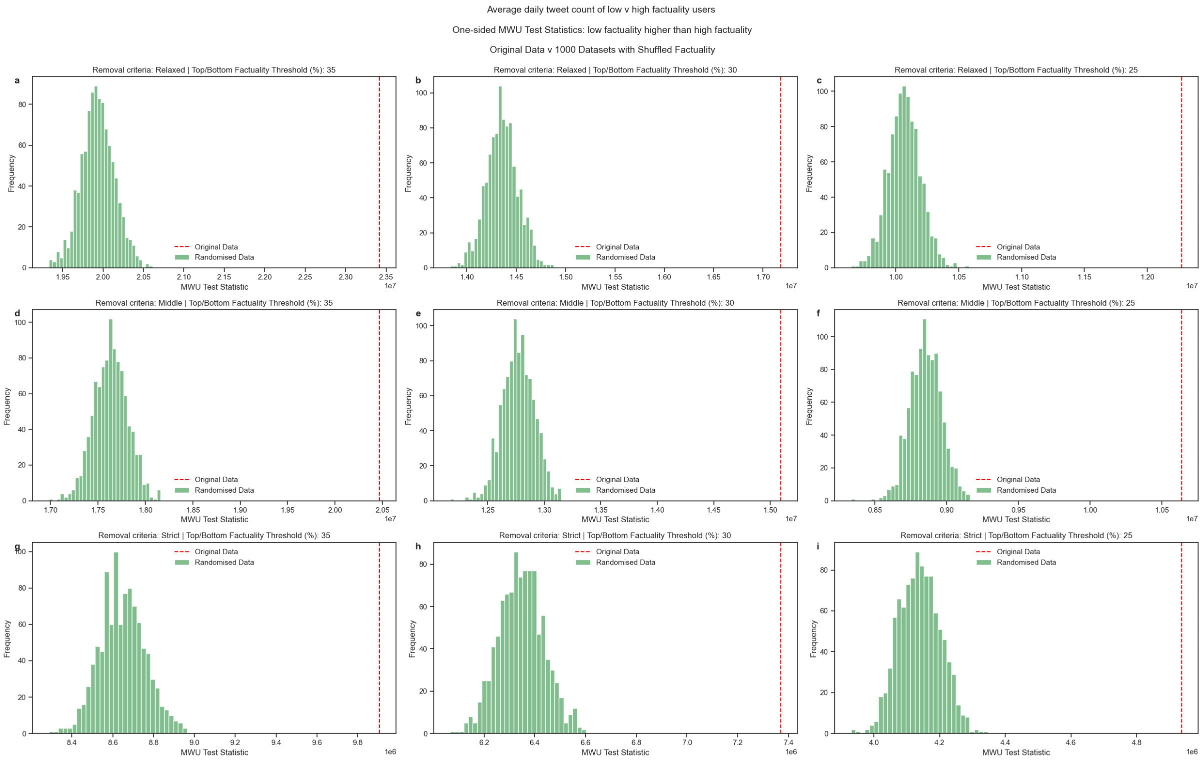


Figure A.5: **Low factuality users tend to have higher tweet count than high factuality users.** Panel b: The MWU test score obtained from the empirical data (red dotted line) is higher than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars) in each of the tested datasets.

Distributions of followed account count

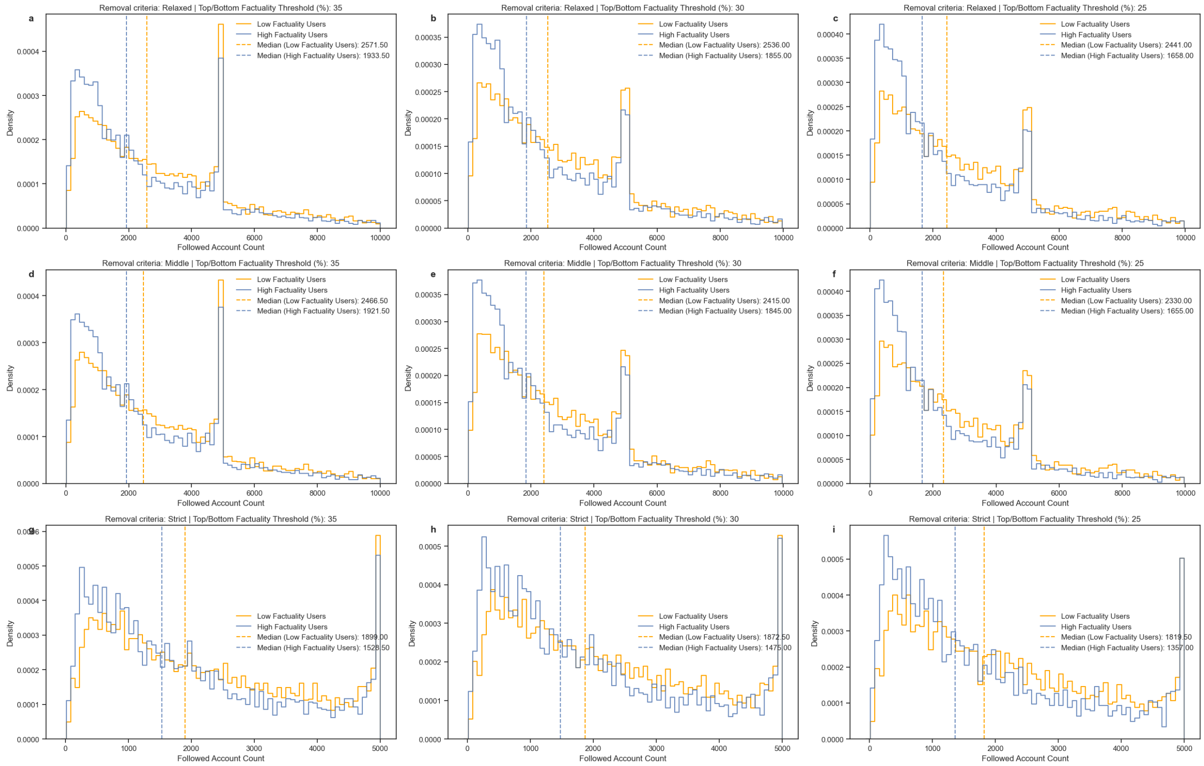


Figure A.6: **Low factuality users tend to have higher followed account count than high factuality users.** Comparing the distributions of followed account count between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly higher number of followed accounts compared to high factuality users (p-value $<$ 0.0001) in each of the tested datasets. The median value for low factuality users (orange dotted line) is higher than for high factuality users (blue dotted line) in each of the tested datasets. The peak around 5000 is due to an X policy that limits the number of new followed accounts until the user obtains more followers.

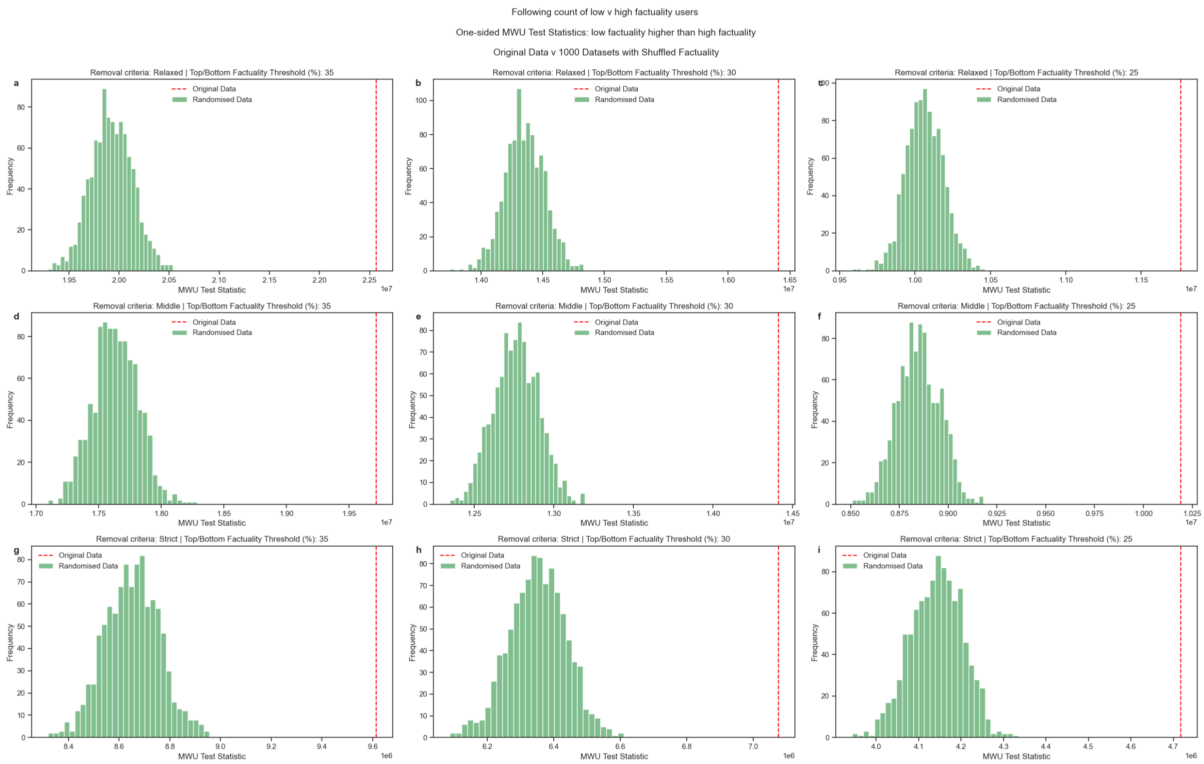


Figure A.7: **Low factuality users tend to have higher followed account count than high factuality users.** The MWU test score obtained from the empirical data (red dotted line) is higher than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars) in each of the tested datasets.

Distributions of account age

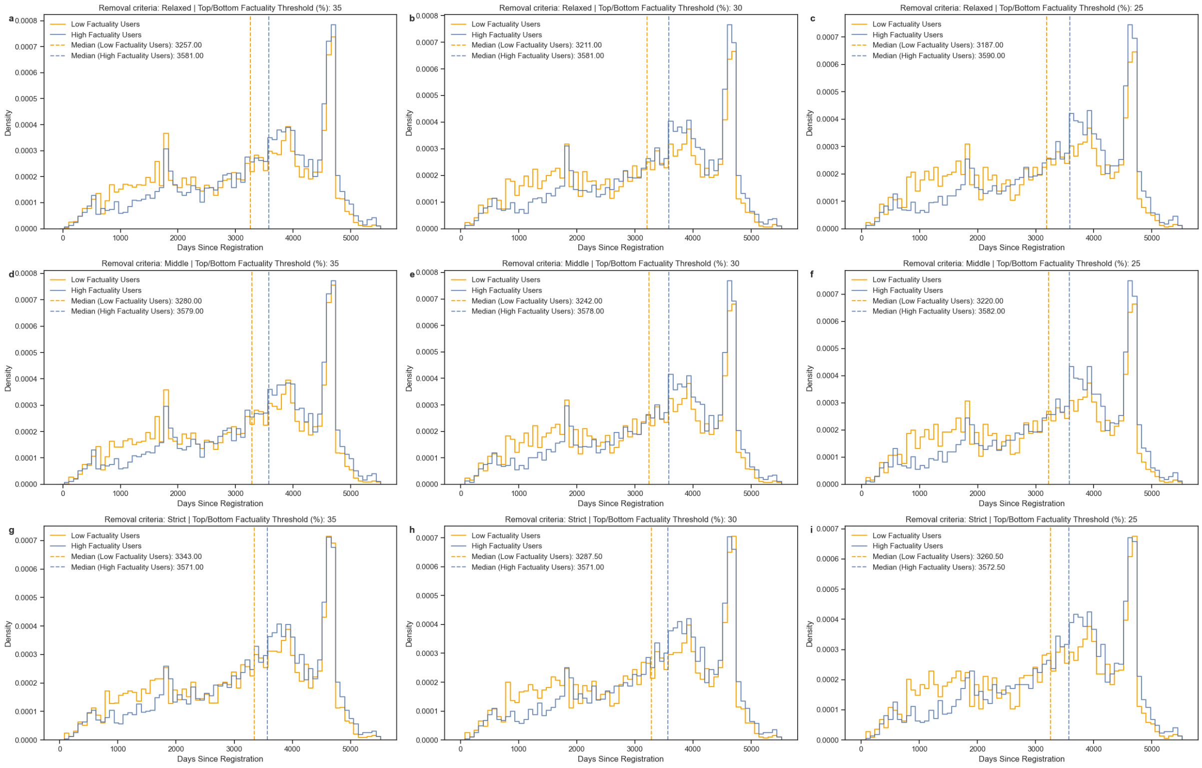


Figure A.8: **Low factuality users tend to have lower number of days since registration than high factuality users.** Comparing the distributions of the number of days since registration between low (orange curve) and high factuality (blue curve) users, we find that low factuality users have a significantly lower number of days compared to high factuality users (p-value $<$ 0.0001) in each of the tested datasets. The median value for low factuality users (orange dotted line) is lower than for high factuality users (blue dotted line) in each of the tested datasets.

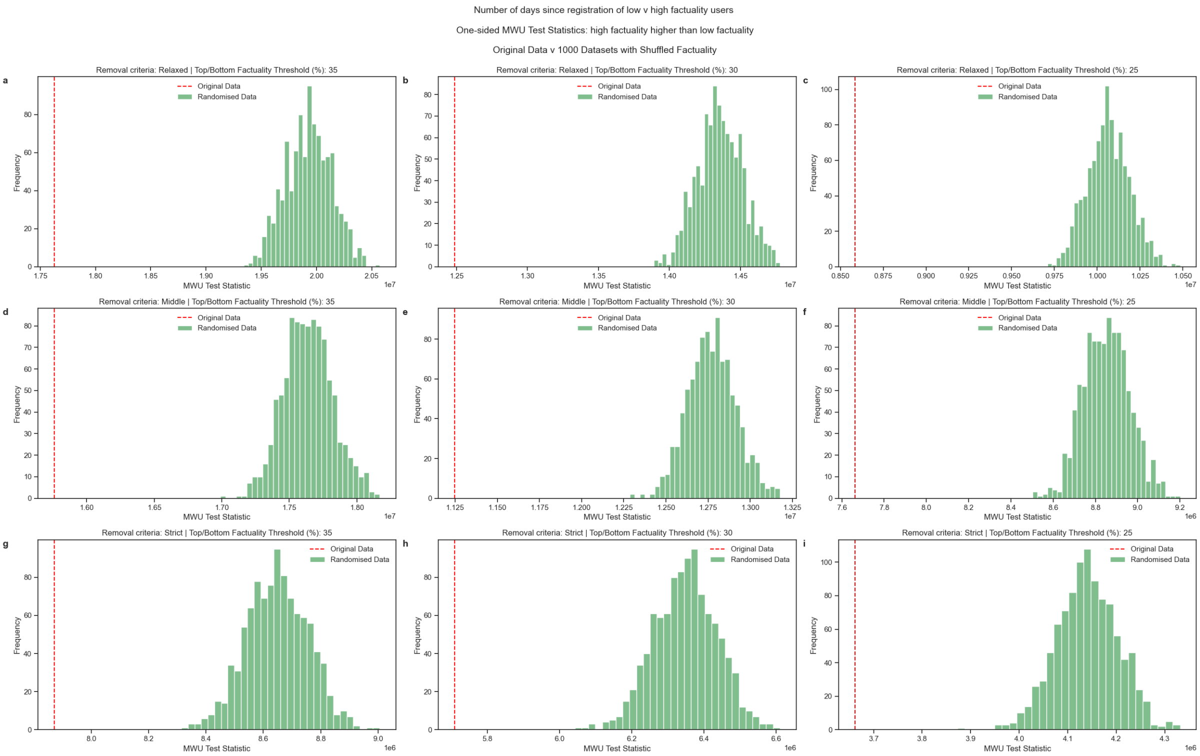


Figure A.9: **Low factuality users tend to have lower number of days since registration than high factuality users.** The MWU test score obtained from the empirical data (red dotted line) is lower than all the MWU test scores calculated on the 1,000 shuffled datasets (green bars) in each of the tested datasets.

Regression results

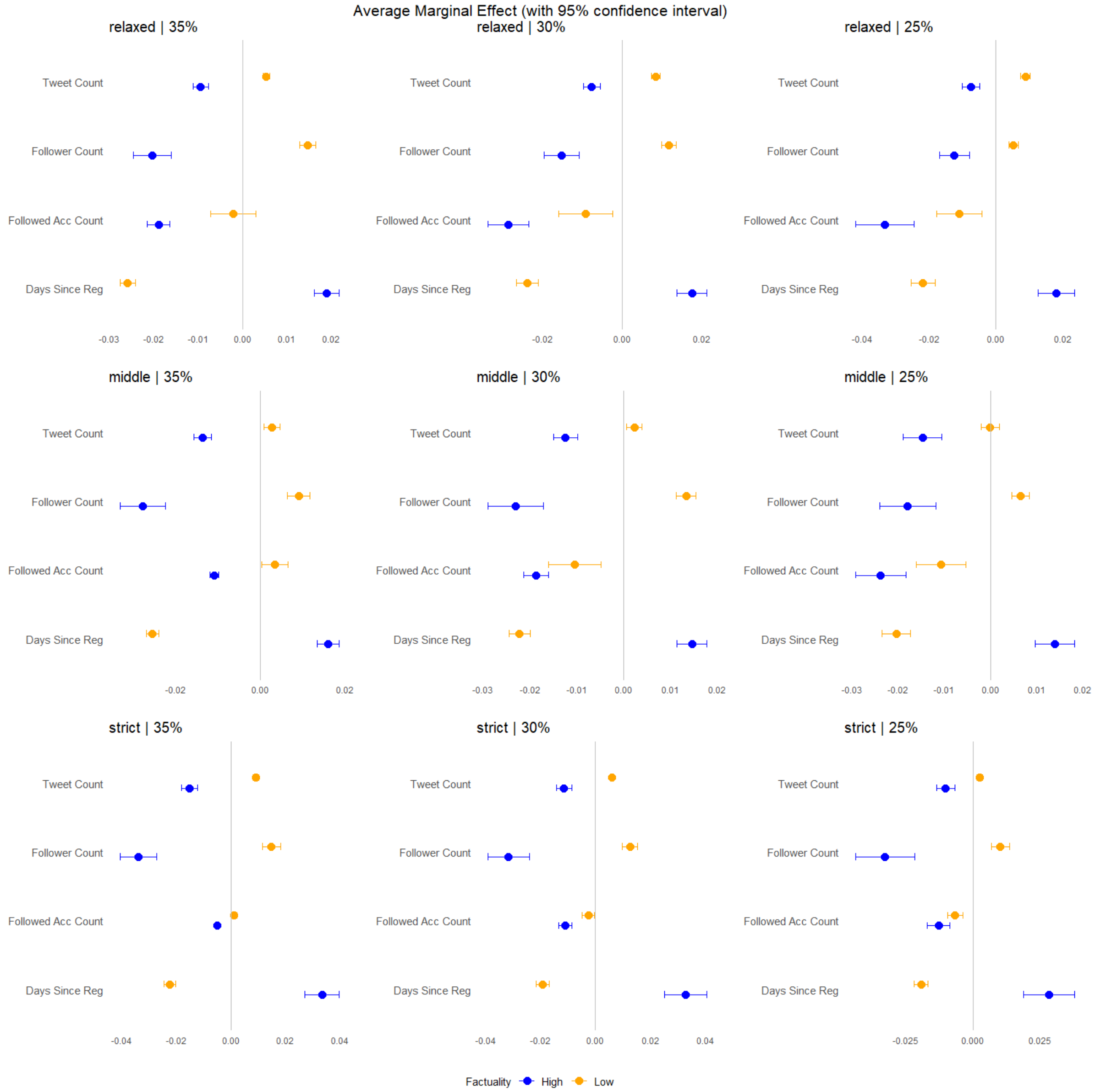


Figure A.10: Average marginal effects of social network characteristics on factuality for the 9 datasets for checking robustness of results.

Table A.1: Average Marginal Effects from Multinomial Logit Models

Specification	Category	Variable	AME	SE	p-value
Middle — 25%	High	Days since registration	0.014	0.002	0.000
Middle — 25%	High	Follower count	-0.018	0.003	0.000
Middle — 25%	High	Following count	-0.024	0.003	0.000
Middle — 25%	High	Tweets per day	-0.015	0.002	0.000
Middle — 25%	Low	Days since registration	-0.020	0.002	0.000
Middle — 25%	Low	Follower count	0.007	0.001	0.000
Middle — 25%	Low	Following count	-0.011	0.003	0.000
Middle — 25%	Low	Tweets per day	0.000	0.001	0.992
Middle — 30%	High	Days since registration	0.015	0.002	0.000
Middle — 30%	High	Follower count	-0.023	0.003	0.000
Middle — 30%	High	Following count	-0.019	0.001	0.000
Middle — 30%	High	Tweets per day	-0.012	0.001	0.000
Middle — 30%	Low	Days since registration	-0.022	0.001	0.000
Middle — 30%	Low	Follower count	0.013	0.001	0.000
Middle — 30%	Low	Following count	-0.010	0.003	0.000
Middle — 30%	Low	Tweets per day	0.002	0.001	0.006
Middle — 35%	High	Days since registration	0.016	0.001	0.000
Middle — 35%	High	Follower count	-0.028	0.003	0.000
Middle — 35%	High	Following count	-0.011	0.001	0.000
Middle — 35%	High	Tweets per day	-0.014	0.001	0.000
Middle — 35%	Low	Days since registration	-0.025	0.001	0.000
Middle — 35%	Low	Follower count	0.009	0.001	0.000
Middle — 35%	Low	Following count	0.003	0.002	0.030
Middle — 35%	Low	Tweets per day	0.003	0.001	0.005
Relaxed — 25%	High	Days since registration	0.018	0.003	0.000
Relaxed — 25%	High	Follower count	-0.012	0.002	0.000
Relaxed — 25%	High	Following count	-0.033	0.004	0.000
Relaxed — 25%	High	Tweets per day	-0.008	0.001	0.000
Relaxed — 25%	Low	Days since registration	-0.022	0.002	0.000
Relaxed — 25%	Low	Follower count	0.005	0.001	0.000
Relaxed — 25%	Low	Following count	-0.011	0.003	0.002
Relaxed — 25%	Low	Tweets per day	0.009	0.001	0.000
Relaxed — 30%	High	Days since registration	0.018	0.002	0.000

Table A.1: Average Marginal Effects from Multinomial Logit Models *(continued)*

Specification	Category	Variable	AME	SE	p-value
Relaxed — 30%	High	Follower count	-0.015	0.002	0.000
Relaxed — 30%	High	Following count	-0.029	0.003	0.000
Relaxed — 30%	High	Tweets per day	-0.008	0.001	0.000
Relaxed — 30%	Low	Days since registration	-0.024	0.001	0.000
Relaxed — 30%	Low	Follower count	0.012	0.001	0.000
Relaxed — 30%	Low	Following count	-0.009	0.003	0.008
Relaxed — 30%	Low	Tweets per day	0.008	0.001	0.000
Relaxed — 35%	High	Days since registration	0.019	0.001	0.000
Relaxed — 35%	High	Follower count	-0.020	0.002	0.000
Relaxed — 35%	High	Following count	-0.019	0.001	0.000
Relaxed — 35%	High	Tweets per day	-0.009	0.001	0.000
Relaxed — 35%	Low	Days since registration	-0.026	0.001	0.000
Relaxed — 35%	Low	Follower count	0.015	0.001	0.000
Relaxed — 35%	Low	Following count	-0.002	0.003	0.423
Relaxed — 35%	Low	Tweets per day	0.005	0.000	0.000
Strict — 25%	High	Days since registration	0.028	0.005	0.000
Strict — 25%	High	Follower count	-0.033	0.006	0.000
Strict — 25%	High	Following count	-0.013	0.002	0.000
Strict — 25%	High	Tweets per day	-0.010	0.002	0.000
Strict — 25%	Low	Days since registration	-0.019	0.001	0.000
Strict — 25%	Low	Follower count	0.010	0.002	0.000
Strict — 25%	Low	Following count	-0.007	0.001	0.000
Strict — 25%	Low	Tweets per day	0.003	0.001	0.000
Strict — 30%	High	Days since registration	0.033	0.004	0.000
Strict — 30%	High	Follower count	-0.032	0.004	0.000
Strict — 30%	High	Following count	-0.011	0.001	0.000
Strict — 30%	High	Tweets per day	-0.011	0.001	0.000
Strict — 30%	Low	Days since registration	-0.019	0.001	0.000
Strict — 30%	Low	Follower count	0.013	0.001	0.000
Strict — 30%	Low	Following count	-0.003	0.001	0.029
Strict — 30%	Low	Tweets per day	0.006	0.000	0.000
Strict — 35%	High	Days since registration	0.034	0.003	0.000
Strict — 35%	High	Follower count	-0.034	0.003	0.000

Table A.1: Average Marginal Effects from Multinomial Logit Models *(continued)*

Specification	Category	Variable	AME	SE	p-value
Strict — 35%	High	Following count	-0.005	0.000	0.000
Strict — 35%	High	Tweets per day	-0.015	0.001	0.000
Strict — 35%	Low	Days since registration	-0.022	0.001	0.000
Strict — 35%	Low	Follower count	0.015	0.002	0.000
Strict — 35%	Low	Following count	0.001	0.000	0.010
Strict — 35%	Low	Tweets per day	0.009	0.001	0.000

Multicollinearity diagnostics

Because several account-level activity metrics are conceptually related, we assessed multicollinearity using pairwise correlations (Figure A.11) and variance inflation factors (Table A.2). While most variables exhibited moderate correlations, follower count and following count were strongly correlated (see Figure A.11), resulting in elevated VIF values for these two metrics.

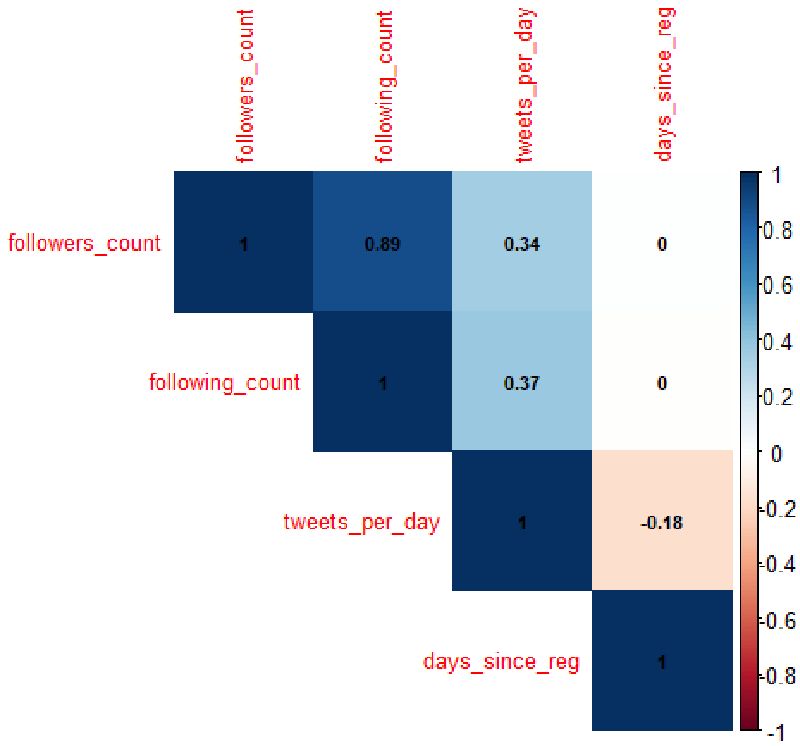


Figure A.11: Pairwise correlations between the four studied social network characteristics.

To assess whether this collinearity materially affects our substantive conclusions, we conducted additional robustness analyses in which we re-estimated the multinomial logistic regression models while excluding (i) followed account count and (ii) follower count, respectively. Figures A.12 and A.13 report the average marginal effects (AMEs) under these alternative

Table A.2: Variance Inflation Factors (VIF) for the Four Studied Social Network Characteristics

Term	**VIF**	**VIF 95% CI**	**Increased SE**	**Tolerance**	**Tolerance 95% CI**
Low Correlation					
Tweets per day	3.69	[3.60, 3.79]	1.92	0.27	[0.26, 0.28]
Days since registration	3.19	[3.11, 3.27]	1.79	0.31	[0.31, 0.32]
High Correlation					
Follower count	13.46	[13.07, 13.86]	3.67	0.07	[0.07, 0.08]
Following count	13.97	[13.56, 14.38]	3.74	0.07	[0.07, 0.07]

model specifications across all dataset constructions and factuality thresholds.

Across all specifications, the direction, magnitude, and statistical significance of the remaining coefficients—most notably tweet frequency and days since registration—remain highly stable. Importantly, the substantive conclusions regarding systematic behavioural differences between users who predominantly share higher- versus lower-factuality sources are unchanged. This indicates that the observed associations are not driven by multicollinearity between follower and following counts.

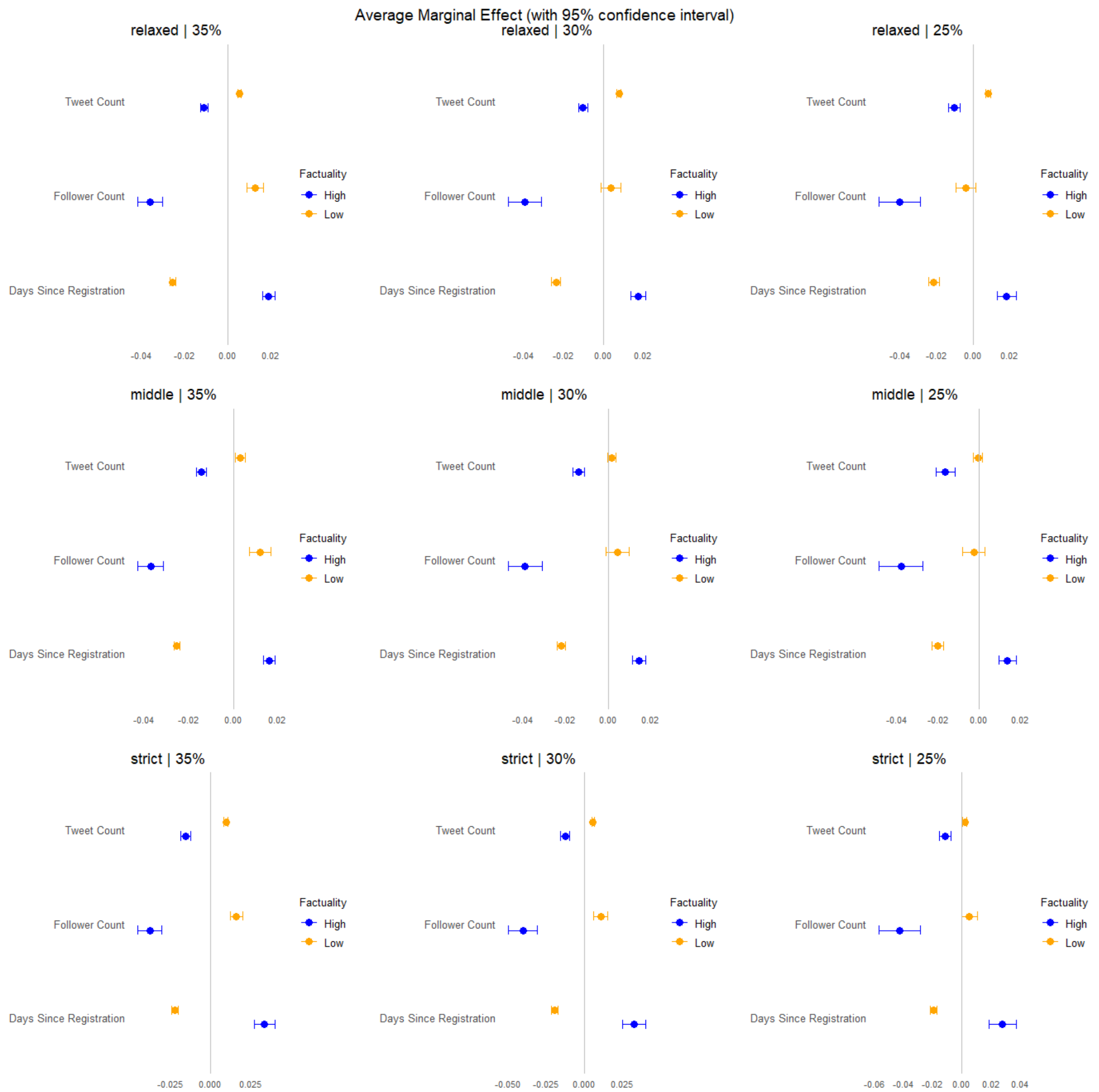


Figure A.12: Average marginal effects from multinomial logistic regression excluding followed account count (to address collinearity with follower count) - all specifications.

Interactions

Through plotting the interactions of follower count with all other metrics, certain differences in the relationships between these metrics and factuality levels across users with varying follower counts have been observed. Specifically, our analysis reveals that in case of users with fewer followers the effect of the followed account count diminishes on high factuality, whereas those with a larger follower base show an increased likelihood of high factuality with more followed accounts (Figure 11, panel a). Thus, the effect of followed account count on high factuality is nonexistent among those with smaller follower base, while positive among those with larger follower base. For users in the low factuality group, we observe a similar trend across both

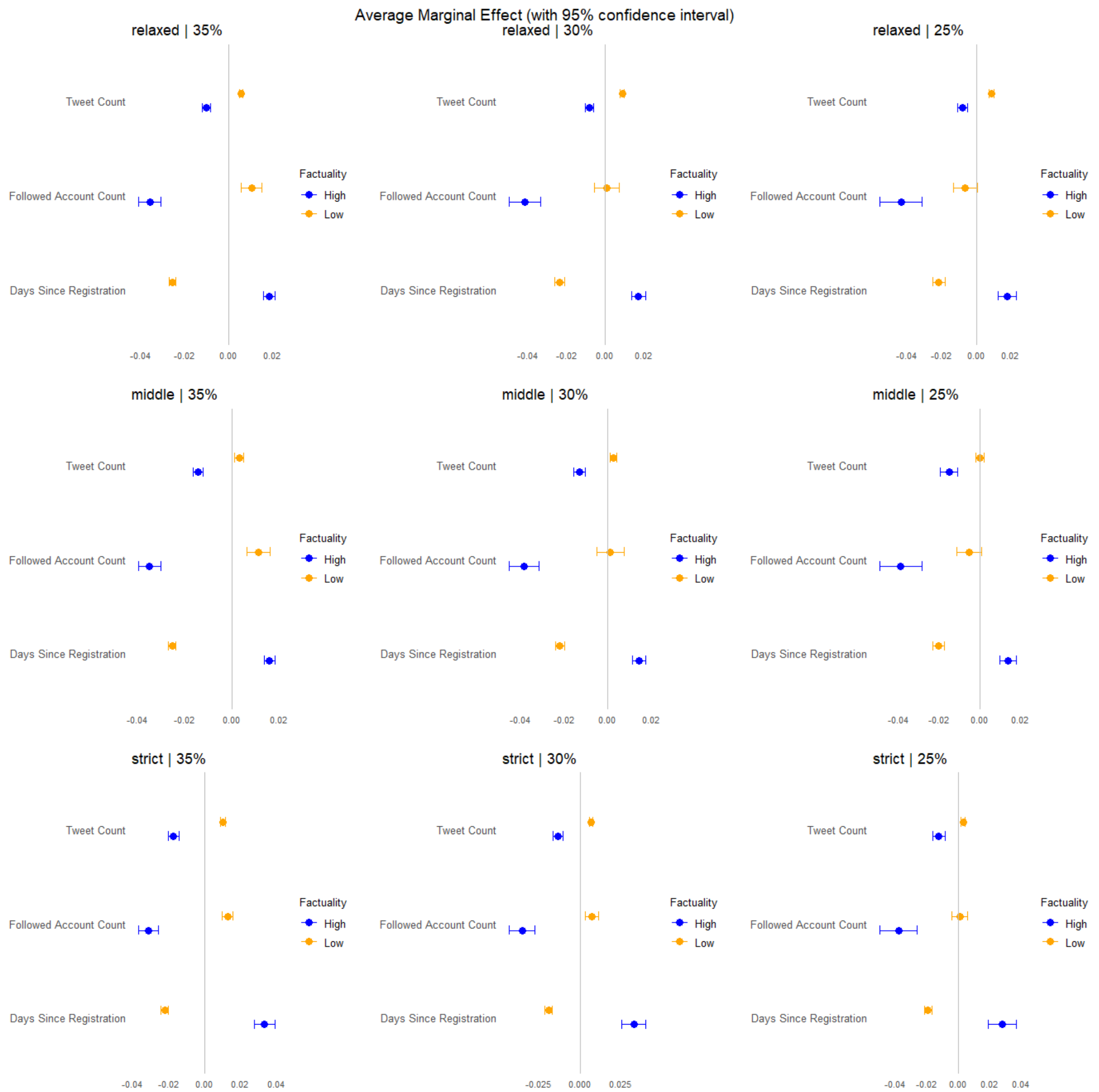


Figure A.13: Average marginal effects from multinomial logistic regression excluding follower count (to address collinearity with follower count) - all specifications.

follower count categories: as users follow more accounts, they become less likely to belong to the low factuality group. This effect is slightly stronger for users with a higher follower count. Among those with fewer followed accounts, popular users are more likely to have low factuality, while in case of those who follow many accounts, there is no significant difference in their likelihood of low factuality by popularity.

In relation to average daily tweet count our analysis (Figure 11, panel b) shows that as users increase their tweet activity, they are less likely to belong to the high factuality category and more likely to belong to the low factuality category, regardless of their follower count. Nevertheless, in the low factuality group, this relationship is less pronounced among users with fewer followers.

Regarding the relationship between factuality and days since registration (Figure 11, panel c) we show that as users spend more time registered on the platform, they are more likely to belong to the high factuality category and less likely to belong to the low factuality category, and this is consistent in direction and magnitude across users, regardless of their follower count. However, the positive effect on high factuality is slightly stronger among users with higher follower counts. For the low factuality group, we observe a negative relationship with days since registration, with a marginally stronger effect among those with more followers. It shows that among younger users popular ones are more likely to have lower factuality compared to less popular ones. However, among older users there is no difference in the likelihood of factuality between more and less popular users.

In examining average daily tweet count across different levels of followed account count and its association with factuality levels (Figure 11, panel d) we find that among those with few accounts followed the average tweet count does not affect the likelihood of belonging to the high factuality group. However, an increase in tweet activity is associated with an increased likelihood of belonging to the high factuality group among those who follow more accounts. For the low factuality group, tweet count is negatively associated with low factuality and this effect is stronger among those who follow more accounts. There is no significant difference in the likelihood of belonging to the low factuality group among those who tweet infrequently, however, among those who tweet frequently, users following few accounts are way more likely to have low factuality than users who follow more accounts. Figure 11, panel d corresponds to Figure 6, panel a in the main text.

Examining the interaction between days since registration and tweet count in relation to factuality levels reveals consistent trends across users with different average daily tweet counts (Figure 11, panel e). Our analysis demonstrates that as the registration period lengthens, users are more likely to exhibit higher factuality levels, and less likely to exhibit lower factuality levels, irrespective of their tweet activity.

Our findings regarding the interaction between days since registration and followed account count (Figure 11, panel f) indicate that among users who follow fewer accounts, those who have been registered on the platform for a longer duration are more likely to be in the high factuality category, while among those who follow more accounts, the trend is the opposite: the longer they are on the platform, the less likely that they belong to the high factuality group. The relationship between days since registration and low factuality does not vary based on followed account count. The relationship is consistently negative between days since registration and low factuality, indicating that longer registration periods are associated with a decreased likelihood of low factuality, however, this effect is somewhat more pronounced among users who follow fewer accounts. Figure 11, panel f corresponds to Figure 6, panel b in the main text.

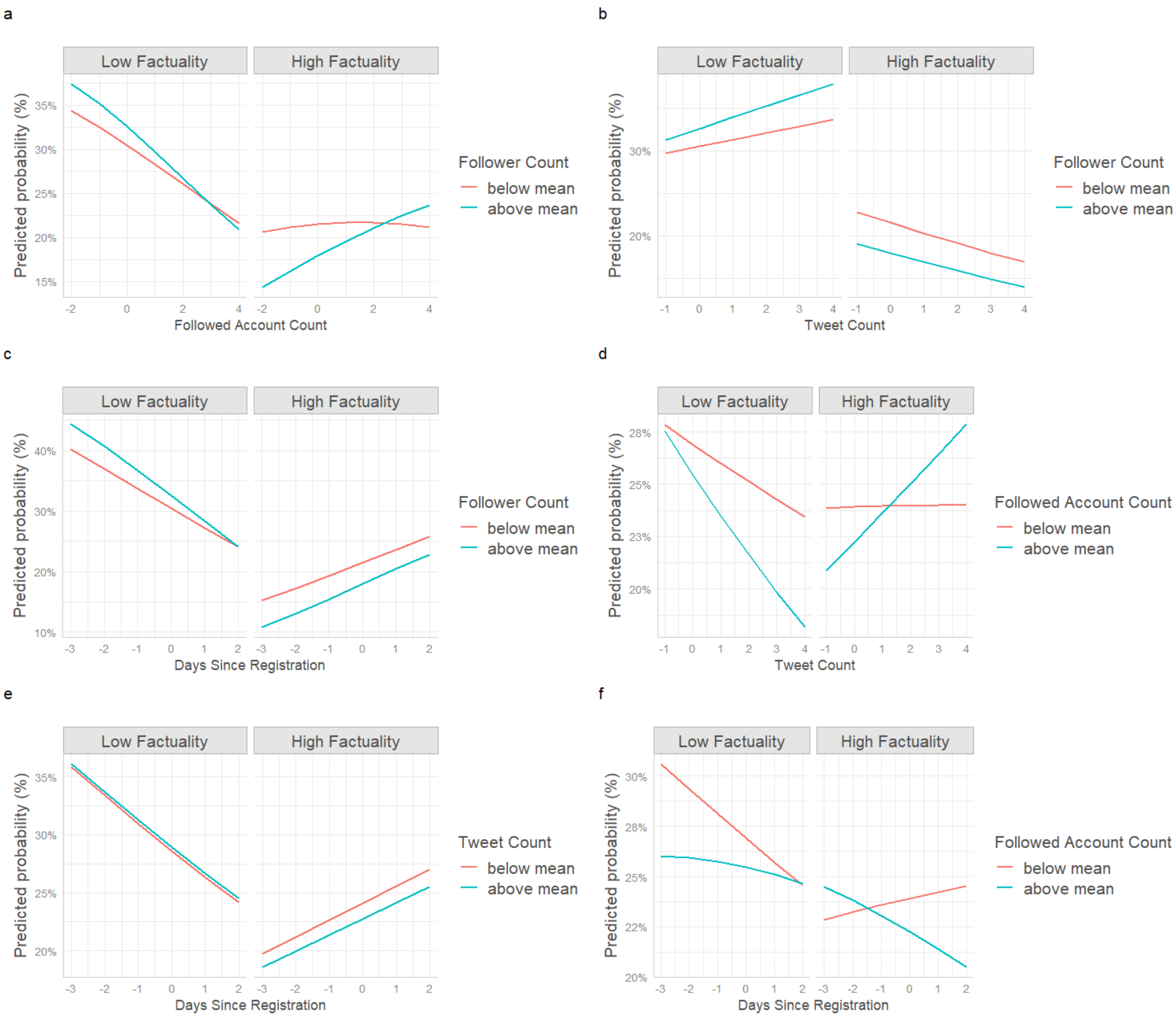


Figure A.14: Interactions between various social network metrics in relation to belonging to low and high factuality groups.

Pseudo-R (McFaddens R)

Pseudo-R (McFaddens R) obtained from shuffled datasets against those derived from the real data to demonstrate that the Pseudo-R obtained from our regression exceeds that expected by random chance.

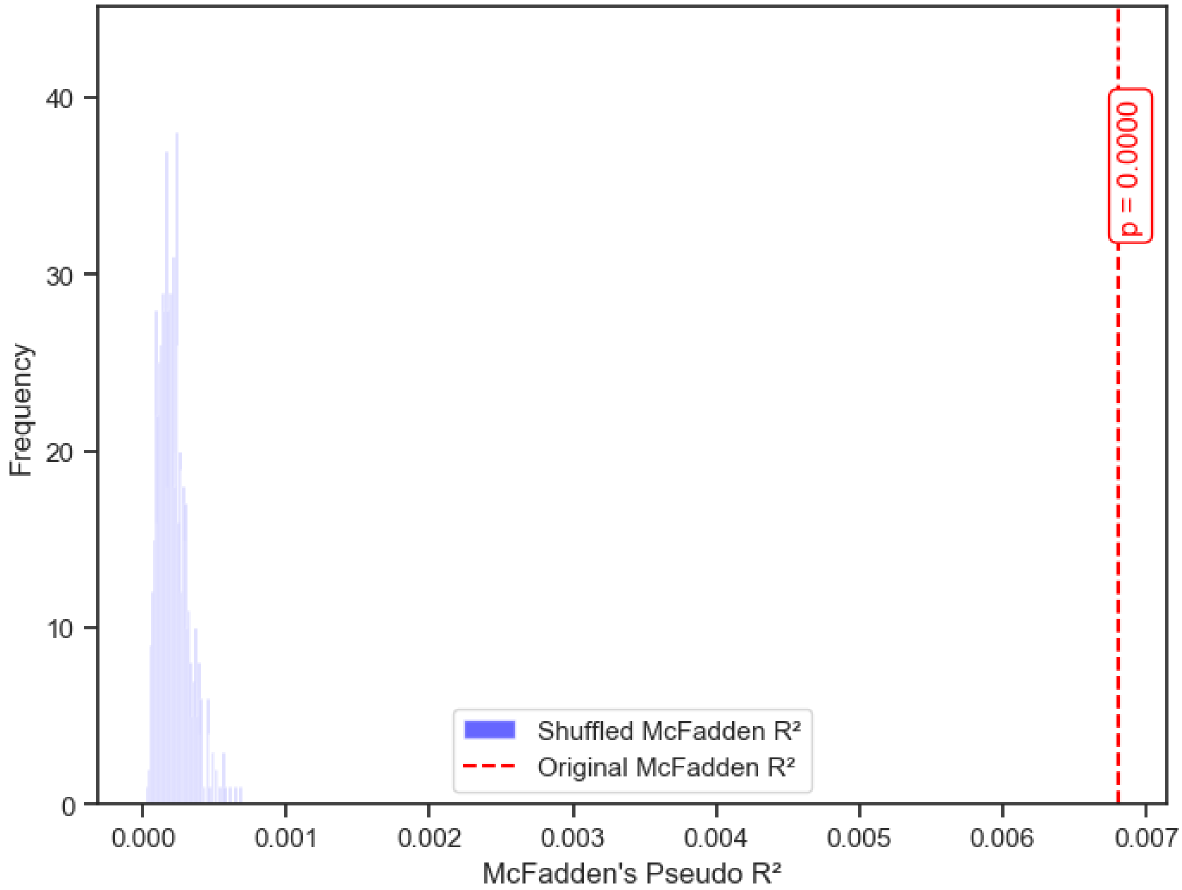


Figure A.15: Pseudo-R (McFaddens R) using the four social network metrics on the real dataset (red dotted line) compared to the accuracy scores obtained on the 1000 randomised datasets (blue bars).

Appendix B

Supplementary Information for Study 2

A. Detailed operationalization of independent variables

Gender is a binary variable that assigns 0 to males and 1 to females. In the U.S. sample, 7 percent of respondents said they belonged to “Other” gender category or did not answer the gender question. We imputed data of these respondents in the main models, but we also run an alternative model where the Other and DK/NA categories are combined and included as a third category in the analysis. We measured the highest level of education of respondent with six categories: from “primary” to “university diploma” in Hungary and from “some high school” to “Ph.D.” in the US. We operationalized age with the year of birth. Subjective wealth was measured with five categories, where the highest means that they live without financial problems, and the lowest means that they live in deprivation.

To control social media usage, we used two variables. The first variable was the frequency of social usage on a 1 to 5 scale, where the lowest means never, and the highest means daily. Here we calculated the maximum of Facebook, Twitter, Instagram, Youtube and TikTok usage. The second variable was the number of social media platforms, on which the respondent is active. Here we asked about the following platforms: Facebook, Instagram, Twitter, Youtube, Linkedin, TikTok, and Spotify.

To measure respondents’ privacy concerns, we used the Internet Users’ Information Privacy Concerns (IUIPC) scale (Malhotra et al., 2004). We applied a confirmatory factor model to extract the three latent dimensions behind the eight validated items. According to our analyses, the model with the three latent variables fit the data well (Hungary CFA:0.99, RMSEA: 0.068; U.S. CFA: 0.99, RMSEA: 0.038). Out of these three dimensions, in the analysis, we only used the ‘control’ and ‘collection’ dimensions of the scale and omitted the ‘awareness’ one, as it highly correlates with the ‘control’ in both samples. High values of these dimensions mean high control over personal information and concerns about the collection of personal data by companies.

For measuring privacy concerns, we calculated the principal component of the following two variables (measured on a 1 to 7 scale):

—"Most businesses handle the personal information they collect about consumers in a proper and confidential way."

—"Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today."

High values here mean high trust in how businesses and organizations protect consumer data.

To measure the respondents' affinity for technology, we used the 9-item version of the Affinity for Technology Interaction Scale (ATI – Franke et al. (2019). We calculated the mean of the items after reverse coding the needed items. The value of the Cronbach alpha was 0.84 in the Hungarian study and 0.89 in the U.S. study. High values here mean a high affinity for technology.

The last group of independent variables was the big five (BF) inventory (John et al., 2008): Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. We used the 15-item version of the scale and applied a confirmatory factor model to extract the dimensions. In both countries, the tested factor model fit the data (Hungary CFA: 0.99, RMSEA: 0.034; US: CFA: 0.938, RMSEA: 0.062). We had to drop out the reversed coded items from the models because of their poor fit in both the Hungarian and U.S. dataset.

B. Manipulated question

The bold parts were the manipulated part of the experiment.

„The various social media sites and platforms (Facebook, Instagram, Twitter, Google) allow users to view and even download information and data about themselves stored on the site. This data is very valuable from a scientific point of view since it captures behavioral patterns not observed elsewhere. Imagine a situation in which you are asked by the Social Science Research Centre to participate in a survey. You are invited to fill in a questionnaire, and you are asked to share your **Facebook** data, **excluding your private messages and pictures/videos**. Downloading the data to your computer and uploading it to the research page would take **less than 1 hour**. For participating in the research, you would receive **3000 HUF** and a **personalized report on your social media usage compared to the rest of the Hungarian internet population**. Once uploaded, the data would be anonymized and only analyzed for research purposes.

Please indicate on a scale of 0-10 how likely would you be to participate in such research! 0 indicates not likely at all, and 10 indicates very likely."

Table A1. Distribution of basic demographic variables

		Hungary	U.S.
Gender	Male	46,2%	57,6%
	Female	53,8%	35,4%
	Other DK/NA		7,0%
Age category	18-29	17,1%	2,9%
	30-44	30,5%	16,7%
	45-59	33,9%	21,3%
	60+	18,5%	59,1%
Education - Hungary	Lower than primary school	0,3%	
	Primary school	5,2%	
	Vocational school	24,1%	
	High school	40,6%	
	College	18,0%	
	University	11,8%	
Education - U.S.	High school graduate or lower		14,8%
	Associate's Degree		12,5%
	B.A. or B.S.		30,3%
	Master's		30,6%
	J.D./M.D.		2,4%
	Ph.D.		9,4%

Table A2

Results of the Vignette Experiment split by the first and second six vignettes – *Hungary (*multilevel mixed-effects linear regression*)*

	Vignettes 1-6			Vignettes 7-12		
(Intercept)	3.77	0.13	<0.001	3.17	0.14	<0.001
Incentive	0.24	0.01	<0.001	0.27	0.01	<0.001
Report	0.01	0.06	0.83	0.03	0.06	0.61
Platform: FB + Google	-0.32	0.06	<0.001	0.01	0.06	0.80
Platform: FB + other	-0.21	0.07	<0.001	-0.10	0.05	0.06
Platform: FB + Google + Other	-0.25	0.06	<0.001	-0.10	0.06	0.10
Time	-0.04	0.05	0.37	-0.09	0.04	0.02

	Vignettes 1-6			Vignettes 7-12		
Type of data: no PM/loc	-0.13	0.05	0.01	0.01	0.07	0.90
Type of data: no PM/vid	-0.21	0.07	<0.001	0.02	0.06	0.79
Type of data: no PM/loc/vid	-0.36	0.07	<0.001	0.03	0.06	0.58
Variances of random effects						
Variance: constant	11.08			11.44		
Variance: residual	2.76			2.13		
Proportion of Level 1 variance	19.9%			15.7%		
Proportion of Level 2 variance	80.1%			84.3%		
Model fit						
Variance explained (Level 1)	9.2%			11.6%		
Variance explained (overall)	1.7%			1.6%		

Table A3

Results of the Vignette Experiment in the subsample of Facebook users and Facebook and Google users– *Hungary (*multilevel mixed-effects linear regression

	Active Facebook users			Active Facebook and Google users		
(Intercept)	3.65	0.12	**<0.001**	3.67	0.13	**<0.001**
Incentive	0.25	0.01	**<0.001**	0.27	0.01	**<0.001**
Report	0.01	0.03	0.829	0.01	0.03	0.793
Platform: FB + Google	-0.20	0.04	**<0.001**	-0.24	0.05	**<0.001**
Platform: FB + other	-0.21	0.04	**<0.001**	-0.25	0.05	**<0.001**
Platform: FB + Google + Other	-0.18	0.04	**<0.001**	-0.21	0.05	**<0.001**
Time	-0.03	0.03	0.308	-0.02	0.03	0.536
Type of data: no PM/loc	-0.05	0.04	0.262	-0.05	0.05	0.305

	Active Facebook users			Active Facebook and Google users		
Type of data: no PM/vid	-0.21	0.04	**<0.001**	-0.20	0.05	**<0.001**
Type of data: no PM/loc/vid	-0.22	0.04	**<0.001**	-0.22	0.05	**<0.001**
Variances of random effects						
Variance: constant	11.03			11.37		
Variance: residual	2.60			2.61		
Proportion of Level 1 variance	19.1%			18.7%		
Proportion of Level 2 variance	80.9%			81.3%		
Model fit						
Variance explained (Level 1)	9.4%			10.0%		
Variance explained (overall)	1.7%			1.8%		
Sample size						
N	924			833		

Table A4

Results of the Vignette Experiment in the subsample of Facebook users and Facebook and Google users– *U.S. (*multilevel mixed-effects linear regression*)*

	Active Facebook users			Active Facebook and Google users		
(Intercept)	1.71	0.15	**<0.001**	1.85	0.16	**<0.001**
Incentive	0.42	0.03	**<0.001**	0.43	0.03	**<0.001**
Report	0.10	0.06	0.133	0.12	0.07	0.084
Platform: FB + Google	0.04	0.07	0.500	0.02	0.07	0.765
Platform: FB + other	0.06	0.07	0.358	0.06	0.07	0.379
Platform: FB + Google + Other	0.13	0.07	0.054	0.12	0.07	0.089
Time	-0.43	0.05	**<0.001**	-0.49	0.05	**<0.001**

	Active Facebook users			Active Facebook and Google users		
Type of data: no PM/loc	-0.08	0.07	0.253	-0.07	0.07	0.328
Type of data: no PM/vid	-0.06	0.07	0.385	-0.04	0.07	0.592
Type of data: no PM/loc/vid	0.07	0.07	0.280	0.11	0.07	0.109
Variances of random effects						
Variance: constant	8.93			9.29		
Variance: residual	1.37			1.36		
Proportion of Level 1 variance	13.3%			15.7%		
Proportion of Level 2 variance	86.7%			84.3%		
Model fit						
Variance explained (Level 1)	11.0%			12.3%		
Variance explained (overall)	1.6%			1.8%		
Sample size						
N	603			549		

Table A5

Results about willingness to donate data – U.S. (multilevel mixed-effects linear regression) – Three-category gender

	Model with vignette dimensions and controls		
(Intercept)	0.35	0.95	0.71
Incentive	0.35	0.02	0.00
Report	0.15	0.05	0.01
Platform: FB + Google	0.14	0.06	0.01
Platform: FB + other	0.10	0.05	0.07
Platform: FB + Google + Other	0.18	0.06	0.00
Time	-0.38	0.04	0.00

	Model with vignette dimensions and controls		
Type of data: no PM/loc	-0.01	0.06	0.80
Type of data: no PM/vid	-0.04	0.06	0.46
Type of data: no PM/loc/vid	0.06	0.05	0.31
Controls			
Gender (Female)	0.55	0.20	0.01
Gender (Other, DK/NA)	-0.10	0.38	0.78
Age	-0.01	0.01	0.24
Education	-0.02	0.07	0.76
Subjective wealth	-0.15	0.13	0.27
IUIPC_control	0.20	0.14	0.18
IUIPC_collect	-0.53	0.18	0.01
Privacy beliefs	0.17	0.09	0.08
Tech attitudes	0.06	0.07	0.39
BF: openness	-0.00	0.19	1.00
BF: conscientiousness	-0.17	0.20	0.42
BF: extroversion	0.02	0.10	0.84
BF: agreeability	0.14	0.20	0.51
BF: neuroticism	0.00	0.03	0.95
Social Media usage frequency	0.25	0.12	0.03
No of platforms	0.29	0.08	0.00
Variances of random effects			
Variance: constant	7.00		
Variance: residual	1.33		
Proportion of Level 1 variance	16.0%		
Proportion of Level 2 variance	84.0%		
Model fit			
Variance explained (Level 1)	7.6%		
Variance explained (Level 2)	15.5%		
Variance explained (overall)	14.3%		

C. Respondents with zero willingness to donate data

A substantial number of respondents (184 out of a total of 758) expressed zero willingness to donate their digital trace data across all vignette conditions. These individuals appear to represent a group that expresses categorical unwillingness to participate in data donation studies, largely independent of specific design features such as incentives or framing. Their presence highlights that non-participation in data donation research is not necessarily only a function of suboptimal study design or excessive burden, but may instead reflect principled opposition, possibly driven by reasons such as heightened privacy concerns, low institutional trust, or strong normative beliefs about data sharing.

Stated willingness to donate digital trace data does not closely replicate realized donation behaviour (235). Nevertheless, examining the characteristics of respondents who express consistently zero willingness remains informative. These responses capture normative orientations and abstract evaluations of privacy trade-offs, even if actual behaviour ultimately emerges from the interaction of such orientations with situational constraints, incentives, and practical considerations. In this sense, respondents with zero willingness are particularly informative for understanding the reluctance component of data donation, as distinct from non-participation driven by capacity-related barriers.

An analysis of this group's characteristics provides insight into which segments of the population are categorically excluded from data donation designs, even under favourable hypothetical conditions. In comparing demographic characteristics of individuals who expressed no willingness to donate their data consistently, we find that the differences between respondents who never expressed willingness to participate and those who expressed at least some willingness are statistically detectable but substantively nuanced. Respondents with consistently zero willingness are, on average, slightly older (approximately three years), though this difference is small in magnitude. While education reaches statistical significance, the median educational level is identical across groups, suggesting that this difference is unlikely to be substantively meaningful. There are no gender differences between the groups. More notable differences emerge in variables related to digital engagement. Respondents who never expressed willingness report lower levels of technological affinity and use fewer online platforms on average. The difference in the number of platforms used is the most pronounced: individuals in the zero-willingness group report using approximately 0.7 fewer platforms. Taken together, these patterns suggest that categorical unwillingness to donate data is not primarily associated with socio-demographic exclusion, but rather with lower digital embeddedness and comfort with technology. As a result, even optimally designed data donation studies are likely to systematically exclude a non-random subset of individuals. This has important implications for representativeness and suggests that data donation should be understood as complementing, rather than replacing, other data collection strategies. Descriptive statistics of the group of respondents with consistently zero willingness to donate data are presented below.

Table B.6: Demographic comparison of respondents who never expressed participation probability with those who did at least once, using t-tests (continuous variables) and χ^2 tests (categorical variables). Significance is indicated by * $p < .05$, ** $p < .01$, *** $p < .001$.

Variable	Sig	Never expressed participation	Expressed participation $\geq$ once
Age	**	48.3 (13.9)	45.1 (13.4)
Education (median)	*	4	4
Female (%)		53.3	53.9
Tech affinity	**	3.28 (1.1)	3.54 (0.85)
Number of platforms	***	2.85 (1.62)	3.51 (1.7)

D. Interactions between vignette level variables

We examined interactions between incentive amount and other key covariates. Specifically, we tested all interactions between incentives and the main respondent-level characteristics included in the models. For space reasons, these analyses were not included in the original manuscript. However, several interactions were statistically significant and are now reported in Figure B.2. The results show that the effect of incentives is stronger for respondents in the highest compared to the lowest education category, for those with more frequent (as opposed to less frequent) Facebook usage, and for respondents with more negative (rather than positive) privacy attitudes. These interaction patterns suggest meaningful heterogeneity in responsiveness to incentives across subgroups, while leaving the main average effects unchanged.

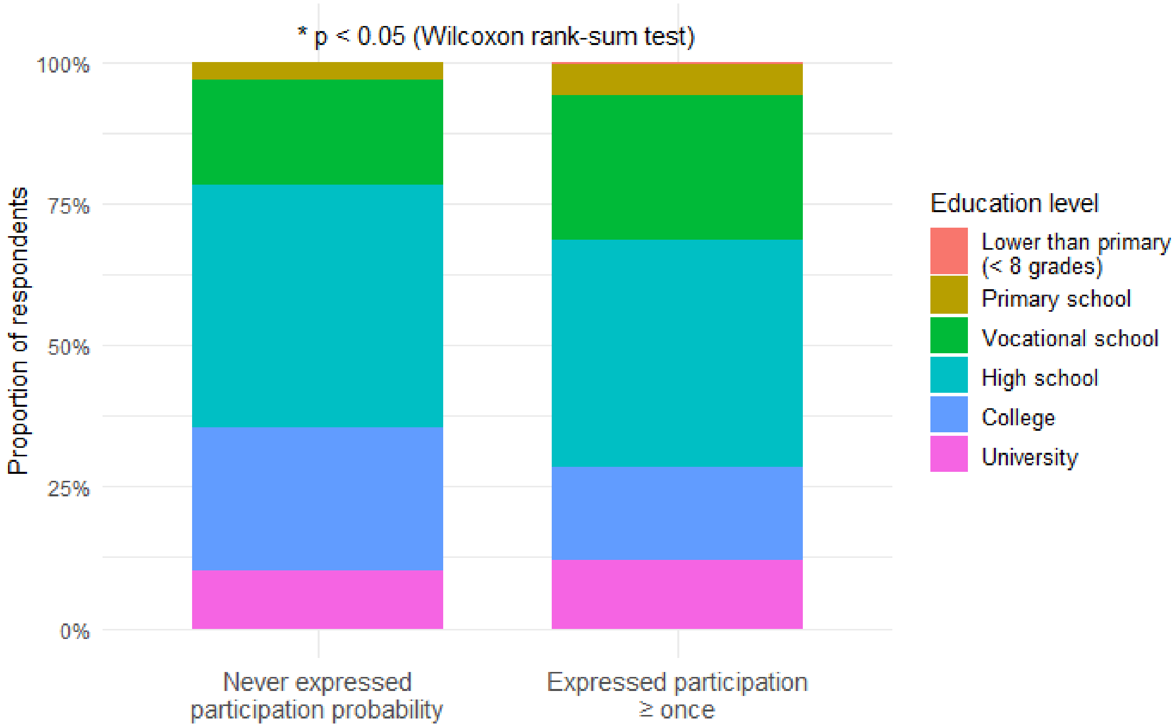


Figure B.1: Comparison of the education levels of respondents who never expressed participation probability with those who did at least once using Wilcoxon rank-sum test. Significance is indicated by * $p < .05$, ** $p < .01$, *** $p < .001$.

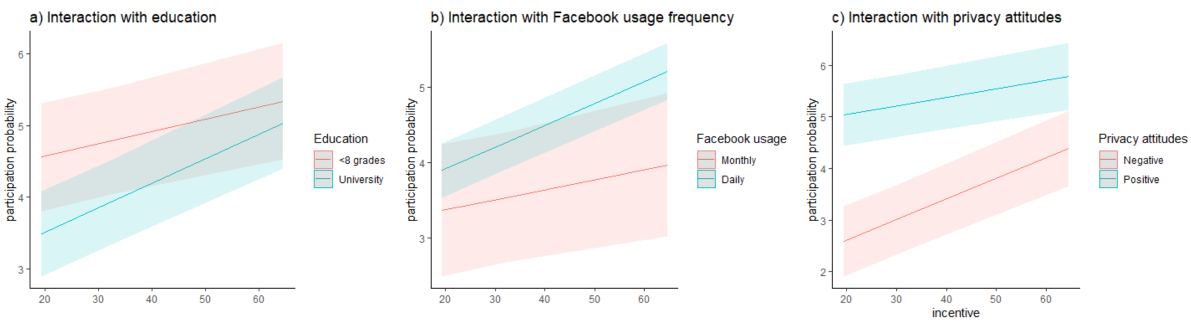


Figure B.2: Interactions between incentive amount and education, Facebook usage, and privacy attitudes, respectively.

Appendix C

Supplementary Information for Study 3

Supplementary Information for Data and Methods

Threshold Selection for Cosine-Similarity Facebook Page Networks

To determine an appropriate cutoff for cosine similarity when constructing the page–page network, we conducted exploratory analysis of network size across different thresholds (Figure C.1). As the threshold increased, the number of edges declined sharply, while the number of nodes remained relatively stable until approximately 0.5, after which larger connected components began to fragment and many pages dropped into small dyads and triads. We therefore selected a threshold of 0.5 as a balance between removing spurious weak ties and retaining a connected network with meaningful community structure.

Alternative Outcome Variable Specifications for Testing Robustness

To facilitate transparency and robustness checks, we constructed four alternative specifications of the misinformation engagement indices, reflecting different assumptions about the relative weight of engagement behaviours. The "main" index is reported in the main text, while results based on the other specifications—emphasizing higher intent, greater exposure, or excluding potentially sceptical engagement—are provided below.

Table C.1 summarises the components and their associated multipliers for each outcome variable alternative.

Table C.1: Component Multipliers in Misinformation Engagement Indices

Component	Main Index	High-Intent Emphasis	Exposure Emphasis	Excl. Sceptical
Posts with Misinformation Links	1.0	1.2	1.0	1.0
Comments with Misinformation Links	1.0	1.2	1.0	1.0
Likes/Follows (Pages and Groups)	0.6	0.6	0.8	0.6
Reactions Pages/Groups	0.3	0.3	0.4	0.1
Comments on Pages/Groups	0.7	0.7	0.7	excluded

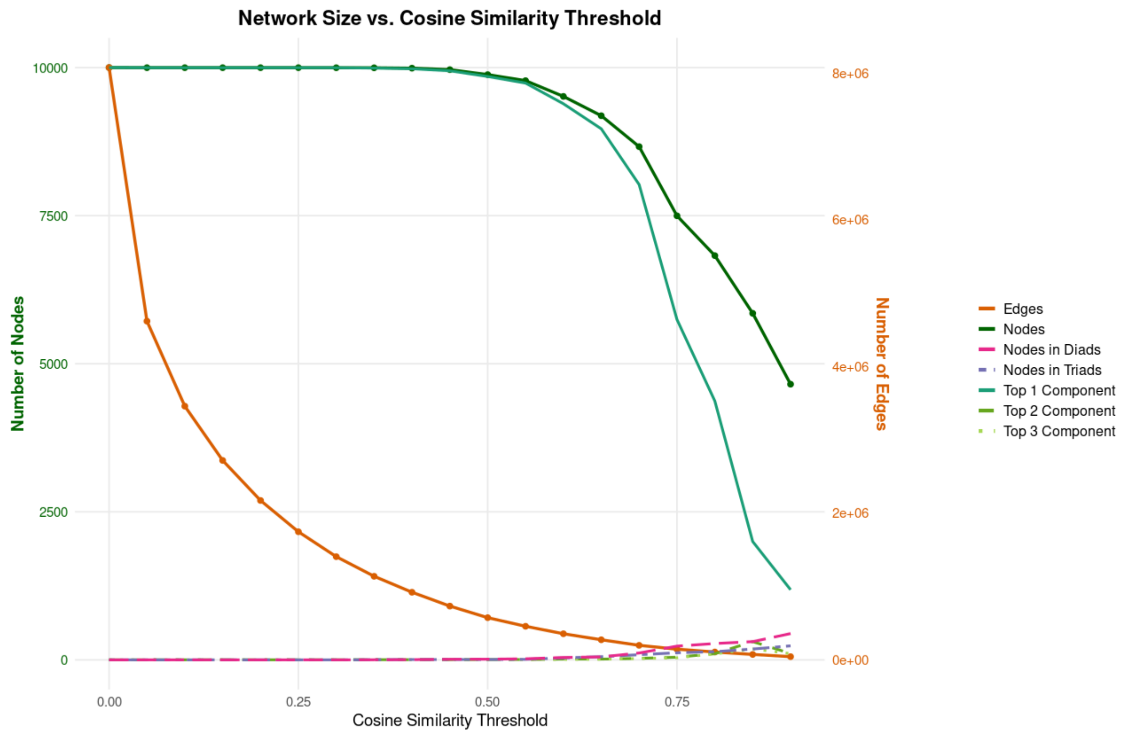


Figure C.1: Network size as a function of cosine similarity threshold.

Each index was computed as the weighted sum of these components. For example, the main index was calculated as:

$$\begin{aligned}\texttt{misinfo\_count\_main} = {} & 1.0 \times \text{misinfo\_posts} \; + \; 1.0 \times \text{misinfo\_comments} \; + \\ & 0.6 \times \left(\text{misinfo\_pages\_liked\_count} + \text{misinfo\_groups\_liked\_count}\right) \; + \\ & 0.3 \times \left(\text{misinfo\_page\_reactions\_count} + \text{misinfo\_group\_reactions\_count}\right) \; + \\ & 0.7 \times \left(\text{misinfo\_page\_comments\_count} + \text{misinfo\_group\_comments\_count}\right).\end{aligned}$$

Analogous formulas were applied to construct the other indices using the multipliers shown in Table C.1.

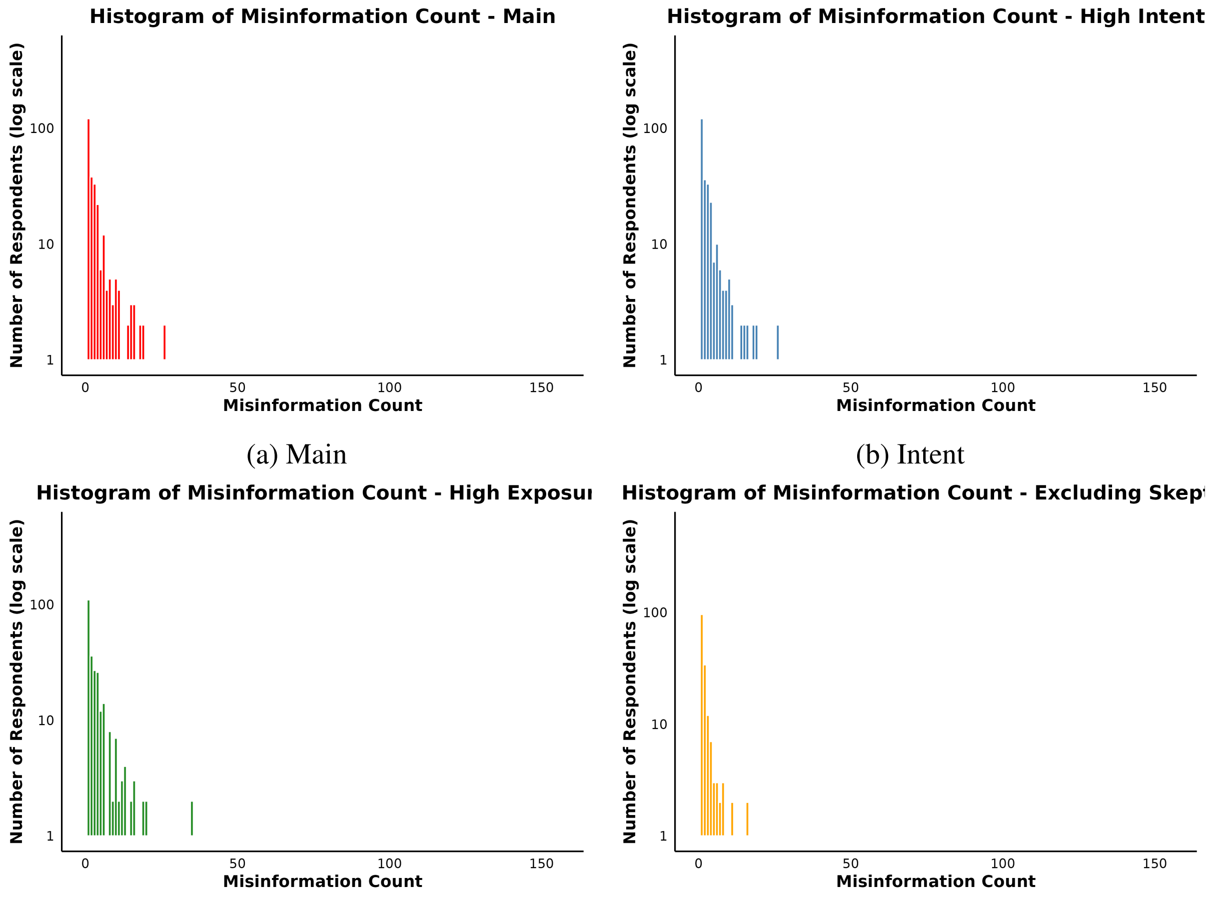
Prevalence of Misinformation in the Dataset by Type of Engagement

Table C.2 reports the counts and proportions of posts and comments containing links to misinformation websites, as well as the counts and proportions of interactions with misinformation Facebook pages and groups, that together constitute our composite outcome variable.

Table C.2: Prevalence of Misinformation in the Dataset by Type of Engagement

Measure	Total (N)	Misinformation (N)	Proportion
Misinformation posts (links)	13,333	80	0.0060
Misinformation comments (links)	7,193	41	0.0057
Page likes/follows	168,789	137	0.0008
Group joins/follows	85,495	2	0.0000
Reactions to pages	2,491,119	2,484	0.0010
Reactions to groups	14,777	317	0.0215
Comments on pages	308,363	748	0.0024
Comments in groups	1,139	99	0.0869

Distributions of the Four Outcome Variable Specifications



(a) Main

(b) Intent

(c) Exposure

(d) Excluding sceptical

Figure C.2: Distribution of the four outcome variable specifications measuring misinformation engagement. A large share of observations take the value zero (no engagement), while the remainder form a long right-skewed tail. This distribution motivates the use of hurdle models, combining a logistic component for predicting any engagement and a Gamma regression with log link for modeling the positive values.

Supplementary Information for Results

Interest Clusters Identified in the Facebook Page Network

As described in the main text, we applied community detection to the Facebook page network to identify clusters of pages with similar patterns of follower overlap. (Figure C.3) shows the distribution of the number of pages in each of the identified interest clusters. The table below provides the full list of 46 clusters obtained through community detection on the Facebook page network. For each cluster, we assigned a descriptive label that captures its dominant theme, supplemented by a brief summary of the content and examples of representative pages. These examples are included for illustrative purposes and highlight the most central or recognizable nodes within each cluster. The descriptive labels are not intended as strict categorical definitions, but rather as interpretive guides to facilitate understanding of the structure of the network.

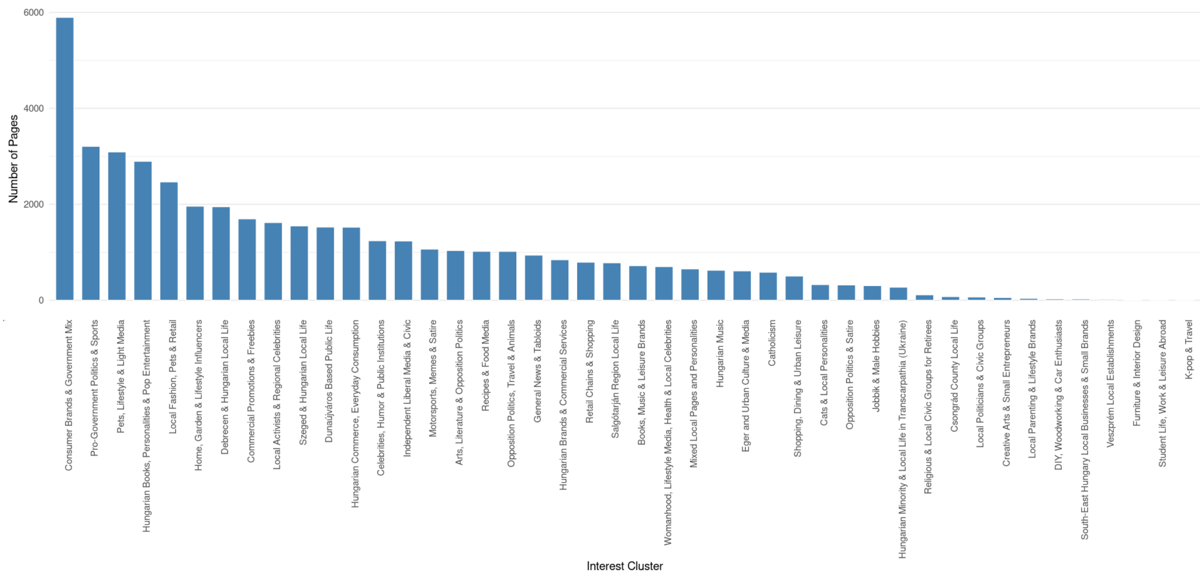


Figure C.3: Distribution of Interest Cluster Sizes.

Table C.3: Facebook Page Clusters

Name	Description	Example Pages	Page Count
Pets, Lifestyle & Light Media	Mix of pet rescue, cute animal content, lifestyle blogs, and popular Hungarian portals.	Imádom a kutyám (I love my dog), Madármentő állomás (Bird Rescue Station), Rádió 1, NLC.hu, Origo	3092

Name	Description	Example Pages	Page Count
General News & Tabloids	Hungarian mass media and tabloids focusing on daily news and sensational stories.	24.hu, ATV, Blikk, Ripost, Mindenegyben Blog (Everything in One Blog), Metropol	942
Consumer Brands & Government Mix	Blend of Hungarian consumer brands, corporate pages, government communication, and personalities.	Müller Magyarország (Müller Hungary), Spar Magyarország (Spar Hungary), Rossmann Magyarország (Rossmann Hungary), Magyarország Kormánya (Government of Hungary), Fidesz	5897
Home, Garden & Lifestyle Influencers	Lifestyle and hobby cluster with baking, gardening, home decor, and influencers.	Torta és Karamell (Cake and Caramel), Bálint Gazda Kertje (Garden of Gardener Bálint), Jocó bácsi világa (Uncle Jocó's World), Filmezzünk.hu (Let's Watch Movies), Rubint Réka	1962
Opposition Politics, Travel & Animals	Citizen-initiated opposition Politics, Travel Interest & Animals.	Nyugati fény (Western Light), Aki Orbán ellen van, az velünk van (Those who are against Orbán are with us), I love London, Nyíregyházi állatpark (Nyíregyháza Zoo)	1021
Local Fashion, Pets & Retail	Local boutiques, pet charities, stop-shops, and small-town retail.	Amnesia Pilisvörösvár, Tamás Ékszer (Tamás Jewelry), Eszkuláp Állatvédő Egyesület (Eszkuláp Animal Protection Association), Turista Magazin (Tourist Magazine), Kutyaovi (Dog Kindergarten)	2468
Arts, Literature & Opposition Politics	Tech/gadget sites, quirky innovation blogs, and science curiosities.	Színház online (Theatre Online), Operaház (Opera House), Donáth Anna, Magyar Narancs (Hungarian Orange), Széll Bernadett	1038
Recipes & Food Media	Hungarian recipe blogs, food magazines, and cooking influencers.	Nosalty, Mindmegette (Tried It All), Street Kitchen, Gasztroangyal (Gastro Angel), Receptneked (Recipe for You)	1022

Name	Description	Example Pages	Page Count
Hungary-Specific Brands & Commercial Services	Hungary-specific versions of brands, commercial services, and corporate campaigns.	Hell Energy, Nescafé Magyarország (Nescafé Hungary), XXXLutz Magyarország (XXXLutz Hungary), Huawei Mobile	846
Womanhood, Lifestyle Media, Health & Local Celebrities	Womanhood, Lifestyle Media, Health & Local Woman Celebrities.	Fem3 Café, Család-barát magazin (Family-Friendly Magazine), She.hu, Szimpatika Gyógyszertár (Szimpatika Pharmacy)	703
Shopping, Dining & Urban Leisure	Budapest-based consumer and leisure culture.	Allee, Westend, Mammut, Pólus Center, Pizza Me, Sugár Mozi (Sugár Cinema)	507
Hungarian Music	Hungarian music performers and music programming.	ByeAlex, Majka, Hooligans, Tankcsapda (Tank Trap), X-Faktor, RTL	627
Hungarian Minority & Local Life in Transcarpathia (Ukraine)	A regional cluster anchored in Kárpátalja (Transcarpathia).	Kárpátaljai Magyar Nagycsaládosok Egyesülete (Association of Hungarian Large Families in Transcarpathia), (Beregovo District Administration), Mukachevo.net	274
Retail Chains & Shopping	Large-scale supermarkets and retail chains.	Tesco Magyarország (Tesco Hungary), Lidl Magyarország (Lidl Hungary), Aldi Magyarország (Aldi Hungary), Auchan, Media Markt	795
Creative Arts & Small Entrepreneurs	Artists, florists, performers, and cultural events.	Winkler Virágműhely (Winkler Flower Workshop), Vagabund Kiadó (Vagabond Publisher), Rákász Gergely, Bonbonka, Benedek Helga	58
Eger and Urban Culture & Media	Regional newspapers, heritage sites, and local history.	Heves Megyei Hírlap (Heves County Journal), Hajdú Online, Mazsihisz (Federation of Hungarian Jewish Communities)	614

Name	Description	Example Pages	Page Count
Catholicism	Catholic pages (churches, religious personalities, related media outlets).	Böjte Csaba (Brother Csaba Böjte), Laci Atya (Father Laci), 777, Esztergom-Budapesti Főegyházmegye (Archdiocese of Esztergom-Budapest), Szent Margit Templom (Saint Margaret Church)	586
Independent Liberal Media & Civic	Opposition-friendly outlets, investigative journalism, and NGOs.	Telex, 444.hu, Átlátszó (Transparent), Partizán, Amnesty International	1237
Cats & Local Personalities	Cat (and other pet) related institutions, with local personalities.	Állat és természetvédelmi közszolgálat (Animal and Nature Protection Service), Cat café, Lovas Zsuzsanna	328
Mixed Local Pages and Personalities	Small Local Pages and Personalities.	Karácsony imádók oldala (Christmas Lovers' Page), Road zenekar (Road Band), Marianne fashion Jászberény	656
Jobbik & Male Hobbies	Jobbik politicians and hobby interests (cars, motorcycles, weapons).	Jobbik Magyarországért (Jobbik for Hungary), Ander Balázs, Russian extreme offroad trucks	307
Salgótarján Region Local Life	Local Life & Animal Welfare in Salgótarján Region.	Salgótarján régen és most (Salgótarján Then and Now), 3100.hu - Salgótarján hírportálja (Salgótarján News Portal), Mancs a szívben Állatvédő Egyesület (Paw in the Heart Animal Protection Association)	782
Opposition Politics & Satire	Opposition parties and satire critical of government.	Tibi Atya (Father Tibi), Kétfarkú Kutya Párt (Two-Tailed Dog Party), Puzsér Róbert, Tudományos Mémek (Scientific Memes), Pénztáros Lőrinc (Cashier Lőrinc)	321

Name	Description	Example Pages	Page Count
Veszprém Local Establishments	Schools, cafés, shops in Veszprém county	Hajrá Veszprém (Go Veszprém), Café Frei Veszprém, Magyar-Angol Tannyelvű Gimnázium Balatonalmádi (Hungarian-English Bilingual High School Balatonalmádi)	22
Local Politicians & Civic Groups	Small-town public figures and civic associations.	Janiczak Dávid, Ózdi Városi Televízió (Ózd City Television), Kovács János	69
Furniture & Interior Design	Furniture, interior design, and home decoration (new and second-hand).	OutletBútor (Outlet Furniture), Használt-konyhabútor (Used Kitchen Furniture)	13
Local Parenting & Lifestyle Brands	Baby products, cosmetics, and lifestyle entrepreneurs.	Minifarm, Babyberry, Boda Zsanett Sminktetoválás (Zsanett Boda Permanent Makeup), Óvodai Nevelés (Preschool Education)	41
DIY, Woodworking & Car Enthusiasts	DIY, carpentry, cosplay, and automotive clusters.	DIY & Crafts, The Dusty Lumber Co., Vesco Dániel, Kamui Cosplay, Car Tuning	29
Student Life, Work & Leisure Abroad	Hungarian universities and student opportunities.	Pécsi Tudományegyetem (University of Pécs), Camp America Hungary, Vendéglátós állások Ausztriában (Hospitality Jobs in Austria)	11
K-pop & Travel	Niche fandom for Korean pop culture mixed with travel blogs.	BTS Kim Seokjin, The Kim Family, Világcsavargó (World Wanderer), Aura Hotel, Diatron	5
Pro-Government Politics & Sports	Fidesz-aligned government pages and sports fandom.	M4 Sport, Orbán Viktor, Nemzeti Sport (National Sport), Szijjártó Péter, Varga Judit	3209
Dunaújváros Based Public Life	Opposition Politics & Local Public Life (Dunaújváros focus, with lifestyle/entertainment overlap).	Jakab Péter, Márki-Zay Péter, Dunaújváros Önkormányzata (Municipality of Dunaújváros), Retro disco zenék (Retro Disco Music)	1528

Name	Description	Example Pages	Page Count
Celebrities, Humor & Public Institutions	Easy Entertainment, Celebrities, Humor & Public Institutions.	Barbara Palvin, Scarlett Johansson Fans, Idézetek és viccek egy helyen (Quotes and Jokes in One Place), Országos Mentőszolgálat (National Ambulance Service), Magyar Rendőrség (Hungarian Police), BRFK (Budapest Police Headquarters)	1242
Motorsports, Memes & Satire	Motorsport fandom, meme culture, parody pages	MKKP (Hungarian Two-Tailed Dog Party), Boxutca (Pit Lane), 9gag, Formula 1 Hungary, Michael Schumacher	1068
Hungarian Books, Personalities & Pop Entertainment	Hungarian authors, cookbooks, FM radios, and entertainers.	Hadházy Ákos, Sláger FM (Hit FM), Könyvmolyok (Bookworms), Természettár (Nature Collection), BritneyJoy	2897
Hungarian Commerce, Everyday Consumption	Hungarian-oriented blend of commerce, local life, entertainment, and inspirational content.	Supershop, Telekom HU, Foodpanda, Wellhello, Fluor Tomi	1525
Debrecen & Hungarian Local Life	Debrecen life—local businesses, family tips, crafts, and culture.	Békésszentandrási Szent András Sörfőzde (Saint Andrew Brewery of Békésszentandrás), Kendőben Jó (Good in a Scarf), Debreceni Egyetem Memes (University of Debrecen Memes), Ugar Brewery	1951
Szeged & Hungarian Local Life	Szeged related local press, politicians, establishments.	Szeged365, Utazómajom (Travel Monkey), Tudatos Életmód (Conscious Lifestyle), Rádió 88 Szeged, Újságmúzeum (Newspaper Museum), Demény	1551
Local Activists & Regional Celebrities	Local activists, small-town figures, and regional media influencers.	Zsombor SMA1, Polcz Péter, Szabó Attila, Rádió 88 Szeged, Helyi Hírek (Local News)	1621

Name	Description	Example Pages	Page Count
Religious & Local Civic Groups for Retirees	Small religious, cultural, and civic associations focused on retirees.	Szeretem a családomat (I Love My Family), Napi útravaló (Daily Spiritual Food), Polgári szertartásvezetés (Civil Ceremony Leading), Nagyszülők (Grandparents)	115
Commercial Promotions & Freebies	Promotions, food service, fast food, and prize giveaways.	KFC Magyarország (KFC Hungary), Auchan Korzó (Auchan Mall), Ingyen Termékminták (Free Product Samples), Babudo Webáruház (Babudo Webshop), Erika Tóth	1698
Books, Music & Leisure Brands	Publishers, bookstores, and music-related brands.	Könyvkaptár (Book Hive), Libri, Alexandra Könyvesbolt (Alexandra Bookstore), Vatera, Zenekar.hu (Band.hu)	722
Csongrád County Local Life	Local arts, gastronomy, community news, and lifestyle in Csongrád.	Kecskeméti Katona József Nemzeti Színház (Katona József National Theatre of Kecskemét), Csongrád / Kiskunfélegyháza Városi Önkormányzat (Municipality of Csongrád / Kiskunfélegyháza), Pompei Pizzéria (Pompei Pizzeria)	78
South-East Hungary Local Businesses & Small Brands	Restaurants, cafés, small manufacturers, and niche Hungarian food brands in the South-East of Hungary.	Nissan, Mirage Café, Poco, Kötöny-Hús Kft. (Kötöny-Meat Ltd.), 3Dmoments	28

Note: Clusters vary in size and thematic cohesion. While larger clusters are dominated by clear lifestyle, cultural, or political themes, some of the smaller ones represent niche interests (e.g., K-pop fandom, celebrity pages). The naming of clusters is therefore partly heuristic, aiming to balance descriptive clarity with interpretive flexibility.

Full Results for the Main Outcome Variable Specification

In this subsection, we present the full regression results for our models. In the main paper, we only reported statistically significant coefficients for the Main outcome. Here, we include the complete coefficient estimates for both the zero models (estimating the likelihood of sharing no misinformation) and the positive models (estimating the volume of misinformation shared, conditional on sharing at least once).

Zero models

Figure C.4 presents the results of the zero models, which estimate the probability that a respondent engages with misinformation at all (versus never sharing any). These models therefore capture the baseline likelihood of misinformation engagement.

Table C.4 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.048 (Model 1), and 0.281 (Model 2), reflecting the significant contribution of interest clusters and behavioural controls to model fit.

Summary. These models show that age, particular Facebook page–based interest clusters, and Facebook activity are significantly associated with individual misinformation engagement.

Figure C.4: Zero models predicting zero vs non-zero sharing (Main outcome).

Figure C.5: Zero models predicting zero vs non-zero sharing (Main outcome), with standardised explanatory variables.

Positive models

Figure C.6 shows the results of the positive models, which estimate the volume of misinformation shared among those who have engaged at least once. These models thus shed light on what drives the intensity of engagement.

Table C.4 presents that adjusted McFadden R^2 values for the positive models are 0.048 (Model 1), and 0.150 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. The positive models show that, conditional on engaging with misinformation, age, Facebook page–based interest clusters (some different from those affecting misinformation engagement likelihood), and platform activity are significantly associated with the volume of individual misinformation engagement.

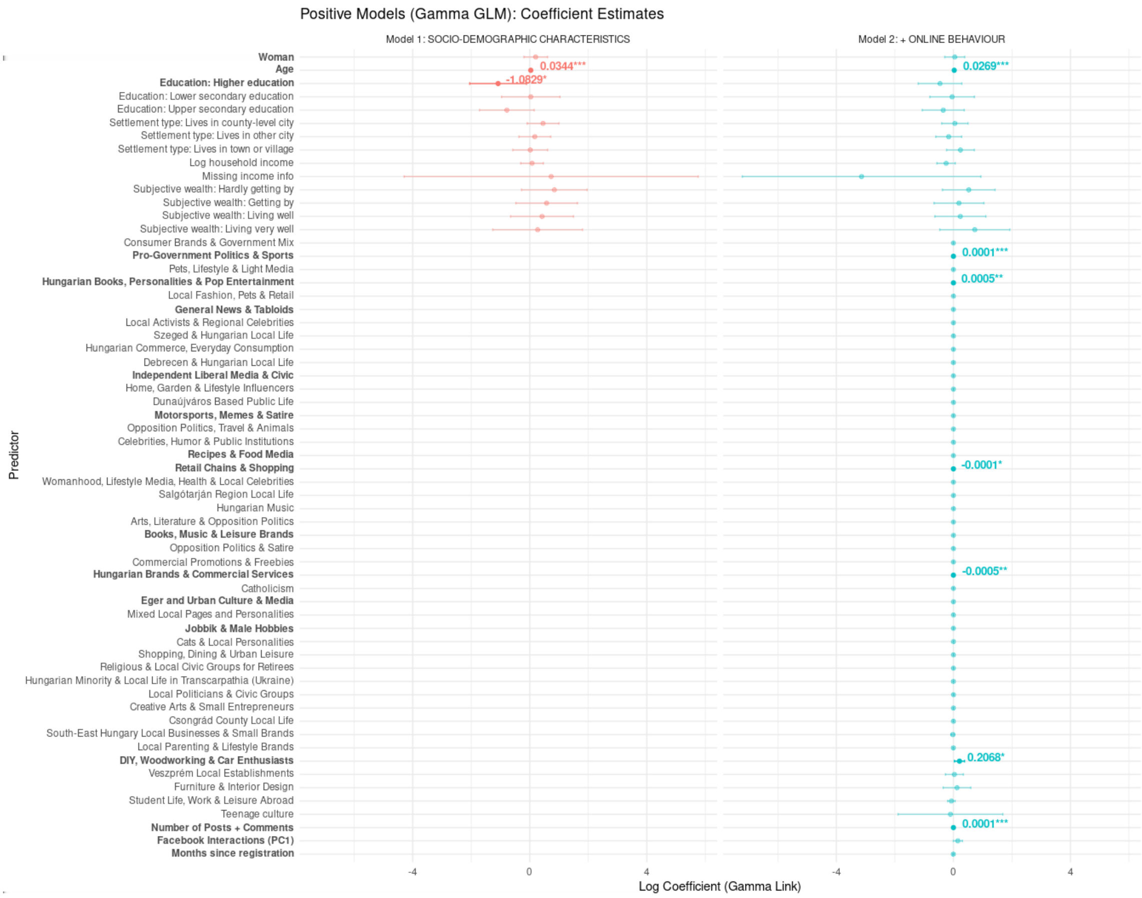


Figure C.6: Positive models predicting volume of sharing (Main outcome).

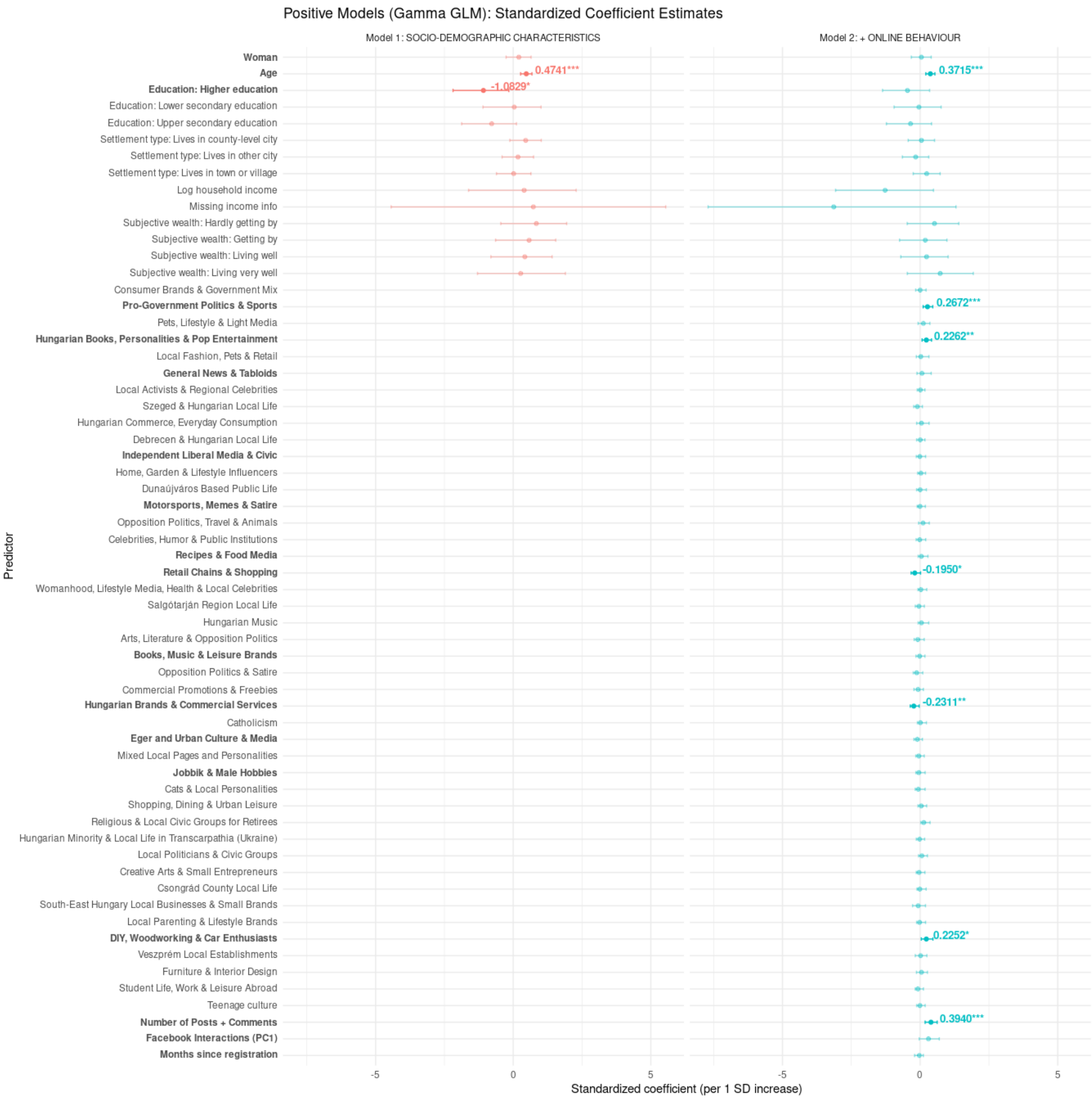


Figure C.7: Positive models predicting volume of sharing (Main outcome), with standardised explanatory variables.

Table C.4: Model Fit: Adjusted McFadden R^2 for Zero and Positive Misinformation Engagement Models with the Main Outcome Variable

Model	**Model 1**	**Model 2**
Zero models	0.048	0.281
Positive models	0.048	0.150

Combined summary of main outcome models

Taken together, the zero-inflation and positive models clarify the two stages of misinformation engagement. The zero models highlight the factors that determine whether individuals engage at all, showing robust roles for age, specific Facebook page–based interest clusters, and Facebook activity measures. The positive models then show that, among those who engage with misinformation, similarly to the zero models, age, Facebook page–based interest clusters–although some different from those in the zero models–, and platform activity are significantly associated with the volume of individual misinformation engagement. Together, these results provide the full picture underlying the significant-only results reported in the main text.

Robustness Checks with the Three Alternative Variable Specifications

In this section we present results for three alternative outcome variables: *High Intent*, *Exposure*, and *Excluding sceptical*. These serve as robustness checks to demonstrate that our findings are not dependent on the exact operationalization of misinformation sharing.

High Intent outcome

Zero models

Figure C.8 presents the results of the zero models for the high-intent outcome, which predict whether a respondent engages with misinformation at all. The patterns closely resemble those reported for the Main outcome (Figure C.4), with age emerging as the most consistent predictor: older respondents are significantly more likely to engage with misinformation. As in the Main outcome models, the effect of gender is significant in the simplest specification, with women more likely to engage, but attenuates once interest clusters and behavioural controls are included. Education and income remain non-robust predictors.

The set of significant Facebook page–based interest clusters closely resembles the Main outcome. Engagement with *General News & Tabloids* and *Eger and Urban Culture & Media* predicts a higher likelihood of misinformation sharing, while affinity with clusters such as *Independent Liberal Media & Civic*, *Jobbik & Male Hobbies*, and *DIY, Woodworking & Car Enthusiasts* predicts lower engagement.

Finally, platform activity variables show the same pattern as in the Main outcome: the number of posts/comments and Facebook interaction intensity are positively associated with a greater likelihood of engagement, while account age shows a negative association.

Table C.5 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.048 (Model 1), and 0.281 (Model 2), reflecting the significant contribution of interest clusters and behavioural controls to model fit.

Summary. These results confirm the robustness of the Main outcome models: while higher activity levels increase the likelihood of misinformation engagement, age and certain Facebook page–based interest clusters robustly predict whether individuals engage with misinformation at all.

Figure C.8: Zero models predicting zero vs non-zero sharing (High Intent outcome).

Positive models

Figure C.9 shows the results of the positive models for the high-intent outcome, which estimate the volume of misinformation shared among those who have engaged at least once. As with the Main outcome models (Figure C.6), age remains the most consistent demographic predictor, with older respondents sharing larger volumes of misinformation. Gender is not significant once engagement is established, while education shows a negative effect in the simplest model but loses significance in the full model. The same Facebook page–based interest clusters remain significant predictors of volume as in the Main outcome. Behavioural predictors again mirror the Main outcome: the number of posts/comments is a significant positive predictor, with the effect size being somewhat stronger than in the Main outcome model. The effect is account age is not significant.

Table C.5 presents that adjusted McFadden R^2 values for the positive models are 0.048 (Model 1), and 0.149 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. The positive models for the high-intent outcome largely replicate the Main outcome findings: conditional on engaging with misinformation, age, Facebook page–based interest clusters, and platform activity are significantly associated with the volume of individual misinformation engagement.

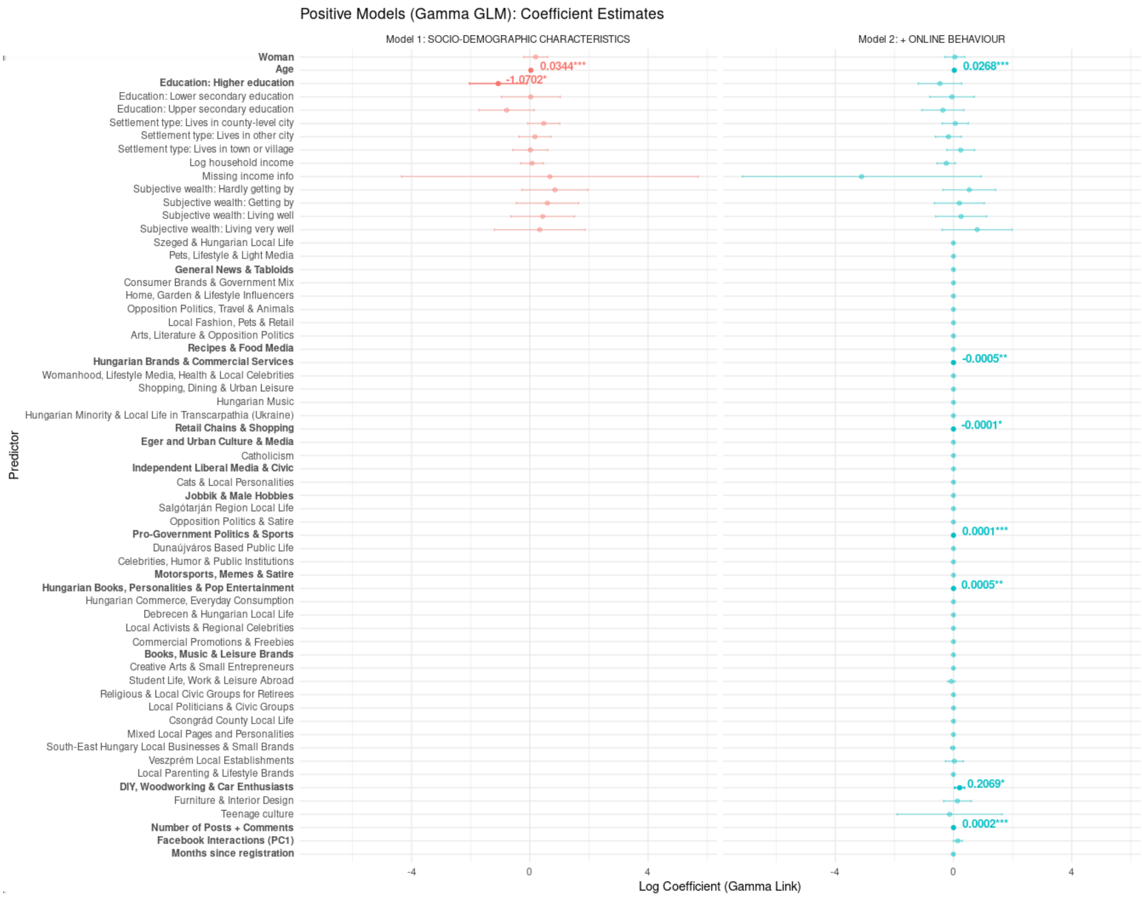


Figure C.9: Positive models predicting volume of sharing (High Intent outcome)

Table C.5: Adjusted McFadden R^2 for Zero and Positive Intent Models

Model	Model 1	Model 2
Zero models	0.048	0.281
Positive models	0.048	0.149

Combined summary of high-intent models

The high-intent outcome models confirm that the Main outcome results are robust to this alternative specification. Both sets of models highlight the same structural patterns: age, certain interest clusters, and platform activity are consistently associated with misinformation engagement. The overlap across outcomes indicates that the Main results are not driven by the choice of outcome variable.

Exposure outcome

Zero models

Figure C.10 presents the zero models predicting whether respondents engage with misinformation at all under the Exposure specification. The demographic patterns largely mirror the Main outcome (Figure C.4) and the High Intent specification (Figure C.8).

Table C.6 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.048 (Model 1), and 0.281 (Model 2), reflecting the significant contribution of interest clusters and behavioural controls.

Summary. The zero model results in the Exposure specification largely replicate the Main and High Intent outcomes: while higher activity levels increase the likelihood of misinformation engagement, age and certain Facebook page–based interest clusters robustly predict whether individuals engage with misinformation at all.

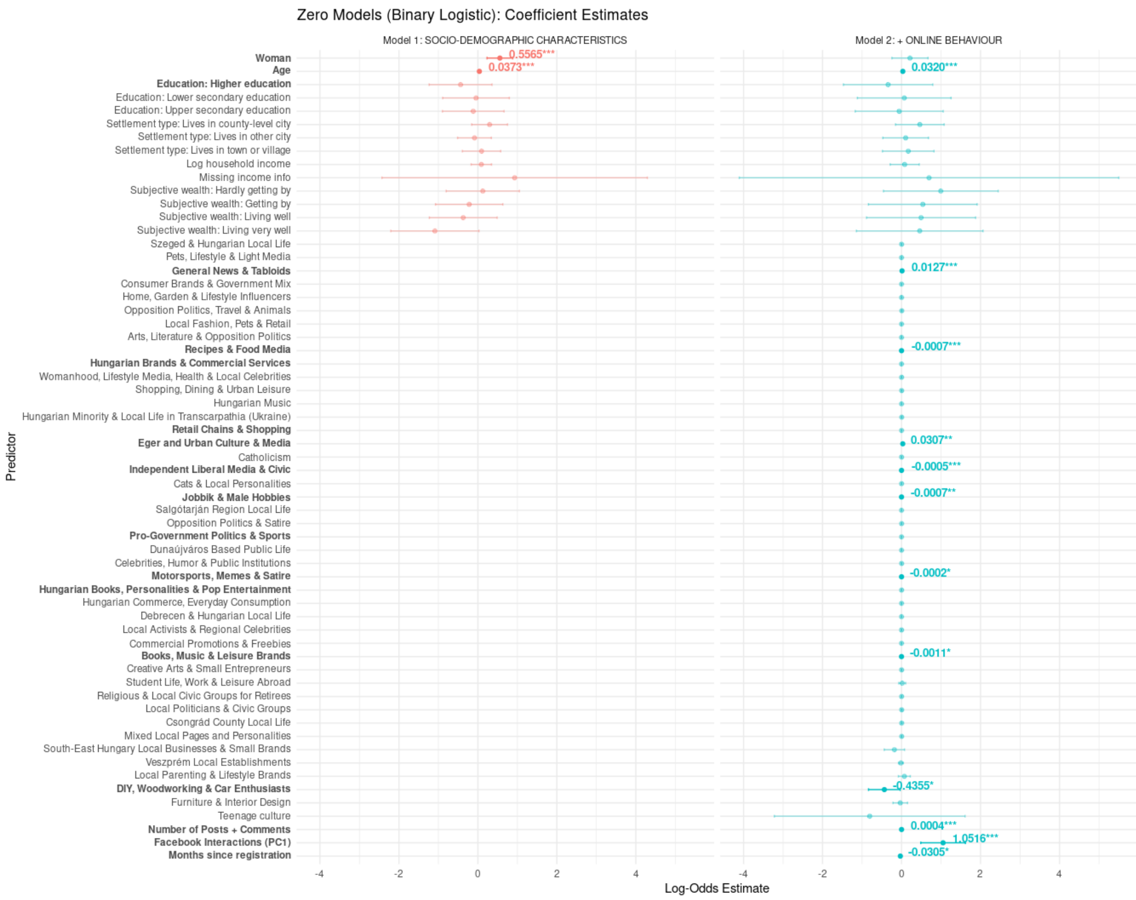


Figure C.10: Zero models predicting zero vs non-zero sharing (Exposure outcome).

Positive models

Figure C.11 shows the positive models estimating the volume of misinformation shared under the Exposure specification. Again, results closely resemble the Main outcome. Age is a consistent positive predictor of sharing volume. Higher education predicts lower sharing in the first model, but this effect weakens after adding behavioural variables. Interest cluster coefficients and the effect of Facebook activity replicate the Main and High Intent outcome patterns.

Table C.6 presents that adjusted McFadden R^2 values for the positive models are 0.045 (Model 1), and 0.149 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. Positive models with the exposure outcome variable reproduce the Main and High Intent outcome findings, confirming the robustness of the patterns across specifications.

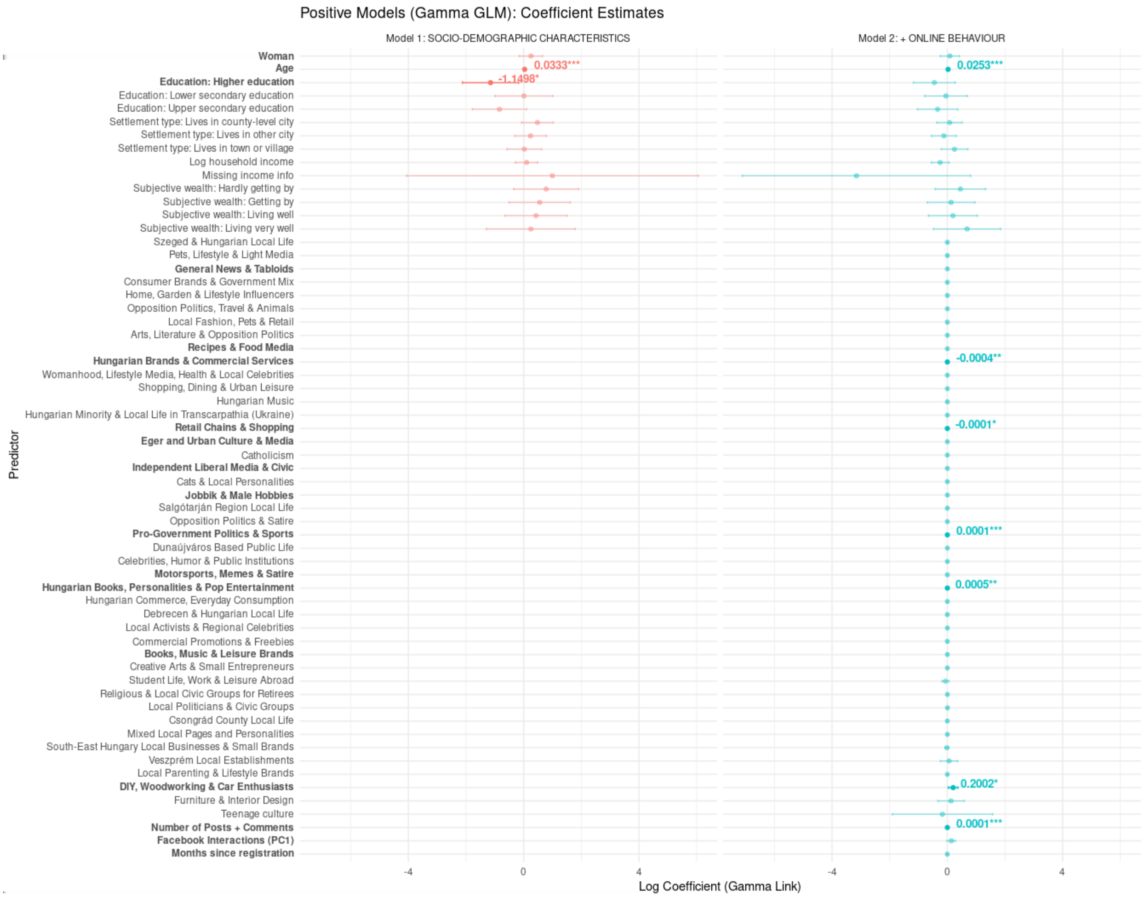


Figure C.11: Positive models predicting volume of sharing (Exposure outcome).

Table C.6: Adjusted McFadden R^2 for Zero and Positive Exposure Models

Model	Model 1	Model 2
Zero models	0.048	0.281
Positive models	0.045	0.149

Combined summary of exposure models

Overall, the Exposure outcome variable specification confirms that the Main findings are robust to alternative modelling of engagement. Both the zero and positive model results show the same demographic and behavioural patterns as the Main and High Intent outcomes. This suggests that the observed predictors of misinformation engagement are not sensitive to minor differences in model specification or sampling.

Excluding sceptical respondents

Zero models

Figure C.12 shows the zero models predicting whether respondents engage with misinformation at all, excluding comments from the outcome variable specification, to avoid ambiguity due to potential scepticism. The main demographic patterns largely replicate the Main outcome Figure C.4), with age emerging again as the strongest predictor of engagement. Older individuals remain significantly more likely to engage. Women show a positive association in the simplest specification, which disappears when Facebook page–based interest clusters and behavioural controls are added. A difference that arises compared to the Main outcome is that Living in a county-level city is positively associated with misinformation engagement. The effect is however only weakly significant.

Regarding Facebook page–based interest clusters, some the same patterns re-emerge as for the other outcome specifications: *General News & Tabloids* and *Eger and Urban Culture & Media* are associated with higher engagement, while *Independent Liberal Media & Civic*, and *Recipes & Food Media* predict lower engagement.

Differently from the other outcome variable specifications, *Consumer Brands & Government Mix*, and *Local Fashion, Pets & Retail* emerge as negative associations with misinformation engagement likelihood, and none of the other interest clusters turn out to be significantly associated with misinformation engagement.

Finally, among behavioural variables, as in the case of the other outcome variable specifications, more posts/comments and higher Facebook interaction intensity strongly predict misinformation engagement, replicating the full-sample findings. However, the effect of account age is not significant.

Table C.7 presents that adjusted McFadden R^2 values for the zero models increase across specifications: 0.039 (Model 1) and 0.257 (Model 2) reflecting the significant contribution of interest clusters and behavioural controls to model fit.

Summary. Excluding potentially sceptical engagements from the outcome specification leaves the zero model results largely unchanged: while controlling for activity, age and set of interest clusters remain key predictors.

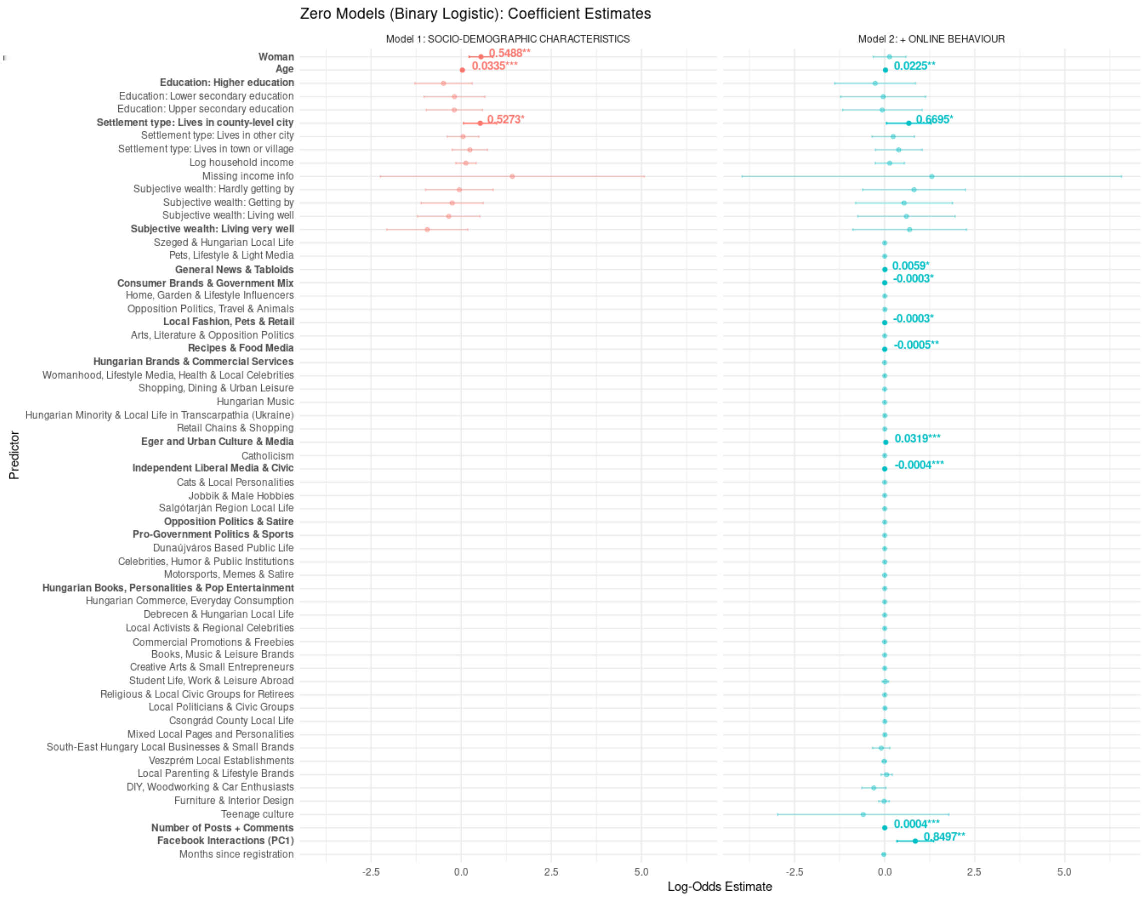


Figure C.12: Zero models predicting zero vs non-zero sharing (Excluding sceptical outcome).

Positive models

Figure C.13 presents the positive models, which estimate the volume of misinformation engagement among those who engage, excluding sceptical engagement. The results again mirror the Main outcome (Figure C.6), with age a consistent and strong predictor of sharing volume. As in the case of the other outcome variables, education shows a negative effect in the simplest specification (higher education associated with less sharing), but this attenuates once behavioural variables are added. As opposed to the other outcome specifications, higher subjective income is associated with higher volume of misinformation engagement in the simplest model. But this effect is explained away once behavioural variables are introduced.

Compared to the Main outcome, most interest cluster effects remain, while few shifts in interest cluster associations emerge. As seen for other outcome variable specifications, *Hungarian Brands & Commercial Services* predicts lower volumes, while *Pro-Government Politics & Sports* and *Hungarian Books, Personalities & Pop Entertainment* show positive associations with volume. Differently from the other outcome variable specifications, *Opposition Politics & Satire* appear negatively associated, and other interest clusters effects are not significant.

Behavioural predictors remain robust. The number of posts/comments and Facebook interaction intensity is a significant positive predictor of volume, consistent with the Main outcome.

Table C.7 presents that adjusted McFadden R^2 values for the positive models are 0.053 (Model

1), and 0.210 (Model 2), showing improved fit with the addition of interest clusters and behavioural measures.

Summary. The positive models excluding sceptical engagement confirm the robustness of the Main findings: while controlling for activity age and some interest clusters remain the strongest predictors, with minor shifts in certain effects.

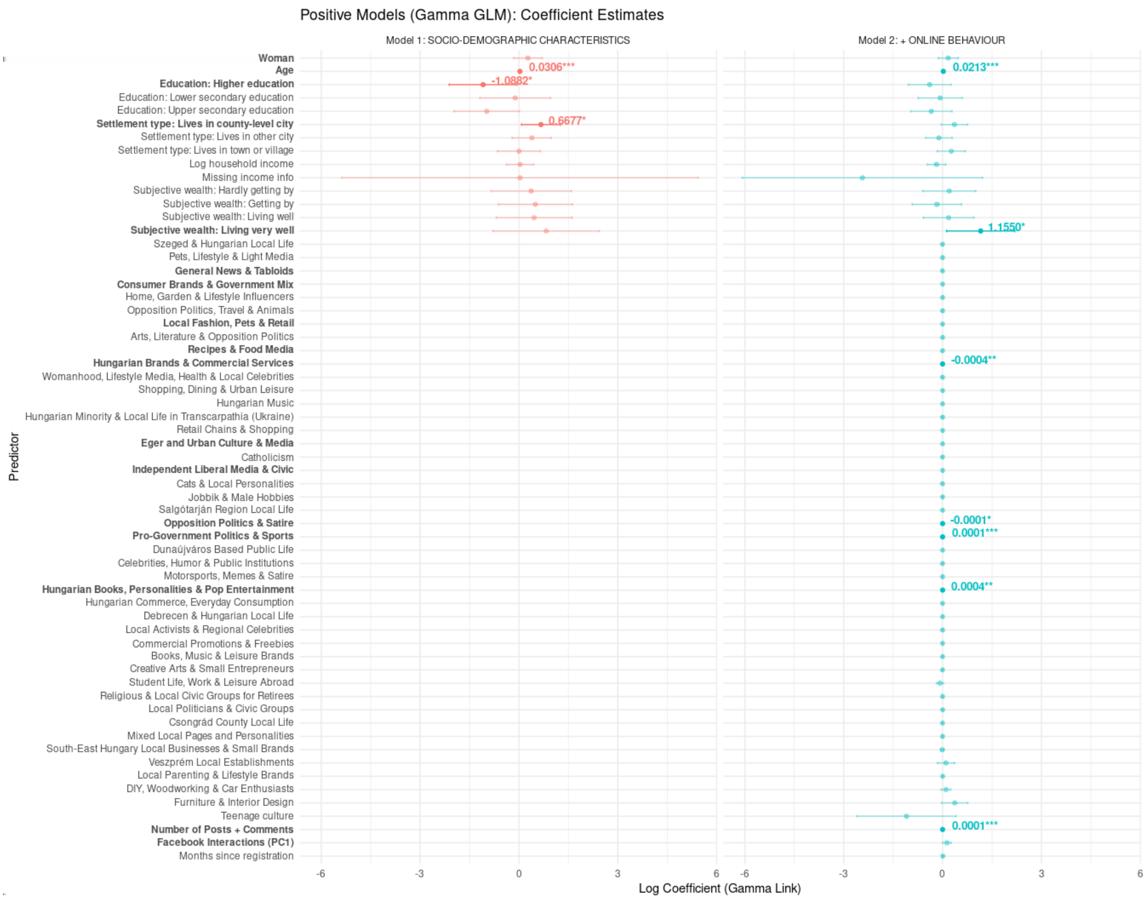


Figure C.13: Positive models predicting volume of sharing (Excluding sceptical outcome).

Table C.7: Adjusted McFadden R^2 for Zero and Positive Excluding Sceptical Models

Model	Model 1	Model 2
Zero models	0.039	0.257
Positive models	0.053	0.210

Combined summary of exclusion models

Overall, excluding sceptical respondents does not materially change the structure of the results. Across both model types, when controlling for Facebook activity, age and a set of Facebook page–based interest clusters remain central predictors of misinformation engagement and volume. A handful of additional or shifted associations emerge, but most cluster associations remain qualitatively similar. These findings indicate that the Main results are not driven by sceptical respondents, further strengthening robustness.

Robustness Checks with the Inclusion of Political Attitudes and Psychological Characteristics

The following figures present results from models that additionally include respondents' political attitudes and psychological characteristics. Figure C.14 shows the zero models, estimating the likelihood of engaging with misinformation, while Figure C.15 shows the positive models, estimating the volume of engagement among participants who shared at least once. These additions allow us to examine the potential role of political and psychological factors alongside socio-demographic and behavioural predictors, and the results indicate that these additional variables are generally not robustly significant, with the main patterns observed in the original models remaining largely unchanged. Table C.8 presents the adjusted McFadden R^2 values for each of the 5 zero and positive models.

Figure C.14: Zero models predicting sharing, including Political Attitudes and Psychological Characteristics.

Figure C.15: Positive models predicting volume of sharing, including Political Attitudes and Psychological Characteristics.

Table C.8: Model Fit: Adjusted McFadden R^2 for Zero and Positive Misinformation Engagement Models with the Main Outcome Variable

Model	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Zero models	0.048	0.093	0.095	0.269	0.298
Positive models	0.048	0.095	0.094	0.170	0.203