Political Elites, Misinformation, and Mobilization: Evidence from Brazil[1]

Frederico Batista Pereira[1], Natália S. Bueno[2], Felipe Nunes[3], João Pedro Oliveira[4], Nara Pavão[5], and Valerie Wirtschafter[6]

[1]Assistant Professor, University of North Carolina at Charlotte

[2]Assistant Professor, Emory University

[3]Associate Professor, Universidade Federal de Minas Gerais

[4]Independent Researcher

[5]Assistant Professor, Universidade Federal de Pernambuco

[6]Senior Data Analyst, Brookings Institution

Work in progress. Please do not cite without permission.

Note to TPBW: Thank you for taking the time to read our work. We appreciate it and we welcome all types of feedback. This is work in progress and, while we believe this manuscript already has plenty of evidence, we have resources for additional rounds of data collection (experimental and observational).

[1]Supplementary material for this article is available in the Online Appendix. Replication files are available in link TBD. The studies reported here were submitted to Emory University, UNC-Charlotte, Fundação Getulio Vargas's IRBs and were deemed exempt by research ethics committees. The research received funding from Facebook's Foundational Integrity Research and FAPEMIG. Our registered pre-analysis plan can be found at https://osf.io/g9tku.

Abstract

Misinformation is a growing concern among the public and political elites. Yet we still lack a good understanding of the political effects of misinformation. We argue that misinformation contains politically motivated content that sends unambiguous signals in favor or against a political group, thus increasing the salience of political identities and making individuals more oriented towards political and partisan goals. Consequently, disseminating misinformation benefits politicians because this type of story is effective at mobilizing voters. We empirically test this argument using novel observational and experimental data from Brazil. We show that politicians' posts that contain misinformation have higher levels of seemingly positive interactions with social media users in Brazil. Furthermore, respondents who are exposed to misinformation show a lower willingness to participate in campaign-related activities in favor of the target of the misinformation and have lower levels of affect regarding that target. Although the effects we find are small, they indicate that misinformation may pay off by damaging the target of misinformation.

Key Words: Misinformation, Global South, Mobilization, Partisanship, Political Elites, Social Media, Brazil

Word count: 11,817 (with references)

1 Introduction

While misinformation is a growing concern among the public and political elites, important studies suggest that its persuasion effects are negligible (Little 2018; Guess et al. 2020) and that this type of content is unlikely to directly affect election outcomes (Allcott and Gentzkow 2017). Others suggest that concerns about post-truth politics are exaggerated (Nyhan 2020). This skepticism is based on abundant evidence indicating that misinformation exposure and consumption is tightly associated with political identities and partisan motivated reasoning (Berinsky 2017; Flynn, Nyhan and Reifler 2017; Nyhan and Reifler 2010; Batista Pereira, Bueno, Nunes and Pavão 2022*a*), such that this content may reinforce existing beliefs but not significantly change voters' minds. Furthermore, existing evidence of large media effects on changing minds are fleeting and contextual, and, as with any piece of media, the effects of misinformation are also challenging to detect (Valkenburg et al. 2016). In light of this evidence, it is puzzling that misinformation, as a business and a political issue, is booming. If not persuasion, what are the electoral implications of elite-driven misinformation? We argue that even if it fails to persuade voters, misinformation could be an effective tool to mobilize them.

Political elites play a key role in disseminating misinformation (Nyhan 2020). Even if politicians are not the main source of political misinformation, they may represent a key "node capable of stemming its spread" (Mele et al. 2017, p. 8) and should be particularly influential in shaping public opinion and political behavior (Mele et al. 2017; Zhao et al. 2012; Zaller 1992). Recent work on the dissemination and effects of misinformation tends to focus on regular social media users in the United States (Guess, Nagler and Tucker 2019; Vosoughi, Roy and Aral 2018; Grinberg et al. 2019) and has overlooked the role of political elites (Nyhan 2018). As a result, we lack a clear understanding of misinformation's consequences for politics (Flynn, Nyhan and Reifler 2017).

We argue that misinformation benefits politicians because it is effective at mobilizing voters, inducing higher levels of political engagement and participation, than other types of content politicians share on social media. Political misinformation is mobilizing because its politically motivated content sends unambiguous signals in favor of or against a political group, thus increasing the salience of political identities and altering the identity-related payoffs of effort choice (Dickson and Scheve 2006; Horz 2018). Voters with political identities are more likely to be primed and mobilized by misinformation shared by political elites.

To empirically test these expectations, we combine novel observational and experimental data from Brazil, one of the largest democracies in the world and a country where political misinformation is abundant (Machado et al. 2019; Resende et al. 2019). The observational study relies on a number of approaches to identify misinformation in over four million posts by about one thousand political leaders in Brazil on Facebook, Twitter, and Instagram between 2018 and 2020. It compares engagement metrics across posts with and without misinformation to measure the association between misinformation shared by elites and online engagement by social media users.

We also conducted a large survey experiment in Brazil in the months leading up to the 2022 presidential election that randomly exposes five thousand study participants to real Facebook posts by politicians containing either misinformation, in the form of hyperpartisan news or false information, or phatic communication, a very common type of social media content that does not convey political information. After seeing these posts, study participants answer a battery of survey questions that assess their willingness to participate in a series of offline and online political activities related to the upcoming elections.

The observational study reveals that political elites very rarely share misinformation, yet about 14% of high-level political officials and candidates have shared at least one false story between 2018 and 2020. Importantly, posts containing misinformation attract substantially more attention in the form of reactions on social media. Furthermore, our evidence suggests that the high levels of engagement with misinformation online reflects support for the politician, and, to a lesser degree, criticism. We do not find social media users' questioning or correcting politicians' posts that contain falsehoods to any relevant degree.

The findings from our experimental study suggest that misinformation depresses participation in support of the group that is targeted by misinformation. We find that respondents exposed to posts shared by political leaders containing both false and hyperpartisan content are less willing to participate in events in favor of Lula and the Workers' Party (PT) – the target of the pieces of misinformation selected for our study – in addition to reducing affect towards supporters of the PT (*petistas*). As a consequence, we find that the political group in opposition to Lula and the PT benefits from sharing misinformation about their opponents.

Furthermore, our survey experiment, which contained a simulated Facebook environment, shows that misinformation posts are associated with higher engagement on several online actions, such as commenting on a post. Interestingly, findings from both forced exposure to misinformation (in the experiment) and selective exposure to misinformation (in the observational study) suggest that misinformation is effective at drawing reactions from individuals online, despite significant differences in the design of each study.

Our study contributes to research on political behavior and misinformation in three ways. First, it is the only study we are aware of that documents the effect of elite-shared misinformation on online engagement and willingness to participate politically. Our findings suggest that, while misinformation may not persuade individuals, it may still have an impact, albeit likely limited, on voters' behavior. Second, we find that both hyperpartisan and false content appears to (de)mobilize voters and we do not find systematic differences between these two components of misinformation; respondents do not appear to react differently to content that is false compared to content that is biased or misleading. Finally, our set of evidence suggests that misinformation pays off: our observational findings strongly suggest an association between elite-shared misinformation and users' online engagement that is, in many ways, beneficial to political leaders. In sum, misinformation is harmful to the target of the false and biased stories, without engendering substantial backlash to the political group that shares falsehoods and hyperpartisan content.

In what follows, we detail why we expect misinformation to increase political mobilization. Then, we present our study describing the relationship between misinformation shared by politicians in social media and online engagement. Finally, we outline our survey experiment designed to test the causal relationship between misinformation shared by political elites and political engagement. We then discuss the main findings and implications for future research.

2 Misinformation and Political Mobilization

Political news are part of a supply-and-demand market in which the media and political elites supply information that voters demand (Allcott and Gentzkow 2017). This market has recently undergone drastic changes. The rapid growth of online information ecosystems has weakened traditional gatekeepers of information—such as mainstream media and professional journalists—and restored political elites' ability to operate as active and direct suppliers of information (Shirky 2008). In social media environments, political elites can easily tailor their discourse to resonate with members of politically homophilic echo chambers (Ecker et al. 2022), and spread (intentionally or unintentionally) highly biased or false content that sends unambiguous signals in favor of or against a political group (Nyhan 2020; Tucker et al. 2018). As a result, the internet has become an ideal medium for the proliferation of misinformation (Ecker et al. 2022; Vosoughi, Roy and Aral 2018). Because political elites have many social media followers (Pennycook and Rand 2021) and are capable of shaping political perceptions and behaviors (Zaller 1992), their ability to spread of misinformation can be particularly consequential (Ecker et al. 2022).

While misinformation shared online by political elites is unlikely to reach non-supporters due to audience fragmentation, algorithmic personalization implemented by social media platforms, and selective exposure (Guess and Reifler 2018; Cinelli et al. 2021; Bakshy, Messing and Adamic 2015), which drastically limits the potential for persuasion, it can reach supporters and nonetheless have important consequences. Studies confirm that misinformation tends to have minimal persuasive effects (Little 2018; Guess et al. 2020; Nyhan 2020), leaving the puzzle of why misinformation is a booming issue in politics unsolved. Recognizing that persuasion might not be the main goal behind the spread of misinformation (Little 2018; Tucker et al. 2018), we propose that the spread of misinformation in politics may be attractive to politicians because this type of content is effective at mobilizing voters.

Misinformation is commonly conceptualized as knowingly false, misleading, or unsupported claims about factual matters (Lazer et al. 2018; Guess, Nagler and Tucker 2019). In academic and public discussions about misinformation, related terms such as disinformation, propaganda, rumors, and fake news are used interchangeably. Each of these terms highlights a different component of communication: disinformation implies intention to deceive, fake-news denotes the mimicking of articles published in established media, and rumors refer to the social transmission of misinformation (Guess and Lyons 2020).

Above these specific variations in the type of communication, political misinformation has two defining components: a false content and appeals to group identity. In political contexts, elite communication does not deviate from reality in a random way; instead, it has a specific direction and is biased to bolster or harm political groups. The false content represents the informational dimension of misinformation and it refers to the presence of lies or statements whose content is knowingly false. The second component of misinformation refers to group rhetoric and appeals to political identities. In practice, the misinformation we see in politics is likely to be a mix of lies and appeals to group identity (Horz 2018). In this paper, we use the term misinformation to refer to communication that contains one of these key defining components. More specifically, misinformation encompasses communication that always has politically motivated appeals, but that has either blatantly false or highly misleading but not undeniable false statements.

But how does misinformation mobilize voters? While a false content is a key defining aspect of misinformation, we argue that the politically motivated nature of such content is the active ingredient that triggers its mobilizing effects. Political engagement and participation require resources like time, cognitive energy, money and civic skills (Dawkins 2017). Voters' decisions to participate in these costly efforts are a function of the benefits they receive from this participation. Voters' effort utility function includes policy outcomes, but also identity-related payoffs. In other words, political identities can serve as important motivation for effort choice (Dickson and Scheve 2006; Horz 2018).[2]

When campaigning, politicians can strategically use political speech and propaganda to mobilize voters by appealing to their political identities. Studies suggest that the activation of political identities by politically-motivated content makes people more oriented towards political and partisan goals (Delton, Petersen and Robertson 2018; Druckman, Peterson and Slothuus 2013). Political motivations, or the "the tendency to strive toward particular end states or goals" related to politics (Jerit and Zhao 2020, p. 79) may influence not only the internal processes through which we seek and evaluate information, known as motivated reasoning (Druckman, Peterson and Slothuus 2013), but also our disposition to engage in behavior that promotes these political goals and end states (Huddy, Mason and Aarøe 2015; Delton, Petersen and Robertson 2018).

This could happen because the salience of political identities determines "the strength of behavioral prescriptions associated with membership in social groups" (Dickson and Scheve 2006, p. 6). In the context of elections, these behavioral prescriptions may instruct voters to act to boost the status and position of their political group. When political

[2]We note that, while political motives are often associated with partisanship, i.e., with identities shaped around political parties, these motives may also be rooted in other forms of political and social identities. These other types of identities also give rise to the political goals and motivations that drive political action. In contexts where partisanship is weak and limited—which is the case of Brazil (Samuels 2006; Samuels and Zucco 2018) and other developing nations (Lupu 2015; Mainwaring and Scully 1999; Mainwaring and Torcal 2006)— individuals' political motivations may be centered around political leaders. In fact, multiple studies indicate that misinformation often activates political and social priors that are not necessary rooted in partisan identities (Flynn, Nyhan and Reifler 2017; Berinsky 2017; Badrinathan 2021).

identities are made salient, individuals' self-esteem is more dependent on their political group and the position it occupies in the competition. Under these conditions, the costs of deviating from one's group prescriptions increase, and voters' will have more incentives to participate in campaign-related activities that promote their preferred candidate (Dickson and Scheve 2006). By activating political identities and enhancing political motives, misinformation spread by politicians is likely to induce voters to participate in electoral activities that may benefit political allies.[3]

There is evidence that mobilization efforts can be effective at changing behaviors (Gerber and Green 1999) and that the activation of social and political identities plays a key role in this process (Bryan et al. 2011; Dickson and Scheve 2006). Also, we know that fake news spreads faster and broader than real news due to human action, as opposed to automatized profiles (bots), indicating that voters might be particularly attracted to this type of story (Vosoughi, Roy and Aral 2018). Guess et al. (2020) find that controlled exposure to a misinformation article reinforced politically congenial beliefs and increased intentions to vote, although it had no effect on other forms of political participation.

If our expectations are correct, we should find support for the following hypotheses:

Hypothesis 1 (H1) *Posts containing misinformation generate more engagement and participation than posts that do not contain misinformation.*

Because the hyperpartisan dimension of misinformation is what drives mobilization and this component is always present in political misinformation (regardless of whether it contains strictly false information), we expect both posts that contain false statements and posts that lack a blatantly false content to promote higher levels of engagement and participation:

[3]Misinformation is more likely to activate political identities than other common types of political messages, like those that focus on policy. However, if policy motivated claims contain strong appeals to group identities, such as policies related to divisive issues (e.g., abortion policy) they can be as mobilizing as misinformation.

Hypothesis 1A: Posts containing false stories generate more participation than posts that do not contain misinformation.

Hypothesis 1B: Posts containing hyperpartisan news generate more participation than posts that do not contain misinformation.

There are reasons to be skeptical of the mobilizing power of the false content of misinformation. We know that, due to limited analytic thinking, people rarely evaluate misinformation in terms of truthfulness (Pennycook and Rand 2019), and that accuracy is not the dominant driver of sharing misinformation online (Chen, Pennycook and Rand 2021). Like other authors, we relax the assumption of Bayesian learning, implying that voters' belief in the content shared by politicians does not condition the effects of misinformation (Little 2017; Horz 2018; Dickson and Scheve 2006). In other words, misinformation can be seen as a form of communication that is capable of changing behavior, promoting engagement and participation, even if voters do not believe it (Horz 2018; Little 2017). If this is correct, we should find that the two types of misinformation (blatantly false and highly misleading) are equally effective at mobilizing voters, contingent on their content being equally perceived as political.[4] However, if we find that false stories are more mobilizing than hyperpartisan news, this should be taken as evidence that the unique *fakeness* component of misinformation accounts for mobilization effects. However, because we lack strong theoretical basis for developing hypotheses about the differential effects of false and hyperpartisan statements, we will answer the following research question:

RQ1: Do posts containing falsehoods have a different effect on levels of mobilization compared to posts that contain hyperpartisan content?

To better understand the electoral consequences of misinformation shared by political elites, we also explore whether false and hyperpartisan content generates net mobilization effects, i.e., whether it promotes political engagement that favors a specific political group

[4]Data from a pilot survey indicate that, on average, the fake and hyperpartisan news stories we use in our experiment present similar levels of perceived political content.

(either the source or target of the post containing misinformation). To assess whether posts containing misinformation generate, on average, more mobilization in favor of the target (PT/Lula) or the source (Bolsonaro) than posts containing phatic communication, we test the following hypotheses:[5]

Hypothesis 2 (H2) *Posts containing misinformation generate higher levels of net mobilization (i.e., more mobilization against Lula or in favor of Bolsonaro) than posts containing phatic communication.*

Hypothesis 2A: Posts containing false stories generate more net mobilization than posts containing phatic communication.
Hypothesis 2B: Posts containing hyperpartisan news generate more net mobilization than posts containing phatic communication.

According to our argument, misinformation mobilizes because it activates political identities, making individuals more motivated to act to promote the status of their preferred political group. If this argument is correct, we should observe an increase in affective polarization among study participants exposed to misinformation shared by politicians.

Hypothesis 3 (H3) *Posts containing misinformation are more likely to increase affective polarization towards political groups than posts containing phatic communication.*

Hypothesis 3A: Posts containing false stories are more likely to increase affective polarization towards political groups than posts containing phatic communication.
Hypothesis 3B: Posts containing hyperpartisan news are more likely to increase affective polarization towards political groups than posts containing phatic communication.

[5]These hypotheses are included in the pre-analysis plan as exploratory.

The theoretical discussion supporting the main hypotheses above also suggests that not all voters will respond equally to misinformation shared by political elites. Political identities should significantly shape how voters respond to this content. More specifically, those who have a strong political identity should be more likely to be mobilized by misinformation than those who are indifferent (or who lack a political identity) towards relevant political groups. If these expectations hold, we should find evidence in support of the following hypotheses:[6]

Hypothesis 4 (H4) *Posts containing misinformation are more likely to increase participation among voters with political identities than among voters without political identities.*

Hypothesis 5 (H5) *Posts containing misinformation are more likely to increase net mobilization among voters who are supporters of Bolsonaro than among voters who are supporters of Lula and/or the Workers' Party.*

In this study, we investigate the mobilizing effects of misinformation shared by political elites by looking at different forms of political participation that range from low-cost activities often referred to as *slacktivism* (such as likes and other forms of reactions available on social media platforms) (Kwak et al. 2018) to willingness to engage in more costly offline and online campaign-related activities. Throughout this paper, we use the term engagement to refer to social media reactions, and the term participation to refer to willingness to participate in a series of online and offline campaign activities.

Next, we present our observational study in which we describe the association between misinformation by political leaders and online engagement.

3 Does Misinformation Promote Online Engagement?

In this section, we empirically describe the extent to which politicians share misinformation online in social media platforms, and whether posts that contain misinformation

[6] H5 was included in the pre-analysis plan as exploratory.

are associated with higher levels of online engagement.

3.1 Data

Our sample of politicians is composed of Brazilian political leaders in different elected and unelected positions: the president, vice-president, cabinet members, governors, vice-governors, federal deputies, senators,[7] and candidates for mayoral office in all state capitals in 2020, resulting in a total of 945 politicians with active accounts in either Facebook, Instagram, or Twitter. We collected, via CrowdTangle, approximately 1.5 and 1.1 million politicians' posts from Facebook and Instagram, respectively, and approximately 1.5 million politicians' tweets from Twitter, via Twitter Academic API.

We rely on four different approaches to identify posts that contain fake and hyperpartisan content, the two components of misinformation we analyze in this paper.[8] To identify posts that contain hyperpartisan news stories, we rely on the "domain approach" and on the "repeated domain approach." To detect misinformation in the form of false stories, that is, content that is both false and hyperpartisan, we use the "text based approach" and the "URL approach."

The first approach—"domain approach"—follows standard practice in existing literature and classifies posts as containing hyperpartisan content if they include a hyperlink to "low-quality/credibility" domains (Allcott and Gentzkow 2017; Guess and Reifler 2018).[9] We obtained these domains from Facebook's URLs dataset[10] that contains URLs shared on the platform that were flagged as false by third-party fact-checkers. We extracted 228

[7]We collect social media data for those elected in 2018. For senatorial seats elected in 2014, we collected data for senators in office in January 2018. For cabinet members, we collected data on all members up to January 2020.

[8]Parts of this text will share similarities with our working paper (Batista Pereira, Bueno, Nunes, Oliveira, Pavão and Wirtschafter 2022) because it described in detail the methods we utilize for measuring misinformation.

[9]This type of measurement is common practice in studies focusing on the United States, where such list of "low-quality/credibility" domains is established.

[10]The original dataset contains data about more than 38 million URLs shared more than 100 times publicly on Facebook. For our analysis, we narrowed in URLs shared widely in Brazil.

domains from these URLs.[11] Anytime a politician shared a link from one of these domains, we classify the post as containing hyperpartisan content. The second approach we use to detect hyperpartisan content shared by politicians—"repeated domain approach"-—offers a more conservative measure that relies on the same principles of the previous approach, but restricts the list of "low-quality/credibility" websites to only those domains flagged more than one time by Facebook's URLs dataset. Although the domains used in these approaches come from a list of websites that at some point shared content identified as false, not all posts from them contain false information, but rather mostly polarizing content that we treat here as hyperpartisan.

To identify posts that contain fake stories, we first rely on the "URL approach." We simply extract the complete URLs—rather than just the domain—flagged as false by third-party fact-checkers on Facebook and classify as fake those posts that contain one of these URLs. Finally, we also rely on a novel text-based approach to detect false stories in posts shared by politicians. In this approach, we compare the text of rumors that circulated in Brazil to the text of politicians' social media posts. We train a classification model (Naive Bayes) on a dataset containing the text of 4,050 false stories and 2,000 social media posts that did not contain any misinformation.[12] We then apply this model to the entire dataset of about 4 million social media posts by political elites in Brazil, which yields a predicted probability of containing a false story at the post level. To further trim our dataset and guard against false positives, we eliminated any posts with a predicted probability of containing false stories below 0.9. We then calculate the cosine similarity between each of these posts "at risk" of containing misinformation with each false story we scraped from the independent fact checking organization (*Boatos*). Finally,

---

[11] We exclude common domains such as youtube.com, twitter.com, yahoo.com, and a main newspaper from this list.

[12] These posts were initially selected based on a supervised learning model of false stories and stories from reputable, well-established news agencies, and human review of each post against of entire dataset of 15,000 fact-checking stories. We built models separately for each social media platform. The F-1 score varied between .940 to .954 and the prediction accuracy ranged from 92.40% and 96.07%.

three coders conducted a human review of each post with a cosine similarity higher than 0.4. We detail these steps, our tests for reliability and validity for this measure elsewhere (Batista Pereira, Bueno, Nunes, Oliveira, Pavão and Wirtschafter 2022).[13] By design, the text-based and the URL approaches are much more likely to detect strictly false content than the domain approaches. Both approaches rely on fact-checked content deemed as false and compare politicians' social media posts against that content. Therefore, by definition, all the posts identified by the text-based and URL approaches have content that was flagged as false by fact-checking organizations.

On the other hand, the domain approaches (domain and repeated domain), also by design, are not capturing strictly false content. For example, we selected a random sample of 200 posts identified by the domain approach. Out of the 139 posts with working links, 131 (95%) stories were not checked by any fact-checking agency. Of the 8 stories that were checked by a fact-checking agency, 7 were deemed false or misleading and 1 story was checked and classified as true. Furthermore, out of a sample of 2,516 posts containing hyperlinks to domains listed in the domain approach, only 5 posts were fact-checked by social media companies (based on visual tags).

Despite not capturing strictly false stories, the domain approach is effective in identifying content that is polarizing, or generally speaking, hyperpartisan. Research assistants classified the news linked in 200 posts identified as containing misinformation by the domain approach according to their level of "politicization/partisanship." On average, 70% of these posts were classified as "politicized/partisan," or "clearly partisan/politicized". About 10% of these posts were classified as "a little partisan/politicized" and about 20% were deemed as having a politically neutral content.

We measure online engagement in multiple ways. We create an aggregate measure

[13]This "text-based" approach has several complementarities to the other approaches described here. The text-based approach does not restrict the detection of misinformation to content shared via hyperlinks, which is very useful since misinformation that circulates in Brazil tends to be unaccompanied by hyperlinks. Furthermore, this approach examines the content posted, and specifically identified false content rather than identifying the presence of hyperlinks at risk of containing misinformation due to its origins in websites and portals that have a history of sharing misinformation.

of reactions to posts across all platforms (Facebook, Instagram, and Twitter). Reactions include the sum of all Facebook and Instagram's interactions (share, comment, love, like, angry, etc., depending on platform), and the sum of Twitter's interactions (reply, quote, like, retweet). Also, since all three platforms included in the analysis have a "like" form of reaction, we examine "likes" separately. Finally, for platform-specific analyses, we break down types of reactions separately.

3.2 Misinformation Is Associated with Online Engagement

A relevant share of politicians have shared either misinformation or hyperpartisan content, according to our measures. About 1.5% of the politicians shared false content (URL approach), while 44% shared hyperpartisan content at least once (domain approach). At the same time, misinformation is a small part of politicians' "online menu:" fewer than 1% of their posts contain misinformation (Table A3). Our first results compare online reactions to posts containing misinformation or hyperpartisan content, measured in our four approaches (fake stories in text and URL approaches, hyper-partisan in the domain and repeated domain approaches), to posts without those two features.

As shown in Tables A4 and A5, posts with misinformation and hyperpartisan content, across three of the four measures, have higher levels of engagement in terms of overall reactions and likes across all platforms. The exception here are posts that were visibly tagged by Facebook as containing misinformation (the URL approach).

Yet the degree to which this content is correlated with online engagement is a function of several factors, including the type of measure of online behavior, platform, user, and the timing of posting. Furthermore, sharing misinformation systematically varies with politicians' partisanship and ideology (Batista Pereira, Bueno, Nunes, Oliveira, Pavão and Wirtschafter 2022). As a consequence, chief among the challenges in estimating the relationship between sharing misinformation and online engagement are the systematic differences between politicians who share misinformation and those who do not and the type of online audiences that interact with them. For example, politicians who are more active on social media and who campaign on polarizing issues may be more likely to both share misinformation and have greater online followings. Also, platforms vary in terms of following and engagement, and if we find systematic differences in the frequency and types of misinformation sharing by platform, it could mask a true underlying relationship between online activity as a response to misinformation. Finally, it is conceivable that misinformation is more likely to be posted during political upheaval, also leading to more online reactions to posts.

To address these potential issues, we rely on fine-grained fixed effects models. Tables 1 and 2 compare posts within the same politician, same platform, the same month, and the same year. The differences between posts with and without misinformation and hyperpartisan content are quite large for three out of four models; a post with false stories (measured via the text approach) is associated with an increase of 172% in online reactions and a post with hyperpartisan stories (captured by the repeated domain approach) is correlated with an increase of 75% in online reactions. The online response to misinformation is the same if we measure "likes" instead of overall reactions, as shown by Table 2. Using cluster robust standard errors at the politician level, we reject the hypothesis of no effect of misinformation on online engagement at the 5% level for all models, but for hyperpartisan stories measured via the domain approach. Furthermore, we find in Table 1 that posts with false stories (text approach) are associated with more reactions (diff = 0.43, p-val = 0.05) than hyperpartisan posts (repeated domain), but the same is not true for likes (diff = 0.14, p-val = 0.33) shown in Table 2.

Table 1: Correlates of Online Engagement (Reactions): Types of Misinformation

	Reactions (Log)	Reactions (Log)	Reactions (Log)	Reactions (Log)
Fake (Text)	1.002***			
	(0.076)			
Fake (URL)		0.653*		
		(0.283)		
Hyper. (Domain)			0.083	
			(0.115)	
Hyper. (Repeat Domain)				0.564**
				(0.200)
Num.Obs.	4 032 907	4 032 907	4 032 907	4 032 907
RMSE	1.19	1.19	1.19	1.19
Std.Errors	Pol.	Pol.	Pol.	Pol.
FE: Pol-Plat-Mon-Year	Y	Y	Y	Y

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimates from Tables 1 and 2 are useful because they allow the investigation of variation in reactions to content while adjusting for politician, platform, and timing factors. Yet using posts as unit of analysis suffers from the high degree of variability of online engagement with posts and it fails to capture more sustained mobilization over time.

To examine whether misinformation is associated with continued online engagement, we compare levels of mobilization between a week in which a politician shared a post containing misinformation or hyperpartisan content to weeks without such posts. All analyses are within the same politician, same month, and same year, and we include (logged) total posts as a control. Figure A1 shows similar results. Across all measures of online engagement, but likes, politicians have about 10% to 25% higher levels of online engagement in weeks in which they share false stories (identified by the text approach) than in weeks without falsehoods. At the same time, across all measures of online engagement, we fail to see any statistically significant association between mobilization and posts containing false stories via the URL approach.[14] The findings for weeks contain-

[14]There is also a lot more uncertainty around these estimates due to small number of weeks with posts identified via the URL approach.

Table 2: Correlates of Online Engagement (Likes): Types of Misinformation

	Likes (Log)	Likes (Log)	Likes (Log)	Likes (Log)
Fake (Text)	0.668***			
	(0.093)			
Fake (Full URL)		0.614**		
		(0.199)		
Hyper (Domain)			−0.009	
			(0.112)	
Hyper (Repeat Domain)				0.473**
				(0.181)
Num.Obs.	4 032 907	4 032 907	4 032 907	4 032 907
RMSE	1.30	1.30	1.30	1.30
Std.Errors	Pol.	Pol.	Pol.	Pol.
FE: Pol-Plat-Mon-Year	Y	Y	Y	Y

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ing hyperpartisan posts are more mixed: we observe associations that are statistically distinguishable from zero for comments (domain approach), shares (domain and repeat domain), and borderline statistical significance for reactions (p-val= 0.11), and estimates range from 2% to 8% more engagement.[15]

Overall, our post-level analysis suggests a large and systematic association between online engagement around posts containing misinformation, whereas our week-level analysis suggests a more muted association between "continued" levels of online engagement and misinformation. This analysis, however, is not informative about the type of engagement that misinformation is associated with on social media. Users may be engaging more with false or hyperpartisan content shared by politicians, but this engagement could reflect efforts to question false information or admonish politicians for sharing it.

To examine the nature of engagement around content shared by political leaders on social media, we break down engagement by different metrics (likes, comments, share, love, wow, angry, sad, care, overall combined reactions as a sum of of all of those) on

[15]In this current analysis, we aggregate at the week-level, meaning that the story could have been posted early or late in that week. As a result, our analysis is likely a combination of reactions issues post or prior to posting false/hyperpartisan stories). Furthermore, controlling for total number of posts could be inducing "post-treatment biases." In future versions of this paper, we will add analyses looking at seven days after posting false/hyperpartisan stories.

Facebook. We follow the empirical strategy from Table 1 by comparing posts with and without misinformation within the same politician, year, and month (platform is held constant because we only look at Facebook for this analysis), but we instead use disaggregated metrics based on the type of online reaction rather than the overall combined measure of reactions.

Figure 1 shows a large positive association between posts containing false stories and reactions such as comments, shares, wow, angry, and sad, especially when false content in the posts is identified via the text-based approach. We also find associations between false stories measured via the text approach and the URL approach (and more uncertainty around estimates for the URL approach given the small number of posts directly tagged by Facebook) for likes, and haha, but the size of the association estimates tend to be smaller. The patterns and magnitudes are similar between posts containing hyperpartisan content and those containing false content identified in the URL approach, even if estimates tend to be slightly more muted for posts identified via the domain approach. Finally, we find no positive association between posts containing misinformation in any measure and the reactions love and care.[16]

While informative, these Facebook reactions are limited in what they can tell us about whether social media users believe misinformation shared by elites, question it, or use it to support their own political group and attack opponents. Furthermore, although there are interesting differences in types of metrics for posts containing falsehoods identified in the text approach, for most of the other measures, even if the point estimates differ, we cannot distinguish between estimates as statistically different from each other.

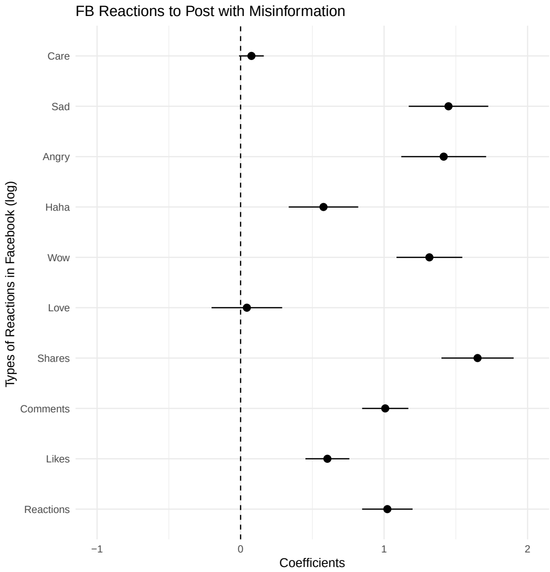
We turn to Twitter as a way to assess the content of reactions to tweets containing falsehoods. We randomly sampled 500 quotes and 500 replies to tweets containing falsehoods identified the text approach.[17] Two research assistants, working independently,

[16]Indeed, we find a negative association between love and posts with hyperpartisan content (domain).

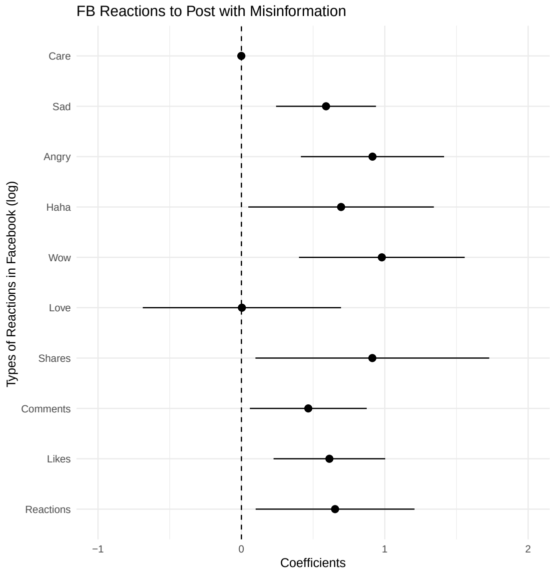
[17]We are in the process of coding quotes and replies for tweets sharing misinformation measure via the domain and the repeated domain approaches.

Figure 1: Association between Posts Containing Misinformation and Different Types of Online Reactions in Facebook

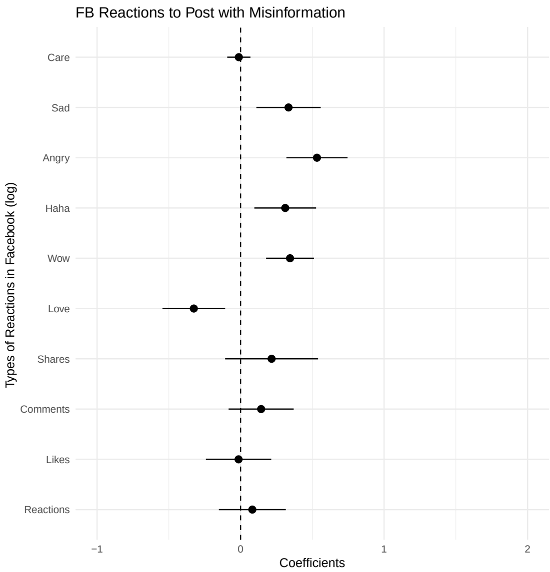
(a) False (Text Approach)



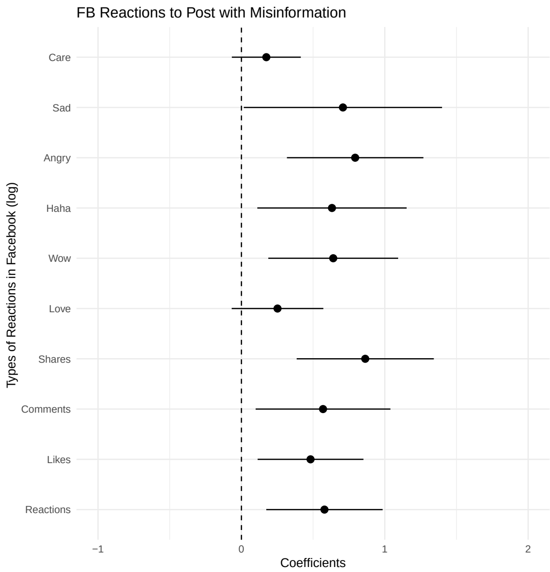
(b) False (URL Approach)



(c) Hyper. (Domain Approach)



(d) Hyper. (Repeated Domain)



Notes: All models contain standard errors clustered at the politician level and politician-month-year fixed effects. Platform is not included because this Figure only includes Facebook posts).

classified each quote or reply by answering several questions about the content of the tweet. We created three aggregate indicators for whether the quote/reply: (1) questions (fact checks, raises a question, or corrects), (2) attacks (attacks the politician, the content of the tweet, or mocks), or (3) supports (shows support for the politician or tweet content) the tweet containing a falsehood.[18] Figure 2 shows the percentage of tweets, according to both coders,[19] in which users either questioned, supported, or attacked the politician or the content of the tweet identified as containing falsehoods via the text approach.

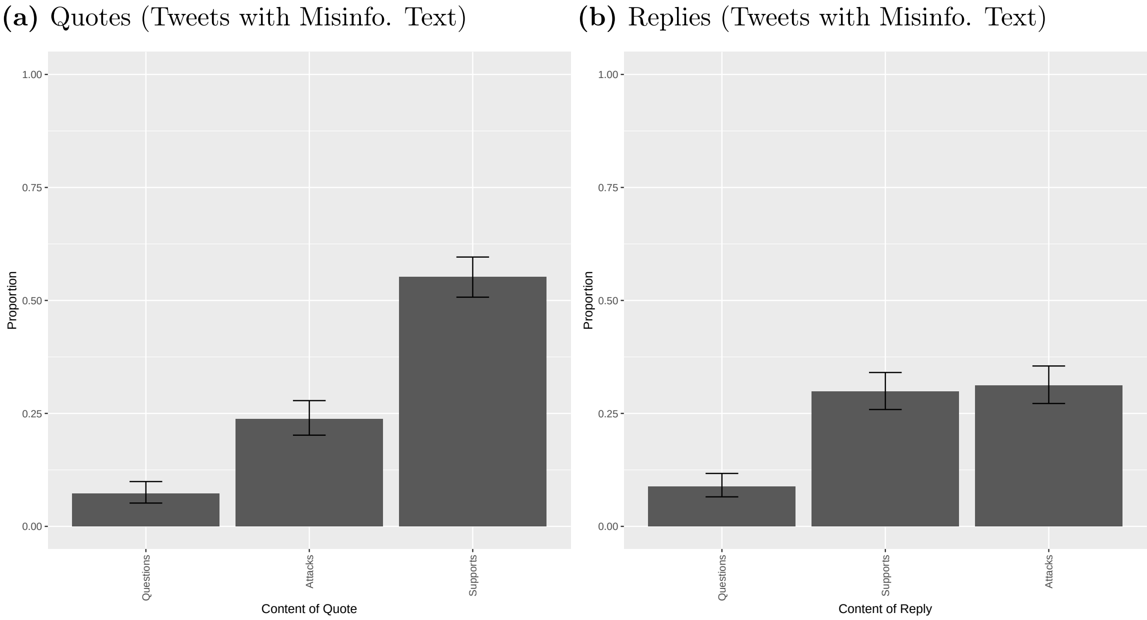
Overall, a small share of users correct or even question tweets containing falsehoods – in fact, this is even likely an overestimate because we allowed for mild questioning that gave the politician the "benefit of doubt." The most common behavior in over 50% of the quotes we analyzed is supporting the politician/tweet. Second, with just under 25% of quotes, we observe their content as attacking the politician/tweet. With replies, we observe a more even split between supporting and attacking the politician/tweet.[20]

[18]We also identified other behavior, such as attacking/defending some other politician unrelated to the tweet. See Figures A2 and A3 for the disaggregated categories.

[19]See A2 and A3 for the separate classifications for each research assistant.

[20]We also analyze "getting ratioed." We find no support for the claim that tweets containing falsehoods are more likely to "getting ratioed" compared to tweets without misinformation (see Table A10).

Figure 2: Types of Quotes to Tweets Containing Fake Stories



3.3 Discussion

Our descriptive findings are consistent with the claims that misinformation is associated with mobilization. Social media users react more to posts containing misinformation and hyperpartisan content, than to posts without those features. Importantly, our findings are not driven by the different types of politicians who share misinformation, the platforms they use, or the moments in which they post misinformation as we conduct most of our analyses within the same politician, platform, year, and month.

Posts containing misinformation, however, are rare. As shown in Table A20, fewer than 1% of posts by political leaders contain misinformation. While politicians rarely post misinformation, they are very active on social media (posting, on average, six times a day), meaning that politicians have shared, at a maximum, 23,227 posts containing misinformation (detected via the domain approach) and at least 421 posts containing falsehoods (identified via the text approach). Furthermore, as with most online behavior data, reactions' to politicians' post are highly right-skewed, and posts containing misinformation not only have higher average reactions, but also medians as much as 10 times larger than posts without misinformation (1,138 vs. 118 reactions), suggesting their impact is not undermined by their low prevalence. Our descriptive findings, rather than focusing on the prevalence of misinformation shared by elites, attempt to uncover the ties between elite-endorsed misinformation and political mobilization.

That said, our claims are limited by threats to causality and measurement. Regarding measurement, online engagement does not necessarily translate into offline or online political mobilization. While distinctions between online and offline behavior may be irrelevant for some political questions, several metrics such as liking, commenting, or even sharing a post may not constitute political actions and may simply be result of exposure on social media without further impact on political attitudes and behavior. Furthermore, even if these metrics of online engagement are associated with some degree of underlying political mobilization, we cannot attribute a causal impact of being exposed to misinformation on political action as posts with and without misinformation vary in many respects, including the political moment in which they are written, their intensity, and tone. We now turn to a survey experiment designed with the goal of measuring whether misinformation can cause electorally-relevant mobilization.

4 Does Misinformation Cause Participation?

We assess the causal effect of misinformation shared by political elites on intention to participate in campaign-related activities via a survey experiment fielded between August 17 and September 2, 2022 in Brazil, two months before the presidential election. *Quaest Consultoria & Pesquisa* conducted the study online with 5,000 respondents. It consisted of presenting respondents with misinformation politicians shared on Facebook[21] and measuring the extent to which it induced higher levels of participation in campaign-related activities.

In order to evaluate the mobilizing potential of misinformation shared by political elites, we compare politicians' social media posts with misinformation (false content) to

[21] While we found evidence of politicians sharing misinformation in Instagram, Twitter, and Facebook, we choose Facebook posts because we found that politicians were more active on Facebook than other platforms.

content that is hyperpartisan but not false and content containing phatic communication, i.e., "speech which is used to express or maintain connection with others in the form of shared feelings, goodwill or general sociability, rather than to impart information exchange." (Miller 2017, p. 3). We chose phatic communications as our reference category for two reasons: (1) this kind of communication is representative of the interactions people frequently have on social media (Berriche and Altay 2020); and (2) it is not politically goal-oriented (Miller 2017, p. 3). Data from a pilot study confirms that phatic posts in our study are more likely to be perceived as having politically neutral content than posts that contain misinformation.

After answering a battery of pre-treatment socio-demographic and political questions, respondents were randomly assigned to one of three experimental conditions, containing either false, hyperpartisan or phatic posts.[22] To ensure that our results are not driven by the particularities of a single source politician (the politician who is posting the information on Facebook), false news story, or phatic post, we selected all three types of posts from five different politicians aligned with president Jair Bolsonaro.[23] All posts selected refer to the target politician/political group (Lula/PT) by name and, if they include someone else, these individuals are well-known close associates of Lula (Lula's spouse and close political ally).[24] Furthermore, we restrict our analysis to negative pieces of misinformation for two reasons: first, most misinformation is negative and second, following the literature on negative campaigns, we expect negative misinformation to be more effective than positive misinformation at mobilizing voters (Batista Pereira, Bueno, Nunes and Pavão 2022*a*).

[22]See posts used in the three conditions in section K in the Appendix.

[23]For example, suppose we select politicians A and B. Then, we select one false stories post from politician A and one for politician B, one hyperpartisan post from politician A and one from politician B, and one phatic post from politician A and one from politician B. Because we will collect three types of posts (false, hyperpartisan, and phatic) from five different politicians, we will have a total of fifteen posts included in the study.

[24]Data from our pilot suggests that posts with false and hyperpartisan stories were perceived, on average, to be more "pro-Bolsonaro" than phatic posts.

We select phatic posts by compiling all posts shared by each of the five politicians whose posts were included for the false and hyperpartisan treatment conditions. We then used unsupervised topic modeling (Latent Dirichlet Allocation) to identify the topic associated with posts that contained phatic communication and to select the five specific posts that were presented to participants assigned to the phatic condition.[25]

Given the null findings in the literature regarding the effects of misinformation, we expose each respondent to five posts – a "high dosage" treatment. For instance, a respondent randomly assigned to the false news condition is exposed to five different posts by different politicians that contain blatantly false stories. Similarly, a participant assigned to the hyperpartisan condition sees five posts containing hyperpartisan content shared by the same five politicians. A participant assigned to the phatic communication condition is presented with five phatic posts shared by the same five politicians. At the end of the survey, respondents assigned to the misinformation and hyperpartisan conditions are also presented with corrective information and tips on how to identify misinformation.[26] Exposing subjects to misinformation raises valid and important ethical concerns that we discuss in detail in section D in the Appendix.

4.1 Survey Measures

We rely on a battery of ten questions that ask respondents if they would be willing to participate in different types of campaign-related activities to support or oppose Lula/PT (the target of posts containing misinformation) or Bolsonaro (which represents the political group that is the source of posts and stands to benefit from Lula/PT losing support).[27]

[25]In the pilot, we tested two different types of content of phatic posts: Mother's Day greetings and cities' anniversary celebrations. Because the results indicated that the former type of content is less likely to be perceived as political, we use politicians' posts that contain Mother's Day greetings as our control.

[26]See Batista Pereira, Bueno, Nunes and Pavão (2022*b*) evidence on the effectiveness of these tips on reducing belief in misinformation.

[27]We also measure participants' willingness to engage directly with the source politicians (i.e., the politicians whose posts are presented to participants) by asking them about their disposition to follow the source politicians on Facebook.

In addition to these survey questions, we also collect participants' online engagement with posts that we present in the experimental manipulations using a stylized newsfeed called Mock Social Media Website Tool (MSMT). The instrument used in the experimental study is included in section K of the Appendix. The table below summarizes the study outcomes and how they are measured.

Table 3: Outcomes

Outcomes	Measurement
Mobilization in favor of target of the post	Willingness to... 1. convince someone to vote for Lula 2. share positive content about Lula on social media 3. attend a campaign rally in favor of Lula 4. display a campaign sign, bumper sticker or flag in favor of Lula 5. join a WhatsApp group to receive information and campaign material about Lula
Mobilization against the target of the post	6. dissuade someone from voting for Lula 7. share negative content about Lula on social media 8. attend a campaign rally in favor of Bolsonaro 9. display a campaign sign, bumper sticker or flag in favor of Bolsonaro 10. join a WhatsApp group to receive information and campaign material about Bolsonaro
Affect	11. Affect towards Petistas 12. Affect towards Bolsonaristas
Online engagement	13. Reactions to posts (like, love, haha, wow, sad, angry, or leaving a comment or text)

To test hypothesis 1, we rely on the overall mobilization index (OMI) that combines all outcomes that tap mobilization in favor of and against the two political groups (questions 1-10). In this index, all outcome measures indicating action—regardless of whether they favor or oppose the two political groups—have positive values. This index is designed to capture participants' general dispositions to engage in political/electoral actions. Testing hypothesis 2 requires measuring the direction of mobilization—whether it favors or opposes a specific political group. We construct a net mobilization index (NMI), using questions 1-10, in a manner such that positive values indicate mobilization against the target politician (PT/Lula, pro-Bolsonaro) or in favor of Bolsonaro, and negative values indicate mobilization in favor of Lula/PT and against Bolsonaro. Both OMI and NMI are created using a mean effects index (Kling, Liebman and Katz 2007).

We also collect measures of online engagement (item 13 in Table 3), comparable to those used in our observational study, from the *Mock Social Media Website Tool (MSMT)* (Jagayat, Arvin, Gurkaran Boparai, Carson Pun, and Becky L. Choma N.d.). This tool simulates an interactive Facebook timeline and allows study participants to react to posts as they would on their own Facebook feeds. We use the MSMT at the end of the questionnaire[28] to expose study participants to the same posts they saw previously (as part of the experimental manipulation) and to collect their online reactions (like, love, haha, wow, sad, angry, or leaving a comment or text) to the posts in the simulated Facebook timeline. We examine all reactions (like, love, haha, wow, sad, angry, or leaving a comment or text) with the posts in the simulated Facebook timeline as a measure of online engagement.[29] We construct a measure of affect towards political groups (*bolsonaristas* and *petistas*) using items 11 and 12 in Table 3 to test hypothesis 3. These items are measured on a scale from 0 to 10 in which 0 means lowest and 10 highest affect regarding Bolsonaro and PT. We measure affective polarization as the absolute value of the difference between the two political groups as a intermediate outcome that could be partially responsible for driving the effects of misinformation on political mobilization.[30]

Our expectations (hypotheses 4 and 5) require we compare the effects of the misinformation treatment across participants with and without a political identity. We classify as having a political identity respondents who declare to either like or dislike

[28]Only a subset of respondents is able to actually see the posts via MSMT within Qualtrics because of browser/phone operating system conflicts. Therefore, we use it at the end of the questionnaire to avoid attrition early in the survey and/or prior to our measure of mobilization.

[29]In later versions of this paper, we will include analysis of the text of the comments and in the timelines. Furthermore, this is a new tool and we present the analysis of these outcomes as exploratory. Our pilots indicated that many survey respondents had no previous experience using MSMT, even though it appeared to be intuitive to many of them.

[30]We do not follow common practice in studies regarding the U.S. as to calculate the difference in the "feeling thermometer score for a respondent's own party and the feeling thermometer score for the opposition party" (Lelkes and Westwood 2017, p.488-9) because many of our respondents are unlikely to identify with either group.

Bolsonaro/Lula/PT in our pre-treatment measures of political group affiliation and we classify as a Lula or Bolsonaro supporter individuals who declare liking Bolsonaro or Lula.[31]

4.2 Misinformation Leads to Lower Levels of Overall Mobilization and Higher Levels of Net Mobilization

In this section we present the results of the main analyses we conduct to test hypotheses 1, 2, 3, 4, and 5, and our research question. A total of 5,014 participants assigned to the false (1,636), hyperpartisan (1,688), and phatic (1,690) conditions, as Table 4 indicates. We conducted balance tests and found no evidence of randomization failure.[32] The results we present in this section are not adjusted by any pre-treatment covariates and rely on OLS regression models with robust (HC2) standard errors.[33]

Table 4: Experimental Assignment

Experimental Condition	N	Percent
False Story	1636	32.63
Hyperpartisan	1688	33.67
Phatic	1690	33.71

The first column of Table 5 examines whether levels of overall political mobilization-—measured as willingness to engage in different types of campaign-related activities-—differ between the misinformation/hyperpartisan and phatic conditions. The results indicate that political engagement levels are lower among individuals exposed to posts containing misinformation than among those assigned to see phatic posts. Therefore, contrary to the expectation stated in hypothesis 1 (that exposure to misinformation would

[31]Per our pre-analysis plan, we conduct the analysis for 5 by dropping respondents who neither Bolsonaro nor Lula/PT supporters. See Appendix K for variable coding.

[32]Individual t-tests and chi-squares for pre-treatment variables, F-tests for the regression of two treatment indicators on all pre-treatment variables, and a multinomial logit regression of three treatment categories on all pre-treatment variables. The results of these tests are presented in section G.

[33]See Appendix I for the correspondence and deviations between our pre-analysis plan and the analyses shown here.

increase mobilization), we find that being exposed to misinformation has a demobilizing effect. The second and third columns of Table 5 show that, while posts with false stories have a negative and statistically significant effect on overall mobilization (the opposite of what we expected in hypothesis 1A), hyperpartisan posts have a smaller and not statistically significant negative effect on overall mobilization (which is not consistent with expectations stated in hypothesis 1B). The difference between these coefficients, however, is not statistically significant (p-value = 0.23).

Table 5: Effects of Misinformation on Overall Mobilization Index (OMI)

	OMI (1)	OMI (2)
(Intercept)	0.000	0.000
	(0.024)	(0.024)
Misinfo.	−0.063*	
	(0.030)	
False		−0.085*
		(0.035)
Hyper.		−0.043
		(0.034)
Num.Obs.	5014	5014

Notes: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Robust standard errors.

We find that exposure to misinformation shared by political elites demobilizes, regardless of its type (if in favor or against a specific political group). To better understand the effects of misinformation on electoral mobilization, we explore whether misinformation shared by political elites generates any net mobilization effects, i.e., whether this type of content is more likely to benefit the political group that is the source of the posts (Bolsonaro) or the target of the information (Lula).

As Table 6 indicates, exposure to politicians' posts containing misinformation promotes political engagement in favor of Bolsonaro and against Lula/PT. The second and third columns of Table 6 examine sub-hypotheses H2A and H2B, respectively. It shows that posts containing either false or hyperpartisan content promotes mobilization in favor of Bolsonaro and against Lula/PT (and the differences between these coefficients is not significant).

Table 6: Effects of Misinformation on Net Mobilization Index (NMI)

	NMI (1)	NMI (2)
(Intercept)	0.000	0.000
	(0.024)	(0.024)
Misinfo.	0.064*	
	(0.030)	
False		0.070*
		(0.035)
Hyper.		0.058+
		(0.034)
Num.Obs.	5014	5014

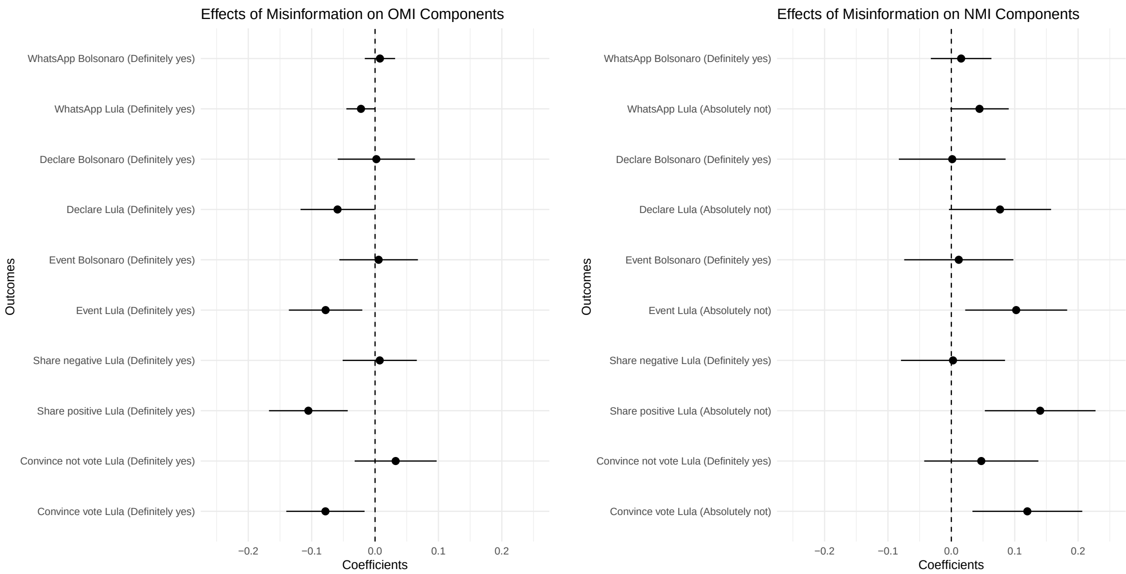
Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors.

Our results suggest that misinformation both demobilizes respondents, according to the OMI, in favor of Lula, and mobilizes respondents in favor of Bolsonaro, according to the NMI. Furthermore, the absolute magnitudes of both OMI and NMI are similar to each other, and somewhat small – around 0.06 standard deviations.

While these findings are seemingly contradictory, the results disaggregated by each component of the overall and net mobilization indices both show that misinformation is demobilizing in terms of actions that support Lula. Figure 3, panel (a), show that subjects exposed to misinformation are less likely to join a WhatsApp group in support of Lula, to declare their support for Lula, to attend a campaign event in support of Lula, share positive news about Lula online, and convince others to vote for Lula, in comparison to those exposed to phatic posts. Figure 3, panel (b), shows the net effect of misinformation on mobilization pro Bolsonaro and against Lula for each component of the NMI, which is consistent with panel (a): respondents exposed to misinformation are more likely to say they would not join a WhatsApp group in support of Lula, to declare their support for Lula, to attend a campaign event in support of Lula, share positive news about Lula online, and convince others to vote for Lula. For example, the component with the largest effects in the OMI corresponds to a 0.10 unit decrease in the willingness to share positive news about Lula in a 4-point Likert scale. For context, for individuals exposed to phatic posts, about 66% said they would absolutely not or possibly not share any positive content about Lula, whereas, in contrast, about 70% of individuals exposed to misinformation said they absolutely not or possibly not share any positive content about Lula.

Figure 3: Effects of Misinformation on OMI and NMI components

(a) OMI Components **(b)** NMI Components



Notes: Bars represent 95% confidence intervals. Robust standard errors. Term in parenthesis is response category with the largest positive value in the scale.

Table 7 shows the effect of misinformation on affective polarization and affect regarding Bolsonaro supporters (*Bolsonaristas*) and Workers' Party supporters (*petistas*). Contrary to our expectation, misinformation is not associated with an increase in affective polarization – widening the gap in feelings toward PT and Bolsonaro supporters. Yet, consistent with the demobilizing effects in pro-Lula/PT campaign activities, respondents exposed to misinformation have lower levels of affect toward PT supporters and their levels of affect toward Bolsonaro supporters remains largely unchanged.

Table 7: Effects of Misinformation on Affect

	Affect Polarization (1)	Affect Bol. Sup. (2)†	Affect PT Sup. (3)†	Affect Bol. Sup. (4)†	Affect PT Sup. (5)†
(Intercept)	4.408***	4.106***	4.122***	4.106***	4.122***
	(0.093)	(0.090)	(0.084)	(0.090)	(0.084)
Misinfo.	−0.098	−0.042	−0.244*		
	(0.115)	(0.110)	(0.103)		
False				−0.014	−0.296*
				(0.128)	(0.119)
Hyper.				−0.069	−0.194
				(0.127)	(0.120)
Num.Obs.	5014	5014	5014	5014	5014

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors. † denotes analyses that were not pre-registered.

Table 8 shows that the effect of misinformation on willingness to participate is not moderated by political identity. Results in Table 8 show that, while partisans in general (model 1) and Bolsonaro supporters (model 2) are more likely to report willingness to participate in campaign-related activities overall (model 1) and in favor of Bolsonaro/anti-Lula (model 2), we fail to detect misinformation as more effective in creating willingness to participate, overall, among partisans, and to participate in activities pro-Bolsonaro/anti-Lula, among Bolsonaro supporters.

Table 8: Effects of Misinformation on OMI and NMI by political identity

	OMI (1)	NMI (2)
(Intercept)	−0.271***	−1.026***
	(0.037)	(0.029)
Misinfo.	−0.069	0.111**
	(0.045)	(0.037)
Partisan	0.481***	
	(0.048)	
Misinfo. * Partisan	0.013	
	(0.059)	
Bol. Sup.		1.857***
		(0.044)
Misinfo. * Bol. Sup.		−0.035
		(0.054)
Num.Obs.	5014	3326
R2	0.059	0.600

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors. Model (2) excludes respondents who support neither Bolsonaro nor Lula.

4.3 Discussion

Misinformation has a limited, albeit detectable, effect on willingness to participate in campaign-related activities. Our survey experiment shows that misinformation demobilizes willingess to perform actions in favor of Lula. Furthermore, we do not find that either fakeness or hyperpartisan content have systematically different effect on mobilization. While estimates for posts with false stories tended to be slightly larger (in absolute values) than estimates for hyperpartisan posts, we fail to find that these estimates are statistically distinguishable from each other.

Contrary to our expectations, misinformation is not more likely to prompt partisans or supporters of Bolsonaro or Lula to be more willing to take part in campaign-related activities. The result regarding Bolsonaro/Lula supporters may partly stem from their more ardent supporters already displaying high levels of willingness to participate in those types of activities: for instance, Bolsonaro's supporters average score for the net mobilization index is at the top 25% of the sample among those assigned to phatic posts. Yet, we do not observe such "ceiling effects" for partisans, which suggests that misinformation does not necessarily influence mobilization by affecting those who already have developed firmer political identities as we had originally hypothesized. While we fail to find that minsinformation's effect is concentrated among partisans or Bolsonaro/Lula supporters, we find that misinformation activates animosity against PT supporters by reducing affect, although feelings of affective polarization remain unchanged.

Both the results from our observational analysis and experimental findings suggest that misinformation pays off. Misinformation leads to demobilization in favor of Lula and reduced affect toward PT supporters, all of which stands to benefit politicians in the opposing camp, largely responsible for spreading these hyperpartisan and false stories. Even though we find that respondents exposed to misinformation are less willing to follow online the politicians who spread misinformation (Table A26), there is no effect of misinformation on their willingness to take part in campaign-related activities related to Bolsonaro or on affect toward Bolsonaro, which suggest that the cost of spreading misinformation is low. Furthermore, our study of online engagement suggests higher levels of engagement with posts containing misinformation and weak evidence that users are, to a meaningful degree, mocking, correcting, questioning, or even attacking the misinformation spreaders.

Yet comparing findings from the observational and experimental analyses requires several caveats. The outcomes and comparisons made in the observational and experimental studies are not directly commensurable: the outcomes of online engagement in our observational study measure online behavior and they differ from outcomes measuring willingness to participate in campaign-related activities in our survey. In the experimental study, we compare willingness to participate between respondents exposed to misinformation and those exposed to phatic content while in the our observational study posts containing misinformation are compared to posts without any misinformation. And, finally, the controlled exposure to misinformation we measure in our survey experiment is an entirely distinct data generating process than the way in which social media users are exposed to misinformation across social media platforms.

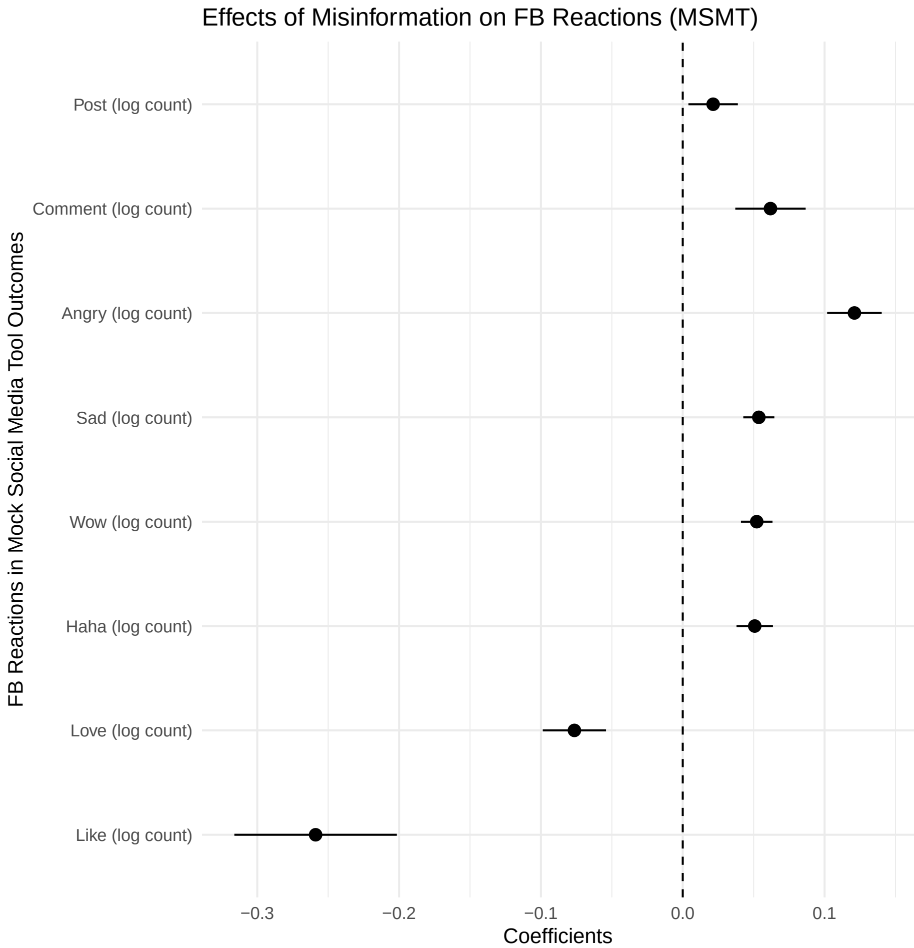
We partly address these questions by examining survey respondents reactions to posts containing misinformation (compared to posts containing phatics) via the *Mock Social Media Tool*, which simulates Facebook's newsfeed environment within the survey, thereby reducing commensurability issues between experimental and observational studies. Therefore, if survey respondents' behavior in the survey bears resemblance to the average social media user included in our observational analysis, we should expect that reactions to posts containing misinformation should also elicit more online reactions than phatic posts, as our observational results imply.[34]

And, indeed, as shown in Figure 4, respondents exposed to a timeline that contained misinformation are more likely to react, across a number of online engagement measures, than respondents exposed to a timeline containing phatic posts: respondents exposed to elite-shared misinformation post and comment more often on politicians' posts, and their reactions have higher levels of "anger," "sad," "wow," and "haha." The two important exceptions are "likes" and "loves," which are more frequent reactions in phatic posts than posts containing misinformation. Interestingly, as shown by Figure 1 in our observational study, "love" is systematically no different or at lower levels in posts containing misinformation compared to posts without misinformation – and we observe the same pattern for the measure of "love" in the simulated Facebook environment in our survey. The main discrepancy between the results from the simulated Facebook environment and the observational study regards "likes" given we observe equal or larger number of "likes" in posts containing misinformation in the observational analysis. Overall, despite significant differences in design and underlying data generating processes, respondents in the survey experiment react to misinformation shared by political leaders in a manner that is largely consistent with the way users of social media react to misinformation posed by political

[34] Note to reader: we plan to address the type of comparisons we are making in the observational study in future versions of this paper. We are in the process of detecting phatic posts to create a comparison group in the observational study.

leaders.

Figure 4: Effects of Misinformation on FB reactions via MSMT



Notes: Bars represent 95% confidence intervals. Robust standard errors. N = 3,467 (see Appendix H.1 for discussion).

Finally, our experimental study focuses on one type of misinformation that has a negative tone and that is critical of Lula/PT and our experimental findings suggest that misinformation affected respondents' attitudes regarding the PT – and not the political group sharing misinformation about PT. An open question is whether other types of misinformation, those with positive tone, and that have different content, could generate different effects. Plausibly, negative false and misleading stories about the PT may be more likely to generate anti-Lula/PT mobilization, as the literature on negative campaigning suggests. However, it is equally plausible that false and misleading stories that praise Bolsonaro could also mobilize – just in a different direction.

5 Conclusion

To the best of our knowledge, this study is the first to provide evidence on the effects of elite-shared misinformation on mobilization. In doing so, it offers an important contribution to the debate about the political consequences of misinformation, and helps us calibrate our concerns about this type of communication and "post-truth politics." Even if misinformation fails to persuade voters, as many recent studies suggest, it can nonetheless shape their political participation, which can have implications for political competition and elections.

More specifically, we find that misinformation shared by politicians is effective, although to a limited extent, at reducing mobilization in favor of the target of the communication. Furthermore, social media users engage more with politicians' posts that contain misinformation than with posts that lack misinformation, and such engagement appears to be, to some extent, positive for the politician rather than pushback against sharing falsehoods or biased content.

Moreover, our findings suggest that posts containing falsehoods tend to have similar mobilizing effects to posts that convey hyperpartisan messages. In this sense, misinformation mobilizes not because it provides false content, but primarily because it conveys such content through hyperpartisan communication that evokes emotions and activates political identities. While one can argue that the polarizing rhetoric of hyperpartisan messages represents a departure from the normative ideal of democratic representation, it does not necessarily resort to falsehoods in order to exert influence. Instead, hyperpartisan messages generally dispute the meaning and interpretations of facts, which is a regular component of political discourse (Gaines et al. 2007). The troublesome aspect of our findings is that entirely fabricated content combined with polarizing messages can take the place of typical of hyperpartisan rhetoric. Hence, under unfavorable conditions in the public debate, political elites can simply resort to false content as subterfuge to produce hyperpartisan rhetoric and mobilize voters.

Our study suggests new directions for research on political misinformation. Future studies should investigate whether the effect we find for negative misinformation is also observed for misinformation that conveys positive messages that praises the target. Moreover, future research could further probe the causal roles of falseness and hyperpartisanship in fostering reactions and participation, as some of our findings suggest that the latter is the active ingredient behind the mobilizing effects of misinformation. Finally, future research should investigate the effects of misinformation on outcomes that tap more costly types of behaviors beyond reactions to politicians' posts and intentions to perform campaign-related activities.

References

Allcott, Hunt and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *The Journal of Economic Perspectives* 31(2):211–235.
URL: *http://www.jstor.org/stable/44235006*

Allcott, Hunt, Matthew Gentzkow and Chuan Yu. 2019. "Trends in the diffusion of misinformation on social media." *Research & Politics* 6(2):2053168019848554.

Badrinathan, Sumitra. 2021. "Educative Interventions to Combat Misinformation: Evidence from a Field Experiment in India." *American Political Science Review* 115(4):1325–1341.

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

Batista Pereira, Frederico, Natália S. Bueno, Felipe Nunes, João Pedro Oliveira, Nara Pavão and Valerie Wirtschafter. 2022. "Detecting Misinformation: The Spread of False News by Political Leaders in the Global South." *https://osf.io/preprints/socarxiv/hu4qr* .

Batista Pereira, Frederico, Natália S. Bueno, Felipe Nunes and Nara Pavão. 2022*a*. "Fake News, Fact Checking, and Partisanship: The Resilience of Rumors in the 2018 Brazilian Elections." *Journal of Politics* .

Batista Pereira, Frederico, Natália S. Bueno, Felipe Nunes and Nara Pavão. 2022*b*. "Inoculation Reduces Misinformation: Experimental Evidence from a Multidimensional Intervention in Brazil." *Available at SSRN: https://ssrn.com/abstract=3144139* .

Berinsky, Adam J. 2017. "Rumors and Health Care Reform: Experiments in Political Misinformation." *British Journal of Political Science* 47(2):241–262.

Berriche, Manon and Sacha. Altay. 2020. "Internet users engage more with phatic posts than with health misinformation on Facebook." *Palgrave Communications* 6(71).

Bryan, Christopher J., Gregory M. Walton, Todd Rogers and Carol S. Dweck. 2011.

"Motivating voter turnout by invoking the self." *Proceedings of the National Academy of Sciences* 108(31):12653–12656.

URL: *https://www.pnas.org/doi/abs/10.1073/pnas.1103343108*

Chen, Cathy X, Gordon Pennycook and David G Rand. 2021. "What Makes News Sharable on Social Media?".

URL: *psyarxiv.com/gzqcd*

Cinelli, Matteo, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi and Michele Starnini. 2021. "The echo chamber effect on social media." *Proceedings of the National Academy of Sciences* 118(9):e2023301118.

URL: *https://www.pnas.org/doi/abs/10.1073/pnas.2023301118*

Dawkins, Ryan. 2017. "Political participation, personality, and the conditional effect of campaign mobilization." *Electoral Studies* 45:100–109.

URL: *https://www.sciencedirect.com/science/article/pii/S0261379416300270*

Delton, A. W., M. B. Petersen and T. E. Robertson. 2018. "Partisan goals, emotions, and political mobilization: The role of motivated reasoning in pressuring others to vote." *The Journal of Politics* 80(3):890–902.

Dickson, Eric S. and Kenneth Scheve. 2006. "Social Identity, Political Speech, and Electoral Competition." *Journal of Theoretical Politics* 18(1):5–39.

URL: *https://doi.org/10.1177/0951629806059594*

Druckman, James N., EriK Peterson and Rune Slothuus. 2013. "How Elite Partisan Polarization Affects Public Opinion Formation." *American Political Science Review* 107(1):57–79.

Ecker, Ullrich K. H., Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K. Vraga and Michelle A. Amazeen. 2022. "The psychological drivers of misinformation belief and its resistance to correction." *Nature Reviews Psychology* 1:13–29.

Flynn, D.J., Brendan Nyhan and Jason Reifler. 2017. "The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics." *Political*

Psychology 38(S1):127–150.

URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/pops.12394*

Gaines, Brian, James Kuklinski, Paul Quirk, Buddy Peyton and Jay Verkuilen. 2007. "Same Facts, Different Interpretations: Partisan Motivation and Opinion on Iraq." *The Journal of Politics* 69(4).

Gerber, Alan S. and Donald P. Green. 1999. "Does canvassing increase voter turnout? A field experiment." *Proceedings of the National Academy of Sciences* 96(19):10939–10942.

URL: *https://www.pnas.org/doi/abs/10.1073/pnas.96.19.10939*

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson and David Lazer. 2019. "Fake news on Twitter during the 2016 U.S. presidential election." *Science* 363(6425):374–378.

URL: *https://science.sciencemag.org/content/363/6425/374*

Guess, Andrew, Dominique Lockett, Benjamin Lyons, Jacob Montgomery, Brendan Nyhan and Jason Reifler. 2020. ""Fake news" may have limited effects on political participation beyond increasing beliefs in false claims." *Harvard Kennedy School Misinformation Review* 1.

Guess, Andrew, Jonathan Nagler and Joshua Tucker. 2019. "Less than you think: Prevalence and predictors of fake news dissemination on Facebook." *Science Advances* 5:1–8.

Guess, Andrew M. and Benjamin A. Lyons. 2020. *Misinformation, Disinformation, and Online Propaganda.* SSRC Anxieties of Democracy Cambridge University Press p. 10–33.

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

Guess, Andrew, Nyhan Brendan and Jason Reifler. 2018. "Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign." *European Research Council* .

Horz, Marco M. 2018. "Informational and Non-Informational Propaganda." *Comparative Politics Newsletter, The Organized Section in Comparative Politics of the American Political Science Association* 28(2):38–43.

Huddy, Leonie, Lilliana Mason and Lene Aarøe. 2015. "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity." *American Political Science Review* 109(1):1–17.

Jagayat, Arvin, Gurkaran Boparai, Carson Pun, and Becky L. Choma. N.d. "Mock Social Media Website Tool.".
URL: *https://hadoop.apache.org*

Jerit, Jennifer and Yangzi Zhao. 2020. "Political Misinformation." *Annual Review of Political Science* 23(1):77–94.
URL: *https://doi.org/10.1146/annurev-polisci-050718-032814*

Kling, Jeffrey R., Jeffrey B. Liebman and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1):83–119.

Kwak, Nojin, Daniel S Lane, Brian E Weeks, Dam Hee Kim, Slgi S Lee and Sarah Bachleda. 2018. "Perceptions of Social Media for Politics: Testing the Slacktivism Hypothesis." *Human Communication Research* 44(2):197–221.
URL: *https://doi.org/10.1093/hcr/hqx008*

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

Lelkes, Yphtach and Sean J Westwood. 2017. "The limits of partisan prejudice." *The Journal of Politics* 79(2):485–501.

Little, Andrew T. 2017. "Propaganda and credulity." *Games and Economic Behavior* 102:224–232.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825616301476*

Little, Andrew T. 2018. "Fake news, propaganda, and lies can be pervasive even if they aren't persuasive." *Comparative Politics Newsletter* 28(2):49–55.

Lupu, Noam. 2015. Partisanship in Latin America. In *The Latin American voter: Pursuing representation and accountability in challenging contexts*, ed. Ryan E. Carlin, Matthew M. Singer and Elizabeth J. Zechmeister. University of Michigan Press.

Machado, Caio, Beatriz Kira, Vidya Narayanan, Bence Kollanyi and Philip N. Howard. 2019. A Study of Misinformation in WhatsApp groups with a focus on the Brazilian Presidential Elections. In *Proceedings of The Web Conference*. WWW'19 San Francisco, USA: .

Mainwaring, Scott and Mariano Torcal. 2006. Party System Institutionalization and Party System Theory After the Third Wave of Democratization. In *Handbook of Political Parties*, ed. Richard S. Katz and William Crotty. Sage.

Mainwaring, Scott and Timothy Scully. 1999. *Building Democratic Institutions: Party Systems in Latin America*. Stanford: Stanford University Press.

Mele, Nicco, David Lazer, Matthew Baum, Nir Grinberg, Lisa Friedland, Kenneth Joseph andWill Hobbs and Carolina Mattsson. 2017. "Combating Fake News: An Agenda for Research and Action.".

Miller, Vincent. 2017. "Phatic culture and the status quo: Reconsidering the purpose of social media activism." *Convergence* 23(3):251–269.
URL: *https://doi.org/10.1177/1354856515592512*

Nyhan, Brendan. 2018. *How Misinformation and Polarization Affect American Democracy*. Available at SSRN: https://ssrn.com/abstract=3144139.

Nyhan, Brendan. 2020. "Facts and Myths about Misperceptions." *Journal of Economic Perspectives* 34(3):220–36.
URL: *https://www.aeaweb.org/articles?id=10.1257/jep.34.3.220*

Nyhan, Brendan and Jason Reifler. 2010. "When Corrections Fail: The Persistence of

Political Misperceptions." *Political Behavior* 32(2):303–330.

URL: *https://doi.org/10.1007/s11109-010-9112-2*

Pennycook, Gordon and David G. Rand. 2019. "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning." *Cognition* 188:39–50. The Cognitive Science of Political Thought.

URL: *https://www.sciencedirect.com/science/article/pii/S001002771830163X*

Pennycook, Gordon and David G. Rand. 2021. "The Psychology of Fake News." *Trends in Cognitive Sciences* 25(5):388–402.

URL: *https://www.sciencedirect.com/science/article/pii/S1364661321000516*

Resende, Gustavo, Philipe Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara Almeida and Fabricio Benevenuto. 2019. (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. In *Proceedings of The Web Conference*. WWW'19.

URL: *https://doi.org/10.1145/3308558.3313688*

Samuels, David. 2006. "Sources of Mass Partisanship in Brazil." *Latin American Politics and Society* 48(2):1–27.

Samuels, David and Cesar Zucco. 2018. *Partisans, Antipartisans, and Nonpartisans: Voting Behavior in Brazil*. Cambridge University Press.

Shirky, Clay. 2008. *Here comes everybody: The power of organizing without organizations*. Penguin.

Tucker, Joshua A., Andrew Guess, Pablo Barbera, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal and Brendan Nyhan. 2018. "Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature.".

Valkenburg, Patti M, Jochen Peter, Joseph B Walther et al. 2016. "Media effects: Theory and research." *Annual review of psychology* 67(1):315–338.

Vosoughi, Soroush, Deb Roy and Sinan Aral. 2018. "The spread of true and false news online." *Science* 359(6380):1146–1151.

URL: *https://www.science.org/doi/abs/10.1126/science.aap9559*

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* Cambridge University Press.

Zhao, Laijun, Qin Wang, Jingjing Cheng, Ding Zhang, Ting Ma, Yucheng Chen and Jiajia Wang. 2012. "The impact of authorities' media and rumor dissemination on the evolution of emergency." *Physica A: Statistical Mechanics and its Applications* 391:3978–3987.

Supporting Information (SI)

Contents

A Summary Statistics Online Engagement Study - Observational Data

Table A1: Summary Statistics: Post-level analysis

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Year	3	0	2019.2	0.8	2018.0	2019.0	2020.0
Month	12	0	6.9	3.2	1.0	7.0	12.0
False (text)	2	0	0.0	0.0	0.0	0.0	1.0
Hyper (domain)	2	0	0.0	0.1	0.0	0.0	1.0
False (URL)	2	0	0.0	0.0	0.0	0.0	1.0
Hyper (repeat domain)	2	0	0.0	0.0	0.0	0.0	1.0
Reactions	58048	0	1796.4	13 755.0	0.0	120.0	2 469 318.0
Reactions (log)	58048	0	4.8	2.3	0.0	4.8	14.7
Reactions (log hyper.)	58048	0	5.5	2.4	0.0	5.5	15.4
“Likes”	49813	0	1298.7	11 163.4	0.0	79.0	2 423 284.0
“Likes” (log)	49813	0	4.3	2.5	0.0	4.4	14.7
“Likes (log hyper.)	49813	0	4.8	2.6	0.0	5.1	15.4

Table A2: Descriptive Statistics by Misinformation

False (text)	False (URL)	Hyper (domain)	Hyper (rep. domain)	Median Likes	Avg. Likes	Median Reactions	Avg. Reactions
0.00	0.00	0.00	0.00	78.00	1293.10	118.00	1783.57
0.00	0.00	1.00	0.00	593.00	1615.16	1228.00	3003.34
0.00	0.00	1.00	1.00	2409.00	5643.67	4315.00	8979.02
0.00	1.00	1.00	0.00	122.00	767.17	271.00	2163.48
0.00	1.00	1.00	1.00	2494.00	2279.20	3867.00	4114.00
1.00	0.00	0.00	0.00	590.00	4989.49	1138.00	9060.86
1.00	0.00	1.00	0.00	696.00	757.80	1540.00	1844.40
1.00	0.00	1.00	1.00	1858.00	3757.40	4950.00	9792.60
1.00	1.00	1.00	0.00	1581.00	1581.00	3368.00	3368.00
1.00	1.00	1.00	1.00	50.00	50.00	92.00	92.00

Table A3: Prevalence of Politicians Sharing False Content and Posts Containing False Content by Detection Approach (Table 2 from Batista Pereira et al 2022)

Detection Approach	Pct. of Politicians Sharing False News	Pct. of Posts from Politicians Sharing False News
Text	15.4497	0.0104
Domain	43.7037	0.5759
Facebook URL	1.4815	0.0009
Repeat Domain	13.6507	0.085
n total	945	4,032,907

B Detecting False and Hyperpartisan Posts

To identify false and hyperpartisan content,[35] we rely on several different strategies, including those commonly deployed in misinformation research in the United States that rely on URLs to hyperpartisan websites and a text-based approach better suited to the Global South, where false content regularly circulates via WhatsApp and other messaging platforms. In total, we use four different approaches to measuring false and hyperpartisan content: (1) a domain-based approach; (2) a repeated domain-based approach; (3) a URL approach; and (4) a novel text-based approach. While the first two approaches are designed to measure "hyperpartisan" content that may not be exclusively false, the latter two approaches measure explicitly false content shared by politicians online.

In research on the spread of misinformation in the United States, scholars have long-relied on an expanding list of low-quality, hyperpartisan domains, where misinformation is most likely to spread (Guess, Nyhan and Reifler 2020; Allcott, Gentzkow and Yu 2019). Although these domain may capture false content, they do not exclusively do so, and are likely to also feature hyperpartisan information that may be politically biased, but not explicitly false. For scholars studying the spread of misinformation in the U.S., this list has been adapted over time to incorporate additional domains based on qualitative knowledge and systematic assessments of quality. In the Brazilian context, we know of no such existing list of domains. However, we are able to compile a list of hyperpartisan domains using Facebook's URL database. To operationalize the "domain-based approach" for the Brazilian context, we collect all URLs shared at least 100 times, fact-checked as false on Facebook by a third-party fact checker, with a geographic concentration in Brazil. From there, we collect all unique root domains, which we consider the "domain-based approach" (N=228). In addition, we develop a "repeated domain-based approach" drawing on this same list of Facebook URLs, but restricting out sample to only root domains that have shared two or more stories fact-checked as false in the Facebook URLs database (N=59). These two lists constitute our measure of hyperpartisan content, which includes both false and politically charged (but not explicitly false) information. We exclude common domains, such as youtube.com and yahoo.com, from this analysis.

To operationalize false content, we also rely on two different measures. In one measure, the "URL approach," we identify posts using the exact URLs fact checked as false in the Facebook URL Database, rather than the root domain (N=365). In addition, we develop a "text-based approach" the relies on the exact content of rumors circulating in Brazil, using supervised machine learning, natural language processing, and manual review. This text-based approach, which we detail in the working paper (Batista Pereira, Bueno, Nunes, Oliveira, Pavão and Wirtschafter 2022), is particularly well-suited to contexts, like Brazil, where false content may circulate on messaging apps or social media websites in the form of large blocks of text or images and does not include hyperlinks to external websites.

To develop this text-based approach, we first collect the exact text of all rumors fact-checked by the external fact checking website *Boatos* (N=4,050). We then train a series of classification model to cull our sample of posts to only those that are among the most likely to contain false content. As part of this process, we first train a Naive Bayes model

[35]Parts of this appendix may replicate or be closely related to our other working papers that use the same data and method.

on a dataframe that includes both the false content from *Boatos* and stories by reputable news outlets in Brazil. We apply this model to over four million posts across Facebook, Twitter, and Instagram to identify politicians' social media posts that are the *least* likely to contain false content, based on a low predicted probability of being classified as false. We then retrain a classification model using both the false content and these identified posts, and apply this model to our full dataset again. From these, we selected posts with a predicted probability of .9 or higher of containing false content, based on the classification model.

As a next step, we calculated the cosine similarity– which measures the similarity of documents between 0 (not similar) and 1 (identical)–between every post predicted to be false and the content of the *Boatos* fact check. This allow us to further cull our sample to social media posts that closely resembles false content assessed by external fact checker and remove unrelated content. We eliminate posts with a cosine similarity of less than .4, meaning that they are unlikely to match or be similar to fact checked content. This leaves us with .1% of our original dataframe, or around 4,000 posts, which were then manually reviewed by at least two independent coders.

C Additional Analyses, Online Engagement Study, Observational Data

Table A4: Average reactions and likes to posts containing false stories and without false stories

	No False		False		False - No False	
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Reactions (Text)	1,795.6	13,752.9	8,957.7	26,080.8	7,162.1	1,271.1
Likes (Text)	1,298.4	1,1162.0	4,890.1	20,170.2	3,591.8	983.1
Reactions (URL)	1,796.4	13,755.0	2,410.3	7,211.6	613.9	1,201.9
Likes (URL)	1,298.7	11,163.4	979.9	2,202.3	-318.9	367.1

Table A5: Average reactions and likes to posts containing hyperpartisan stories and without hyperpartisan stories

	No Hyper		Hyper		Hyper - No Hyper	
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Reactions (Domain)	1,784.3	13,779.4	3,883.2	8,288.1	2,098.9	54.8
Likes (Domain)	1,293.5	11,187.8	2,206.9	5,440.3	913.4	36.1
Reactions (Rep. Domain)	1,790.3	13,752.0	8,971.7	15,313.6	7,181.4	261.6
Likes (Rep. Domain)	1,295.0	11,163.5	5,631.6	10,100.1	4,336.6	172.6

Table A6: Correlates of Online Engagement (Reactions): Week-level Engagement, False and Hyperpartisan Posts

	Reactions (log)	Reactions (log, FB)	Reactions (log)	Reactions (log)
Fake (Text)	0.095**			
	(0.030)			
Fake (URL)		−0.010		
		(0.156)		
Hyper (Domain)			0.017	
			(0.011)	
Hyper (Repeat Domain)				0.020
				(0.020)
Total Posts FB (log)		1.693***		
		(0.025)		
Total Posts (log)	1.113***		1.113***	1.113***
	(0.008)		(0.008)	(0.008)
Num.Obs.	127 643	127 643	127 643	127 643
RMSE	0.50	0.65	0.50	0.50
Std.Errors	Pol.	Pol.	Pol.	Pol.
FE: Pol-M-Y	Y	Y	Y	Y

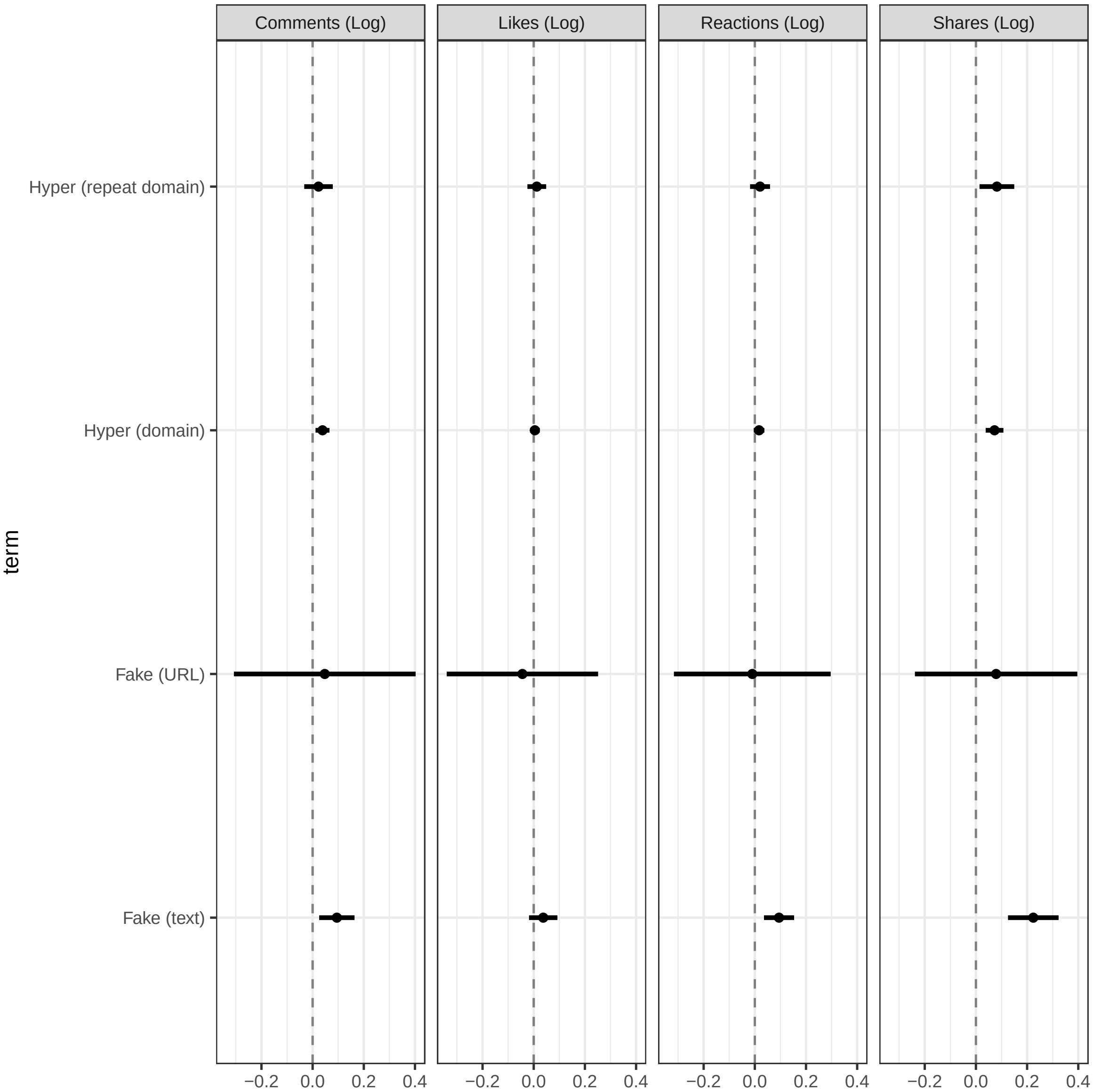
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: Correlates of Online Engagement (Likes): Week-level Mobilization, False and Hyperpartisan Posts

	Likes (Log)	Likes (Log, FB)	Likes (Log)	Likes (Log)
Fake (Text)	0.037			
	(0.028)			
Fake (URL)		−0.044		
		(0.151)		
Hyper (Domain)			0.005	
			(0.010)	
Hyper (Repeat Domain)				0.012
				(0.019)
Total Posts FB (log)		1.636***		
		(0.023)		
Total Posts (log)	1.091***		1.091***	1.091***
	(0.008)		(0.008)	(0.008)
Num.Obs.	127 643	127 643	127 643	127 643
RMSE	0.48	0.60	0.48	0.48
Std.Errors	Pol.	Pol.	Pol.	Pol.
FE: Pol-M-Y	Y	Y	Y	Y

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A1: Correlates of Online Engagement: Week-level Engagement, False and Hyper-partisan Posts



Notes: All models control for total number of posts in a given week. Standard errors clustered at the politician level. Politician-Month-Year fixed effects included in all models.

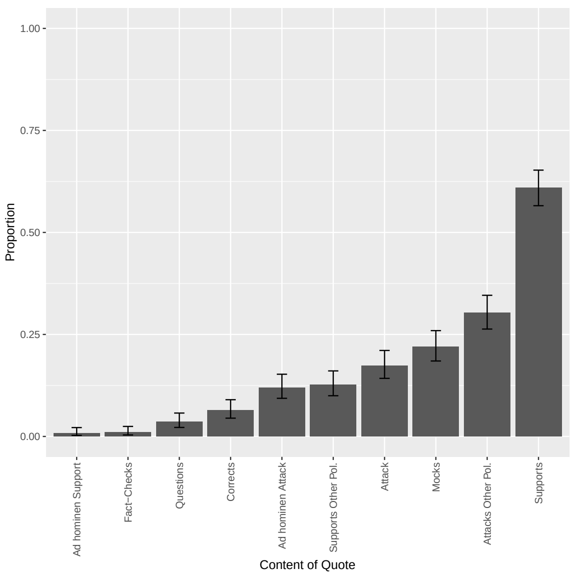
Table A8: Correlates of Online Engagement (Comments): Week-level Engagement, False and Hyperpartisan Posts

	Comments (log)	Comments (log)	Comments (log)	Comments (log)
Fake (Text)	0.095**			
	(0.035)			
Fake (URL)		0.047		
		(0.181)		
Hyper (Domain)			0.039**	
			(0.014)	
Hyper (Repeat Domain)				0.023
				(0.028)
Total Posts FB (log)		1.361***		
		(0.019)		
Total Posts (log)	1.027***		1.027***	1.027***
	(0.008)		(0.008)	(0.008)
Num.Obs.	127 643	127 643	127 643	127 643
RMSE	0.60	0.68	0.60	0.60
Std.Errors	Pol	Pol	Pol	Pol
FE: Pol-M-Y	Y	Y	Y	Y

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Figure A2: Types of Quotes to Tweets Containing Fake Stories

(a) Quotes (Tweets with Misinfo. Text, RA 1)



(b) Quotes (Tweets with Misinfo. Text, RA 2)

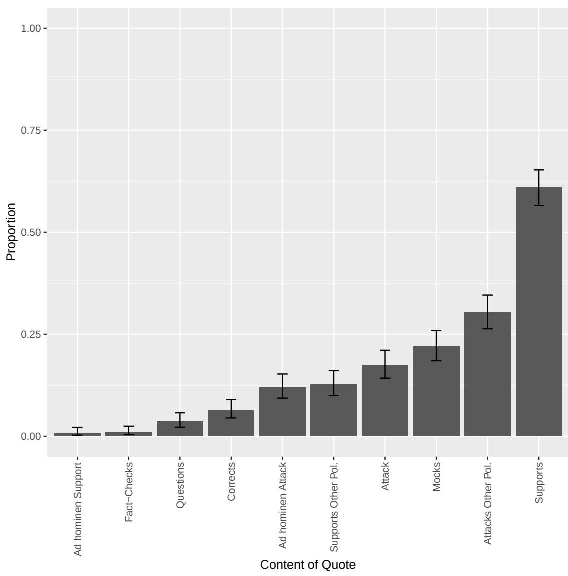


Table A9: Correlates of Online Engagement (Shares): Week-level Engagement, False and Hyperpartisan Posts

	Shares (log)	Shares (log)	Shares (log)	Shares (log)
Fake (Text)	0.224***			
	(0.050)			
Fake (URL)		0.079		
		(0.162)		
Hyper (Domain)			0.073***	
			(0.018)	
Hyper (Repeat Domain)				0.082*
				(0.035)
Total Posts FB (log)		1.509***		
		(0.017)		
Total Posts (log)	1.293***		1.292***	1.293***
	(0.013)		(0.013)	(0.013)
Num.Obs.	127 643	127 643	127 643	127 643
RMSE	0.81	0.65	0.81	0.81
Std.Errors	Pol	Pol	Pol	Pol
FE: Pol-M-Y	Y	Y	Y	Y

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

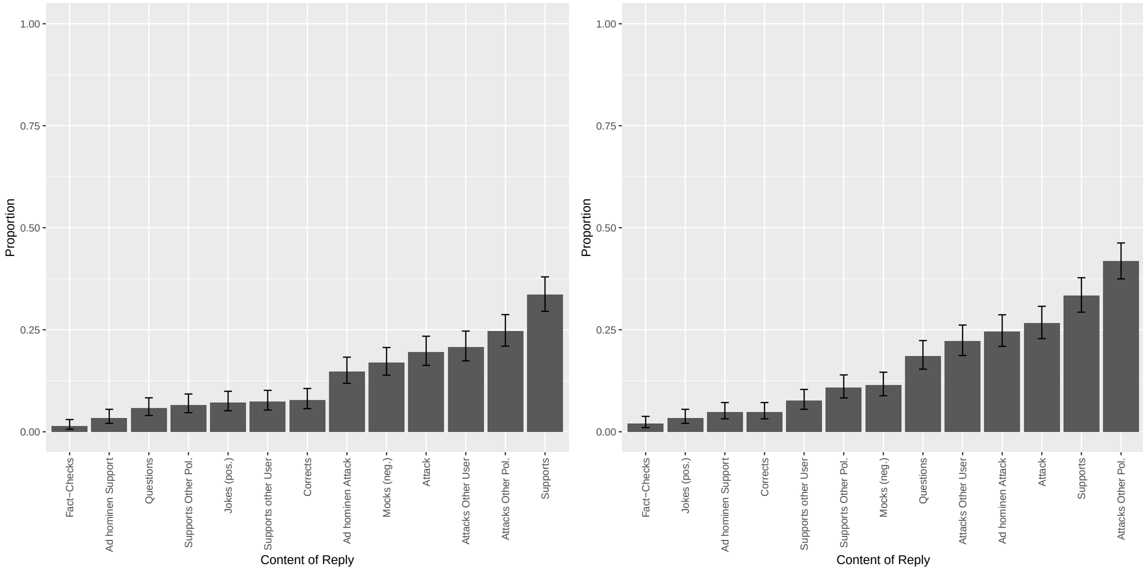
Table A10: Correlation between Misinformation and "Getting Ratioed" on Twitter

	Ratioed	Ratioed	Ratioed
Fake (Text)	−0.093*		
	(0.041)		
Hyper (Domain)		0.180	
		(0.202)	
Hyper (Repeat Domain)			0.471+
			(0.268)
Num.Obs.	1 278 309	1 278 309	1 278 309
RMSE	0.54	0.54	0.54
Std.Errors	Pol.	Pol.	Pol.
FE: Pol-M-Y	Y	Y	Y

Notes: Twitter data only. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. "Ratioed" is ratio of the sum of quotes and replies subtracted by retweets over the sum of quotes, replies, and retweets.

Figure A3: Types of Replies to Tweets Containing Fake Stories

(a) Replies (Tweets with Misinfo. Text, RA 1) **(b)** Replies (Tweets with Misinfo. Text, RA 2)



D Ethical Statement

Our experimental study was reviewed and approved by Institutional Review Boards in two U.S. institutions and one Brazilian institution. In any case, exposing subjects to misinformation raises ethical issues. We addressed these concerns by taking the following precautions: we used false stories that had large circulation in Brazil and were not fabricated by the researchers and all participants were debriefed at the end of the study. We also exposed subjects to an awareness-raising list of eight steps in detecting misinformation that we found to be effective in a previous study (Batista Pereira, Bueno, Nunes and Pavão 2022*b*).

Importantly, all subjected consented to taking the survey and it did not contain deception. The researchers did not have access to identifiable data.

Our observational study contains data that is proprietary and we will be unable to share the raw data. Furthermore, it identifies individuals who shared misinformation, even if all their posts being public.

In any case, we can share all of our code used to create the data and the handles and id lists necessary to re-create our datasets using CrowdTangle, Twitter API, and Facebook's URL dataset.

E Pilot Study - Main Results

Table A11: Proportion of Respondents Saying the Post is Fake News

Source	Fake Post	Hyperpartisan Post	Phatic Post (Mother)	Phatic Post (Local)
Kicis	0.53	0.36	0.15	0.24
Martins	0.43	0.58	0.14	0.25
Melo	0.61	0.30	0.16	0.17
Jordy	0.51	0.35	0.15	0.18
Augusto	0.46	0.37	0.13	0.30
Average	0.51	0.39	0.14	0.23

Table A12: Proportion of Respondents Saying that They Believe in the Content of the Post

Source	Fake Post	Hyperpartisan Post	Phatic Post (Mother)	Phatic Post (Local)
Kicis	0.39	0.61		
Martins	0.61	0.49		
Melo	0.32	0.58		
Jordy	0.39	0.77		
Augusto	0.60	0.52		
Average	0.46	0.59		

Note: the question was not asked for phatic posts.

Table A13: Proportion of Respondents Saying that They Have Seen the Post Before

Source	Fake Post	Hyperpartisan Post	Phatic Post (Mother)	Phatic Post (Local)
Kicis	0.23	0.15	0.13	0.05
Martins	0.22	0.35	0.08	0.07
Melo	0.20	0.26	0.07	0.11
Jordy	0.21	0.27	0.07	0.11
Augusto	0.61	0.38	0.09	0.10
Average	0.29	0.28	0.09	0.09

Table A14: How Political is the Post (0-1 Scale)

Source	Fake Post	Hyperpartisan Post	Phatic Post (Mother)	Phatic Post (Local)
Kicis	0.78	0.83	0.47	0.57
Martins	0.81	0.77	0.32	0.46
Melo	0.81	0.81	0.33	0.65
Jordy	0.77	0.86	0.30	0.38
Augusto	0.82	0.88	0.46	0.61
Average	0.80	0.83	0.37	0.53

Table A15: How Pro-Bolsonaro is the Post (0-1 Scale)

Source	Fake Post	Hyperpartisan Post	Phatic Post (Mother)	Phatic Post (Local)
Kicis	0.46	0.52	0.54	0.55
Martins	0.71	0.64	0.50	0.52
Melo	0.71	0.65	0.52	0.59
Jordy	0.56	0.44	0.51	0.50
Augusto	0.76	0.67	0.53	0.64
Average	0.64	0.58	0.52	0.56

F Survey – Summary Statistics

Table A16: Survey summary statistics

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Misinfo. (Dummy)	2	0	0.7	0.5	0.0	1.0	1.0
OMI index	1884	0	−0.0	1.0	−1.7	−0.1	4.2
Convince vote Lula (Def. yes)	4	0	1.9	1.0	1.0	2.0	4.0
Convince not vote Lula (Def. yes)	4	0	2.1	1.1	1.0	2.0	4.0
Share positive Lula (Def. yes)	4	0	2.0	1.1	1.0	2.0	4.0
Share negative Lula (Def. yes)	4	0	2.0	1.0	1.0	2.0	4.0
Event Lula (Def. yes)	4	0	1.8	1.0	1.0	1.0	4.0
Event Bolsonaro (Def. yes)	4	0	1.9	1.1	1.0	1.0	4.0
Declare Lula (Def. yes)	4	0	1.7	1.0	1.0	1.0	4.0
Declare Bolsonaro (Def. yes)	4	0	1.8	1.0	1.0	1.0	4.0
WhatsApp Lula (Def. yes)	2	0	0.2	0.4	0.0	0.0	1.0
WhatsApp Bolsonaro (Def. yes)	2	0	0.2	0.4	0.0	0.0	1.0
NMI index	1884	0	0.0	1.0	−2.1	−0.0	2.0
Convince vote Lula (Abs. no)	4	0	0.8	1.5	−2.0	1.0	2.0
Convince not vote Lula (Def. yes)	4	0	−0.5	1.5	−2.0	−1.0	2.0
Share positive Lula (Abs. no)	4	0	0.7	1.5	−2.0	1.0	2.0
Share negative Lula (Def. yes)	4	0	−0.8	1.4	−2.0	−1.0	2.0
Event Lula (Abs. no)	4	0	1.0	1.4	−2.0	2.0	2.0
Event Bolsonaro (Def. yes)	4	0	−0.8	1.5	−2.0	−2.0	2.0
Declare Lula (Abs. no)	4	0	1.0	1.4	−2.0	2.0	2.0
Declare Bolsonaro (Def. yes)	4	0	−0.9	1.4	−2.0	−2.0	2.0
WhatsApp Lula (Abs. no)	2	0	0.6	0.8	−1.0	1.0	1.0
WhatsApp Bolsonaro (Def. yes)	2	0	−0.6	0.8	−1.0	−1.0	1.0
Affect polarization	11	0	4.3	3.9	0.0	4.0	10.0
Affect Bolsonaristas	11	0	4.1	3.7	0.0	4.0	10.0
Affect Petistas	11	0	4.0	3.5	0.0	4.0	10.0
Age	65	0	37.9	14.0	18.0	36.0	82.0
Sex (Female = 1)	2	0	0.6	0.5	0.0	1.0	1.0
Religion (Cath. = 1)	2	0	0.4	0.5	0.0	0.0	1.0
Education (HS = 1)	2	0	0.6	0.5	0.0	1.0	1.0
Facebook (Yes = 1)	2	0	0.9	0.3	0.0	1.0	1.0
Partisan (Yes = 1)	2	0	0.6	0.5	0.0	1.0	1.0
Lula suppoter (Yes = 1)	2	0	0.3	0.5	0.0	0.0	1.0
Bolsonaro supporter (Yes = 1)	2	0	0.4	0.5	0.0	0.0	1.0
PT supporter (Yes = 1)	2	0	0.3	0.4	0.0	0.0	1.0

Table A17: Balance Tests

		Phatic		Misinfo			
		Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p
Age		38.27	13.92	37.66	13.98	-0.61	0.15
Sex		0.54	0.50	0.57	0.50	0.03	0.06
Religion		0.41	0.49	0.41	0.49	0.01	0.63
Education		0.59	0.49	0.60	0.49	0.01	0.62
Facebook User		0.86	0.34	0.86	0.35	-0.01	0.44
Lula Sup.		0.32	0.47	0.31	0.46	-0.01	0.59
Bols. Sup.		0.36	0.48	0.36	0.48	0.00	0.82
PT Sup.		0.29	0.45	0.28	0.45	-0.01	0.36
		N	Pct.	N	Pct.		
Income	1	216	12.8	437	13.1		
Categories	2	423	25.0	803	24.2		
	3	434	25.7	810	24.4		
	4	447	26.4	931	28.0		
	5	170	10.1	343	10.3		
Interest.	1	343	20.3	641	19.3		
Politics	2	631	37.3	1149	34.6		
	3	513	30.4	1049	31.6		
	4	203	12.0	485	14.6		

G Balance Tests - Experimental Study

Table A18: Balance Test, Omnibus Test

	Misinformation
(Intercept)	0.688***
	(0.041)
Age	−0.001
	(0.001)
Sex	0.025+
	(0.014)
Education	−0.003
	(0.015)
Facebook user	−0.006
	(0.019)
Religion	0.012
	(0.014)
Lula Sup.	0.001
	(0.020)
Bolso Sup.	0.004
	(0.015)
PT Sup.	−0.014
	(0.021)
Income cat. 2	−0.021
	(0.023)
Income cat. 3	−0.033
	(0.024)
Income cat. 4	−0.015
	(0.025)
Income cat. 5	−0.025
	(0.031)
Interest 2	−0.008
	(0.019)
Interestr 3	0.016
	(0.020)
Interest 4	0.051*
	(0.025)
Num.Obs.	5014
R2	0.003
R2 Adj.	0.000

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

The p-value of f-tests for the joint tests of significance is 0.31 (for a regression of misinformation treatment indicator on all pre-treatment covariates). Other joint tests of significance – for regressions of fake indicator (vs hyper) or fake indicator (vs phatic) or hyper indicator (vs phatic) on all pre-treatment covariates – have p-values above 0.3.

We also assessed balance by conducting a multinomial regression of our three treatment categories on all pre-treatment covariates, compared with a null model of a multinomial regression of treatment categories on a constant and found that these models are statistically indistinguishable.

H Additional Experimental Analysis

In this section, we present additional information regarding: i) attrition for the MSMT analysis, ii) survey taker attentiveness, iii) robustness tests for analyses and additional information.

H.1 Attrition MSMT

In our pilots and pre-testing, we learned that many respondents were unable to properly see and interact with MSMT and, when that happened, they were excluded from the survey. As a consequence, MSMT was placed at the end of the survey so that we would minimize respondents' attrition. Furthermore, while we know which respondents were assigned to each treatment condition in the MSMT (same condition they were exposed to earlier in the survey), we are unable to connect the MSMT data (which is stored in a separate server and in Qualtrics' dataset) to the other respondents' reponses in that survey.[36]

Out of 5,104 respondents, 3,467 were able to access MSMT in Qualtrics. As A19 shows, respondents in different treatment conditions are equally like to take part in MSMT, even though attrition is post-treatment. Unfortunately, since we do not have access to covariate data, we are unable to conduct balance tests and to test if attrition is associated with specific attributes. Our pilot tests suggests that attrition is associated with older phone operating systems, which are independent of treatment assignment, but the information available to us about operating system is not precise enough to explain attrition.

Table A19: Comparing Treatment Assignment MSMT and Full Survey

Experimental Condition	N MSMT	Percent MSMT	N Full Survey	Percent Full Survey
1	1167	33.66	1688	33.67
2	1150	33.17	1636	32.63
3	1150	33.17	1690	33.71

H.2 Attentiveness

We have three attention checks, two measured pre-treatment and one measured post-treatment. We find overall high levels of attentiveness both pre- and post-treatment.

[36]Connecting other questionnaire questions to the MSMT would require the users to type in their own id, which is error-prone, costly, and would lead to even higher rates of attrition.

Table A20: Percent of Respondents Passing Attention Checks

Attention Check	Pct.
Attention Check 1 (pre-treat)	96.71
Attention Check 2 (pre-treat)	91.22
Attention Check 3 (post-treat)	81.73
Sum of Attention Checks Passed = 0	1.26
Sum of Attention Checks Passed = 1	5.54
Sum of Attention Checks Passed = 2	15.48
Sum of Attention Checks Passed = 3	77.72
Average	945

H.3 Additional Experimental Analyses

Table A21: Respondents' perceptions on whether source politician's political orientation

	Pro-Bolsonaro Politician	Pro-Lula Politician	Neither Pro-Lula nor Pro-Bolsonaro
False	52.20	19.38	28.42
Hyper	49.67	24.00	26.32
Phatic	38.29	20.27	41.44

I Pre-Registered Hypotheses and Deviations from Pre-Analysis Plan

Tables A22, A23, and A24 provide more information about how our hypotheses relate to our pre-registered hypotheses and how our results directly address our hypotheses.

We deviate from our pre-analysis plan because we show results unadjusted by covariates. We will include the results adjusted by covariates in future versions of the paper. Furthermore, we will employ corrections for multiple comparisons (showing both nominal and corrected p-values) for the PAP hypotheses (H1A, H1B, H2A, H2B, H3A, H3B, H4A, H4B, H5A, H5B, H6A, H6B, H7A, H7B, H8A, H8B) in future versions of the paper. We also discussed in the PAP to rescale all outcomes between 0 and 1 so they have a common scale. For this version, we decided to keep the outcomes in their scales for ease of interpretation, but we will consider rescaling them if it makes interpretation more intuitive and comparison across estimates easier.

Table A22: Connections between Hypotheses, Pre-Registered Hypotheses, and Findings

Hypothesis in Paper	**Pre-Registered Hypothesis**	**Studies**	**Findings**	**Table/Figure**
H1: Posts containing misinformation generate more engagement and participation than posts that do not contain misinformation.	H1: Posts containing misinformation generate more mobilization than posts that contain phatic communication	Observ., Experimental, and MSMT	Posts containing misinfo. are asociated with more reactions (obs.) than posts without minsfo. Compared to phatic posts, misinfo. posts decrease OMI and increase NMI (exp.), and are associated with more reactions, excepts likes and love (MSMT).	Table 1 Table 2 Figure 1 Table 5 Table 6 Figure 3 Figure 4
H1A: Posts containing false stories generate more participation than posts that do not contain misinformation	H1A: Posts containing fake news generate more mobilization than posts that contain phatic communication	Observ. and Experimental	Posts containing false stories have more engagement than posts without false stories (obs.). Compared to phatic posts, posts with false stories decrease OMI and increase NMI.	Table 1 Table 2 Table 5 Table 6
H1B: Posts containing hyperpartisan news generate more participation than posts that do not contain misinformation.	H1B: Posts containing hyperpartisan news generate more mobilization than posts that contain phatic communication	Observ. and Experimental	Posts containing hyperpartisan stories have more engagement than posts without false stories (obs.). Compared to phatic posts, posts with hyperpartisan stories increase NMI (p-val ≈ 0.1), but do not affect OMI.	Table 1 Table 2 Table 5 Table 6
H2: Posts containing misinformation generate higher levels of net mobilization (i.e., more mobilization against Lula or in favor of Bolsonaro) than posts containing phatic communication.	H4 (labeled as exploratory) Posts containing misinformation generate higher levels of net mobilization (i.e., more mobilization against Lula or in favor of Bolsonaro) than pots that contain phatic communication.	Exp.	Compared to phatic posts, misinformation leads to higher levels of net mobilization	Table 6 Figure 3
H2A: Posts containing false stories generate more net mobilization than posts containing phatic communication.	H4A (labeled as exploratory): Posts containing fake news generate more net mobilization than posts that contain phatic communication	Exp.	Compared to phatic posts, false posts lead to higher levels of net mobilization	Table 6
H2B: Posts containing hyperpartisan news generate more net mobilization than posts containing phatic communication.	H4B (labeled as exploratory): Posts containing hyperpartisan news generate more net mobilization than posts that contain phatic communicatio	Exp.	Compared to phatic posts, false posts lead to higher levels of net mobilization (p-vale ≈ 0.1)	Table 6
H3: Posts containing misinformation are more likely to increase affective polarization towards political groups than posts containing phatic communication.	H3: Posts containing misinformation are more likely to increase affect polarization towards political groups than posts that contain phatic communication	Exp.	Compared to phatic posts, we fail to find an increase in affective polarization, but we find a decrease in affect towards PT supporters (unplanned) and no effect towards Bolsonaro supporters (unplanned)	Table 7

Table A23: Connections between Hypotheses, Pre-Registered Hypotheses, and Findings

Hypothesis in Paper	**Pre-Registered Hypothesis**	**Studies**	**Findings**	**Table/Figure**
H3A: Posts containing false stories are more likely to increase affective polarization towards political groups than posts containing phatic communication	H3A: Posts containing fake news are more likely to increase affect polarization towards political groups than posts that contain phatic communication	Exp.	Compared to phatic posts, we fail to find an increase in affective polarization, but we find a decrease in affect towards PT supporters (unplanned) and no effect towards Bolsonaro supporters (unplanned)	Table 7
H3B: Posts containing hyperpartisan news are more likely to increase affective polarization towards political groups than posts containing phatic communication	H3B: Posts containing hyperpartisan news are more likely to increase affect polarization towards political groups than posts that contain phatic communication	Exp.	Compared to phatic posts, we fail to find an increase in affective polarization, and no effects towards Bolsonaro and PT supporters (unplanned)	Table 7
H4: Posts containing misinformation are more likely to increase participation among voters with political identities than among voters without political identities.	H2: Posts containing misinformation are more likely to increase overall mobilization levels among partisans than among nonpartisans	Exp.	We fail to find a statistically significant difference in the effect of misinformation between partisan and non-partisans	Table 8
H5: Posts containing misinformation are more likely to increase net mobilization among voters who are supporters of Bolsonaro than among voters who are supporters of Lula and/or the Workers' Party	H6 (labeled as exploratory): Misinformation is more likely to increase net mobilization among Bolsonaristas than among petistas	Exp.	We fail to find a statistically significant difference in the effect of misinformation between Bolsonaro and Lula supporters	Table 8
RQ1: Do posts containing falsehoods have a different effect on levels of mobilization compared to posts that contain hyperpartisan content?	RQ1: Do posts containing fake news have a different effect on levels of overall mobilization compared to posts that contain hyperpartisan content? and RQ2: Do posts containing fake news have a different effect on levels of net mobilization compared to posts that contain hyperpartisan content?	Obs. and Exp.	We fail to find that the effects of posts containing false content are systematically statistically distinguishable from the effects of posts containing hyperpartisan content (even if coefficients for false content are somewhat larger in absolute value)	Differences reported in text
Unnamed in paper: examining the differences in responses to tweets containing falsehoods	Not part of PAP	Obs.	Replies and quotes suggest a small share of social media users questioning/fact checking/criticizing tweets with falsehoods (clearer in quotes than replies)	Figure 2
Unnamed in paper: examining the differences in MSMT reactions	H8 (labeled as exploratory): Posts containing misinformation generate more reactions than posts that contain phatic communication	MSMT	When aggregating all reactions, we find that posts with misinformation reduce the number of reactions (p-value < 0.1). But when breaking down by types of reactions, there is substantial heterogeneity suggesting negative effects for likes and love and positive effects for the other reactions.	Figure 4 Table A27

Table A24: Connections between Hypotheses, Pre-Registered Hypotheses, and Findings

Hypothesis in Paper	**Pre-Registered Hypothesis**	**Studies**	**Findings**	**Table/Figure**
Not in paper: examining the differences in MSMT reactions (false)	H8A (labeled as exploratory): Posts containing fake news generate more reactions than posts that contain phatic communication	MSMT	When aggregating all reactions, we find that posts with false stories reduce the number of reactions (p-value $<$ 0.1). But when breaking down by types of reactions, there is substantial heterogeneity suggesting negative effects for likes and love and positive effects for the other reactions.	Figure A4
Not in paper: examining the differences in MSMT reactions (hyper)	H8B (labeled as exploratory): Posts containing hyperpartisan generate more reactions than posts that contain phatic communication	MSMT	When aggregating all reactions, we fail to find that posts with hyperpartisan are associated with number of reactions (p-value $>$ 0.1). But when breaking down by types of reactions, there is substantial heterogeneity suggesting negative effects for likes and love and positive effects for the other reactions.	Figure A4
Not in paper	H5 (labeled as exploratory): Misinformation is more likely to increase overall mobilization levels among Bolsonaristas than among Petistas/Lulistas	Exp.	We fail to find different effects of posts with misinformation between Bolsonaro and Lula supporters	Table A25
Not in paper	H5A (labeled as exploratory): Posts containing fake news are more likely to increase overall mobilization levels among Bolsonaristas than among Petistas/Lulistas.	Exp.	We fail to find different effects of posts with false stories between Bolsonaro and Lula supporters	Table A25
Not in paper	H5B (labeled as exploratory): Posts containing hyperpartisan news are more likely to increase overall mobilization levels among Bolsonaristas than among Petistas/Lulistas.	Exp.	We fail to find different effects of hyperpartisan posts between Bolsonaro and Lula supporters	Table A25
Not in paper	H7 (labeled as exploratory): Posts containing misinformation are more likely to increase disposition to engage with source politicians than posts that contain phatic communication.	Exp.	Compared to phatic posts, posts with false stories decrease the willingness to follow the source politicians	Table A26
Not in paper	H7A (labeled as exploratory): Posts containing fake news are more likely to increase disposition to engage with source politicians than posts that contain phatic communication	Exp.	Compared to phatic posts, posts with false stories decrease the willingness to follow the source politician	Table A26
Not in paper	H7B (labeled as exploratory): Posts containing hyperpartisan news are more likely to increase disposition to engage with source politicians than posts that contain phatic communication.	Exp.	Compared to phatic posts, we fail to find a statistically significant effect of hyperpartisan posts on following source politician	Table A26

I.1 Analyses in PAP but not shown in paper

Table A25: The Effect of Misinformation by Political Support on OMI (PAP H5, H5A, H5B)

	OMI	OMI (FN vs Phatic)	OMI (HP vs Phatic)
(Intercept)	0.158***	0.158***	0.158***
	(0.046)	(0.046)	(0.046)
Misinfo.	−0.110+	−0.126+	−0.094
	(0.058)	(0.068)	(0.067)
Bolso. Supporter	0.106+	0.106+	0.106+
	(0.061)	(0.061)	(0.061)
Misinfo. * Bolso. Supporter	0.108	0.127	0.088
	(0.076)	(0.089)	(0.089)
Num.Obs.	2810	1885	1877
R2	0.010	0.010	0.007
AIC	7811.6	5221.1	5198.0
BIC	7841.3	5248.8	5225.7
RMSE	0.97	0.96	0.96

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A26: The Effect of Misinformation on Following Source Politician (H7, H7A, H7B)

	Source (NMI)	Source NMI (FN vs Phatic)	Source NMI (HP vs Phatic)
(Intercept)	−0.405***	−0.405***	−0.405***
	(0.034)	(0.034)	(0.034)
Misinfo.	−0.169***	−0.310***	−0.031
	(0.042)	(0.049)	(0.049)
Num.Obs.	5014	3326	3378
R2	0.003	0.012	0.000
AIC	17 846.8	11 713.1	12 020.0
BIC	17 866.3	11 731.4	12 038.4
RMSE	1.43	1.41	1.43

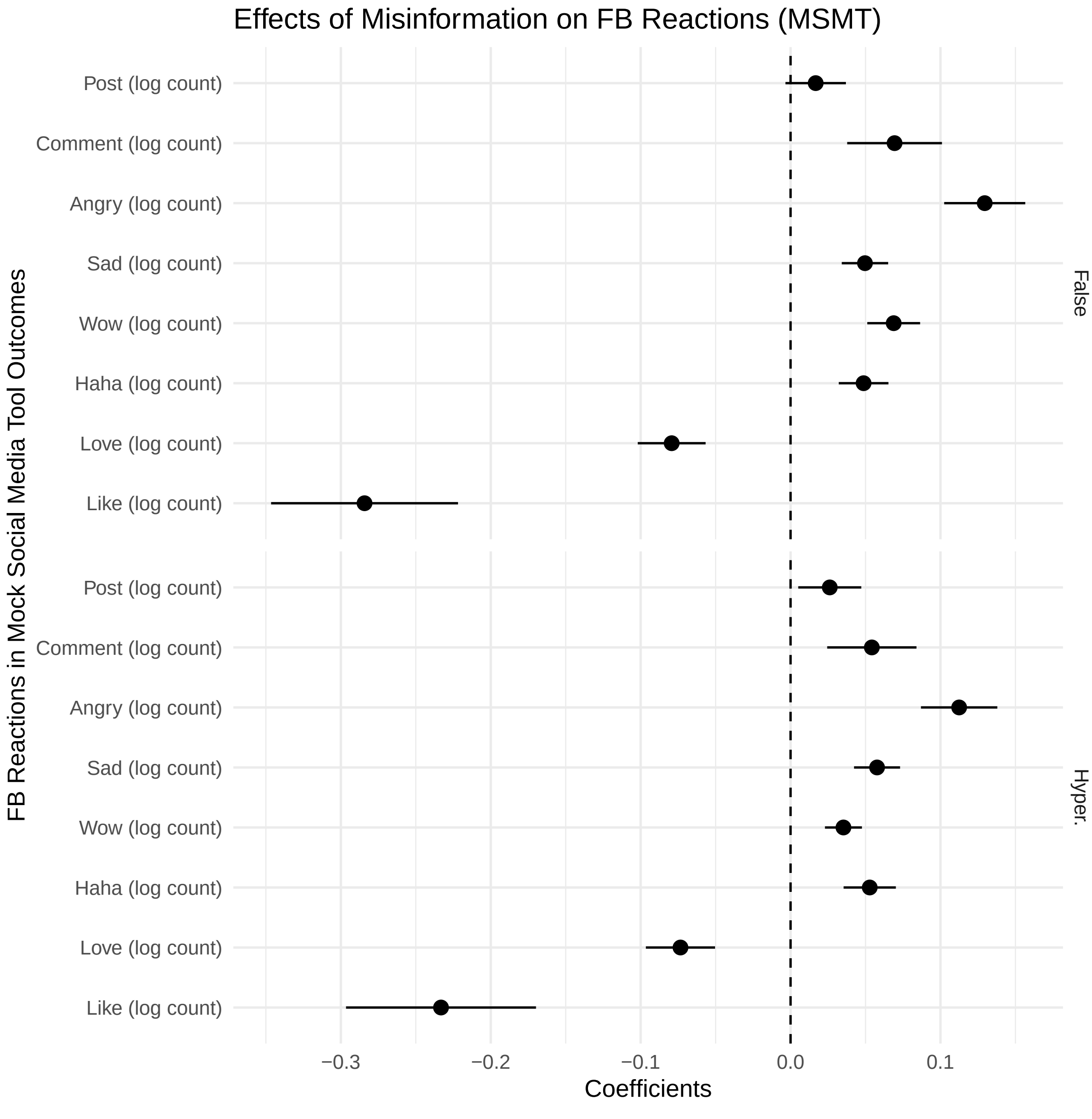
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A27: The Effect of Misinformation on MSMT reactions (log)

	Model 1	Model 2
(Intercept)	0.678***	0.678***
	(0.029)	(0.029)
Misinfo.	−0.061+	
	(0.035)	
False		−0.073+
		(0.040)
Hyper.		−0.050
		(0.040)
Num.Obs.	3467	3467
R2	0.001	0.001
AIC	9467.4	9469.1
BIC	9485.9	9493.7
RMSE	0.95	0.95

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Figure A4: Effects of False and Hyperpartisan Posts on MSMT Reactions



J Data Sources & Transparency

1. Politicians' handles on Twitter, Instagram, and Facebook: collected by research assistants
2. Facebook and Instagram posts collected via CrowdTangle
3. Twitter data collected via academic Twitter API
4. Fact-checking data collected via scraping
5. Survey data collected with survey company Quaest and Qualtrics platform
6. Facebook URL dataset collected via partnership/access established via Social Science One
7. Survey pilot data collected with survey company Quaest

K Full Instrument

In later versions of the paper, we will include all relevant questions translated to English here.

To code respondents as having a political identity, we identify respondents who chose the following answer categories, “I like, but I don’t feel like a [Bolsonarista][Lulista][petista]/I am a [Bolsonarista][Lulista][petista].”, “I hate [Bolsonaro][Lula][PT]”, “I don’t like [Bolsonaro] [Lula][PT]”, and as not having a political identity those respondents who chose “I neither like nor dislike [Bolsonaro][Lula].” We measure supporter of Bolsonaro (Lula) as those who chose “I like, but I don’t feel like a [bolsonarista][lulista]/I am a [bolsonarista][lulista] as answer to questions about support for Bolsonaro (Lula).

In Portuguese: “Gosto, mas não me sinto um(a) [bolsonaristas][lulista][petista]/Sou um(a) [bolsonarista][lulista][petista].”, “Eu detesto o [Bolsonaro][Lula][PT]”, “Não gosto do [Bolsonaro] [Lula][PT]” and as not having a political identity those respondents who chose “Eu não gosto nem desgosto do [Bolsonaro][do Lula].” We measure supporter of Bolsonaro (Lula) as those who chose “Gosto, mas não me sinto um(a) [bolsonaristas][lulista]/Sou um(a) [bolsonarista][lulista].”

See the full instrument (in Portuguese) here: https://www.dropbox.com/s/eop44x9i7bszsd0/instrument.pdf?dl=0