"I Don't Think That's True, Bro!"

Social Corrections of Misinformation in India

Sumitra Badrinathan
American University

Simon Chauchard
University Carlos III Madrid *

Abstract

Fact-checks and corrections of falsehoods have emerged as effective ways to counter misinformation online. But in contexts with encrypted messaging applications (EMAs), corrections must necessarily emanate from peers. Are such social corrections effective? If so, how substantiated do corrective messages need to be? To answer these questions, we evaluate the effect of different types of social corrections on the persistence of misinformation in India (N=5100). Using an online experiment, we show that social corrections substantially reduce beliefs in misinformation, including in beliefs deeply anchored in salient group identities. Importantly, these positive effects are not systematically attenuated by partisan motivated reasoning, highlighting a striking difference from Western contexts. We also find that the presence of a correction matters more relative to how sophisticated this correction is: substantiating a correction with a source only improves its effect in a minority of cases; besides, when social corrections are effective, citing a source does not drastically improve the size of their effect. These results have implications for both users and platforms and speak to countering misinformation in developing countries that rely on private messaging apps.

Keywords: Misinformation; Social Media; Peer Correction; Motivated Reasoning; WhatsApp; India

*This study was registered with Evidence in Governance and Policy (20191008AB) and received IRB approval from Columbia University and IE University (IRB-AAAS3860). We thank D.J. Flynn for help and involvement in the planning, design and implementation phases of this study. For research assiatance we thank Ipsa Arora, Ritubhan Gautam, and Hanmant Wanole. This research was funded by Facebook. The authors thank Alex Leavitt and Devra Moelher, as well as participants at the Facebook Integrity Research Workshop. For comments on the manuscript we are grateful to seminar participants at Columbia University, Leiden University, and the American Political Science Association conference.

Introduction

Over the past decade, a vast research agenda has tested the effect of corrective messages to fight misinformation on social media. The majority of such studies suggest that corrective interventions lead to small and beneficial effects (Chan et al. 2017; Porter and Wood 2021). Since users sometimes also correct each other, a part of this literature has, in addition, explored the effect of "social corrections" and found them to be similarly efficient (Bode and Vraga 2018).

However, these encouraging findings remain largely limited to Western democracies and Global North contexts, with little scholarship, by contrast, focusing on misinformation in the developing world (Pereira et al. 2021; Rosenzweig et al. 2021; Guess et al. 2020). In studies with cross-country comparisons that move beyond the Global North (Arechar et al. 2022), the role of social corrections still remains unexplored. Not only are countries in the Global South understudied, but also misinformation in these contexts can have disastrous offline consequences such as fueling political violence and ethnic riots.

In this article, we build on the emerging literature on social corrections – i.e. peer-to-peer corrective messages – and explore the extent to which such corrections are effective in contexts in which they are most sorely needed. In countries like India, pernicious misinformation builds on longstanding ethnic fractures and can spark conflict and violence. For example, during the coronavirus pandemic, conspiracy theories suggesting that minority religious groups were intentionally conspiring to spread the virus resulted in discrimination, harassment and violence towards India's Muslim community, intensifying an already fraught communal divide (Yasir 2020). Further, since a large part of such misinformation circulates through encrypted messaging applications (EMAs) where corrections can *only* come from other users (as opposed to platform-based warnings and labels), it is especially crucial to understand whether social corrections can play a role in combating misinformation in such contexts.

Are social corrections effective in these contexts? If so, what type of corrective messages work best at altering misinformed beliefs? To answer these questions, this study tests the effect of the presence of social corrections and the type of correction on belief in electoral misinformation in India. We implement a large-scale online experiment conducted in Hindi with over 5,100 respondents in the aftermath of the contentious 2019 general elections, which saw a dramatic surge in political misinformation on social media (Sinha, Sheikh, and Sidharth 2019; Chauchard and Garimella 2022). We show respondents a series of hypothetical WhatsApp group chat conversations. In each conversation, a user posts a false story to which a peer reacts, with or without a correction, and with or without citing evidence.

Results demonstrate that the presence of a social correction significantly reduces belief in misinformation. Relative to a no correction condition, witnessing *any* type of social correction reduces the perceived accuracy of beliefs in misinformation. Importantly, this effect is robust to respondents' partisan identity and persists across different types of misinformation stories, including deep-rooted beliefs. We find that partisan motivated reasoning does not systematically attenuate these corrective effects, suggesting important differences in the mechanisms through which misinformation spreads in contexts like India. Finally, we show that the source and the sophistication of corrective messages do not strongly condition their effect: in our experiment, brief, unsourced and unsubstantiated corrective messages often perform just as well relative to corrections citing evidence from a variety of credible sources.

We note two key implications of our findings. On the one hand, our results suggest that social corrections have a positive effect in the Indian context. Hence, incentivizing users to engage in correcting misinformation or verbalizing their fact-checks on homogeneous networks may help in the reduction of misinformed beliefs. On the other hand, because respondents emerge – in this context – as influenceable (minimal corrections are relatively effective), encouraging users to sound off as easily as possible may also have perverse consequences if bad faith actors themselves use such strategies to peddle more misinformation. Indeed, some evidence shows that debunking true stories may also increase belief in misperceptions (Vraga, Bode et al. 2022). This study hopes to spark a robust research agenda on solutions to misinformation in the Global South, focusing on the challenges that encryption, low digital literacy, and social cleavages bring to information processing.

Social Corrections Across Platforms and Contexts

Social corrections have become a key element in combating misinformation, with platforms increasingly encouraging users to check on each other. This is especially relevant to EMAs: in 2019, faced with public outcry over misinformation going viral on the platform in the lead-up to the general election, WhatsApp launched a large-scale ad campaign to encourage its Indian users to check on their peers' claims.[1] If users indeed followed through with this recommendation and fact-checked their peers, to what extent should we expect such a strategy to reduce the uptake of misinformation in contexts like India?

Existing research focused on the American context suggests that social corrections might indeed be beneficial. Bode and Vraga (2018) compare algorithmic corrections with peer corrections and find that they are both equally effective at dispelling misinformation. van der Meer and Jin (2020) demonstrate that peer corrections are effective at reducing misinformation relative to a control condition.

The literature on source credibility provides a mechanism explaining why social corrections may be effective. Individuals have limited time and cognitive resources to comprehend complex topics such as policy or current affairs, and may therefore use the perceived credibility of sources as a heuristic to guide their evaluation of what is true

[1] See Appendix A.

or false (Eagly and Chaiken 1993). Further, while both expertise and trustworthiness are components of source credibility, the latter is found to be more effective in persuasion than the former (Swire and Ecker 2018). Thus, arguably, peers should be seen as more trustworthy than unknown or distant individuals; indeed, research shows that corrections from peers as well as corrections from those in the same social network are more likely to be accepted (Margolin, Hannak, and Weber 2018).

Additionally, homophily, or the extent to which a person perceives similarities between the way they think and another person does, is often seen as a determinant of source credibility (Housholder and LaMarre 2014). Networks on social media, including in India, are likely to be comprised of likeminded individuals who share political and other views (Tokita, Guess, and Tarnita 2021). Indian WhatsApp users are often part of several large WhatsApp groups, formed to rally around shared issues and interests including political and civic causes (Mont'Alverne et al. 2022). Though they may not personally know every member in these groups, they share a keen connection with like minded ingroup members: many users in closed groups may feel pressure not to believe or say something that runs against the dominant view that anchors the group identity (Davies 2020). Recent research also demonstrates that misinformation may be more rampant in homophilic networks such as these (Acemoglu, Ozdaglar, and Siderius 2021). Thus, we argue that while social corrections may emanate from strangers, these strangers could often be seen as "peers" in the Indian WhatsApp context, making their corrections effective. Accordingly, we hypothesize:

Hypothesis 1: Exposure to corrective messages emanating from peers will reduce the perceived accuracy of misinformation, relative to a no correction condition.

We note here that we test the effect of corrections on those who witness those corrections, as opposed to respondents being corrected themselves. In order to preserve external validity and present corrections in the format respondents would normally see them, we choose to show corrections in a group EMA setting where respondents observe a peer correcting another peer.

Despite the relative success of fact-checking efforts, the empirical literature on misinformation demonstrates that the success of corrections is a function of individuals' preexisting beliefs. In this regard, a primary factor influencing the efficacy of corrections is partisan motivated reasoning (Thorson 2016).[2] In India, however, that partisan motivated reasoning will affect corrections is not a foregone conclusion (Badrinathan 2021). On the one hand, partisan affiliations in India have been shown to be traditionally weaker and less stable, sometimes forming for non-ideological reasons (Bussell 2019). On the other hand, reports show that misinformation is largely political on WhatsApp, and majority of political content comes from groups allied with the Bharatiya Janta Party (BJP), the right-wing, Hindu nationalist government of India (Garimella and Eckles 2020). Further, since we conduct this experiment after a contentious election, during which attachments to parties were arguably heightened (Michelitch and Utych 2018), we might expect motivated reasoning to play a role in information processing.

Despite these contrasting priors, in keeping with findings from the literature on fact-checking, we hypothesize that partisan motivated reasoning should attenuate the effects of corrective messages (Taber and Lodge 2006). Specifically, we hypothesize:

Hypothesis 2: Peer corrections will be more effective when misinformation is attributed to an ideologically dissonant politician (compared to when it is unattributed).

Hypothesis 3: Peer corrections will be more effective when misinformation originates from a dissonant media outlet (compared to unattributed or neutral outlet).

Hypothesis 4: Peer corrections will be less effective when misinformation is ideologically congruent (compared to non-ideological stories).

Finally, we also consider whether corrections backed by evidence are more effective. Research demonstrates that the persuasive quality of an argument is a function of

[2] We note, however, that a small body of research now points to factors other than motivated reasoning to understand belief in misinformation, such as analytic thinking and inattention to accuracy (Pennycook and Rand 2019).

whether or not it is substantiated (Stiff and Mongeau 2016). With social media, Vraga and Bode (2018) test the effect of social corrections on Facebook and Twitter and find that corrections substantiated with a source are more effective at countering misinformation. We thus hypothesize:

Hypothesis 5: Exposure to unsubstantiated corrections (relative to substantiated corrections) will be less effective at countering misinformation.

Design

To test these hypotheses, we designed and fielded an online experiment in India (N $\approx$ 5,100) in the aftermath of the contentious 2019 general elections (registered with OSF). In our experiment, respondents were recruited through Facebook and were shown (in random order) a series of nine hypothetical conversations on WhatsApp group chats, seven of which contained misinformation.[3]

We choose to show respondents the treatment through WhatsApp screenshots as WhatsApp is the most common social media platform in India. With over 480 million users, India is also WhatsApp's largest market. Moreover, in countries like India, political parties and others are known to organize voters into WhatsApp groups curated by religious and other social identities, a communication strategy that gained popularity after the 2014 elections (Chauchard and Garimella 2022). Given the platform affordances WhatsApp offers in terms of sorting users into private and encrypted homophilic groups, WhatsApp users are likely accustomed to hearing from strangers in homophilic networks, a strategy we lean on for this experiment. Further, research shows that WhatsApp users in India are particularly likely to forward messages in groups, constituting a common method for the spread of misinformation (Aneez et al. 2019).

[3] This ratio (7/9) is not meant to be representative of the prevalence of misinformation on online platforms. Rates of misinformation online in a similar context are far smaller, as documented in Chauchard and Garimella (2022). In this design we choose to focus on a larger sample of misinformation to be able to experimentally test the effects of different types of corrections.

The theme of these nine WhatsApp conversations – and hence of the misinformation stimulus respondents were shown – was chosen following a pretest with an online Indian sample.[4] These misinformation stimuli in the WhatsApp conversations, along with their veracity and partisan slant, are listed in Table 1. Every conversation containing a misinformation stimulus was experimentally manipulated to take on several different variations. Hence, respondents took part in nine successive experiments, presented in random order.

The final selection of stories was the product of several constraints and choices. To avoid prompting respondents to systematically reject the veracity of rumors, we included some true stories (2 out of 9). But simultaneously, our goal was to maximize respondent exposure to corrections for controversial fake political rumors that spread widely during the run up to the 2019 elections in India, hence our distribution skewed in favor of false stories (7 out of 9). We selected stories encompassing a broad variety of topics including current electoral politics (stories 8 and 9), health (stories 5 and 6), historical conspiracies (story 7), and in order to test the effect of social corrections on identity-related misinformation, religion and minorities (stories 3 and 4).

Table 1: Veracity and Slant of Dependent Variable Stories

	Story	**Veracity**	**Slant**
1	Australia has won the ICC cricket world cup the most often	True	Neutral
2	There is no cure for HIV/AIDS	True	Neutral
3	The Muslim population in India will overtake the Hindu population	False	BJP-congruent
4	Polygamy is very common in the Muslim population	False	BJP-congruent
5	MMR vaccines are associated with autism and retardation	False	Neutral
6	Drinking cow urine (gomutra) can help build one's immune system	False	BJP-congruent
7	Netaji Bose did not die in a plane crash in 1945	False	Neutral
8	The BJP has hacked electronic voting machines	False	BJP-incongruent
9	UNESCO declared PM Modi best Prime Minister in 2016	False	BJP-congruent

[4] Data from the pretest is presented in Appendix I.

The experimental manipulations for our stories are described below:

1. **Type of correction**. This included four possible conditions: Control (no correction), Domain Expert correction, Fact Checker correction, and Unsubstantiated Correction. Respondents were randomized into one of these 4 groups with equal probability. The between-subject randomization of participants into Control, Domain Expert, Fact Checker, and Unsubstantiated Correction re-occurs prior to each of the 9 successive screenshots respondents are exposed to. As a result, respondents in the control group during the first story need not remain in the control group throughout, as they may be assigned to another experimental group in a subsequent story. In practice, this is equivalent to 9 successive experiments, whose order is randomized (in addition to assignment to treatment being randomized). As a result of this strategy, respondents are typically assigned to a variety of corrections across different rumors, though some of them also likely see similar corrections back to back. Respondents in the Domain Expert treatment (25% of the sample) read a substantiated correction pointing to a domain expert as the source of the correction (for example, the Election Commission of India for electoral misinformation, or the Census Bureau of India for demographic misinformation). Respondents in the Fact Checker treatment (25% of the sample) read a substantiated correction pointing to a verified fact-checker in India as having debunked the misinformation posted.[5] Respondents in the Unsubstantiated Correction treatment (25% of the sample) read a correction that was a simple rebuttal by the second user, devoid of substantiation or a source of correction. This included a one-line simple correction (for instance saying "I don't think that's true, bro!") in response to the misinfor-

[5]The Fact Checker correction was further subdivided into 5 variations. That is, within the 25% of our sample that were randomized into the Fact Checker condition, respondents saw one of 5 versions with equal probability: a message that the online platform WhatsApp had factchecked the misinformation, or the online platform Facebook, or India's oldest print newspaper The Times of India, or third party fact-checking service AltNews, or third party fact-checking service Vishwas News. We do not focus in this manuscript on comparisons between fact checkers; instead in our key analyses we aggregate them up to compare the fact-checker conditions to the three aforementioned conditions.

mation stimulus. The remaining 25% of the sample was randomized into a control group that received no correction.

2. **Media outlet reporting the story**. This manipulation included 3 possible variations: India TV (a relatively right-leaning private news channel), NDTV Hindi (a relatively left-leaning private news channel), DD News (a relatively neutral public channel).

3. **Politician from whom the claim originated**. This manipulation included 3 possible variations: a BJP politician, an Indian National Congress (INC) politician, or an unspecified source. Note that in each story in which this manipulation was employed, we compared either a BJP or an INC politician to our neutral option (an absence of clear partisan source).

Within each of the 4 correction conditions, respondents had an equal probability of being assigned to each of the possible variations of media outlets x politicians. Altogether, this gives us in total 48 possible variations for each story: 8 corrections (including 5 types of Fact Checkers), 3 media outlets, 2 politicians. However, note that this is not a fully crossed design: not every story had 48 variations because we exclude unrealistic manipulations (for instance, non-political rumors originating from politicians). Thus, some stories only had 24 variations. In the table below we detail the precise variations that were included and excluded for each story, along with the total number of conditions per story. The full text of each experimental manipulation, along with samples of the experimental stimuli, is included in Online Appendix C.

Table 2: Experimental Variations for False Stories

Manipulation	Condition	MuslimPop	Polygamy	MMR	Cow Urine	Bose	EVM	UNESCO
Correction source	Control (no correction)	✓	✓	✓	✓	✓	✓	✓
	Unsubstantiated correction	✓	✓	✓	✓	✓	✓	✓
	Domain Expert correction	✓	✓	✓	✓	✓	✓	✓
	Fact Checker correction	✓	✓	✓	✓	✓	✓	✓
Media Source	NDTV Hindi	✓	✓	✓	✓	✓	✓	✓
	India TV	✓	✓	✓	✓	✓	✓	✓
	DD News	✓	✓	✓	✓	✓	✓	✓
Politician Source	INC	×	×	×	×	×	✓	×
	BJP	✓	✓		✓		×	✓
	Unspecified	✓	✓		✓		✓	✓
Total Number of Variations		48	48	24	48	24	48	48

Notes: X's indicate certain variations were excluded for some rumors because of lack of rumor-variation congruence.
Blank spaces mean the entire manipulation was not included for some rumors.
The Fact Checker correction accounts for 5 types of fact-checking sources.

To evaluate our main effects, we aggregate up to the level of receiving *any* correction (75% of the sample) compared to receiving no correction (25% of the sample). To further evaluate the effect of the *type* of correction, we compare the four key correction conditions to each other (unsourced correction vs. expert correction vs. *any* fact-checker correction vs. control).

Procedure and Outcome Variable

After reading each WhatsApp conversation, respondents were asked to evaluate the veracity of the misinformation stimulus included the conversation with a single outcome question:

How accurate is the following statement? [Statement of the rumor]
(not at all accurate, not very accurate, somewhat accurate, very accurate)

Our design took several steps to increase external validity and realism. First, as noted above, we selected a diverse sample of stories. Further, we excluded highly unrealistic manipulations and tailored domain expert corrections to each rumor (e.g., we attribute expert corrections of voter fraud rumors to the Election Commission of

India). Finally, given that respondents each saw nine screenshots, we slightly varied the specific text of the messages in each screenshot to ensure realism (see Appendix C).

Sample

Participants in this study were Hindi speakers recruited through Facebook. The ad used to recruit respondents is in Appendix B. To be eligible to participate, respondents were required to be adult residents of India who used WhatsApp.[6] While we recruited over 5100 participants, the actual N presented in our analyses varies slightly for each dependent variable story (+ or - 1%), as we include observations from respondents who exit the survey before reading all 9 rumors.[7] The experiment was conducted entirely in Hindi, and sample characteristics of our respondents are available in Appendix K. While our experiment does exclude non-Hindi speakers, given the diverse language background of the Indian population (including 22 official languages and several hundred dialects), having a survey that represents all language groups is a challenging prospect. The Hindi speaking share represents the largest language group in the country, including 57% of the population amounting to about 700 million people. We also note that we recruit respondents through Facebook as we estimate this better represents the online population in India relative to those on survey panels.[8]

[6]100% of those initially recruited said they used WhatsApp. This is consistent with social media usage in India where WhatsApp is the most popular platform, hence those with access to Facebook naturally also used WhatsApp.

[7]Only 75 respondents, or less than 1.5% of the sample dropped out through the course of the experiment. There was no scope for differential attrition between the four conditions as respondents are re-randomized into a different condition for each headline they see; i.e. they do not stay in the same condition for each story.

[8]While this recruitment strategy excludes respondents who use WhatsApp but not Facebook, we believe our sampling method is an improvement over using online panels in India as these are strictly limited to English-speaking users, who constitute a very small section of the population.

Results

For each of the 9 claims we asked about, Figure 1 plots the share of respondents in the sample who believed each story to be true. Our findings demonstrate the high salience of false stories in the Indian context. Despite the fact that 75% of all screenshots (across conditions) contained a correction, 6 of the 7 false rumors were rated as accurate or somewhat accurate by over 50% of the sample, with the top two prevalent rumors believed by over 75% of the sample, underscoring the tenacity of misinformation in the Indian context.

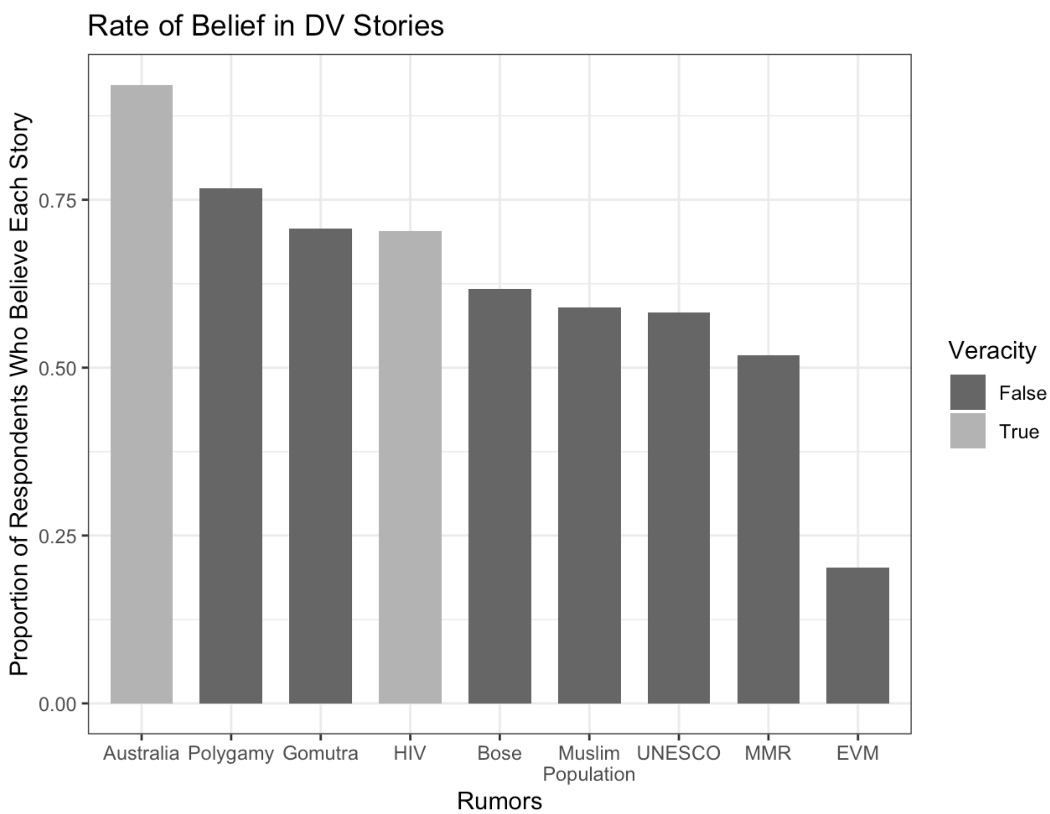


Figure 1: Overall rate of belief in DV stories (true and false) across experimental conditions

The Effect of Social Corrections

Do corrections impact these high rates of belief? We first present results for Hypothesis 1, which tests whether exposure to *any* correction reduces the perceived accuracy of misinformation. To test Hypothesis 1, we pool together all the different types of social corrections such that the primary comparison of interest is between having received a correction (of any kind) and not having received one. The dependent variable measures the self-reported accuracy rating that respondents give to each story on a 4-point scale, with higher values representing greater perceived accuracy (i.e., for false rumors this amounts to higher values meaning false stories are rated as true). Since our design amounts to running several experiments successively, we first estimate separate bivariate OLS models for each of the seven false stories in our experiment, represented by the seven columns in Table 3. We also visually represent our main result in Figure 2.

Our results demonstrate that corrections are effective at reducing beliefs in false stories. Exposure to social corrections significantly reduces the likelihood that respondents report false rumors to be accurate, relative to not receiving any correction. We do not obtain a significant result for only one (out of 7) false story, the rumor that electronic voting machines (EVMs) were hacked by the BJP ahead of the elections. As Figure 1 demonstrates, belief in this rumor was low to begin with, possibly making it harder for the treatment to have an impact. In contrast, a consistent negative effect appears for the remaining stories, although effect sizes vary across rumors. Particularly, we see effects of larger magnitude (greater than 0.4 on a scale from 1 to 4) on two of the stories: the MMR vaccine rumor and the UNESCO rumor.[9] Thus our results show that social corrections, the only suitable techniques for private online spaces, significantly reduce overall rates of beliefs in patently false rumors circulating on WhatsApp in India.

[9]The size of the effect across rumors does not appear related to the prior salience of these rumors in our sampled population. As shown in Appendix Figures I.1 and I.2, many respondents in our pretest had heard of the widely circulated UNESCO rumor, while fewer had heard of the rumor about MMR vaccines. Yet both led to comparatively large corrective effects.

Figure 2: Average Treatment Effect (Any Correction vs. Control). Error bars show 95% confidence intervals

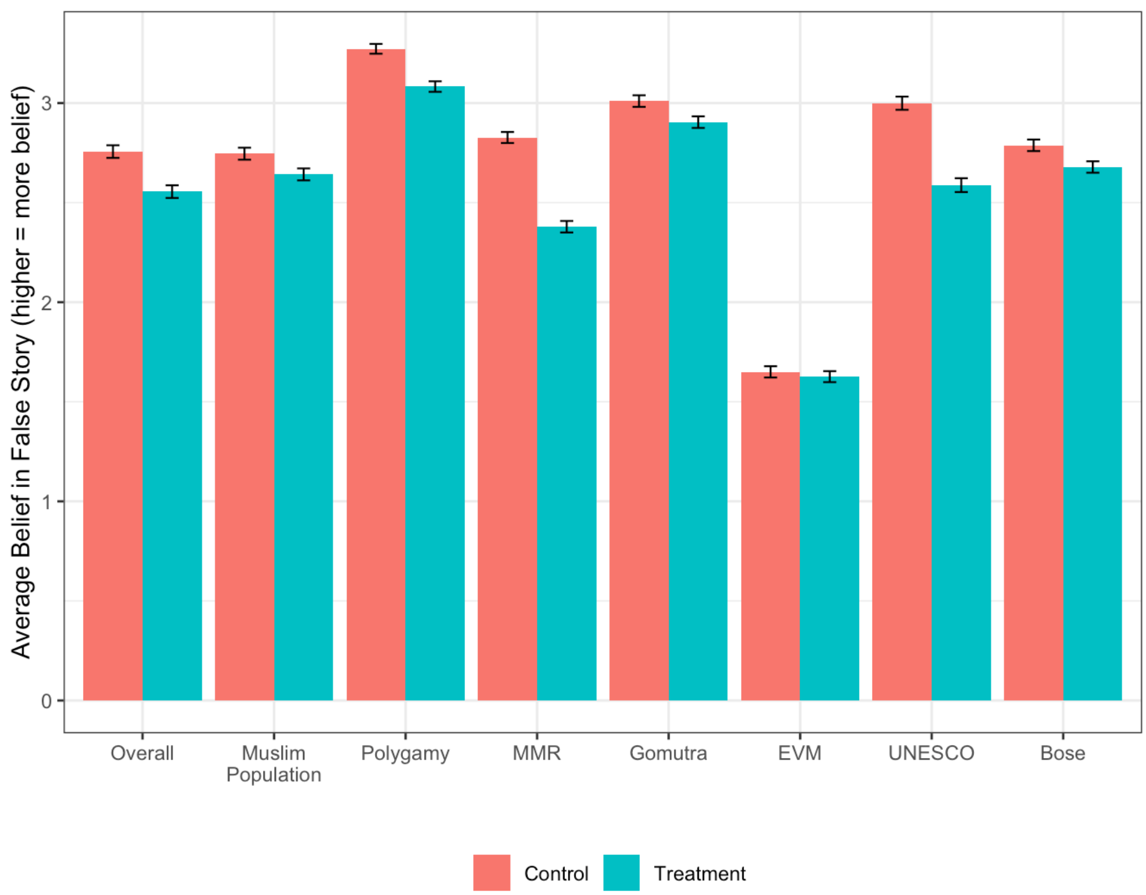


Table 3: Main Effect of Any Correction With Controls

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Any Correction	-0.106^{***}	-0.182^{***}	-0.428^{***}	-0.112^{***}	-0.018	-0.413^{***}	-0.116^{***}
	(0.037)	(0.031)	(0.033)	(0.033)	(0.032)	(0.040)	(0.032)
Dissonant Media	0.084^{*}	-0.011	-0.107^{**}	0.059	-0.087^{*}	0.016	0.030
	(0.050)	(0.044)	(0.046)	(0.049)	(0.046)	(0.059)	(0.047)
Congruent Media	0.041	0.101^{**}	-0.119^{***}	0.048	-0.119^{**}	-0.020	0.037
	(0.049)	(0.045)	(0.045)	(0.048)	(0.046)	(0.059)	(0.045)
Copartisan Politician	-0.022	-0.008		0.091^{*}	0.434^{***}	0.091	
	(0.051)	(0.046)		(0.050)	(0.061)	(0.060)	
Outpartisan Politician	-0.190^{***}	-0.271^{***}		-0.297^{***}	-0.141^{***}	-0.157^{**}	
	(0.064)	(0.057)		(0.065)	(0.047)	(0.080)	
Constant	2.747^{***}	3.275^{***}	2.840^{***}	3.008^{***}	1.662^{***}	2.999^{***}	2.785^{***}
	(0.033)	(0.026)	(0.028)	(0.028)	(0.027)	(0.034)	(0.025)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.004	0.013	0.037	0.008	0.015	0.022	0.003
Adjusted R^2	0.003	0.012	0.037	0.007	0.014	0.021	0.002
Significance (Main effect)	Yes	Yes	Yes	Yes	No	Yes	Yes
Res. Std. Er.	1.094	0.943	1.038	1.057	1.007	1.250	1.048
F Statistic	3.599^{***}	13.555^{***}	65.656^{***}	8.161^{***}	15.753^{***}	23.216^{***}	4.443^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

We find that our results are robust to controls: controlling for the media source or politician being congruent or dissonant, results hold.[10] Finally, we also present a pooled model averaging across all rumors to calculate an overall correction effect for all rumors (Table 6). In this model we also include robust two-way clustered standard errors at respondent and headline levels, and find that our results hold.

Motivated Reasoning and Social Corrections

To what extent are the corrective effects obtained above affected by motivated reasoning? To answer this question, we look at motivated reasoning in three ways, as hypothesized in H2, H3 and H4.

Tables 4 and 5 test Hypothesis 4, that is, whether the effect of the correction is a function of the slant of the story itself. While Table 4 looks at whether corrections are less effective for ideologically congruent stories, Table 5 looks are whether corrections are more effective for ideologically dissonant stories. We limit our analyses to the subset of rumors that are clearly political (rumors 3, 4, 6, 8,and 9). We code rumors as congruent or dissonant ex-ante as a function of participants own ideological inclinations and as a function of our observations of the two parties' campaign platforms. For this analysis, we coded claims == 1 when they were political claims congruent to respondent ideology (for example, if the respondent is a BJP supporter and the claim at hand is a rumor saying BJP politician Modi is the best PM in the world, this is coded as 1), and 0 otherwise, i.e. 0 represents both dissonant claims as well as neutral claims. When participants self-report being close or very close to the BJP, Rumors 3 (Muslim population growth), 4 (polygamy in the Muslim population), 6 (belief about the virtues of cow urine), and 9 (Modi and UNESCO) are coded as congruent rumors. By contrast, Rumor 8 (EVMs) is coded as

[10]Given that we report highly significant results ($p < 0.001$ on 3 of the 6 significant coefficients, and $p < 0.005$ on three more) in Table 6, we test whether these results are robust to a Benjamini-Hochberg adjustment. We show that our results remain significant when we perform a Benjamini-Hochberg adjustment fixing the false discovery rate at 5% (see Appendix L).

dissonant, while Rumors 1, 2, 5 and 7 are coded as neither congenial nor dissonant.

Table 4: Effect of Correction * Congruent Claim on Belief in Rumor

	Dependent variable: Belief in Rumor				
	MuslimPop	Polygamy	Gomutra	EVM	UNESCO
	(1)	(2)	(3)	(4)	(5)
AnyCorrection	-0.092	-0.162^{***}	-0.116^{*}	0.0002	-0.378^{***}
	(0.059)	(0.048)	(0.052)	(0.038)	(0.063)
CongruentClaim	0.238^{***}	0.246^{***}	0.369^{***}	0.520^{***}	0.366^{***}
	(0.067)	(0.053)	(0.055)	(0.056)	(0.069)
AnyCorrection * CongruentClaim	-0.025	-0.043	0.013	-0.057	-0.068
	(0.076)	(0.061)	(0.066)	(0.066)	(0.081)
Constant	2.602^{***}	3.120^{***}	2.784^{***}	1.478^{***}	2.780^{***}
	(0.052)	(0.041)	(0.043)	(0.032)	(0.053)
Observations	5,105	5,104	5,100	5,137	5,110
R^2	0.011	0.020	0.032	0.049	0.036
Adjusted R^2	0.010	0.019	0.031	0.049	0.035
Residual Std. Error	1.090	0.940	1.044	0.989	1.241
F Statistic	18.886^{***}	34.360^{***}	56.184^{***}	88.534^{***}	62.916^{***}
Note:			*p<0.05; **p<0.01; ***p<0.001		

Across Tables 4 and 5, the interaction between the treatment (AnyCorrection) and the slant of the story corrected produces a null result, suggesting that the effect of corrections may not be a function of the ideological slant of the story. We also repeat these analyses with a pooled model (where the dependent variable is the average perceived accuracy across all headlines). This analysis is presented in Appendix E and shows that congruent claims might further be corrected better than non-congruent claims, suggesting that social corrections can decrease belief even in congenial news. The divergence in results between our pooled model and these rumor-specific models, however, do point to a inconclusive finding: we do not find clear evidence that the effect of corrections is conditional on the ideological slant of the story; while it is the case in our pooled model, this is not true for any specific story that we test. While they will require confirmation through additional analyses or additional samples, these findings in our pooled model are nonetheless worth highlighting, as they depart from salient works implying that greater resistance to corrections exists for concordant news (Tappin, Pennycook, and Rand 2021).

Similar results emerge when we look at motivated reasoning in two other ways, as per the politician to whom a story is attributed (Hypothesis 2), or the news outlet reporting the story (Hypothesis 3). We find that interacting the correction with the identity of the politician or media outlet does not reduce or change the effect of corrections. Results from these tests are reported in Appendix Tables D.1 and D.2 (effect of the identity of the politician) and Tables E.1 and E.2 (effect of media outlet). Tables D.3 and E.3 pool across rumors to present average effects and find overall a lack of evidence for corrections being more or less effective as a function of ideology.

Table 5: Effect of Correction * Dissonant Claim on Belief in Rumor

	Dependent variable: Belief in Rumor				
	MuslimPop	Polygamy	Gomutra	EVM	UNESCO
	(1)	(2)	(3)	(4)	(5)
AnyCorrection	-0.127^{***}	-0.195^{***}	-0.088^{**}	-0.004	-0.032
	(0.045)	(0.036)	(0.039)	(0.049)	(0.040)
DissonantClaim	-0.206^{***}	-0.188^{***}	-0.267^{***}	-0.608^{***}	-0.267^{***}
	(0.069)	(0.055)	(0.058)	(0.053)	(0.060)
AnyCorrection* DissonantClaim	0.062	0.016	-0.060	-0.022	-0.058
	(0.078)	(0.064)	(0.069)	(0.062)	(0.070)
Constant	2.816^{***}	3.334^{***}	3.097^{***}	2.022^{***}	3.058^{***}
	(0.040)	(0.031)	(0.033)	(0.042)	(0.035)
Observations	5,104	5,103	5,099	5,136	5,099
R^2	0.006	0.015	0.021	0.090	0.019
Adjusted R^2	0.006	0.015	0.020	0.089	0.019
Res. Std. Er.	1.093	0.942	1.050	0.968	1.051
F Statistic	10.658^{***}	26.475^{***}	36.056^{***}	168.534^{***}	33.051^{***}
Note:					*p<0.1; **p<0.05; ***p<0.01

Taken together, these results demonstrate that corrections can in some cases be effective despite partisan ties or partisan motivated reasoning, and perhaps even for concordant news, contrary to what we hypothesized in H2-4, and contrary to what much of the literature suggests, based on evidence from other cases. Importantly, our results exclude the possibility that these findings might owe to our sample being a low-effort sample: as visible in Tables 4 and 5, our respondents did strongly react - in the expected direction - to the slant of rumors, to their cited sources and to the presence of a correction. This, however, did not systematically mitigate their reaction to the correction. Besides, while our sample tilts towards BJP supporters and educated men – and thus reflects the population of high-frequency social media users in India – the relatively large size of this sample (N>5,000) makes it unlikely that these results owe to insufficient statistical power. These factors suggest that partisan motivated reasoning plays a comparatively less systematic role in India, breaking with results frequently obtained in the American context (Nyhan and Reifler 2010).

Table 6: Pooled Model Averaging Across All Rumors

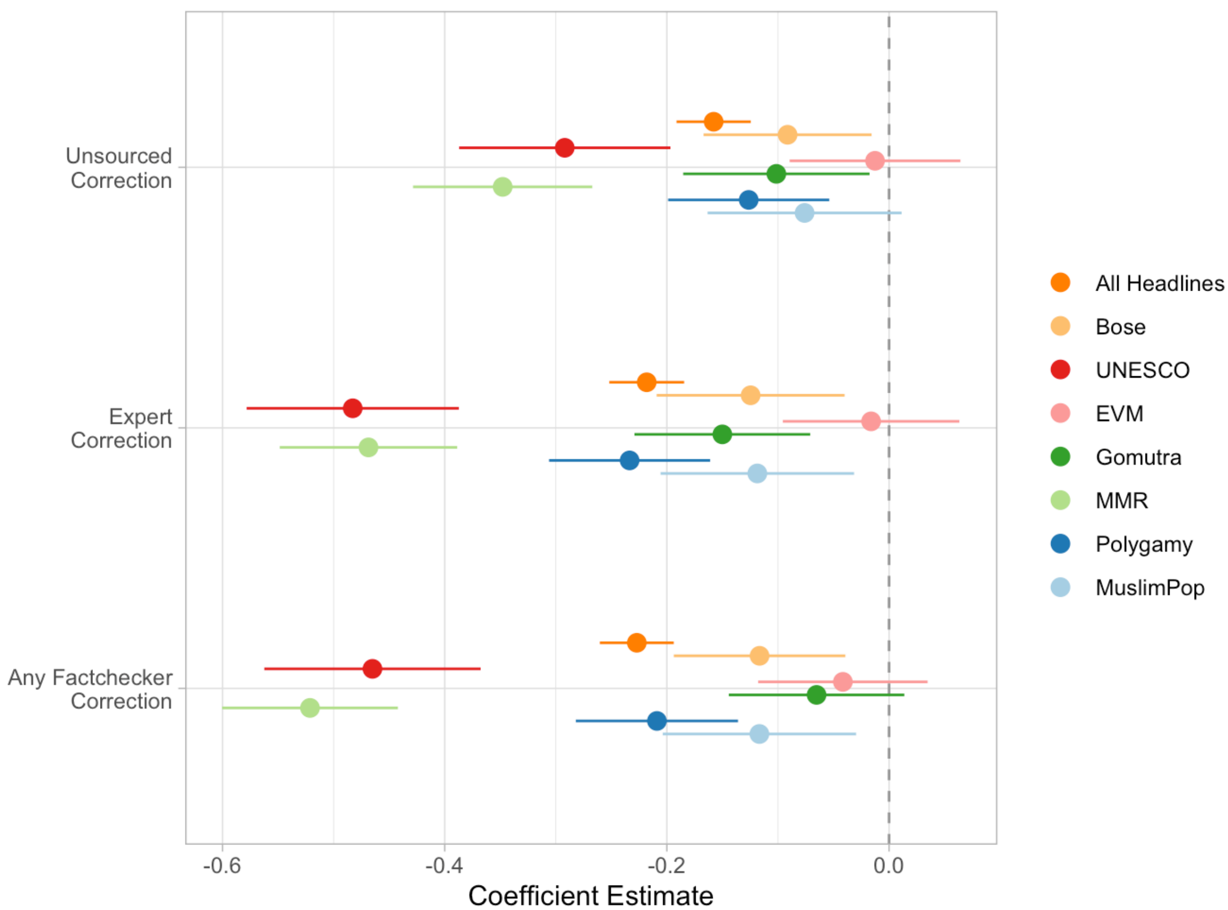
	Dependent variable: Average Belief Across All Rumors	
	(1)	(2)
Any Correction	-0.2010^{***}	
	(0.0598)	
Unsourced Correction		-0.1579^{***}
		(0.0423)
Expert Correction		-0.2182^{***}
		(0.0713)
Any Factchecker Correction		-0.2271^{***}
		(0.0701)
Constant	2.7563^{***}	2.7563^{***}
	(0.1951)	(0.1951)
Observations	35,729	35,729
R^2	0.0060	0.0065
Adjusted R^2	0.0059	0.0064
Residual Std. Error	1.1581 (df = 35727)	1.1578 (df = 35725)
Note:	All regressions include robust SEs clustered at respondent and headline levels	

Are Substantiated Corrections More Effective?

Our main effect in this paper demonstrates that receiving *any* correction (relative to control) can reduce the perceived accuracy of misinformation. While this analysis pools together all corrections, we now examine which types of corrections are most effective. Particularly, we compare substantiated to unsubstantiated messages, and determine whether the source of substantiation plays a role in persuasion.

Figure 3 evaluates the effect of different types of corrections on belief in misinformation, compared to the control condition (no correction). The first row of coefficients represents the size of the effect in corrections without any substantiation. In this case, a peer in the group chat expresses skepticism with the story but does not cite a source to justify their skepticism (merely stating a version of ”I don’t think that’s true, bro!”). The remaining rows represent substantiated corrections, but each with a different source of substantiation.[11] In each treatment group we distinguish by story, including all seven stories that contained a misinformation stimulus, as well as an overall coefficient pooling across all stories.

Figure 3: The Effect of Different Types of Social Corrections, compared to Control Condition



[11]We further disentangle these results in Appendix J.

Several striking findings emerge from these results. Visually, it appears that the corrective effect is larger on some stories when the correction is sourced; this appears particularly clear for the UNESCO and MMR rumors. Critically, however, potential differences in effect sizes remain, at best, very small across sub-types of corrections, and in most cases do not appear to exist at all, as confidence intervals between corrections largely overlap across rows. Given how highly powered our experiment is, we can say with relative confidence that the sophistication of social corrections did not strongly change respondents' reactions: the unsourced correction (a simple, short dismissal of the claim made by the first user) is in fact often as effective as the longer and more clearly sourced corrections in this design. An unidentified participant merely expressing incredulity about a rumor is therefore often as likely to reduce belief in a falsehood as a respondent engaging in a longer correction.

The type of source cited appears to make even less of a difference: corrective messages substantiated with a domain expert do not make the correction more persuasive than other types of sourced corrections: in all cases, respondents are as likely to react to the correction when it is said to originate from a professional fact-checking organization, a prominent newspaper, or the platforms themselves, as opposed to a domain expert. This further implies that respondents open to belief change do not require much expertise in order for their beliefs to be moved.

We do note a potential limitation with the sourced corrections, however. We underscore that messages in these conditions do not actually cite fact-checkers; respondents are merely referred to a fact-checking source, though the actual source is in our case the user. This might constitute a weaker correction than one where a link was provided in the message (Vraga and Bode 2018). In future research, the effect of stronger corrections - i.e. corrections including an actual link to a fact-check - could be tested to assess whether users are better able to correct their beliefs with sourced links. Finally, when further disentangling these results by fact-checker (Appendix J), no source emerges as consistently more persuasive or effective relative to others. Overall, we thus find that the content of social corrections counts less than the mere presence of one. Related research on correction format, tone and style support this findings that correction specifics are less important relative to issuing a correction in the first place (Bode, Vraga, and Tully 2020; Swire and Ecker 2018).

Discussion and Conclusion

Taken together, our results demonstrate that social corrections are effective in contexts like India: in our experiment, exposure to a corrective message posted by an unidentified peer – regardless of the source or substantiation of that message – significantly reduces beliefs in misinformation. Insofar as respondents in our experiment were not incentivized to pay attention to the message, and by design neither knew nor by could identify the individuals posting corrections, these results may be seen as conservative estimates. Arguably, peer corrections on more homophilic networks (friends, colleagues, or likeminded partisans) may achieve a much larger effect.

While additional research will be needed to test this hypothesis, our results point to a potential flip side of comparatively low levels of digital literacy. India has lower rates of formal education and digital literacy, as well a large share of users who are new to the Internet. These factors likely imply that news received via the Internet might automatically have more value, given the unfamiliarity and fascination the medium inspires (Badrinathan 2021). While this may lead misinformation to be more easily believed in the first place, the same may apply to corrections, which may more easily become effective in such contexts.

Beyond our main result, tests on our other hypotheses (H2 to H5) point to another reason why social corrections matter: forces such as partisan motivated reasoning– normally expected to reduce the effect of corrective measures–may not play as systematic a role in this context. While we can only speculate as to the causes of this divergence, one explanation may lie in the nature of partisanship in India. Despite much report of political polarization and transformation of partisanship into a social identity (Chhibber and Verma 2018), India is a country that has traditionally had weaker partisan ties, and politics is thought to be more clientelistic rather than programmatic (Auerbach et al. 2021). This relative weakness of partisanship - at least in the American sense of the term - may imply that motivated reasoning would *not* constitute as big an obstacle to correcting beliefs, or more likely in our view, that partisanship may not be the basis for motivated reasoning in this context–it may instead exist in another, non-partisan form, for instance, along the lines of religious identities (Badrinathan and Chauchard 2021).

Additionally, we show that respondents often do not react very differently to substantiated and unsubstantiated corrections, and that the presence of a correction, rather than its degree of substantiation, appears sufficient to change beliefs. Both Munger (2017) and Siegel and Badaan (2020) show that online hatred and harassment can be significantly reduced by simple nudges, especially if these come from ingroups.

While these multiple findings mutually reinforce our confidence that social corrections matter, we now outline limitations in our design. First, we note that we measure our dependent variable in close proximity to the treatment, and thus cannot speak to the durability of these effects. While some recent findings demonstrate that fact-checks may not persist over time despite repeated exposure (Carey et al. 2022), future studies should nevertheless consider a longer gap between treatment and outcomes to measure whether social correction effects decay over time. We also note that while corrections emanate from peers, we blur out all source names and affiliations in our WhatsApp screenshots. Thus we are unable to test whether the efficacy of corrections is a function of the particular individual posting the corrective message.

Next, we underscore that our results might be a function of our sample: Hindi speaking, Facebook users with enough digital literacy to take an online survey. Individualized access to social media implies that the individual is wealthy enough (and has enough freedom) to obtain a smartphone for personal use and that the person is functionally literate. Thus, these conditions are more likely met in younger, male populations in the more developed parts of India, which are hence overrepresented in our sample (Badrinathan et al. 2021). We hope in future work to expand our sampling recruitment to include populations in India that speak other regional languages, as well as populations that only use WhatsApp or are new to the internet (and may consequently more likely be women, older users, or those residing in non-urban areas).

In addition, a limitation of our study was the partisan imbalance in the mix of stories that we use as our dependent variable measures. This is because political misinformation in India seems to emanate largely from BJP sources (Chauchard and Garimella 2022). This is underscored in previous data by pro-BJP stories being believed to a much greater extent than anti-BJP stories (Badrinathan 2021). In this study, too, we choose stories that are more salient (through a pretest) in the hope that our treatment is able to move respondent attitudes on the misinformation they believe the most. However, since the most salient stories are also those that benefit one side of the ideological spectrum, we end up with an imbalance, alluding to the fact that pro-BJP stories are more salient in the minds of respondents. We also note that, as a result, we over-sampled false stories in our outcome. Thus, while we chose rumors that we believed were normatively important to correct, this subset may not represent the sum of misinformation online in India. Further, over-sampling false stories may be less externally valid as a respondent may not encounter large amounts of misinformation in a single online browsing session (Chauchard and Garimella 2022).

Third, a challenging prospect for future research is to be able to examine the effect of corrections in a more naturalistic setting, outside of a survey or online experiment. We acknowledge that our experiment tests the effect of witnessing a correction rather than being directly corrected. While the private nature of WhatsApp groups makes incorporating these elements logistically and ethically difficult, measuring the impact of misinformation and solutions to counter it within the ecosystem of groups that individuals are a part of will allow us to ascertain the true impact of group solidarity, conformity pressures, and ingroup norms. Relatedly, since our study thus only compares different types of social corrections to each other, it is difficult to separate out whether a correction is effective *because* it is social or not, since we do not have a non-social correction condition. We aim to look at this question in future work.

Beyond chat apps and other messaging applications, our study opens up broader avenues for research on misinformation in developing countries. Much remains to be uncovered about the ability of misinformation to persuade, and to be corrected, in settings of low education, accelerating Internet, and private online spaces. The weakness of the partisan form of motivated reasoning detected in our study suggests that more comparative work on misinformation is needed. Future work should explore the psychological mechanisms leading to belief change, and potentially to offline behaviors, especially in countries where the stakes are as high as violence. Such research should also look into information and misinformation processing on encrypted and personal social media networks such as WhatsApp. The findings from this study have implications not only for developing countries that widely use EMAs, but also for more developed contexts where polarized users are sorted into homophilic networks online.

References

Acemoglu, Daron, Asuman Ozdaglar, and James Siderius. 2021. "Misinformation: Strategic Sharing, Homophily, and Endogenous Echo Chambers." *Working Paper*.

Aneez, Zeenab, Ahmed T Neyazi, Antonis Kalogeropoulos, and Rasmus Kleis Nielsen. 2019. "India Digital News Report." *Reuters Institute for the Study of Journalism*. https://bit.ly/3x3Gpnl.

Arechar, Antonio Alonso, Jennifer Nancy Lee Allen, Rocky Cole, Ziv Epstein, Kiran Garimella, Andrew Gully, Jackson G Lu, Robert M Ross, Michael Stagnaro, Jerry Zhang et al. 2022. "Understanding and Reducing Online Misinformation Across 16 Countries on Six Continents.".

Auerbach, Adam Michael, Jennifer Bussell, Simon Chauchard, Francesca R. Jensenius, Gareth Nellis, Mark Schneider, Neelanjan Sircar, Pavithra Suryanarayan, Tariq Thachil, Milan Vaishnav, and et al. 2021. "Rethinking the Study of Electoral Politics in the Developing World: Reflections on the Indian Case." *Perspectives on Politics*: 1–15.

Badrinathan, Sumitra. 2021. "Educative Interventions to Combat Misinformation: Evidence From A Field Experiment in India." *American Political Science Review*: 1-17.

Badrinathan, Sumitra, Deepaboli Chatterjee, Devesh Kapur, and Neelanjan Sircar. 2021. "Partisan Disagreement: The Role of Media, Personal Networks and Gender in Forming Political Preferences." *Urbanisation* 6 (1): 141–157.

Badrinathan, Sumitra, and Simon Chauchard. 2021. "Leveraging Religiosity Against COVID-19 Misinformation: Experimental Evidence from India." *Working Paper*.

Bode, Leticia, and Emily K Vraga. 2018. "See Something, Say Something: Correction of Global Health Misinformation on Social Media." *Health Communication* 33 (9): 1131–1140.

Bode, Leticia, Emily K Vraga, and Melissa Tully. 2020. "Do the right thing: Tone may not affect correction of misinformation on social media." *Harvard Kennedy School Misinformation Review*.

Bussell, Jennifer. 2019. *Clients and Constituents: Political Responsiveness in Patronage Democracies*. New York: Oxford University Press.

Carey, John M, Andrew M Guess, Peter J Loewen, Eric Merkley, Brendan Nyhan, Joseph B Phillips, and Jason Reifler. 2022. "The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, Great Britain and Canada." *Nature Human Behaviour* 6 (2): 236–243.

Chan, Man-pui Sally, Christopher R Jones, Kathleen Hall Jamieson, and Dolores Albarracín. 2017. "Debunking: A Meta-Analysis of the Psychological Efficacy of Messages Countering Misinformation." *Psychological Science* 28 (11): 1531–1546.

Chauchard, Simon, and Kiran Garimella. 2022. "What Circulates on Partisan WhatsApp in India? Insights from an Unusual Dataset." *Forthcoming. Journal of Quantitative Description: Digital Media*. .

Chhibber, Pradeep K, and Rahul Verma. 2018. *Ideology and Identity: The Changing Party Systems of India*. New York: Oxford University Press.

Davies, William. 2020. "What's Wrong With WhatsApp." *The Guardian*. July 2, 2020. https://www.theguardian.com/technology/2020/jul/02/whatsapp-groups-conspiracy-theories-disinformation-democracy.

Eagly, Alice H, and Shelly Chaiken. 1993. *The psychology of attitudes.* Harcourt Brace Jovanovich College Publishers.

Garimella, Kiran, and Dean Eckles. 2020. "Images and misinformation in political groups: evidence from WhatsApp in India." *Harvard Kennedy School Misinformation Review*.

Guess, Andrew M, Michael Lerner, Benjamin Lyons, Jacob M Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjan Sircar. 2020. "A digital media literacy intervention increases discernment between mainstream and false news in the United States and India." *Proceedings of the National Academy of Sciences* 117 (27): 15536–15545.

Housholder, Elizabeth E, and Heather L LaMarre. 2014. "Facebook Politics: Toward a Process Model for Achieving Political Source Credibility Through Social Media." *Journal of Information Technology & Politics* 11 (4): 368–382.

Margolin, Drew B, Aniko Hannak, and Ingmar Weber. 2018. "Political fact-checking on Twitter: When do corrections have an effect?" *Political Communication* 35 (2): 196–219.

Michelitch, Kristin, and Stephen Utych. 2018. "Electoral Cycle Fluctuations in Partisanship: Global Evidence from Eighty-Six Countries." *The Journal of Politics* 80 (2): 412–427.

Mont'Alverne, Camila, Sumitra Badrinathan, Amy Ross Arguedas, Benjamin Toff, Richard Fletcher, and Rasmus Kleis Nielsen. 2022. "The trust gap: how and why news on digital platforms is viewed more sceptically versus news in general." *Reuters Institute for the Study of Journalism*. September 22, 2022. .

Munger, Kevin. 2017. "Tweetment effects on the tweeted: Experimentally reducing racist harassment." *Political Behavior* 39 (3): 629–649.

Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32 (2): 303–330.

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.

Pereira, Frederico Batista, Natália S Bueno, Felipe Nunes, and Nara Pavão. 2021. "Motivated Reasoning Without Partisanship? Fake News in the 2018 Brazilian Elections." *Working Paper*.

Porter, Ethan, and Thomas J Wood. 2021. "The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom." *Proceedings of the National Academy of Sciences* 118 (37).

Rosenzweig, Leah R, Bence Bago, Adam J Berinsky, and David G Rand. 2021. "Happiness and surprise are associated with worse truth discernment of COVID-19 headlines among social media users in Nigeria." *Harvard Kennedy School Misinformation Review*.

Siegel, Alexandra A, and Vivienne Badaan. 2020. "# No2Sectarianism: Experimental approaches to reducing sectarian hate speech online." *American Political Science Review* 114 (3): 837–855.

Sinha, Pratik, Sumaiya Sheikh, and Arjun Sidharth. 2019. *India Misinformed: The True Story*. Noida: HarperCollins India.

Stiff, James B, and Paul A Mongeau. 2016. *Persuasive Communication*. New York: Guilford Publications.

Swire, Briony, and Ullrich K H Ecker. 2018. "Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication." *Misinformation and Mass Audiences*: 195–211.

Taber, Charles S, and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50 (3): 755–769.

Tappin, Ben M, Gordon Pennycook, and David G Rand. 2021. "Rethinking the link between cognitive sophistication and politically motivated reasoning." *Journal of Experimental Psychology: General* 150 (6): 1095.

Thorson, Emily. 2016. "Belief Echoes: The Persistent Effects of Corrected Misinformation." *Political Communication* 33 (3): 460–480.

Tokita, Christopher K, Andrew M Guess, and Corina E Tarnita. 2021. "Polarized information ecosystems can reorganize social networks via information cascades." *Proceedings of the National Academy of Sciences* 118 (50).

van der Meer, Toni GLA, and Yan Jin. 2020. "Seeking Formula for Misinformation Treatment in Public Health Crises: The Effects of Corrective Information Type and Source." *Health Communication* 35 (5): 560–575.

Vraga, Emily K, and Leticia Bode. 2018. "I do not believe you: how providing a source corrects health misperceptions across social media platforms." *Information, Communication & Society* 21 (10): 1337–1353.

Vraga, Emily, Leticia Bode et al. 2022. "Correcting What's True: Testing Competing Claims about Health Misinformation on Social Media.".

Yasir, Sameer. 2020. "India Is Scapegoating Muslims for the Spread of the Coronavirus." *Foreign Policy*. April 22, 2020. .

Supplementary Information File

"I Don't Think That's True, Bro!"

Social Corrections of Misinformation in India

Contents

A 2019 WhatsApp Campaign Promoting User-driven Corrections

Figure A.1: WhatsApp ad in Indian newspapers

B Advertisement Used to Recruit Respondents

The image below is a screenshot of the advertisement we used to recruit respondents into our survey via Facebook.

Figure B.1: Sample recruitment ad

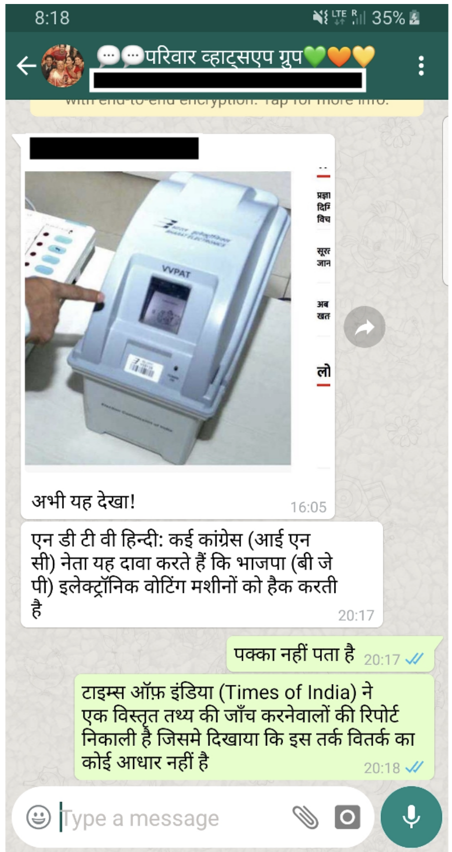
C Full Text of Experimental Manipulations

The text displayed on our WhatsApp screenshots to treatment group respondents varied as a function of the type of rumor, the politician / media outlet it is attributed to, and the nature of the correction. Figure C.1 below describes the full text associated with each hypothetical WhatsApp screenshot in each condition. Figure C.2 shows an example of the treatment stimuli.

			SOURCES VARIATIONS				CORRECTION VARIATIONS		
Rumor	Type of claim	Message 1	Message 2 (part 1)	Message 2 (part 2)		Message 3a (control)	Message 3b (Domain	M3b (Factchecker correction)	M3c (random person
1	Congenial to BJP/dissonant to non-BJP	image	Check this out.	In this article published on the NDTV Hindi/Doordarshan/India TV website, a BJP leader claims that the Muslim population in India will overtake the Hindu population in the future.	In this article published on the NDTV Hindi/Doordarshan/India TV website, it is suggested that the Muslim population in India will overtake the Hindu population in India in the future.	thanks -- will read!	The other day I saw a report by prominent demographers that said this was extremely unlikely to happen. So I don't think that's right actually.	The other day I saw a fact-checking report by ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp that said this was extremely unlikely to happen. So I don't think that's right actually.	The other day I saw something that said this was extremely unlikely to happen. So I don't think that's right actually.
2	Congenial to BJP/dissonant to non-BJP	image	Just saw this!	This appeared on the NDTV Hindi/Doordarshan/India TV website --- A BJP leader explains that polygamy (having several wifes) is very common in the Muslim population.	This appeared on the NDTV Hindi/Doordarshan/India TV website --- it says that polygamy (having several wifes) is very common in the Muslim population.	Thanks. Will look when I have a second.	Hmm, not sure. The Census of India did a report and showed it to be unsubstantiated, though.	Hmm, not sure. A fact-checking report done by ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp showed this to be unsubstantiated, though.	Hmm, not sure. I saw something showed this to be unsubstantiated, though.
3	Neither congenial nor dissonant to BJP/non-BJP	image	Just came across this article	-	This comes from the NDTV Hindi/Doordarshan/India TV website. Apparently M-R vaccines are associated with autism and retardation.	Wow, ok. will get into this.	Hey I don't think that's true actually. I just saw a report from doctors from AIIMS, there appears to be no basis for this claim...	Hey I don't think that's true actually. I just saw a fact-checking report done by ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp, there appears to be no basis for this claim...	Hey I don't think that's true actually. Someone told me there was not basis for this claim...
4	Congenial to BJP/dissonant to non-BJP	image	This is worth looking at.	The NDTV Hindi/Doordarshan/India TV website just published this. A bunch of BJP leaders said that drinking cow urine (gomutra) helps build one's immune system.	The NDTV/NDTV Hindi/Doordarshan/Republic TV/India TV website just published this. Claims that Drinking cow urine (gomutra) helps build one's immune system.	Got it, thanks for sending :)	Actually not sure about this, brother. I saw a report from doctors from AIIMS explaining why this is not correct.	Actually not sure about this, brother. I saw a fact-checking report done by ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp explaining why this is not correct.	Actually not sure about this, brother. I saw somewhere that this is not correct.
5	Neither congenial nor dissonant to BJP/non-BJP	image	Relevant as the ICC world cup approaches...	-	This comes from the NDTV Hindi/Doordarshan/India TV website. I had forgotten that Australia has more ICC cricket world cup wins than any country!	Great. Thanks for sending :)	-	-	-
6	Neither congenial nor dissonant to BJP/non-BJP	image	Important stuff...	-	the NDTV/NDTV Hindi/Doordarshan/Republic TV/India TV website published this. sad that there's still no cure for HIV/AIDS	thanks. will definitely read.	-	-	-
7	Congenial to non-BJP/dissonant to BJP	image	Just saw this!	NDTV Hindi/Doordarshan/India TV: several INC leaders claim that the BJP hacks electronic voting machines.	NDTV Hindi/Doordarshan/India TV: some people suggesting that the BJP hacks electronic voting machines.	ok! reading now...	Not sure about this... the Election Commission released a serious report saying there's no basis for this claim	Not sure about this claim. ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp has come up with a detailed fact-checking report that showed there was no basis for this argument.	Not sure about this claim. I saw somewhere there is no basis for this argument.
8	Congenial to BJP/dissonant to non-BJP	image	Wow	Just saw this on the NDTV Hindi/Doordarshan/India TV weblsite.. This BJP guy said UNESCO declared PM Modi best Prime Minister in 2016.	Just saw this on NDTV Hindi/Doordarshan/India TV website.. UNESCO declared PM Modi best Prime Minister in 2016!	Thanks, boss :)	Haha that's not right actually.. UNESCO put out a release saying they didn't come up with rankings like that.	Haha that's not right actually.. ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp published a fact-checking thing saying that UNESCO didn't come up with rankings like that.	Haha that's not right actually..
9	Neither congenial nor dissonant to BJP/non-BJP	image	Have a look at this!	From the NDTV Hindi/Doordarshan/India TV website ... Netaji Bose did NOT die in a plane crash in 1945!		wow - thanks for sharing!	This theory has been debunked, I think. I read a report by Delhi University historians explaining there was no ground to believe any of this.	This theory has been debunked, I think. I read a fact-checking report by ALTNEWS/Vishwasnews.com/Times of India/Facebook/WhatsApp explaining there was no ground to believe any of this.	I think this theory has been debunked, though.

Figure C.1: Text for experimental manipulations

Message 1 including misinformation about EVM machines, reference to NDTV Hindi as the news source and Congress as the politician source



Message 2 including correction citing Times of India as a fact-checker

Figure C.2: Sample Treatment Stimuli

D Tests For Hypothesis 2

Hypothesis 2a: Peer corrections will be more effective when misinformation is attributed to an ideologically dissonant politician (compared to when it is unattributed).

To test this hypothesis, we run the following model:

$$\begin{aligned} BeliefAccuracy_i = \alpha + \beta_1 AnyCorrection_i + \beta_2 DissonantPol_i \\ + \beta_3 AnyCorrection * DissonantPol_i + \epsilon_i \end{aligned} \tag{D.1}$$

As noted in the body of the article, we limit our analyses to the subset of rumors that are clearly partisan in nature (rumors 3, 4, 6, 8, and 9) and code whether the claim was attributed in the prompt to a congenial or dissonant politician. We code a politician as congenial or dissonant as a function of the respondent's partisan inclination towards the BJP (the ruling party), relying on the respondent's expressed closeness to this party. A BJP politician is deemed congenial if the respondent describes herself as close or very close to the party and dissonant if the respondent describes herself as far or very far from the party. By contrast, a INC politician is deemed congenial if the respondent describes herself as far or very far to the BJP and dissonant if the respondent describes herself as close or very close to the BJP. Note that we are pooling members of both major parties in each category (e.g., "dissonant" takes the value of 1 for BJP identifiers who read an anti-BJP claim and for INC identifiers who read a pro-BJP claim).

In Table D.1 and D.2 we show models for each false story. In Table **??** we show a pooled model collapsing across all rumors with standard errors clustered at respondent and headline levels.

Table D.1: Effect of Any Correction * Dissonant Speaker on Belief in Rumor

	Dependent variable: Belief in Rumor				
	MuslimPop	Polygamy	Gomutra	EVM	UNESCO
	(1)	(2)	(3)	(4)	(5)
AnyCorrection	-0.127^{***}	-0.178^{***}	-0.092^{***}	-0.007	-0.405^{***}
	(0.038)	(0.031)	(0.033)	(0.033)	(0.040)
DissonantPol	-0.564^{***}	-0.310^{*}	-0.242	-0.229^{**}	-0.184
	(0.149)	(0.164)	(0.154)	(0.108)	(0.203)
AnyCorrection* DissonantPol	0.482^{***}	0.062	-0.055	0.008	0.016
	(0.163)	(0.174)	(0.168)	(0.118)	(0.218)
Constant	2.775^{***}	3.280^{***}	3.018^{***}	1.666^{***}	3.005^{***}
	(0.034)	(0.026)	(0.028)	(0.028)	(0.034)
Observations	5,104	5,103	5,099	5,136	5,109
R^2	0.005	0.012	0.006	0.005	0.022
Adjusted R^2	0.004	0.011	0.006	0.005	0.021
BH Significance (interaction term)	Yes	No	No	No	No
Res. Std. Er.	1.093	0.944	1.058	1.011	1.250
F Statistic	7.934^{***}	20.647^{***}	10.792^{***}	9.195^{***}	37.700^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Hypothesis 2b: Peer corrections will be less effective when misinformation is attributed to an ideologically congenial politician (compared to when it is unattributed).

$$\begin{aligned} Belief\,Accuracy_i = \alpha + \beta_1 AnyCorrection_i + \beta_2 CongenialPol_i + \\ \beta_3 AnyCorrection * CongenialPol_i + \epsilon_i \end{aligned} \tag{D.2}$$

Table D.2: Effect of Any Correction * Congenial Speaker on Belief in Rumor

	Dependent variable: Belief in Rumor				
	MuslimPop	Polygamy	Gomutra	EVM	UNESCO
	(1)	(2)	(3)	(4)	(5)
AnyCorrection	-0.102^{***}	-0.181^{***}	-0.131^{***}	-0.043	-0.400^{***}
	(0.039)	(0.031)	(0.034)	(0.032)	(0.041)
CongenialPol	0.070	0.167	-0.011	0.518^{***}	0.409^{***}
	(0.132)	(0.107)	(0.119)	(0.180)	(0.157)
AnyCorrection* CongenialPol	-0.054	-0.155	0.176	-0.120	-0.349^{**}
	(0.141)	(0.116)	(0.129)	(0.191)	(0.167)
Constant	2.741^{***}	3.262^{***}	3.011^{***}	1.638^{***}	2.979^{***}
	(0.034)	(0.027)	(0.028)	(0.027)	(0.035)
Observations	5,104	5,103	5,099	5,136	5,109
R^2	0.002	0.008	0.004	0.010	0.022
Adjusted R^2	0.001	0.008	0.004	0.009	0.022
BH Significance (interaction term)	No	No	No	No	No
Res. Std. Error	1.095	0.946	1.059	1.009	1.250
F Statistic	2.747^{**}	14.136^{***}	7.207^{***}	16.946^{***}	38.594^{***}
Note:					*p<0.1; **p<0.05; ***p<0.01

Table D.3: Pooled Model for H2a and H2b

	Dependent variable: Average Belief Across All Rumos	
	(1)	(2)
Any Correction	-0.1552^{*}	-0.1759^{**}
	(0.0774)	(0.0650)
Outpartisan Speaker	-0.5281^{***}	
	(0.0850)	
Any Correction * Outpartisan Speaker	0.1591	
	(0.0900)	
Copartisan Speaker		0.3453
		(0.1989)
AnyCorrection * Copartisan Speaker		-0.1356
		(0.1135)
Constant	2.7552^{***}	2.7163^{***}
	(0.2876)	(0.3003)
Observations	25,551	25,551
R^2	0.0113	0.0073
Adjusted R^2	0.0112	0.0072
Residual Std. Error (df = 25547)	1.1932	1.1956
Note:	Includes robust SEs clustered at respondent and headline levels	

E Tests For Hypothesis 3-5

Hypothesis 3a: Peer corrections will be more effective when misinformation originates from a dissonant media outlet (compared to unattributed or neutral outlet).

Hypothesis 3b: Peer corrections will be less effective when misinformation originates from a congenial media outlet (compared to an unattributed or neutral outlet).

To test these hypotheses, we code a media outlet as congenial or dissonant as a function of the respondent's expressed proximity to the BJP. Concretely, we code the "pro-BJP" outlet (here, India TV) as congenial and the "anti-BJP" outlet (here, New Delhi TV or NDTV) as dissonant when the respondent reports feeling close or very close to the BJP. By contrast, we code the "pro-BJP" outlet (India TV) as dissonant and "anti-BJP" outlet (NDTV) as congenial when the respondent reports feeling far or very far to the BJP.

We test this hypothesis with the following model:

$$\begin{aligned} BeliefAccuracy_i = \alpha + \beta_1 AnyCorrection_i + \beta_2 CongenialMedia_i + \beta_3 DissonantMedia_i + \\ \beta_4 AnyCorrection * CongenialMedia_i + \beta_5 AnyCorrection * DissonantMedia_i + \epsilon_i \end{aligned} \tag{E.1}$$

In Table E.1 we show models for each of the 7 false stories. In Table **??** we show a pooled model collapsing across all rumors with standard errors clustered at respondent and headline levels.

Table E.1: Effect of Any Correction * Media Outlet Source on Belief in Rumor

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AnyCorrection	-0.150^{***}	-0.167^{***}	-0.419^{***}	-0.128^{***}	-0.006	-0.399^{***}	-0.110^{***}
	(0.041)	(0.033)	(0.036)	(0.036)	(0.035)	(0.043)	(0.034)
Congenial Media	-0.224^{*}	0.230^{**}	-0.078	0.080	-0.163	0.036	0.186
	(0.126)	(0.113)	(0.121)	(0.108)	(0.111)	(0.151)	(0.114)
Dissonant Media	-0.132	0.061	-0.050	-0.184	0.051	0.094	-0.048
	(0.121)	(0.108)	(0.121)	(0.116)	(0.118)	(0.143)	(0.116)
AnyCorrection* CongenialMedia	0.271^{**}	-0.194	-0.049	-0.056	0.080	-0.067	-0.174
	(0.135)	(0.122)	(0.131)	(0.119)	(0.121)	(0.162)	(0.125)
AnyCorrection* DissonantMedia	0.219^{*}	-0.134	-0.067	0.272^{**}	-0.134	-0.092	0.091
	(0.131)	(0.116)	(0.130)	(0.126)	(0.127)	(0.155)	(0.127)
Constant	2.773^{***}	3.256^{***}	2.834^{***}	3.015^{***}	1.658^{***}	2.992^{***}	2.781^{***}
	(0.036)	(0.027)	(0.030)	(0.029)	(0.029)	(0.036)	(0.027)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.003	0.009	0.038	0.003	0.002	0.021	0.003
Adjusted R^2	0.002	0.008	0.037	0.002	0.001	0.020	0.002
BH Significance (Correction*Congenial)	No	No	No	No	No	No	No
BH Significance (Correction*Dissonant)	No	No	No	No	No	No	No
Res. Std. Er.	1.095	0.945	1.038	1.060	1.013	1.251	1.048
F Statistic	3.055^{***}	9.650^{***}	39.456^{***}	3.394^{***}	1.663	21.697^{***}	3.190^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table E.2: Pooled Model for H3

	Dependent variable: Average Belief Across All Rumors
Any Correction	-0.1984^{***}
	(0.0564)
Congenial Media	-0.0038
	(0.0673)
Dissonant Media	-0.0242
	(0.0442)
Any Correction * Congenial Media	-0.0172
	(0.0625)
Any Correction * Dissonant Media	0.0160
	(0.0743)
Constant	2.7580^{***}
	(0.1932)
Observations	35,729
R^2	0.0060
Adjusted R^2	0.0059
Residual Std. Error	1.1581 (df = 35723)
Note:	Includes robust SEs clustered at respondent and headline levels

We also provide below an additional test for Hypotheses 4a and 4b, this time with a pooled model collapsing across all rumors with standard errors clustered at respondent and headline levels.

Table E.3: Pooled Model for H4a and H4b

	Dependent variable: Average Belief Across All Rumors	
	(1)	(2)
Any Correction	-0.1300	-0.1860^{**}
	(0.0723)	(0.0695)
Congenial Claim	0.5899^{**}	
	(0.2066)	
Any Correction * Congenial Claim	-0.0735^{*}	
	(0.0331)	
Dissonant Claim		-0.5627^{*}
		(0.2458)
Any Correction * Dissonant Claim		0.0400
		(0.0374)
Constant	2.4079^{***}	2.9527^{***}
	(0.3689)	(0.1767)
Observations	25,551	25,551
R^2	0.0531	0.0504
Adjusted R^2	0.0530	0.0503
Residual Std. Error (df = 25547)	1.1678	1.1694
Note:	Includes robust SEs clustered at respondent and headline levels	

Finally, we provide an additional test for Hypothesis 5 with the pooled model collapsing across all rumors. The omitted (comparison category) in this model is the Peer (unsourced) correction.

Table E.4: Pooled Model for H5

	Dependent variable: Average Belief Across All Rumors
Control	0.1486^{***}
	(0.0422)
Expert Correction	-0.0750^{**}
	(0.0355)
Any Factchecker Correction	-0.0839^{**}
	(0.0382)
Constant	2.6132^{***}
	(0.1777)
Observations	35,729
R^2	0.0062
Adjusted R^2	0.0061
Residual Std. Error	1.1580 (df = 35725)
Note:	Include robust SEs clustered at respondent and headline levels

F Heterogeneous Effects of BJP Support

To complement our tests of motivated reasoning (based on the congeniality/dissonance of the information presented and the source of the information), we present OLS results from models that test whether BJP voters react differently to corrective information. We code BJP support as a dummy variable where 1 = support for the BJP and 0 = otherwise.

$$\begin{aligned} Belief\,Accuracy_i = \alpha + \beta_1 AnyCorrection_i + \beta_2 BJPSupport_i \\ + \beta_3 AnyCorrection * BJPSupport_i + \epsilon_i \end{aligned} \tag{F.1}$$

Table F.1: Effect of BJP Support * Correction

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AnyCorrection	-0.078	-0.168^{***}	-0.452^{***}	-0.027	-0.123^{**}	-0.405^{***}	-0.058
	(0.064)	(0.053)	(0.058)	(0.056)	(0.052)	(0.068)	(0.054)
BJP Support	0.424^{***}	0.338^{***}	0.092	0.578^{***}	-0.870^{***}	0.491^{***}	0.237^{***}
	(0.069)	(0.055)	(0.060)	(0.057)	(0.054)	(0.071)	(0.053)
AnyCorrection *	-0.043	-0.027	0.006	-0.114^{*}	0.151^{**}	-0.017	-0.081
BJP Support	(0.078)	(0.064)	(0.070)	(0.068)	(0.064)	(0.083)	(0.066)
Constant	2.460^{***}	3.040^{***}	2.765^{***}	2.616^{***}	2.238^{***}	2.669^{***}	2.629^{***}
	(0.057)	(0.045)	(0.050)	(0.047)	(0.045)	(0.058)	(0.044)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.029	0.032	0.037	0.051	0.124	0.052	0.009
Adjusted R^2	0.029	0.032	0.037	0.050	0.123	0.051	0.009
Res. Std. Er.	1.080	0.934	1.038	1.034	0.949	1.231	1.045
F Statistic	51.459^{***}	56.702^{***}	65.187^{***}	90.418^{***}	241.189^{***}	93.117^{***}	16.129^{***}
Note:					*p<0.1; **p<0.05; ***p<0.01		

G Main Effect of Congenial / Dissonant Claim

In this section, we show that the claims we code as congenial to respondents are more likely to be believed (G.1) and that the claims we code as dissonant to respondents are less likely to be believed (G.2). In each case we run a simple bivariate OLS model:

$$Belief = \alpha + \beta_1(CongenialClaim/DissonantClaim) + \epsilon \tag{G.1}$$

Table G.1: Effect of Rumor Congeniality on Belief

	Dependent variable: Belief in Rumor				
	MuslimPop	Polygamy	Gomutra	EVM	UNESCO
	(1)	(2)	(3)	(4)	(5)
CongenialClaim	0.218***	0.214***	0.378***	0.480***	0.309***
	(0.031)	(0.027)	(0.030)	(0.029)	(0.036)
Constant	2.530***	3.001***	2.702***	1.478***	2.506***
	(0.025)	(0.021)	(0.024)	(0.017)	(0.028)
Observations	5,104	5,103	5,099	5,136	5,109
R^2	0.009	0.012	0.030	0.049	0.014
Adjusted R^2	0.009	0.012	0.030	0.049	0.014
Res. Std. Er.	1.091 (df = 5102)	0.944 (df = 5101)	1.045 (df = 5097)	0.989 (df = 5134)	1.255 (df = 5107)
F Statistic	48.141***	62.228***	157.494***	264.374***	73.238***
Note:					*p<0.1; **p<0.05; ***p<0.01

Table G.2: Effect of Rumor Dissonance on Belief

	Dependent variable: Belief in Rumor				
	MuslimPop	Polygamy	Gomutra	EVM	UNESCO
	(1)	(2)	(3)	(4)	(5)
DissonantClaim	−0.157***	−0.176***	−0.309***	−0.625***	−0.231***
	(0.033)	(0.028)	(0.031)	(0.028)	(0.038)
Constant	2.716***	3.190***	3.035***	2.019***	2.772***
	(0.019)	(0.016)	(0.018)	(0.022)	(0.021)
Observations	5,104	5,103	5,099	5,136	5,109
R^2	0.004	0.008	0.019	0.090	0.007
Adjusted R^2	0.004	0.007	0.018	0.089	0.007
Res. Std. Er.	1.093 (df = 5102)	0.946 (df = 5101)	1.051 (df = 5097)	0.967 (df = 5134)	1.259 (df = 5107)
F Statistic	22.998***	38.607***	96.136***	505.256***	37.676***
Note:					*p<0.1; **p<0.05; ***p<0.01

H Summary Statistics

Table H.1: Summary Statistics for Muslim Population Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,104	2.665	1.096	1	3	4
Any Correction	5,104	0.781	0.414	0	1	1
Outpartisan Speaker	5,104	0.069	0.253	0	0	1
Copartisan Speaker	5,104	0.126	0.332	0	0	1
Congenial Media	5,104	0.134	0.341	0	0	1
Dissonant Media	5,104	0.127	0.333	0	0	1
Congenial Claim	5,104	0.616	0.486	0	1	1
Dissonant Claim	5,104	0.326	0.469	0	0	1
BJP Partisan	5,104	0.678	0.467	0	1	1
Congress Partisan	5,104	0.057	0.233	0	0	1
Pure Control	5,104	0.023	0.148	0	0	1
Peer Correction	5,104	0.258	0.438	0	0	1
Expert Correction	5,104	0.261	0.439	0	0	1
Alt News	5,104	0.051	0.220	0	0	1
Vishwas	5,104	0.053	0.225	0	0	1
TOI	5,104	0.048	0.213	0	0	1
Facebook	5,104	0.054	0.226	0	0	1
WhatsApp	5,104	0.056	0.229	0	0	1

Table H.2: Summary Statistics for Polygamy Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,103	3.133	0.949	1	3	4
Any Correction	5,103	0.735	0.441	0	1	1
Outpartisan Speaker	5,103	0.063	0.243	0	0	1
Copartisan Speaker	5,103	0.118	0.323	0	0	1
Congenial Media	5,103	0.116	0.321	0	0	1
Dissonant Media	5,103	0.127	0.333	0	0	1
Congenial Claim	5,103	0.618	0.486	0	1	1
Dissonant Claim	5,103	0.323	0.468	0	0	1
BJP Partisan	5,103	0.680	0.467	0	1	1
Congress Partisan	5,103	0.058	0.234	0	0	1
Pure Control	5,103	0.022	0.145	0	0	1
Peer Correction	5,103	0.247	0.431	0	0	1
Expert Correction	5,103	0.247	0.431	0	0	1
AltNews	5,103	0.045	0.208	0	0	1
Vishwas	5,103	0.050	0.219	0	0	1
TOI	5,103	0.048	0.215	0	0	1
Facebook	5,103	0.047	0.212	0	0	1
WhatsApp	5,103	0.050	0.218	0	0	1

Table H.3: Summary Statistics for MMR Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,061	2.500	1.058	1	3	4
Any Correction	5,061	0.729	0.444	0	1	1
Congenial Media	5,061	0.125	0.331	0	0	1
Dissonant Media	5,061	0.121	0.326	0	0	1
BJP Partisan	5,061	0.680	0.466	0	1	1
Congress Partisan	5,061	0.058	0.234	0	0	1
Pure Control	5,061	0.022	0.146	0	0	1
Peer Correction	5,061	0.234	0.424	0	0	1
Expert Correction	5,061	0.243	0.429	0	0	1
AltNews	5,061	0.054	0.227	0	0	1
Vishwas	5,061	0.053	0.224	0	0	1
TOI	5,061	0.050	0.218	0	0	1
Facebook	5,061	0.046	0.210	0	0	1
WhatsApp	5,061	0.049	0.215	0	0	1

Table H.4: Summary Statistics for Gomutra Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,099	2.935	1.061	1	3	4
Any Correction	5,099	0.706	0.455	0	1	1
Outpartisan Speaker	5,099	0.061	0.240	0	0	1
Copartisan Speaker	5,099	0.121	0.326	0	0	1
Congenial Media	5,099	0.128	0.334	0	0	1
Dissonant Media	5,099	0.127	0.334	0	0	1
Congenial Claim	5,099	0.618	0.486	0	1	1
Dissonant Claim	5,099	0.323	0.468	0	0	1
BJP Partisan	5,099	0.680	0.467	0	1	1
Congress Partisan	5,099	0.058	0.233	0	0	1
Pure Control	5,099	0.023	0.151	0	0	1
Peer Correction	5,099	0.204	0.403	0	0	1
Expert Correction	5,099	0.251	0.433	0	0	1
AltNews	5,099	0.053	0.224	0	0	1
Vishwas	5,099	0.051	0.221	0	0	1
TOI	5,099	0.048	0.214	0	0	1
Facebook	5,099	0.052	0.222	0	0	1
WhatsApp	5,099	0.046	0.209	0	0	1

Table H.5: Summary Statistics for EVM Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,136	1.633	1.014	1	1	4
Any Correction	5,136	0.728	0.445	0	1	1
Outpartisan Speaker	5,136	0.125	0.331	0	0	1
Copartisan Speaker	5,136	0.063	0.243	0	0	1
Congenial Media	5,136	0.125	0.331	0	0	1
Dissonant Media	5,136	0.125	0.331	0	0	1
Congenial Claim	5,136	0.323	0.468	0	0	1
Dissonant Claim	5,136	0.618	0.486	0	1	1
Congress Partisan	5,136	0.057	0.232	0	0	1
BJP Partisan	5,136	0.680	0.467	0	1	1
Pure Control	5,136	0.022	0.148	0	0	1
Peer Correction	5,136	0.250	0.433	0	0	1
Expert Correction	5,136	0.221	0.415	0	0	1
AltNews	5,136	0.051	0.220	0	0	1
Vishwas	5,136	0.053	0.224	0	0	1
TOI	5,136	0.051	0.220	0	0	1
Facebook	5,136	0.055	0.227	0	0	1
WhatsApp	5,136	0.048	0.214	0	0	1

Table H.6: Summary Statistics for UNESCO Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,109	2.697	1.264	1	3	4
Any Correction	5,109	0.734	0.442	0	1	1
Outpartisan Speaker	5,109	0.059	0.235	0	0	1
Copartisan Speaker	5,109	0.118	0.322	0	0	1
Congenial Media	5,109	0.119	0.324	0	0	1
Dissonant Media	5,109	0.119	0.323	0	0	1
Congenial Claim	5,109	0.619	0.486	0	1	1
Dissonant Claim	5,109	0.324	0.468	0	0	1
Congress Partisan	5,109	0.057	0.233	0	0	1
BJP Partisan	5,109	0.681	0.466	0	1	1
Pure Control	5,109	0.010	0.099	0	0	1
Peer Correction	5,109	0.253	0.435	0	0	1
Expert Correction	5,109	0.249	0.433	0	0	1
AltNews	5,109	0.049	0.215	0	0	1
Vishwas	5,109	0.044	0.204	0	0	1
TOI	5,109	0.050	0.218	0	0	1
Facebook	5,109	0.047	0.211	0	0	1
WhatsApp	5,109	0.042	0.202	0	0	1

Table H.7: Summary Statistics for Bose Rumor

Statistic	N	Mean	St. Dev.	Min	Median	Max
Belief in Rumor	5,117	2.716	1.049	1	3	4
Any Correction	5,117	0.663	0.473	0	1	1
Congenial Media	5,117	0.126	0.331	0	0	1
Dissonant Media	5,117	0.114	0.317	0	0	1
BJP Partisan	5,117	0.681	0.466	0	1	1
Congress Partisan	5,117	0.057	0.232	0	0	1
Pure Control	5,117	0.023	0.149	0	0	1
Peer Correction	5,117	0.252	0.434	0	0	1
Expert Correction	5,117	0.175	0.380	0	0	1
AltNews	5,117	0.047	0.211	0	0	1
Vishwas	5,117	0.045	0.207	0	0	1
TOI	5,117	0.047	0.212	0	0	1
Facebook	5,117	0.048	0.214	0	0	1
WhatsApp	5,117	0.048	0.214	0	0	1

I Pretest Data

We ran a pretest on a panel of Facebook-recruited Indian respondents in early May 2019 (N=640) to measure the salience and rate of belief in 37 different rumors commonly disseminated on social media in India. These rumors were:

1. In the future, the Muslim population in India will overtake the Hindu population in India.

2. Polygamy is very common in the Muslim population.

3. Papaya leaf juice is a good way to cure dengue fever.

4. The food prepared by menstruating women is contaminated and rots faster.

5. M-R vaccines are associated with autism and retardation.

6. M-R vaccines are sometimes used by the government to control the population growth amongst certain groups.

7. One must sleep on the left side after having food, as any other sleeping position could be harmful to the digestive tract.

8. Drinking cow urine (gomutra) can help build one's immune system.

9. Gandhi did not try to save Baghat Singh and may even have been a co-conspirator in his death.

10. Indira Gandhi converted to Islam after marrying Feroze Gandhi.

11. Netaji Bose did NOT die in a plane crash in 1945.

12. Arvind Kejriwal has a drinking problem and makes videos while drunk.

13. Sonia Gandhi smuggled Indian treasures to Italy.

14. The BJP has hacked electronic voting machines.

15. NRIs will be able to vote online during the 2019 elections.

16. New Indian notes have a GPS chip to detect black money.

17. UNESCO declared PM Modi best Prime Minister in 2016.

18. WhatsApp profile pictures can be used by ISIS for terror activities.

19. People with cancer shouldn't eat sugar as it feeds cancer cells.

20. Biopsy causes a tumour to turn cancerous.

21. One should not take the P/500 paracetamol, as doctors have shown it to contain machupo, one of the most dangerous viruses in the world.

22. Dengue can be prevented with coconut oil, cardamom seeds, and eupatorium perfoliatum.

23. Amul Kulfi has some pig contents.

24. Drinking Pepsi after eating Polo or Mentos can cause instant death.

25. The BJP is in league with Facebook to remove anti-BJP pages and advertisements.

26. PM Modi hired a makeup artist for 15 lakh monthly salary.

27. Amit Shah personally ordered the assassination of Judge Loya.

28. Arun Jaitley is the current minister of Finance of the Government of India.

29. Scientists warn that current air quality in Delhi shortens lifespan by several years on average.

30. Pryanka Chopra married an American singer in 2018.

31. Mukesh Ambani's residence in Mumbai is the largest private home in the world.

32. India is now the fifth largest economy in the world.

33. Sachin Tendulkar owns the record number of runs record in the ICC cricket world cup.

34. Australia is the country that has won the ICC cricket world cup the most often.

35. According to the 2011 census, Sikhs represent less than 2% of the total Indian population.

36. There is no vaccine that cures HIV/AIDS.

37. Gandhi started his political career in South Africa before coming back to India.

In Figure I.1 we plot we percent of the pretest sample who said they heard each rumor. In Figure I.2 we plot the percent of the sample who said a given rumor was very accurate or somewhat accurate. We highlight the rumors from this list that we selected for the final experiment.

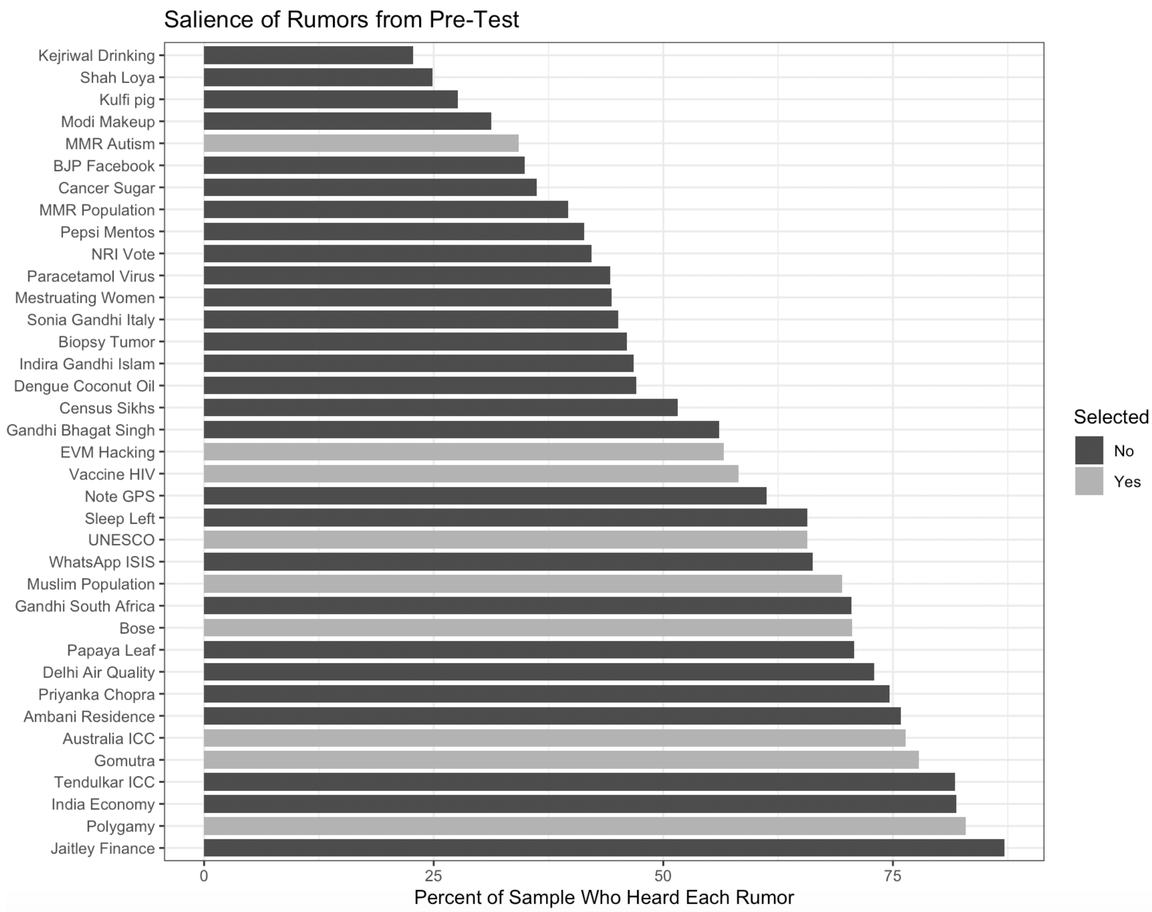


Figure I.1: Salience of Pretest Rumors

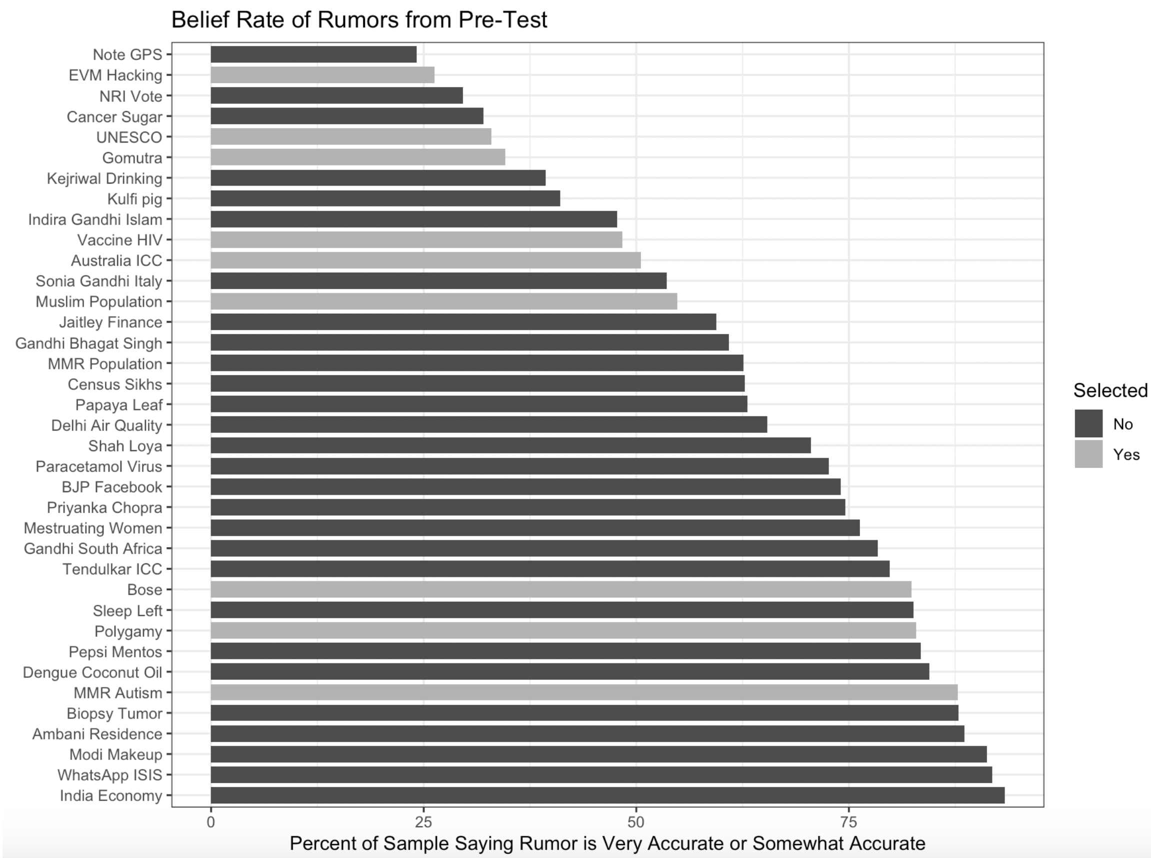


Figure I.2: Belief in Pretest Rumors

J Comparing the Effects of Different Types of Social Corrections

Table J.1: The Effect of Different Types of Social Corrections

	Dependent variable:						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unsourced Correction	-0.076	-0.126^{***}	-0.348^{***}	-0.101^{*}	-0.013	-0.292^{***}	-0.091^{*}
	(0.045)	(0.037)	(0.041)	(0.043)	(0.039)	(0.049)	(0.039)
Domain Expert Correction	-0.119^{**}	-0.234^{***}	-0.469^{***}	-0.150^{***}	-0.016	-0.483^{***}	-0.125^{**}
	(0.044)	(0.037)	(0.041)	(0.040)	(0.041)	(0.049)	(0.043)
Any Factchecker Correction	-0.117^{**}	-0.209^{***}	-0.521^{***}	-0.065	-0.042	-0.465^{***}	-0.117^{**}
	(0.044)	(0.037)	(0.040)	(0.040)	(0.039)	(0.050)	(0.039)
Constant	2.746^{***}	3.272^{***}	2.827^{***}	3.010^{***}	1.650^{***}	2.999^{***}	2.788^{***}
	(0.033)	(0.026)	(0.028)	(0.027)	(0.027)	(0.034)	(0.025)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.002	0.009	0.039	0.003	0.0002	0.024	0.003
Adjusted R^2	0.001	0.009	0.038	0.002	-0.0004	0.024	0.002
Residual Std. Error	1.095	0.945	1.037	1.060	1.014	1.249	1.048
F Statistic	3.045^{*}	16.280^{***}	68.117^{***}	4.889^{**}	0.398	42.110^{***}	4.338^{**}

Note: *p<0.05; **p<0.01; ***p<0.001

Table J.2: Comparing Sourced Corrections and Control to Unsourced Corrections (Omitted Category)

	Dependent variable:						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Control	0.100*	0.125***	0.322***	0.068	0.023	0.251***	0.147***
	(0.045)	(0.037)	(0.041)	(0.042)	(0.039)	(0.049)	(0.039)
Expert Correction	−0.036	−0.113**	−0.148***	−0.072	0.001	−0.226***	−0.024
	(0.042)	(0.037)	(0.041)	(0.043)	(0.041)	(0.048)	(0.043)
Any Factchecker Correction	−0.034	−0.088*	−0.201***	0.012	−0.024	−0.208***	−0.016
	(0.042)	(0.037)	(0.041)	(0.043)	(0.039)	(0.049)	(0.039)
Constant	2.663***	3.152***	2.507***	2.932***	1.633***	2.742***	2.687***
	(0.029)	(0.026)	(0.029)	(0.031)	(0.027)	(0.033)	(0.025)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.002	0.009	0.037	0.002	0.0003	0.022	0.004
Adjusted R^2	0.002	0.009	0.036	0.002	−0.0003	0.022	0.004
Residual Std. Error	1.095	0.945	1.038	1.060	1.014	1.250	1.047
F Statistic	3.715*	16.214***	64.747***	3.891**	0.480	38.838***	7.292***

Note: *p<0.05; **p<0.01; ***p<0.001

Figure J.1: The Effect of Sourced Corrections and Control Condition, compared to Unsourced Correction (Omitted Category)

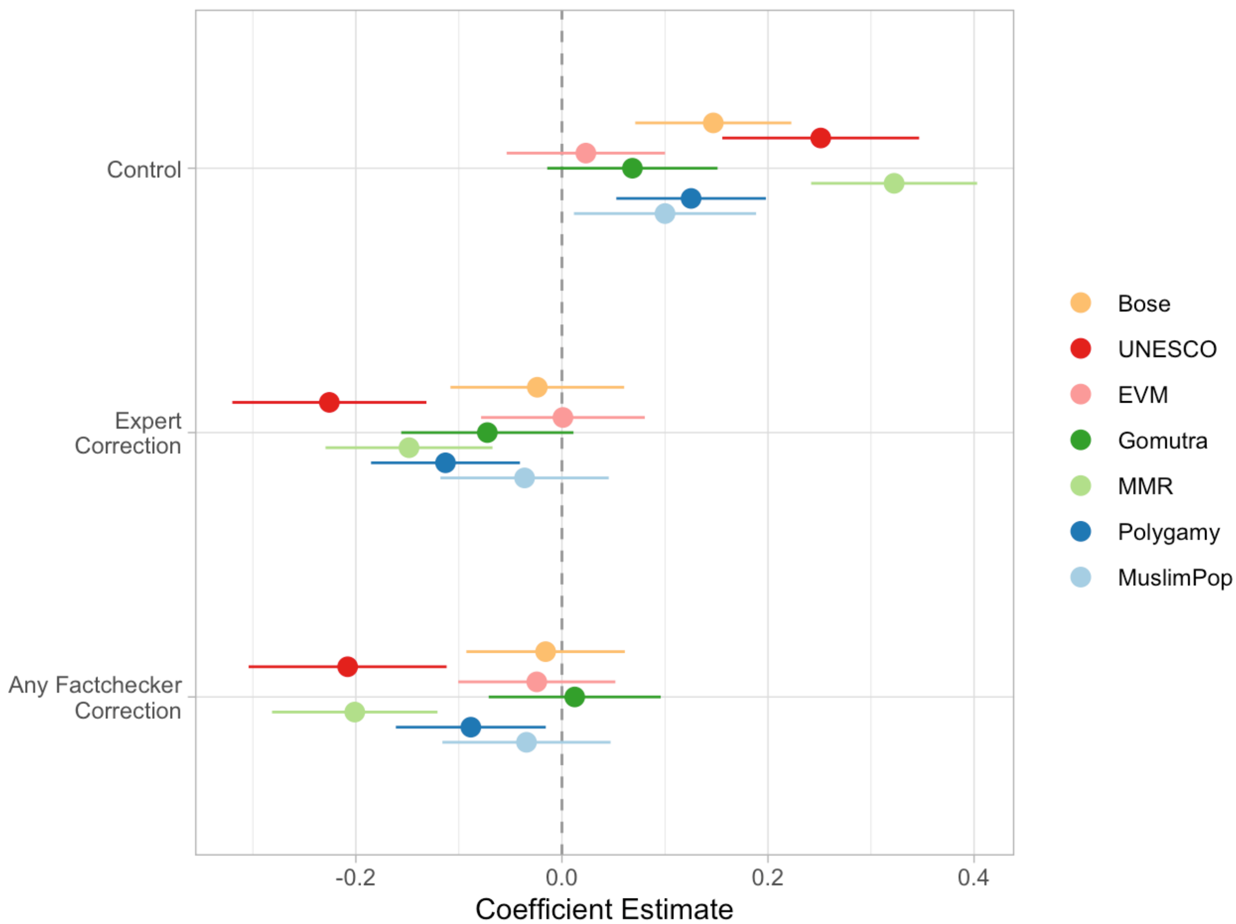


Figure J.2: Comparing Sourced Corrections (pooled) and Control to Unsourced Corrections (Omitted Category)

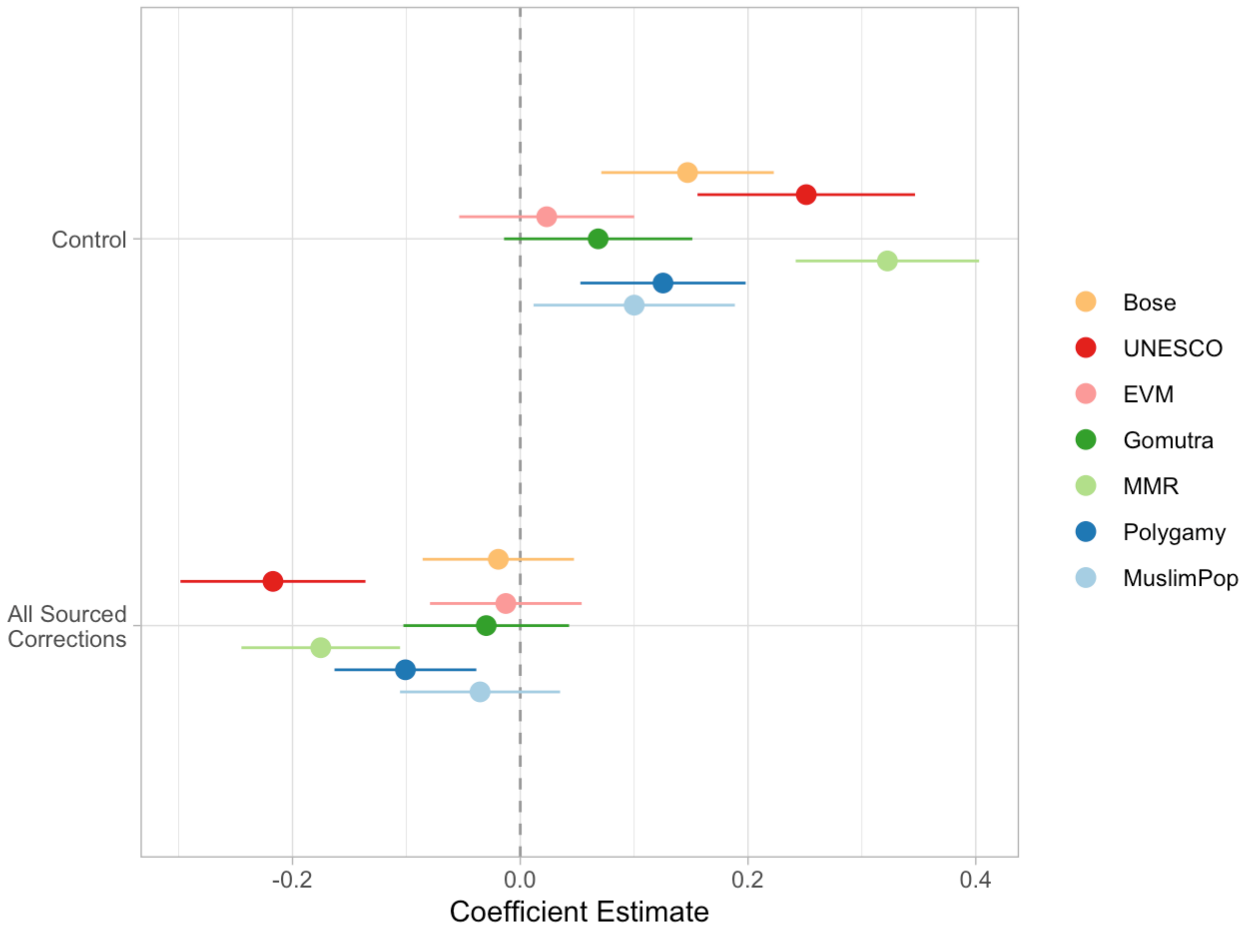


Table J.3: Disentangling by Identity of the Fact-checking Authority

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unsourced Correction	-0.076^{*}	-0.126^{***}	-0.348^{***}	-0.101^{**}	-0.013	-0.292^{***}	-0.091^{**}
	(0.045)	(0.037)	(0.041)	(0.042)	(0.039)	(0.049)	(0.039)
Domain Expert	-0.119^{***}	-0.234^{***}	-0.469^{***}	-0.149^{***}	-0.016	-0.483^{***}	-0.125^{***}
	(0.044)	(0.037)	(0.041)	(0.040)	(0.041)	(0.049)	(0.043)
AltNews	-0.079	-0.199^{***}	-0.591^{***}	-0.092	-0.040	-0.487^{***}	-0.171^{**}
	(0.075)	(0.067)	(0.069)	(0.069)	(0.068)	(0.086)	(0.072)
Vishwas News	-0.134^{*}	-0.276^{***}	-0.468^{***}	-0.101	-0.038	-0.367^{***}	-0.075
	(0.074)	(0.064)	(0.069)	(0.071)	(0.067)	(0.090)	(0.074)
Times of India	-0.115	-0.248^{***}	-0.522^{***}	-0.046	-0.051	-0.398^{***}	-0.207^{***}
	(0.077)	(0.065)	(0.071)	(0.073)	(0.068)	(0.085)	(0.072)
Facebook	-0.186^{**}	-0.160^{**}	-0.529^{***}	-0.115	-0.040	-0.547^{***}	-0.102
	(0.074)	(0.066)	(0.073)	(0.070)	(0.066)	(0.088)	(0.072)
WhatsApp	-0.070	-0.158^{**}	-0.494^{***}	0.034	-0.040	-0.529^{***}	-0.028
	(0.073)	(0.065)	(0.072)	(0.074)	(0.070)	(0.091)	(0.071)
Constant	2.746^{***}	3.272^{***}	2.827^{***}	3.010^{***}	1.650^{***}	2.999^{***}	2.788^{***}
	(0.033)	(0.026)	(0.028)	(0.027)	(0.027)	(0.034)	(0.025)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.002	0.010	0.039	0.004	0.0002	0.025	0.003
Adjusted R^2	0.001	0.009	0.038	0.002	-0.001	0.024	0.002
Res. Std. Er.	1.095	0.945	1.037	1.060	1.014	1.249	1.048
F Statistic	1.589	7.424^{***}	29.487^{***}	2.585^{**}	0.174	18.591^{***}	2.520^{**}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

K Sample Characteristics

Table K.1: Summary Statistics of Key Variables

Statistic	N	Mean	St. Dev.	Min	Median	Max
Age	4,948	29.68	9.43	18	27	76
Male	5,136	0.86	0.34	0	1	1
Education	5,136	6.93	0.97	1	7	8
Hindu	5,136	0.87	0.33	0	1	1
Upper Caste (General)	5,136	0.57	0.49	0	1	1
SC / ST	5,136	0.13	0.34	0	0	1
BJP Partisan	5,136	0.65	0.47	0	1	1
Facebook Use Frequency	5,136	5.40	0.99	1	6	6
WhatsApp Use Frequency	5,136	5.65	0.83	1	6	6

L Benjamini-Hochberg Adjustment

L.1 Adjustment for Table 3

In this table, we present results from a Benjamini-Hochberg adjustment, based on a False Discovery Rate of 5% (0.05).

As seen in the table, the six effects that were significant in Table 3 (main paper) remain significant under Benjamini-Hochberg's adjustment strategy (Benjamini and Hochberg 1995).

Rumor	**Corrective Effect**	**St. Err.**	**p-value**	**Adjusted α**	**Benj.-Hoch Significance**
MuslimPop	0.106	0.037	0.005	0.035	Yes
Polygamy	0.182	0.031	2.29e-9	0.021	Yes
MMR	0.428	0.033	<2e-16	0.014	Yes
Gomutra	0.112	0.033	0.001	0.042	Yes
EVM	0.018	0.032	0.566	0.05	No
Unesco	0.413	0.040	<2e-16	0.014	Yes
Bose	0.116	0.032	0.0002	0.028	Yes

L.2 Adjustment for Table D.1

In this table, we present results from a Benjamini-Hochberg adjustment, based on a False Discovery Rate of 5% (0.05).

We are interested here in the significance of the interaction term between correction and the identity of the politician.

As seen in the table, the one effect that was significant in Table D1 (appendix) remains significant under Benjamini-Hochberg's adjustment strategy (Benjamini and Hochberg 1995).

Rumor	Corrective Effect	St. Err.	p-value	Adjusted α	Benj.-Hoch Significance
MuslimPop	.482	.163	.003	0.01	Yes
Polygamy	.062	.174	.722	0.02	No
Gomutra	-.055	.168	.744	0.03	No
EVM	.008	.118	.947	0.05	No
Unesco	.016	.218	.943	0.04	No

L.3 Adjustment for Table D.2

In this table, we present results from a Benjamini-Hochberg adjustment, based on a False Discovery Rate of 5% (0.05).

We are interested here in the significance of the interaction term between correction and the identity of the politician.

As seen in the table, the one interaction that was significant in Table D2 (appendix) - the interaction related to the UNESCO rumor - does *not* remain significant under Benjamini-Hochberg's adjustment strategy (Benjamini and Hochberg 1995).

Rumor	Corrective Effect	St. Err.	p-value	Adjusted α	Benj.-Hoch Significance
MuslimPop	-.054	.141	.702	0.05	No
Polygamy	-.155	.116	.179	0.03	No
Gomutra	.176	.129	.172	0.02	No
EVM	-.120	191	.529	0.04	No
Unesco	-.349	.167	.036	0.01	No

L.4 Adjustment for Table E1

L.4.1 Correction on *AnyCorrection*CongenialMedia* interaction

In this table, we present results from a Benjamini-Hochberg adjustment, based on a False Discovery Rate of 5% (0.05).

We are interested here in the significance of the interaction term between correction and congenial media.

As seen in the table, the one effect that was significant in Table E1 (appendix) - the one on *MuslimPop* - does not remain significant under Benjamini-Hochberg's adjustment strategy (Benjamini and Hochberg 1995).

Rumor	Corrective Effect	St. Err.	p-value	Adjusted α	Benj.-Hoch Significance
MuslimPop	.271	.135	.045	0.007	No
Polygamy	-.194	.122	.111	0.014	No
MMR	-.049	.131	.709	0.05	No
Gomutra	-.056	.119	.634	0.035	No
EVM	.080	.121	.510	0.028	No
Unesco	-.067	.162	.678	0.042	No
Bose	-.174	.125	.162	0.021	No

L.4.2 Correction on *AnyCorrection*DissonantMedia* interaction

In this table, we present results from a Benjamini-Hochberg adjustment, based on a False Discovery Rate of 5% (0.05).

We are interested here in the significance of the interaction term between correction and dissonant media.

As seen in the table, the two effects that was significant in Table E1 (appendix) - the ones on *MuslimPop* and *Gomutra*- do not remain significant under Benjamini-Hochberg's adjustment strategy (Benjamini and Hochberg 1995).

Rumor	**Corrective Effect**	**St. Err.**	**p-value**	**Adjusted α**	**Benj.-Hoch Significance**
MuslimPop	.219	.131	.096	0.014	No
Polygamy	-.134	.116	.250	0.021	No
MMR	-.067	.130	.606	0.05	No
Gomutra	.272	.126	.031	0.007	No
EVM	-.134	.127	.289	0.028	No
Unesco	-.092	.155	.550	0.042	No
Bose	.091	.127	.476	0.035	No

M Demographic Heterogeneous Effects

In this section we analyze whether demographic characteristics of the sample interact with the effectiveness of the treatment. The tables below present heterogeneous treatment effect models for age, gender, and education (appendix X). We see that while age and education do not interact with the treatment for most rumors, gender might play a role in the uptake of such corrections. For 3/7 of the rumors, men who received a correction became worse at perceived accuracy (i.e. they were more likely to rate a false story as true after receiving the treatment).

Table M.1: Heterogeneous Correction Effects by Age

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AnyCorrection	−0.154	−0.076	-0.370^{***}	-0.218^{*}	−0.066	-0.305^{*}	0.001
	(0.123)	(0.101)	(0.108)	(0.106)	(0.105)	(0.132)	(0.102)
Age	−0.0003	0.005	0.002	−0.001	−0.002	0.003	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
AnyCorrection * Age	0.002	−0.004	−0.003	0.004	0.001	−0.004	−0.004
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Constant	2.753^{***}	3.118^{***}	2.771^{***}	3.055^{***}	1.703^{***}	2.899^{***}	2.729^{***}
	(0.108)	(0.087)	(0.093)	(0.088)	(0.090)	(0.114)	(0.083)
Observations	5,104	5,103	5,061	5,099	5,136	5,109	5,117
R^2	0.002	0.009	0.036	0.002	0.0002	0.021	0.003
Adjusted R^2	0.001	0.008	0.035	0.002	−0.0004	0.020	0.002
Residual Std. Error	1.095	0.945	1.039	1.060	1.014	1.251	1.048
F Statistic	2.816^{*}	14.690^{***}	62.127^{***}	4.075^{**}	0.329	36.206^{***}	4.580^{**}

Note: *p<0.05; **p<0.01; ***p<0.001

Table M.2: Heterogeneous Correction Effects by Gender

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AnyCorrection	-0.250^{*}	-0.259^{**}	-0.498^{***}	-0.373^{***}	-0.041	-0.624^{***}	-0.270^{**}
	(0.100)	(0.083)	(0.091)	(0.087)	(0.087)	(0.104)	(0.085)
Male	0.068	-0.076	0.201^{*}	-0.236^{**}	0.017	-0.213^{*}	-0.056
	(0.095)	(0.077)	(0.084)	(0.078)	(0.080)	(0.095)	(0.075)
AnyCorrection * Male	0.168	0.081	0.058	0.309^{***}	0.018	0.249^{*}	0.185^{*}
	(0.108)	(0.089)	(0.097)	(0.094)	(0.093)	(0.112)	(0.092)
Constant	2.688^{***}	3.341^{***}	2.651^{***}	3.214^{***}	1.636^{***}	3.180^{***}	2.837^{***}
	(0.088)	(0.072)	(0.078)	(0.072)	(0.075)	(0.087)	(0.070)
Observations	5,080	5,077	5,039	5,075	5,111	5,084	5,092
R^2	0.006	0.008	0.042	0.004	0.0002	0.022	0.004
Adjusted R^2	0.005	0.007	0.042	0.004	-0.0003	0.021	0.003
Residual Std. Error	1.092	0.944	1.035	1.059	1.013	1.250	1.047
F Statistic	10.115^{***}	13.508^{***}	74.211^{***}	7.330^{***}	0.409	37.618^{***}	6.421^{***}

Note: *p<0.05; **p<0.01; ***p<0.001

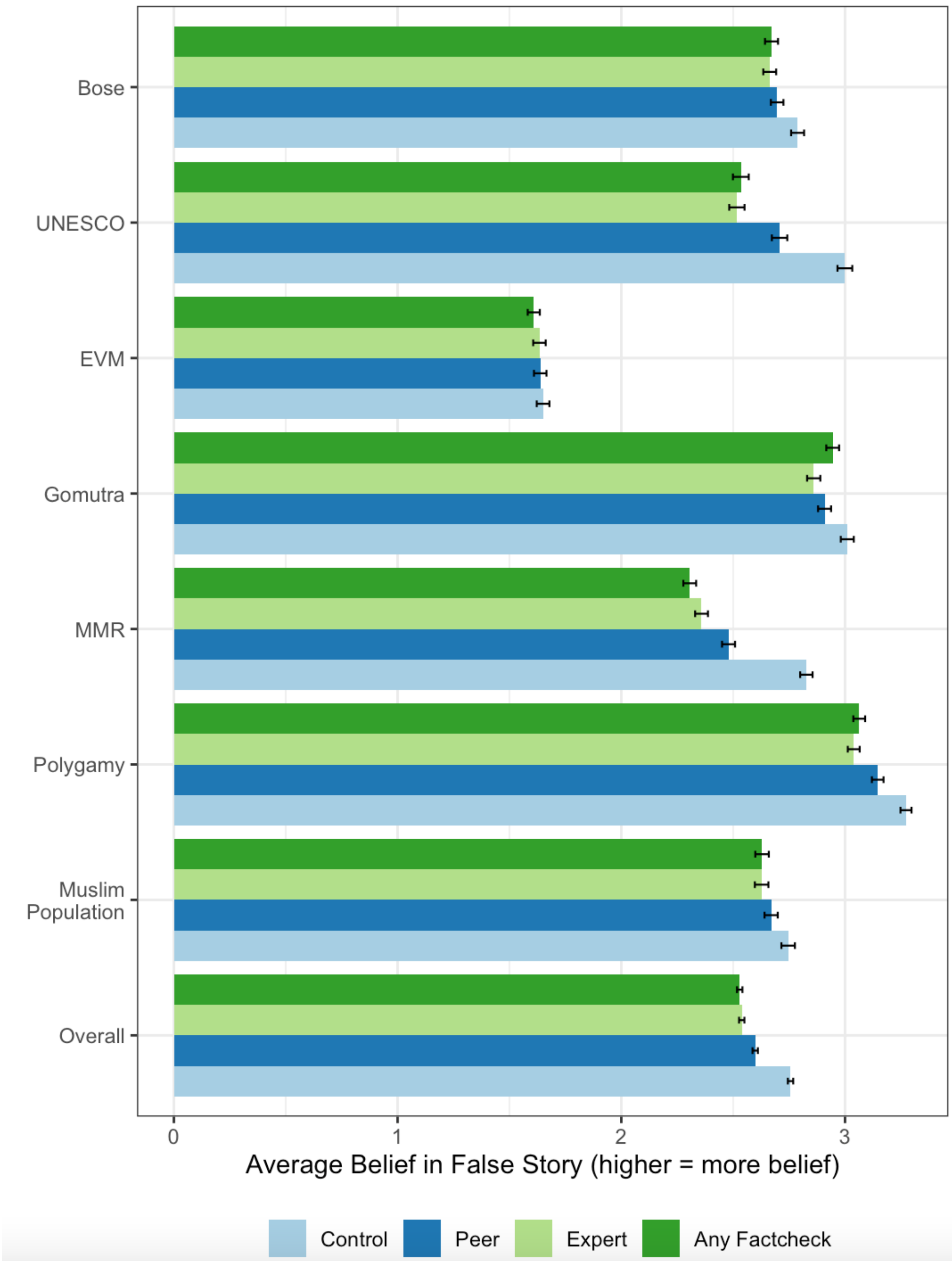
Table M.3: Heterogeneous Correction Effects by Education

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AnyCorrection	-0.006	-0.249	0.038	-0.067	0.523^{*}	-0.265	-0.363
	(0.272)	(0.219)	(0.238)	(0.241)	(0.236)	(0.295)	(0.231)
Education	-0.033	-0.024	-0.068^{*}	-0.001	0.010	-0.130^{***}	-0.065^{*}
	(0.034)	(0.027)	(0.029)	(0.029)	(0.029)	(0.037)	(0.027)
AnyCorrection * Education	-0.014	0.009	-0.066	-0.005	-0.079^{*}	-0.019	0.038
	(0.039)	(0.031)	(0.034)	(0.034)	(0.034)	(0.042)	(0.033)
Constant	2.979^{***}	3.435^{***}	3.272^{***}	3.010^{***}	1.577^{***}	3.887^{***}	3.240^{***}
	(0.240)	(0.186)	(0.202)	(0.203)	(0.202)	(0.256)	(0.191)
Observations	4,844	4,842	4,805	4,836	4,876	4,846	4,854
R^2	0.003	0.008	0.044	0.002	0.003	0.032	0.004
Adjusted R^2	0.003	0.007	0.044	0.001	0.003	0.031	0.003
Residual Std. Error	1.092	0.945	1.031	1.061	1.013	1.247	1.043
F Statistic	5.282^{**}	12.759^{***}	74.457^{***}	3.246^{*}	5.402^{**}	52.619^{***}	5.897^{***}

Note: *p<0.05; **p<0.01; ***p<0.001

N Visualizing All Correction Groups

Figure N.1: The Effect of Different Types of Corrections, Compared to Control



O Order Effects

Table O.1: Main Effect (H1) With Controls and Order Effects

	Dependent variable: Belief in Rumor						
	MuslimPop	Polygamy	MMR	Gomutra	EVM	UNESCO	Bose
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AnyCorrection	-0.141^{**}	-0.165^{***}	-0.403^{***}	-0.132^{**}	-0.020	-0.393^{***}	-0.116^{**}
	(0.048)	(0.039)	(0.043)	(0.042)	(0.042)	(0.051)	(0.040)
Dissonant Media	0.074	-0.007	-0.167^{**}	0.023	-0.089	0.074	0.067
	(0.062)	(0.056)	(0.058)	(0.061)	(0.060)	(0.075)	(0.059)
Congenial Media	0.013	0.099	-0.135^{*}	0.053	-0.104	-0.013	0.060
	(0.062)	(0.056)	(0.057)	(0.060)	(0.059)	(0.075)	(0.056)
Copartisan Speaker	0.011	-0.005		0.117	0.472^{***}	0.073	
	(0.063)	(0.057)		(0.062)	(0.075)	(0.075)	
Outpartisan Speaker	-0.115	-0.237^{**}		-0.248^{**}	-0.164^{**}	-0.149	
	(0.080)	(0.075)		(0.083)	(0.059)	(0.105)	
Order of Story	-0.005	0.001	-0.005	-0.021^{*}	0.003	-0.005	0.015
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)
Constant	2.802^{***}	3.256^{***}	2.856^{***}	3.126^{***}	1.651^{***}	2.999^{***}	2.690^{***}
	(0.063)	(0.051)	(0.058)	(0.058)	(0.056)	(0.069)	(0.055)
Observations	3,165	3,170	3,135	3,161	3,173	3,156	3,167
R^2	0.004	0.010	0.036	0.010	0.019	0.020	0.004
Adjusted R^2	0.002	0.008	0.034	0.008	0.017	0.018	0.003
Residual Std. Error	1.090	0.942	1.039	1.057	1.015	1.248	1.043
F Statistic	2.037	5.478^{***}	28.870^{***}	5.177^{***}	10.267^{***}	10.637^{***}	3.021^{*}
Note:							$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$