

Monk Prayogshala Working
Paper # 2023-7

Hansika Kapoor
Sarah Rezaei
Swanaya Gurjar
Anirudh Tagat
Denny George
Yash Budhwar
Arathy Puthillam

August, 2023

DOES INCENTIVIZATION PROMOTE SHARING "TRUE" CONTENT ONLINE?

Does Incentivization Promote Sharing “True” Content Online?

Hansika Kapoor

Department of Psychology, Monk Prayogshala, Mumbai, India
Neag School of Education, University of Connecticut, Storrs, USA

Sarah Rezaei

Department of Psychology, Monk Prayogshala, Mumbai, India

Swanaya Gurjar

Department of Psychology, Monk Prayogshala, Mumbai, India
³Department of Psychology, Cleveland State University, Cleveland, USA

Anirudh Tagat

Department of Psychology, Monk Prayogshala, Mumbai, India

Denny George

Tattle Civic Tech, India

Yash Budhwar

Tattle Civic Tech, India

Arathy Puthillam

Department of Psychology, Monk Prayogshala, Mumbai, India

Address correspondence to Hansika Kapoor at hk@monkprayogshala.in

The Department of Psychology Working Papers are a series of ongoing research outputs from the Department of Department of Psychology, Monk® Prayogshala®. The purpose of making these papers publicly available is to initiate dialogue and receive feedback on the preliminary work presented. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of Monk Prayogshala, any supporting agencies, or any of its associated entities.

Monk® Prayogshala® is a Section 8 company registered under the Indian Companies Act, 2013. The contents of this document are the Intellectual Property of Monk® Prayogshala® (Sec. 25), a company incorporated under the provisions of the Companies Act, 1956 and having its registered office at 4114, Oberoi Garden Estates, C Wing, Next to Chandivali Studios, Powai, Mumbai 400 072, India. (C) Monk® Prayogshala®, 2011-2020. All rights reserved. The recipient of this document is not permitted to copy, make available, sell, disclose, publish, disseminate or otherwise transmit the information contained in this document without prior permission from Monk® Prayogshala®.

The authors have no conflicts of interest to disclose. Author contributions: HK: Conceptualization, Methodology, Formal data analysis, Writing - Review and Editing, Supervision, Project administration, Funding acquisition. SR: Formal data analysis, Writing - Original draft preparation, Writing - Review and Editing. SG: Methodology, Writing - Original draft preparation, Writing - Review and Editing; AT: Formal data analysis; DG: Software, Funding acquisition; YB: Operations management; AP: Conceptualization, Project administration, Supervision, Funding acquisition, Writing - Review and Editing. This project was funded by the Grant for the Web. The study has been preregistered in the OSF repository.

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

Registration DOI: <https://doi.org/10.17605/OSF.IO/K6MVU>. The data that support the findings of this study are openly available in the OSF repository: <https://osf.io/pfscj/>. This study was approved by the ethics committee at Monk Prayogshala in March, 2022(#086-022).

The authors are grateful to participants at the Tattle Open Call, Behavioural Research in Economics Workshop (BREW)-ESA 2022 for valuable comments and suggestions on the paper. This version of the paper is a pre-publication version of the paper titled 'Does incentivization promote sharing "true" content online?' published in the Harvard Kennedy School Misinformation Review. Please cite the published version, available at [this link](#).

Does Incentivization Promote Sharing “True” Content Online?

Abstract

Information sharing is often driven by social, legal, and financial incentives. To understand online information-sharing behaviour, the present study assessed whether incentivizing people to share “good” and factual content and disincentivizing the sharing of misinformation influenced information sharing. A mock social media platform was created where participants were shown posts that they could share, react to (using emojis), or choose ‘read more’ to get more information. Participants ($N = 908$) were divided into two groups where they received either financial (micropayments) or social (followers) incentives for each true information they shared. Results showed that incentivization, regardless of the type, encouraged people to share more true information; however, the two incentivised groups did not differ in their sharing behaviour, of all posts, regardless of veracity. Further, older and more educated individuals and those with a right-leaning political ideology were more likely to share posts, with the latter being more likely to react to posts as well. Limitations and future directions are discussed.

Keywords: incentives, information-sharing, misinformation, social media, web monetization

Does Incentivization Promote Sharing “True” Content Online?

1. Introduction

The online information ecosystem is influenced by what news gets shared, which has significant implications for people’s behaviour by manipulating their implicit emotions and attitudes (Hudson et al., 2015; Wang et al., 2021). Sharing information gives people confidence and makes them feel more informed and knowledgeable (Bastick, 2021); however, there are dangers to such freedom of creating and sharing information, which includes the proliferation of misinformation (Safieddine et al., 2017) and distortion of reality (Flaxman et al., 2016) among others.

Any attempt to reconceptualize the Internet requires reviewing incentives, not only for content creation but also for content discovery and amplification. Such incentives can apply to financial (real money or tokens) and social (engagement via likes, followers, shares, etc.) Academic and civic attention has been devoted to platforms’ amplification algorithms. But in countries such as India, content is amplified in closed messaging apps such as WhatsApp, even in absence of platform-enabled amplification (Mitra, 2020). The lack of centralised moderation makes information consumers the primary line of defence against low-quality information. With WhatsApp’s ability to sow misinformation easily in this manner, it has attracted attention from researchers focusing on the Global South (Latin America, Asia, Africa, and Oceania) where this app is popularly used. In Brazil, where WhatsApp is a popular mode of communication, Resende et al. (2019) studied the content of political WhatsApp group texts during the 2018 presidential campaign. They found that messages containing false information were more likely to spread quickly within groups, but took longer to go beyond group boundaries. A content analysis (Narayanan et al., 2019) of information shared on online platforms in India leading up to the 2019 election found that more than 25 percent of the Facebook content shared by the Bharatiya Janata

Party (BJP) and a fifth of the Indian National Congress's content was classified as misinformation. Cross-platform comparison showed that misinformation on WhatsApp is generally visual, while on Facebook it was generally in the form of links to news sites with conspiratorial or extremist content.

1.1 Information-sharing as an Economic Game

Considering social networks as public goods allows us to create an experimental task following economic experimental paradigms. Here, public service, such as voluntarily flagging bad information has a personal cost but generates an overall net benefit for the network. Therefore, it is possible to view misinformation sharing from an economic game perspective. In this case, it may be argued that a rewarding mechanism that incentivizes the sharing of good content and punishes the sharing of bad content could be formulated using the volunteer's dilemma (see for example Ehsanfar & Mansouri, 2018).

Within this paradigm, the sharing of good or true information can be considered a cooperative act and can be rewarded, whereas the sharing of divisive or false information can be considered as defection and can be punished. This is because the sharing of divisive information often results in a higher individual payoff of amassing more status or social influence on social media. For example, Rathje et al. (2021) show that divisive content often goes viral on social media. On the other hand, sharing "good" information often is better for the group (here network) as a whole.

1.2 Motives for Information Sharing

While extant literature has explored why people believe in false information (Ecker et al., 2022) and conspiracy theories at large (Douglas et al., 2017), relatively less attention has been given to why people share misinformation. Our understanding of the sharing of misinformation can be informed by our knowledge of the sharing of conspiracy theories. One such common

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

factor is the role of social motives in deciding information sharing on the internet. Douglas and colleagues (2017) postulate that social motives help maintain a positive and competent perceived image of the self and the group. These motives encourage individuals to comprehend information and share misinformation in ways that defend or align with their social identities (Van Bavel et al., 2021). Moreover, misinformation may also be shared, even without belief, to foster stronger ties within the in-group by derogating the out-group (Van Bavel et al., 2021; Osmundsen et al., 2020).

Social engagement, the feedback people receive on their posts online via likes and comments, may act as “online social currency” (Ren et al., 2022). As social engagement affects well-being and perceptions of support, it may heavily influence information-sharing decisions. As high-arousal content is more likely to be shared when people want to create social connections, people may anticipate that sharing misinformation (that also produces high arousal) leads to higher social engagement. Specifically, calculated tradeoffs are made between accuracy and social engagement, with individuals knowingly sharing false information to satisfy their social motives (Ren et al., 2022).

Previous work has indicated that social incentives are ecologically valid. Specifically, they may be as effective, if not more, as monetary rewards (Deaner et al., 2005; Lehner et al., 2017), on tasks involving cognitive control (Licen et al., 2016), attentional orienting (Hayward, et al., 2018), and attentional control (Licen et al., 2019). On the other hand, other studies have shown that social incentives may not be as effective as monetary ones (Radke et al., 2016; Rademacher et al., 2014). An important caveat here might be using a method that is especially sensitive to social incentives.

Other work has shown that monetary incentives induce a performance focus and social incentives induce a normative focus (Mehta et al., 2017). Along with believability, financial incentives have a positive but diminishing effect on sharing information online regardless of

validity. Specifically, financial incentives strongly motivate the spread of health misinformation (Au et al., 2021). Similarly, legislation about misinformation sharing may deter the sharing of information perceived to be true, but not that which is perceived to be false, which was attributed to rebellion and anti-governmental stance (Au et al., 2021).

1.3 Motives for Engagement

1.3.1 Emotional Reactions

Fake headlines may be created to elicit emotional responses and urge audiences to interact with the content in ways that earn profits for the creators (by sharing or accessing the full article; Horner et al., 2022). The advent of social media has made it easier to study these emotional reactions in multiple contexts. Users can judge the accuracy of a social media post based on other endorsements such as the number of likes, comments, and reactions received (Scherer et al., 2021). The inclusion of emojis reflecting anger, sadness, love, laughter, and surprise on social media platforms furthered audience interaction and became another tool of non-verbal and minimal-effort communication (Al-Rawi, 2019). For instance, during the COVID-19 pandemic, news headlines evoked negative emotions like fear, sadness, and anger, with continued exposure to distressing information fueling anxiety (Aslam et al., 2020, Hamidein et al., 2020). Happiness and surprise are also associated with higher belief in COVID-19 misinformation and anger was associated with rejecting said misinformation (Al-Zaman, 2021; Rosenzweig et al., 2021). This emotional reactivity relates to ideological biases that dictate the proliferation of fake news related to opposing parties and suppressing fake news related to one's own party, with political ideology predicting anxious and enthusiastic reactions (Osmundsen et al., 2021; Horner et al., 2022; Yang et al., 2022). The role of emotions in discerning facts from factoids has implications for information processing, sharing, and intervention approaches to combat misinformation.

1.3.2 Reading More About Posts

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

As most people do not read an entire article and/or verify the content of the news they are reading, it is important to use headlines as a reference for the bigger story and lure users to continue reading (Parikh & Atrey, 2018). A balance between reading only the headline and having to read the entire article may be to provide an option to access a short, one or two-sentence summary. The low word length allows users to know enough about the post and is likely to affect their decision of engaging with it. To the best of our knowledge, this is one of the first studies to include a 'read more' option to provide more context about a post.

1.3.3 Political Ideology

Political ideology, which refers to our beliefs about the world and ways to achieve it (Erikson & Tedin, 2003, p. 64), may partially explain beliefs in propaganda and sharing of fake news. Existing research suggests that bottom-up psychological processes make right-leaning individuals more susceptible to conspiracies with greater severity (Enders et al., 2022). Furthermore, conservatives often tend to display more general conspiratorial thinking than liberals (van der Linden et al., 2021), with even the perception of news as being fake differing based on the ideology and affiliations imbibed (Michael & Breaux, 2021). All of this condenses into people on the right sharing more fake news online (Calvillo et al., 2020; Guess et al., 2019; Baptista et al., 2021), with extreme conservatives being more likely to endorse conspiracy theories than extreme liberals (Imhoff et al., 2022). Similarly, supporters of a national right-leaning party in India were less likely than others to accurately recognize fake stories related to their party (Badrinathan, 2021). This suggests that people with certain political ideologies are likely to be more susceptible to believing and sharing misinformation, potentially because of alignments with their social and political identities and biases. Moreover, promoting one's political agenda, partisanship, and animosity towards opposing ideologies also influence why fake news is shared (Osmundsen et al., 2021). However, sweeping generalisations alienate individuals who do not engage in disseminating

misinformation and risk further polarization. Last, some work for reducing misinformation in an Indian context has found support for peer-to-peer correction (Badrinathan et al., 2022). However, other nonsignificant findings hint at the lack of research and raise concerns about the consequences of online actions on offline behaviours leading to riots and violence (Badrinathan, 2021). An examination of the role of political ideology in a multiparty country like India needs to be explored to identify and curb contributing factors of misinformation sharing.

1.4 The Present Study

Against this background, the present study aims to investigate whether and how financial or social incentivization can curb the spread of misinformation online. Specifically, we seek to understand whether using web monetization technologies (with real money) can incentivize people to share accurate information. Can (micro) financial incentives countervail other incentives and nudge individuals to be more critical media consumers and amplify high-quality content? We aim to investigate whether incentivizing people to share “good” and factual content and disincentivizing the sharing of misinformation, either through micropayments or through social feedback, reduces the sharing of misinformation.

The following research question and hypotheses were pre-registered¹:

RQ1: Do social or financial incentives affect the likelihood of sharing content?

H1. There is a significant difference in terms of sharing true content between the baseline and the incentivized condition for all participants.

H1a. There is a significant difference in the amount of true content shared between the baseline and the financial incentive condition.

H1b. There is a significant difference in the amount of true content shared between the baseline and the social incentive condition.

¹ <https://osf.io/k6mvu>

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

H2. There is a significant difference in terms of sharing true and plausible content between the incentivized conditions.

H2a. There is a significant difference in terms of sharing true content between participants who earn social incentives versus those who earn financial incentives.

H2b. There is a significant difference in terms of sharing plausible content between participants who earn social incentives versus those who earn financial incentives.

H3. There is a significant difference in terms of sharing false and implausible content between the incentivized conditions.

H3a. There is a significant difference in terms of sharing false content between participants who earn social incentives versus those who earn financial incentives.

H3b. There is a significant difference in terms of sharing implausible content between participants who earn social incentives versus those who earn financial incentives.

H4. Those in the social incentive condition are more likely to share wholesome content compared to those in the financial incentive condition.

Additionally, two exploratory research questions were asked:

RQ2: Does additional information beyond a headline affect the decision to share a post?

RQ3: Do different types of post content elicit different socio-emotional reactions (measured through three emojis)?

2. Method

The study received ethical approval from the Institutional Review Board at Monk Prayogshala in March 2022 (#086-022).

2.1 Participants

Power analysis was computed using G*Power (latest ver. 3.1.9.7; Faul et al., 2007; Faul et al., 2009). Ren et al. (2021) find that in the like condition, 45.2% of participants shared conspiracy

beliefs. Accordingly, we used a two-proportions z-test to calculate the sample size. The proportions of .45 and .55 were used along with a power of .80, alpha of .05, and allocation ratio of 1 in G*Power; the sample size was estimated to be 784 (392 in each condition: financial and social incentives). Keeping in mind an attrition rate of 25%, we aimed to recruit 908 participants.

A total of 2464 participants filled out a screening survey (described subsequently), out of which the data of 908 participants² (women = 336, men = 552, other = 20; $M_{\text{age}} = 28.8$, $SD = 65.61$, age range:18–69 years, Hindus = 712, Muslims = 54, Christians = 28, Atheists = 51, Other religions = 63) were retained. Invalid responses were discarded based on nationality (not Indian), age (below 18 years), self-reported honesty (< 7 on a 10-point scale), not passing at least one of the two attention checks, not completing all 3 days of the study, and if they had duplicate responses. About 50% of the sample were employed, 26% were students, 11% were self-employed and less than 5% were either unemployed or retired. Of the 908 participants, 441 were part of the financial condition and 467 were part of the social condition (Table 1).

2.2 Measures

2.2.1 Political ideology scale (Puthillam et al., 2021)

This is a 34-item scale that measures political ideology among Indians on a 7-point Likert scale. Its three factors include purity-based cultural norms ($\alpha = 0.89$), obedience to hierarchical authority ($\alpha = 0.94$), and economic ideology. Higher scores on the PI scale and subscales indicate right-leaning political ideology. For the purpose of this study, only the purity and obedience subscales were used, as the economic ideology subscale showed poor internal consistency in earlier work ($\alpha = 0.38$).

2.2.2 Post-task Measures

² Of the 2464 that filled the survey, 1557 individuals fulfilled the inclusion criteria and were invited to participate in the study. From the 1557 individuals, there were participants who dropped out at different stages of the experiment and only 940 individuals completed all 3 days of the experiment. Data from these participants were further cleaned based on other inclusion criteria (passing the attention checks) and finally valid data from 908 participants were retained.

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

The following measures were part of the follow-up Qualtrics form the participants completed 3 days of the experiment.

Moral Emotions. Based on scholarship related to moral emotions (Haidt, 2003; Tangney et al., 2007), participants were asked to rate nine moral emotions they felt while participating in the study over three days. These moral emotions include positive ones, specifically pride, awe/elevation, and gratitude/thankfulness ($\alpha = 0.75$) and negative emotions, i.e., guilt, shame, embarrassment, anger, disgust, and pity ($\alpha = 0.95$).

Post-task Questionnaire. Participants were also asked a series of questions with respect to their participation in the study (refer to Supplementary Table S1). The responses to these questions were rated on a 7-point scale ranging from 1 = *Not at all* to 7 = *All the time*. Some of the questions were “*How often did you contemplate the message?*”, “*To what extent did you like the messages shown to you?*”, “*To what extent were you familiar with the messages shown to you?*”.

Participants were also asked questions about their understanding of the tasks they performed and whether they were aware of how they earned incentives. For example, those in the social incentive condition were asked questions like: “*To what extent did you understand how you gained followers?*”, “*To what extent were you thinking about followers while sharing the messages?*”. Similarly, those in the financial incentive condition were asked questions like “*To what extent did you understand how you gained money?*”. Responses to these questions were also rated on a 7-point scale (ranging from 1 = *Not at all* to 7 = *Perfectly well*).

The post-task measures also included questions about their motivations to share posts: “*While engaging in the experiment, what motivated you to share the messages?*” and “*While engaging in the experiment, what motivated you to not share the messages?*”. Participants were asked to describe their responses in a few sentences.

2.3 Procedure

The study link for the screening survey was posted on social media platforms such as Twitter and Instagram. The link directed the participants to a Qualtrics form where they had to respond to a few demographic questions, a short questionnaire assessing their political ideology, and a bot test. Those who fulfilled the inclusion criteria (nationality, age, self-reported honesty, 1/2 attention checks) and consented to participate were invited via email to participate in the experiment.

The experimental task involved a social media platform *Meshi* created for the study where participants were given unique login IDs to sign in, as in any social media platform. The experiment was done over three days for each participant to reduce fatigue. Participants were shown a total of twenty-five posts: five posts on the first day (baseline) and ten each on the following two days. Each message shown belonged to one of the five types—true, false, plausible, implausible, and wholesome.

The posts under the ‘plausible’ category were phrased to indicate the probability or likelihood of a certain piece of factual information being true (e.g. “There is a 90% chance that India will not experience a fourth wave of the coronavirus”). The ‘implausible’ category posts were also created with the probability/likelihood phrasing (e.g., “There is a high chance that India’s third COVID-19 wave was not the last one”); however, the posts here were not factual. The posts under the true (e.g., “Scientists believe that new COVID-19 infections are not a major concern anymore”) and false categories (e.g., “There will not be any new variants of the coronavirus after omicron”) contained the same information as those in the plausible and implausible categories but without the probability or likelihood phrasing. Wholesome posts were those that were likely to elicit joy; for example “Dog that knows 40 commands gets a job at a children’s hospital in the US.” Details about content creation and the conditions are included in the [Appendix](#).

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

Participants could choose to (a) share the post, and/or (b) react to it using one of the three emojis displayed depicting the emotions of anger, happiness, and disgust, and/or (c) press 'read more' to get further information on the message displayed. Further, participants were randomly assigned to one of two conditions—financial or social incentives.

In the financial incentive condition, participants were informed that they have to read the message thoroughly and use their judgement to determine which content is true, or false, and which would spread joy (wholesome). Further, on the second and third days of the study, participants were rewarded for sharing true/plausible content by receiving ₹10 in their *Meshi* wallet and penalised for sharing false/implausible content by deduction of ₹4 from their wallet. The instructions for the social incentive condition were the same; however, with respect to rewards, on the second and third days of the study, participants were rewarded for sharing true content by getting 100 followers for each true content shared and will be penalised for sharing false content by deduction of 40 followers. Sharing wholesome posts did not entail receiving any incentives or disincentives in both conditions. All participants received initial endowments of ₹ 40 and 400 followers at the start of the experiment.

At the end of Day 3, participants responded to a Qualtrics form with a few post-task questions measuring moral emotions associated with the experimental task, their engagement (emotional and attention) with that task, understanding of the incentives associated with sharing, and open-ended questions about motivations for sharing/not sharing the posts.

3. Results

Data were analysed using the RStudio software version 22.7.01 (RStudio team, 2022) and STATA 17. The analysis code is available on OSF. Sample descriptives and zero-order correlations are displayed in Table 2. Individuals with more education were less likely to share all content, except false posts; negative correlations were found between education and sharing true $r(908) = -$

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

.08, $p < .05$, plausible $r(908) = -.08$, $p < .05$, implausible $r(908) = -.08$, $p < .05$, and wholesome posts $r(908) = -.10$, $p < .05$. Those with more right-leaning political ideology were less likely to choose to 'read more' about a post. Similarly, the subcomponents of the PI scale (obedience to hierarchical authority and purity-based cultural norms) also negatively correlated with reading more about a post. Individuals with a more right-leaning political ideology were also more likely to share posts, except for false posts. Those more likely to show obedience towards hierarchical authority were likely to share all types of posts; however, those who upheld more purity-based cultural norms were more likely to share true and plausible posts only.

The correlations for the subsample ($n = 627$) that responded to the post-task measures are displayed in Table 3. Younger individuals were more likely to give angry reactions to false and implausible posts, whereas older individuals were more likely to give happy reactions to false and implausible posts. Those who experienced more positive moral emotions during the span of the experiment were also likely to give more happy reactions to true and plausible posts, and angry reactions to all posts (except wholesome). They were also less likely to give disgusted reactions to true and plausible posts and happy reactions to false and implausible posts. A similar pattern of results was found for those who experienced more negative moral emotions during the experiment as well. Moreover, those who experienced more negative moral emotions were less likely to give happy reactions to wholesome content.

A series of Bonferroni corrected t-tests were computed on participant demographics in both conditions to check for demographic distribution across the two financial and social conditions (Table S2). Although there were more males and employed persons in the social condition, and more students and atheists in the financial condition, these differences were not significant at the .002 level (Bonferroni-corrected p-value, for 23 comparisons), indicating that both groups were equivalent on demographic characteristics.

3.1 Sharing of True and Plausible Posts

The number of true posts shared in any incentivised condition was significantly higher than that of the baseline (H1), $t(907) = -47.78, p < 0.001, d = 1.59$. Significantly more true posts were shared in both incentivised conditions compared to baseline (H1a and H1b): financial, $t(440) = -33.56, p < 0.001, d = 1.6$, and social, $t(466) = -33.99, p < 0.001, d = 1.57$. However, the two incentivized groups did not differ significantly in the mean number of true (H2a), $t(900.5) = 0.06, p = 0.95, d = 0.01$, and plausible posts shared (H2b), $t(903.11) = 1.34, p = 0.18, d = 0.09$. The sharing of all factual (true and plausible) messages also did not differ across the two incentivized conditions (H2), $t(902.93) = 0.79, p = 0.43, d = 0.05$.

3.2 Sharing of False and Implausible Posts

The type of incentive also did not influence the number of false posts shared (H3a), $t(905.78) = -1.56, p = 0.12, d = 0.1$ or the number of implausible posts shared (H3b) in the two incentivised conditions, $t(905.26) = -1.75, p = 0.08, d = 0.11$. Similarly, the sharing of all false messages (false and implausible) did not differ significantly across the two incentivized conditions (H3), $t(905.51) = -1.75, p = 0.08, d = 0.12$.

3.3 Sharing of Wholesome Posts

Similar results were found for wholesome posts shared in the financial and social conditions (H4). The type of incentive did not influence the number of wholesome posts shared, $t(902.05) = -0.47, p = 0.64, d = 0.03$.

3.4 Reactions and 'Read More'

A factorial ANOVA was conducted to examine the effect of participant's demographics (age, gender, religion, education, political ideology), reading additional information (clicking 'read more'), the type of incentives (financial/social) and the type of message (true, false, plausible, implausible, wholesome) on the decision to share/not share a post at the message level (RQ2).

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

Results indicated that this decision differed significantly based on these variables, $F(18, 22281) = 214.89, p < 0.001, R^2 = 0.15$.

Sharing a message was coded as 1 and not sharing it was coded as 0. Generally, false ($M = .29, SD = .45$) and implausible messages ($M = .29, SD = .45$) were significantly less likely to be shared as compared to true ($M = .61, SD = .49$), plausible ($M = .59, SD = .49$), and wholesome messages ($M = .61, SD = .49$). Those who clicked on “read more” were more likely to share messages, regardless of type.

The type of incentive did not influence sharing behaviour, $F(1, 22281) = 0.14, p = 0.70$.

Among the covariates, older and more educated individuals were more likely to share messages; there was no difference between men and women in post-sharing behaviour, and those with more right-leaning political beliefs (with respect to purity and obedience to authority) were likely to share more posts³.

Similar ANOVAs were run to assess the effects of demographics and post-related behaviour on happy, disgusted, and angry reactions to the messages (RQ3, Figure 1). Happy reactions were predicted by this combination of variables, $F(18, 22281) = 127.38, p < 0.001, R^2 = 0.09$. Factual posts—true ($M = .49, SD = .50$) and plausible ($M = .47, SD = .50$)—were more likely to elicit happy reactions than false ($M = .26, SD = .44$) and implausible posts ($M = .27, SD = .45$); wholesome posts ($M = .64, SD = .48$) were more likely to elicit happy reactions than any other type of posts. Those who clicked on “read more” were also more likely to give a happy reaction across conditions. Among the covariates, there was no difference between men and women when giving happy reactions to posts, and age did not influence happy reactions to posts, $F(18, 22281) = 4.30, p = 0.01$. Only those with more right-leaning political beliefs related to obedience to hierarchical authority gave more happy reactions to posts.

³ Differences in post sharing among religious groups was not inferred as nearly 78% of the sample identified as being Hindu and the remaining categories each comprised less than 6% of the data.

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

Similar results were found for disgust reactions where a combination of demographics and post-related behaviour was seen to influence such reactions to posts, $F(18, 22281) = 47.05, p < 0.001, R^2 = 0.04$. Contrary to happy reactions, disgust reactions were more likely to be elicited by false ($M = .18, SD = .38$) and implausible ($M = .17, SD = .38$) posts than true posts; however, true ($M = .10, SD = .30$) and plausible ($M = .10, SD = .31$) posts were more likely to elicit disgust reactions than wholesome posts ($M = .03, SD = .18$). Reading more about the post elicited more disgust reactions, $F(18, 22281) = 6.32, p = 0.01$; participants in the social condition ($M = .13, SD = .33$) gave significantly more disgust reactions than those in the financial condition ($M = .11, SD = .31$). There was no difference between men and women when giving disgusted reactions to posts; age was not a significant predictor, and those with more right-leaning political beliefs gave more disgusted reactions.

For anger reactions, both demographics and post-related behaviour showed significant influence on anger reactions to posts, $F(18, 22281) = 95.85, p < 0.001, R^2 = 0.07$. Similar to disgust reactions, false and implausible posts ($M_s = .23, SD_s = .42$) were more likely to elicit anger reactions than true and plausible posts ($M_s = .14, SD_s = .35$); these in turn were reacted to more angrily than wholesome posts ($M = .04, SD = .20$). Reading more about the post elicited more anger reactions, $F(18, 22281) = 20.22, p < 0.001$; participants in the financial condition ($M = .16, SD = .37$) gave significantly more anger reactions than those in the social condition ($M = 0.15, SD = .36$). Older participants, $F(18, 22281) = 33.01, p < 0.001$, and those with more right-leaning political beliefs gave more angry reactions. No gender difference between men and women was observed.

3.5 Post-task Measures

A sub-sample of 627 individuals who participated in the experiment over the 3 days responded to the follow-up Qualtrics form with the post-task measures (Table S1). Results show

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

that on average, the sample contemplated the messages shown to them over the 3 days and also thought of the consequences to their friends/family and others, including those they dislike/hate if the messages were true. With respect to incentives, participants in the two incentivised groups (social and financial) on average were aware of how they gained and lost their incentives (followers/money). However, more individuals from the social incentive group thought about their incentives while sharing messages than those from the financial group.

4. Discussion

The present study aimed to examine if people's information-sharing behaviour on social media can be influenced by social or financial incentives. Specifically, we attempted to understand whether incentivization can curb misinformation sharing on social media platforms. For this purpose, we created a social media platform '*Meshi*' where participants were shown a total of twenty-five posts over a period of three days that they could share, react to, or choose 'read more' to get further information. Our results indicate that incentivization, regardless of the type, encourages people to share more true information. However, the two incentivised groups (financial and social) did not differ in the mean number of factual posts (true and plausible) shared.

4.1 Role of Incentives in Sharing Posts

Sharing posts when there are incentives involved can be explained by rational choice theory, which posits that choices are made through a cost-benefit analysis intended to maximise personal utility, with any form of incentive increasing the likelihood of participation (Becker, 1976; Logan et al., 2018; Zou, 2016). However, the lack of significant differences between social and financial incentives for sharing true information suggests that a particular incentive did not take precedence over the other. It is likely that participants were motivated solely by the presence of an incentive and a desire to earn more money or create greater social engagement. Moreover,

participants may have been unable to decipher the nuances between true and plausible messages as sharing of all factual information also did not differ between the two incentivized conditions.

Similarly, individuals were equally likely to share false information regardless of the nature of the incentive. This suggests that the sharing of (mis)information is not governed only by social or monetary incentives or conversely, that the sharing of (mis)information may be equally likely to be governed by both. It is also plausible that the kind of social incentives matter; in the current study, users gained or lost followers, whereas, in the real world, social engagement is often compounded and can take many forms (likes, reactions, shares, etc.). Similarly, the amount of money to be gained or lost can influence sharing behaviour. A positive implication of this null result may be that curbing misinformation may not require substantial monetary resources. Individuals who choose not to engage with false content (passive) or flag false content (active) can be socially incentivized for being responsible netizens.

4.2 Post Reactions

News is known to elicit emotional reactions in people (Gross and D'Ambrosio 2004), with earlier studies indicating how such news can elicit a myriad of emotional responses (Ryu 1982; Grabe et al., 2001). Results from the present study indicated that false and implausible messages were more likely to elicit negative reactions of anger and disgust than any other type of message. It is likely that false messages elicit more anger reactions than other messages because fake news is often laden with emotions of anger; for example, for political attacks (Higgins, 2017). Similarly, previous studies have shown how emotional responses dictate individuals' information processing and seeking behaviour (Erhardt et al., 2021), alter their attitudes or opinions (Marcus et al., 2019; Vasilopoulos, 2019), as well as influence behavioural changes (e.g., Weeks, 2015).

Conversely, wholesome messages elicited more happy reactions than any other type of message. Wholesome messages included information that would generally bring joy to people

including acts of kindness, contributions made to society and the environment, and the like.

Interestingly enough, in a study by Al-Rawi (2019) on networked emotional news stories, participants clicked positive news stories the most and reacted positively ('love' emoji) to them.

Despite the popular idea that people pay more attention to bad news than good news (Shoemaker & Cohen, 2005), people clearly seek and react to good news. Varied emotional reactions to false, true, and wholesome messages also provide content validation for the nature of posts used in the current study.

4.3 Demographics and Sharing Behaviour

Sharing behaviour on social media, especially when it comes to inaccurate information, is also largely influenced by age (Guess et al., 2019). Our results suggest that older individuals were more likely to share messages, possibly associated with relatively lower digital literacy, to be able to navigate through misinformation or be able to successfully judge the reliability of posts. Hence, they may be more likely to share posts unsparingly regardless of the type of content. More educated individuals in our sample were likely to share posts. It is likely that educated individuals are more inclined towards engaging with information online and sharing helps them disseminate information to the community.

When comparing graduate and undergraduate students in their motivations for sharing misinformation, Chen et al. (2015) found that for graduate students, the motivations to share included reasons related to the characteristics of the information. For example, graduate students were more likely to share misinformation if it was interesting, new and eye-catching, or was a good topic for conversation. Although Chen et al.'s (2015) study focused on the sharing of misinformation, it is possible that the current sample of educated individuals was generally more likely to share posts regardless of their content type because they found them interesting or perceived them as great conversation starters.

4.4 Political Ideology and Engagement with Posts

While clicking “read more” was related to sharing more posts in general, right-leaning individuals were less likely to click “read more” and more likely to share true and wholesome posts but not false posts. Further, the nuances in sharing behaviors among the subcategories of right-leaning individuals offer more insights into the complexities of political ideology in India. Only obedience towards hierarchical authority, and not upholding purity-based cultural norms, predicted sharing of false posts; thus, specific beliefs of discrimination, traditionalism, and compliance with authority are likely to be strong drivers of disseminating misinformation. As a result, findings on the role of political ideology in predicting sharing behaviours are only partially consistent with past research, which suggests that conservatives are more susceptible to sharing misinformation online. However, this may be explained by past studies being overrepresented by samples from the United States, with investigations of sharing behaviours and political ideology related to political misinformation and fake news in the COVID-19 context (Calvillo et al., 2020; Garrett & Bond, 2021; Guess et al., 2019; Havey, 2020; Pennycook & Rand, 2019). Our unique findings also highlight the pitfalls of generalising findings in WEIRD societies to non-WEIRD ones and underscore the need to account for varying cultural influences.

In a recent study, Yang et al., (2022) investigated whether reactions to misinformation and corrective messages regarding COVID-19 differed. They found misinformation and political ideology of the post’s sources to be more predictive of people’s anxious and enthusiastic reactions. Affective polarisation generally leads individuals to politicise neutral or apolitical issues such as COVID-19 (Druckman et al., 2020) which in turn leads them to react to messages about such issues in a certain way. Conservatives’ animus towards the other party may have led them to ignore or doubt messages about COVID-19. In the present study, right-leaning individuals were more likely to give reactions to posts. It is possible that these individuals are more likely to find

messages happy, disgusting, or anger-inducing depending on the nature of the political meaning they attach to these messages.

4.6 Limitations and Future Directions

Studies of misinformation sharing are scant in the Global South, with this being one of the first to study web monetization in this context. That said, this study was not without limitations. First, the study only used posts containing text, which did not provide insights into misinformation across different formats. Second, the manipulation of incentives did not affect sharing behaviour. Future research can extend this work, by altering incentivization methods as well as adapting the content of posts to another format. Additionally, future work could also highlight the interaction between emotions and sharing behavior online. Our study also did not explore how the different themes of messages (Refer to table A1) are shared. For example, future research could focus on whether certain socio demographics are associated with sharing a certain type (theme) of message. Moreover, our study did not take into account the reaction time of participants (i.e. how long each participant spent evaluating the messages before engaging with them), which could potentially contribute to our understanding of sharing behavior. Additionally, given the literature on how political misinformation is widespread on social media platforms especially in the Global South (Narayanan et al., 2019; Resende et al., 2019), researchers could further look into how spreading misinformation can be a way of partisan cheerleading, if incentivization can further help our understanding of this phenomenon (for example Peterson & Lyengar, 2021), and how it can be curbed.

As the social incentives in the present study were purely simulated, it is likely that participants were less concerned with whether they gained or lost “imaginary” followers. The effectiveness of social incentives can be assessed through a real-world experiment where users of a social media platform, like Facebook, can receive a visible marker such as a star against their

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

profile if they consistently share true information/factual posts. In this hypothetical experiment, with say 100 users, fact checkers can be recruited for every 10 participants who would verify shared posts. Based on a predetermined threshold, users can be awarded a star denoting that they are credible sources of information on the platform. To ensure that individuals posting true information are prompted to engage with it, there can be an additional requirement of lateral reading (reading more about the information on the open internet; Wineburg et al., 2022). Further, they can supplement their post with additional sources of the information, to enhance its credibility to other readers (see also McGlynn et al., 2020). This profile 'star' can act as a social incentive for users to engage with and share more true information online. Additionally, a leaderboard system can be put in place that displays top users of the month who consistently share true information. It can be argued that applying social incentives and criteria for sharing information online may be seen as curbing freedom of speech. However, individuals who share misinformation will not be actively penalized for sharing false news; neither will they be socially incentivized for sharing the same. Furthermore, algorithmically, information from individuals with more "social stars" can be made more prominent in social media feeds, driving engagement with more true news than false.

As Ha et al. (2022) suggest, redesigning the online space to provide users with necessary tools to self-moderate can also be attempted. For platforms like Whatsapp and YouTube, group admins or users can be responsible for moderating content shared within groups and comment sections. With respect to handing out social incentives (like stars) to users who share true content, a trustworthy entity such as a fact checker can first evaluate content shared by group admins, and then entrust them to provide the incentives. Indeed, with any form of incentivization involved, the question of whether social media platforms should be transparent with their users arises. In case that they employ social incentives surreptitiously, it's likely to be business-as-usual, with users

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

unaware of the algorithms underlying the platform. On the other hand, being transparent about incentivization can lead to beneficial outcomes (such as a better and more truthful information landscape) or unintended consequences (such as users gaming the incentives themselves for personal gain over the common good). In our opinion, incentive mechanisms should be transparently applied to inform and engage users as well as potentially lead to a ripple effect among social media platforms.

Based on the results of our study, behavior-based interventions may work just as well as financial incentives to promote the sharing of true information online. Although a recent study identified a significant effect of financial incentives on truth discernment and the subsequent sharing of factual information (Rathje et al., 2023), non-financial (and more scalable) interventions also played a role in truth discernment, albeit a smaller one. Taken together, perhaps a combined intervention may be more practical and effective (see Bak-Coleman et al., 2022), where a social incentive structure is employed as a behavior-based intervention along with other psychological interventions that have successfully demonstrated impact in curbing misinformation (see Gwiazdzinski et al., 2023).

Lastly, the study sample was recruited online and represented a fairly digitally literate sample. Future work can attempt to replicate these findings with low digital literacy populations. Moreover, the present study can also be replicated in countries of the Global North (Adetula et al., 2022) to understand how individuals from more digitally sophisticated and literate countries respond to incentives and how that influences their sharing of (mis)information online.

The results of our study point to the eminent need for further analysis and research in the area of information-sharing on social media, especially in the Global South. These have provided multiple potential directions for future work with one such area being understanding the influences of linguistic features of a post on decisions of sharing. Recognizing the potential of social media

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

users to influence others through the information they choose to share is an important first step to kickstarting a digital literacy movement. The start of healthier online information ecosystems may rely on exploring ways to encourage users to be more critical of the information they see, and inevitably, share.

References

- Adetula, A., Forscher, P. S., Basnight-Brown, D., Azouaghe, S., & IJzerman, H. (2022). Psychology should generalize from — not just to — Africa. *Nature Reviews Psychology*, 1(7), 370–371. <https://doi.org/10.1038/s44159-022-00070-y>
- Al-Rawi, A. (2020). Networked emotional news on social media. *Journalism Practice*, 14(9), 1125–1141. <https://doi.org/10.1080/17512786.2019.1685902>
- Al-Zaman, M. S. (2021). Social media and COVID-19 misinformation: How ignorant Facebook users are? *Heliyon*, 7(5), e07144. <https://doi.org/10.1016/j.heliyon.2021.e07144>
- Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities and Social Sciences Communications*, 7(1), 1–9. <https://doi.org/10.1057/s41599-020-0523-3>
- Au, C. H., Ho, K. K., & Chiu, D. K. (2021). Stopping healthcare misinformation: The effect of financial incentives and legislation. *Health Policy*, 125(5), 627–633. <https://doi.org/10.1016/j.healthpol.2021.02.010>
- Badrinathan, S., Chauchard, S., & Flynn, D. J. (2020). I don't think that's true, bro!. An experiment on fact-checking misinformation in India [Manuscript submitted for publication]. <https://sumitrabadrinathan.github.io/Assets/paper-whatsapp.pdf>
- Badrinathan, S. (2021). Educative interventions to combat misinformation: Evidence from a field experiment in India. *American Political Science Review*, 115(4), 1325–1341. <https://doi.org/10.1017/S0003055421000459>
- Bak-Coleman, J. B., Kennedy, I. M., Wack, M., Beers, A., Schafer, J. A., Spiro, E. S., Starbird, K., & West, J. D. (2022). Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour*, 6(10), 1372–1380. <https://doi.org/10.1038/s41562-022-01388-6>

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Baptista, J. P., Correia, E., Gradim, A., & Piñeiro-Naval, V. (2021). The influence of political ideology on fake news belief: The Portuguese case. *Publications*, 9(2), 23.
<https://doi.org/10.3390/publications9020023>
- Bastick, Z. (2021). Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior*, 116, 106633.
<https://doi.org/10.1016/j.chb.2020.106633>
- Becker, G. S. (1976). *The economic approach to human behavior* (Vol. 803). University of Chicago press.
- Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., & Rutchick, A. M. (2020). Political ideology predicts perceptions of the threat of covid-19 (and susceptibility to fake news about it). *Social Psychological and Personality Science*, 11(8), 1119–1128.
<https://doi.org/10.1177/1948550620940539>
- Chen, X., Sin, S. C. J., Theng, Y. L., & Lee, C. S. (2015). Why Students Share Misinformation on Social Media: Motivation, Gender, and Study-level Differences. *The Journal of Academic Librarianship*, 41(5), 583–592. <https://doi.org/10.1016/j.acalib.2015.07.003>
- Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2020). How Affective Polarization Shapes Americans' Political Beliefs: A Study of Response to the COVID-19 Pandemic. *Journal of Experimental Political Science*, 8(3), 223–234.
<https://doi.org/10.1017/xps.2020.28>
- Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. *Current Biology*, 15(6), 543–548.
<https://doi.org/10.1016/j.cub.2005.01.044>

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

Douglas, K. M., Sutton, R. M., & Cichocka, A. (2017). The Psychology of Conspiracy Theories.

Current Directions in Psychological Science, 26(6), 538–542.

<https://doi.org/10.1177/0963721417718261>

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P.,

Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1), 13–29.

<https://doi.org/10.1038/s44159-021-00006-y>

Ehsanfar, A., & Mansouri, M. (2017). Incentivizing the dissemination of truth versus fake news in social networks. *12th System of Systems Engineering Conference (SoSE)*, 1-6.

<https://doi.org/10.1109/SYSE.2017.7994981>.

Enders, A., Farhart, C., Miller, J., Uscinski, J., Saunders, K., & Drochon, H. (2022). Are republicans

and conservatives more likely to believe conspiracy theories? *Political Behavior*. Advance online publication. <https://doi.org/10.1007/s11109-022-09812-3>

Erhardt, J., Freitag, M., Filsinger, M., & Wamsler, S. (2021). The Emotional Foundations of Political

Support: How Fear and Anger Affect Trust in the Government in Times of the Covid-19

Pandemic. *Swiss Political Science Review*, 27(2), 339–352. <https://doi.org/10.1111/spsr.12462>

Erikson R. S. & Tedin K. L. (2003). *American public opinion: its origins content and impact* (6th ed.).

Longman.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power

analysis program for the social, behavioral, and biomedical sciences. *Behavior Research*

Methods, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power

3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(S1), 298-320.
<https://doi.org/10.1093/poq/nfw006>
- Garrett, R. K., & Bond, R. M. (2021). Conservatives' susceptibility to political misperceptions. *Science Advances*, 7(23). <https://doi.org/10.1126/sciadv.abf1234>
- Grabe, M. E., Zhou, S., & Barnett, B. (2001). Explicating sensationalism in television news: Content and the bells and whistles of form. *Journal of Broadcasting & Electronic Media*, 45(4), 635-655. http://doi.org/10.1207/s15506878jobem4504_6
- Gross, K., & D'ambrosio, L. (2004). Framing emotional response. *Political psychology*, 25(1), 1-29.
<https://doi.org/10.1111/j.1467-9221.2004.00354.x>
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), eaau4586.
<https://doi.org/10.1126/sciadv.aau4586>
- Gwiaździński, P., Gundersen, A. S., Piksa, M., Krysińska, I., Kunst, J. R., Noworyta, K., Olejnik, A., Morzy, M., Rygula, R., Wójtowicz, T., & Piasecki, J. (2023). Psychological interventions countering misinformation in social media: A scoping review. *Frontiers in Psychiatry*, 13.
<https://doi.org/10.3389/fpsy.2022.974782>
- Ha, L., Graham, T., & Gray, J. E. (2022). Where conspiracy theories flourish: A study of YouTube comments and Bill Gates conspiracy theories. *Harvard Kennedy School (HKS) Misinformation Review*, 3(5).
- Haidt, J. (2003), The moral emotions, in R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (eds.) *Handbook of affective sciences*. [Online]. Oxford: Oxford University Press. 852–870

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Hamidein, Z., Hatami, J., & Rezapour, T. (2020). How people emotionally respond to the news on COVID-19: An online survey. *Basic and Clinical Neuroscience*, 11(2), 171.
<http://doi.org/10.32598/bcn.11.covid19.809.2>
- Havey, N. F. (2020). Partisan public health: how does political ideology influence support for COVID-19 related misinformation? *Journal of Computational Social Science*, 3(2), 319–342.
<https://doi.org/10.1007/s42001-020-00089-2>
- Hayward, D. A., Pereira, E. J., Otto, A. R., & Ristic, J. (2018). Smile! Social reward drives attention. *Journal of Experimental Psychology: Human Perception and Performance*, 44(2), 206–214.
<https://doi.org/10.1037/xhp0000459>
- Higgins, M. (2017). Mediated populism, culture and media form. *Palgrave Communications*, 3(1), 1-5. <https://doi.org/10.1057/s41599-017-0005-4>
- Horner, C. G., Galletta, D., Crawford, J., & Shirsat, A. (2021). Emotions: The Unexplored Fuel of Fake News on Social Media. *Journal of Management Information Systems*, 38(4), 1039-1066.
<https://doi.org/10.1080/07421222.2021.1990610>
- Hudson, S., Roth, M. S., Madden, T. J., & Hudson, R. (2015). The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism Management*, 47, 68–76.
<https://doi.org/10.1016/j.tourman.2014.09.001>
- Imhoff, R., Zimmer, F., Klein, O., António, J. H. C., Babinska, M., Bangertner, A., Bilewicz, M., Blanuša, N., Bovan, K., Bužarovska, R., Cichočka, A., Delouvé, S., Douglas, K. M., Dyrendal, A., Etienne, T., Gjoneska, B., Graf, S., Gualda, E., Hirschberger, G., ... van Prooijen, J.-W. (2022). Conspiracy mentality and political orientation across 26 countries. *Nature Human Behaviour*, 6(3), 392-403. <https://doi.org/10.1038/s41562-021-01258-7>

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Lehner, R., Balsters, J. H., Herger, A., Hare, T. A., & Wenderoth, N. (2017). Monetary, food, and social rewards induce similar Pavlovian-to-instrumental transfer effects. *Frontiers in Behavioral Neuroscience*, 10, 247. <https://doi.org/10.3389/fnbeh.2016.00247>
- Ličen, M., Hartmann, F., Repovš, G., & Slapničar, S. (2016). The impact of social pressure and monetary incentive on cognitive control. *Frontiers in Psychology*, 7, 93. <https://doi.org/10.3389/fpsyg.2016.00093>
- Ličen, M., Slapničar, S., Slana Ozimič, A., Hartmann, F. G., Anticevic, A., & Repovš, G. (2019). Social Incentives Shape Neural Circuits Underlying Cognitive Control. *Available at SSRN* 3448389.
- Logan, K., Bright, L. F., Grau, S. L., 2018. Unfriend me, please!: Social media fatigue and the theory of rational choice. *Journal of Marketing Theory and Practice*, 26(4), 357-367. <https://doi.org/10.1080/10696679.2018.1488219>
- Marcus, G. E., Valentino, N. A., Vasilopoulos, P., & Foucault, M. (2019). Applying the Theory of Affective Intelligence to Support for Authoritarian Policies and Parties. *Political Psychology*, 40(S1), 109–139. <https://doi.org/10.1111/pops.12571>
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 1-20. <https://doi.org/10.1186/s41235-020-00252-3>
- McGlynn, J., Baryshevtsev, M., & Dayton, Z. A. (2020). Misinformation more likely to use non-specific authority references: Twitter analysis of two COVID-19 myths. *Harvard Kennedy School Misinformation Review*, 1(3). <https://doi.org/10.37016/mr-2020-37>
- Mehta, R., Dahl, D. W., & Zhu, R. J. (2017). Social-Recognition versus Financial Incentives? Exploring the Effects of Creativity-Contingent External Rewards on Creative Performance. *Journal of Consumer Research*, 44(3), 536–553. <https://doi.org/10.1093/jcr/ucx062>

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Michael, R. B., & Breaux, B. O. (2021). The relationship between political affiliation and beliefs about sources of “fake news.” *Cognitive Research: Principles and Implications*, 6(1), 1-15.
<https://doi.org/10.1186/s41235-021-00278-1>
- Mitra, A. (2020., June). *Why Do WhatsApp Users End Up Spreading Misinformation In India?* Forbes. <https://www.forbes.com/sites/anandamitra/2020/06/29/why-do-whatsapp-users-end-up-spreading-misinformation-in-india/>
- Narayanan, V., Kollanyi, B., Hajela, R., Barthwal, A., Marchal, N., & Howard, P. N. (2019., May). News and information over Facebook and WhatsApp during the Indian election campaign.
<https://demtech.oii.ox.ac.uk/research/posts/news-and-information-over-facebook-and-whatsapp-during-the-indian-election-campaign/>
- Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American Political Science Review*, 115(3), 999-1015.
<https://doi.org/10.1017/S0003055421000290>
- Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter. *American Political Science Review*, 115(3), 999–1015.
<https://doi.org/10.1017/S0003055421000290>
- Osmundsen, M., Petersen, M. B., & Bor, A. (2021, May). *How partisan polarization drives the spread of fake news*. Brookings. <https://www.brookings.edu/techstream/how-partisan-polarization-drives-the-spread-of-fake-news>
- Parikh, S. B., & Atrey, P. K. (2018, April). Media-rich fake news detection: A survey. *In 2018 IEEE conference on multimedia information processing and retrieval (MIPR)* (pp. 436-441).
<https://doi.org/10.1109/MIPR.2018.00093>.

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39-50.
<https://doi.org/10.1016/j.cognition.2018.06.011>
- Peterson, E. J., & Iyengar, S. (2021). Partisan Gaps in Political Information and Information-Seeking Behavior: Motivated Reasoning or Cheerleading? *American Journal of Political Science*, 65(1), 133–147. <https://doi.org/10.1111/ajps.12535>
- Puthillam, A., Kapoor, H., & Karandikar, S. (2021, February 17). Beyond Left and Right: A Scale to Measure Political Ideology in India. <https://doi.org/10.31234/osf.io/fg387>
- Rademacher, L., Salama, A., Gründer, G., & Spreckelmeyer, K. N. (2014). Differential patterns of nucleus accumbens activation during anticipation of monetary and social reward in young and older adults. *Social Cognitive and Affective Neuroscience*, 9(6), 825-831.
<https://doi.org/10.1093/scan/nst047>
- Radke, S., Seidel, E. M., Eickhoff, S. B., Gur, R. C., Schneider, F., Habel, U., & Derntl, B. (2016). When opportunity meets motivation: Neural engagement during social approach is linked to high approach motivation. *NeuroImage*, 127, 267–276.
<https://doi.org/10.1016/j.neuroimage.2015.12.014>
- Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences*, 118(26), e2024292118.
<https://doi.org/10.1073/pnas.2024292118>
- Rathje, S., Roozenbeek, J., Pretus, C., & Van Der Linden, S. (2023). Accuracy and social motivations shape judgements of (mis)information. *Nature Human Behaviour*, 7, 892-903.
<https://doi.org/10.1038/s41562-023-01540-w>

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

- Ren, Z. (Bella), Dimant, E., & Schweitzer, M. E. (2022). Beyond belief: How social engagement motives influence the spread of conspiracy theories. *Journal of Experimental Social Psychology, 104*, 104421. <https://doi.org/10.1016/j.jesp.2022.104421>.
- Resende, G. G., Melo, P., Sousa, H., Messias, J., Vasconcelos, M., Almeida, J. M., & Benevenuto, F. (2019). (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. *The Web Conference*. <https://doi.org/10.1145/3308558.3313688>
- Rosenzweig, L.R, Bago, B., Berinsky, A. J., & Rand, D. G. (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, 2*(4).
<https://doi.org/10.37016/mr-2020-75>
- Ryu, J.S. (1982). Public affairs and sensationalism in local TV news programs. *Journalism & Mass Communication Quarterly, 59*, 137 - 74. <https://doi.org/10.1177/107769908205900111>
- Safieddine, F., Dordevic, M., & Pourghomi, P. (2017). Spread of misinformation online: Simulation impact of social media newsgroups. *2017 Computing Conference*, 899-906.
<https://doi.org/10.1109/SAI.2017.8252201>
- Scherer, L. D., McPhetres, J., Pennycook, G., Kempe, A., Allen, L. A., Knoepke, C. E., ... & Matlock, D. D. (2021). Who is susceptible to online health misinformation? A test of four psychosocial hypotheses. *Health Psychology, 40*(4), 274.
<https://doi.org/10.1037/hea0000978>
- Shoemaker, P. J., & Cohen, A. A. (2005). *News around the world: Content, practitioners, and the public*. Routledge. <https://doi.org/10.4324/9780203959091>
- Tangney, J.P., Stuewig, J. & Mashek, D.J. (2007), Moral emotions and moral behavior, *Annual Review of Psychology, 58*(1), 345–372.

- Van Bavel, J. J., Harris, E. A., Pärnamets, P., Rathje, S., Doell, K. C., & Tucker, J. A. (2021). Political Psychology in the Digital (mis)Information age: A Model of News Belief and Sharing. *Social Issues and Policy Review*, 15(1), 84–113. <https://doi.org/10.1111/sipr.12077>
- van der Linden, S., Panagopoulos, C., Azevedo, F., & Jost, J. T. (2021). The paranoid style in American politics revisited: An ideological asymmetry in conspiratorial thinking. *Political Psychology*, 42(1), 23–51. <https://doi.org/10.1111/pops.12681>
- Vasilopoulos, P. (2019). Affective Intelligence and Emotional Dynamics in Voters' Decision-Making Processes. *Oxford Research Encyclopedia of Politics*.
<https://doi.org/10.1093/acrefore/9780190228637.013.767>
- Wang, Y., Dai, Y., Li, H., & Song, L. (2021). Social Media and Attitude Change: Information Booming Promote or Resist Persuasion? *Frontiers in Psychology*, 12.
<https://doi.org/10.3389/fpsyg.2021.596071>
- Weeks, B. E. (2015). Emotions, Partisanship, and Misperceptions: How Anger and Anxiety Moderate the Effect of Partisan Bias on Susceptibility to Political Misinformation. *Journal of Communication*, 65(4), 699–719. <https://doi.org/10.1111/jcom.12164>
- Yang, A., Jieun, S., Kim, H. M., Zhou, A., Liu, W., Huang-Isherwood, K., Jang, E., Sun, J., Lee, E., Yafei, Z., & Chuqin, D. (2022). Who says what in which networks: What influences social media users' emotional reactions to the covid-19 vaccine infodemic? *Social Science Computer Review*, 0(0). <https://doi.org/10.1177/08944393221128940>
- Zou, Y. (2016). *Buying love through social media: How different types of incentives impact consumers' online sharing behavior*. [Unpublished doctoral dissertation]. Old Dominion University.

Tables and Figures

Table 1. Sample Descriptives for Financial and Social Incentive Conditions

<i>Variable</i>	<i>Financial</i>		<i>Social</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1. Gender	1.46	0.58	1.38	0.54
2. Age	26.4	6.26	31.08	91.27
3. Education	4.8	1.22	4.84	1.05
	4			
4. Religion	3.94	1.32	4.06	1.13
5. True posts shared	2.52	1.34	2.51	1.31
6. Plausible posts shared	2.52	1.36	2.4	1.36
7. False posts shared	1.1	1.4	1.25	1.46
8. Implausible posts shared	1.06	1.41	1.22	1.45
12. Wholesome posts shared	2.42	1.59	2.47	1.57
13. All true posts shared (true + plausible)	5.0	2.42	4.91	2.42
	4			
14. All false posts shared (false + implausible)	2.16	2.66	2.48	2.75
15. Happy reaction (true posts)	1.95	1.21	1.94	1.24
16. Happy reaction (false posts)	1.02	1.12	0.98	1.02
17. Happy reaction (plausible posts)	1.93	1.17	1.83	1.22
18. Happy reaction (implausible posts)	1.02	1.08	1.04	1.07
19. Happy reaction (wholesome posts)	2.44	1.4	2.5	1.38
20. Disgust reaction (true posts)	0.3	0.62	0.38	0.67
	4			
21. Disgust reaction (false posts)	0.61	0.83	0.75	0.89
22. Disgust reaction (plausible posts)	0.3	0.68	0.42	0.67
	9			
23. Disgust reaction (implausible posts)	0.6	0.79	0.7	0.92
	3			
24. Disgust reaction (wholesome posts)	0.14	0.4	0.13	0.38
25. Anger reaction (true posts)	0.5	0.85	0.53	0.8
	6			
26. Anger reaction (false posts)	0.9	1.25	0.9	1.21
	5			

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

27. Anger reaction (plausible posts)	0.5 4	0.76	0.53	0.75
28. Anger reaction (implausible posts)	0.9 3	1.23	0.93	1.25
29. Anger reaction (wholesome posts)	0.17	0.5	0.15	0.44
30. True posts 'read more'	1.1	1.53	0.87	1.42
31. False posts 'read more'	1.17	1.54	0.87	1.41
32. Plausible 'read more'	1.15	1.56	0.82	1.39

INCENTIVIZATION FOR SHARING TRUE CONTENT ONLINE

33. Implausible 'read more'	1.16	1.54	0.8 2	1.39
34. Wholesome 'read more'	1.15	1.55	0.8 6	1.42

Note. *M and SD are used to represent mean and standard deviation. Gender was coded 1 = Man, 2 = Woman, 3 = Non binary/Trans. Religion was coded 1 = Atheism, 2 = Buddhism, 3 = Christianity, 4 = Hinduism, 5 = Islam, 6 = Jainism, 7 = Sikhism, 8 = Zoroastrianism, 10 = Self describe, 11 = Prefer not to disclose. Education was linearly measured.*

Table 2. Sample Descriptives and correlations for demographics, PI scales, posts shared and read more (Experimental)

<i>Variable</i>	<i>M</i>	<i>SD</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>	<i>15</i>	<i>16</i>
1. Gender	1.42	0.56																
2. Age	28.81	65.61	.21**															
3. Education	4.84	1.13	-0.01	0.02														
4. Religion	4	1.23	0	0.01	.10**													
5. Political Ideology (PI)	120.94	34.85	-.26**	0.05	0.03	0.0												
6. PI Purity Subscale	49.63	16.23	-.21**	0.06	0	0.02	.94**											
7. PI Obedience Subscale	71.31	20.49	-.27**	0.04	0.06	.08*	.96**	.80**										
8. True posts shared	2.52	1.32	0.04	-0.02	-.08*	0	.22**	.28**	.14**									
9. False posts shared	1.18	1.44	-0.01	0.06	-0.05	0	0.06	.16**	-0.03	.37**								
10. Plausible posts shared	2.45	1.36	.07*	-0.02	-.08*	0.01	.19**	.27**	.12**	.62**	.35**							
11. Implausible posts shared	1.14	1.43	0.02	0.06	-.08*	0.05	.08*	.18**	-0.01	.41**	.78**	.40**						
12. Wholesome posts shared	2.44	1.58	.08*	0.02	-.10**	0.01	.12**	.21**	0.04	.60**	.44**	.60**	.48**					
13. True posts 'read more'	0.99	1.48	.10**	-0.03	-0.06	0	-.41**	-.35**	-.43**	-0.04	0.01	-.11**	-0.02	.09**				
14. False posts 'read more'	1.02	1.48	.10**	-0.03	-0.06	-0.01	-.39**	-.33**	-.41**	-0.06	0.05	-.10**	-0.02	.10**	.88**			
15. Plausible posts 'read more'	0.98	1.48	.13**	-0.03	-0.06	-0.03	-.41**	-.34**	-.43**	-.07*	0.03	-0.06	-0.02	.09**	.85**	.85**		
16. Implausible posts 'read more'	0.98	1.47	.11**	-0.03	-0.05	-0.03	-.42**	-.35**	-.43**	-.08*	0.01	-.10**	-0.03	.08*	.87**	.87**	.88**	
17. Wholesome posts 'read more'	1	1.49	.12**	-0.02	-0.04	-0.04	-.40**	-.34**	-.42**	-0.05	0	-.10**	-0.04	.12**	.89**	.86**	.87**	.88**

Note. *M and SD are used to represent mean and standard deviation. * indicates $p < .05$. ** indicates $p < .01$. Gender was coded 1 = Man, 2 = Woman, 3 = Non binary/Trans. Education was linearly measured*

Table 3. Sample Descriptives and correlations for demographics, ME scale, and post reactions (experimental)

Variable	M	SD	1	2	3	4	5	6	7	8	9	10
1. Gender	1.34	0.51										
2. Age	26.67	5.43	-0.05									
3. Education	4.98	0.93	0.07	.30**								
4. Religion	4.01	1.07	.10*	0.04	0.05							
5. Positive Moral Emotions	15.29	4.29	-0.05	0.03	0.07	-						
						0.02						
6. Negative Moral Emotions	23.4	11.76	-0.04	0.06	.09*	0.02	.65**					
7. Total Moral Emotions	38.7	14.9	-0.05	0.06	.09*	0.01	.80**	.98**				
8. Happy reaction (true posts)	2.09	1.18	0	0	0.03	0.06	.23**	.27**	.28**			
9. Angry reaction (true posts)	0.6	0.88	-0.01	-0.05	-0.07	-	.12**	.10**	.12**	-.32**		
						0.03						
10. Disgust reaction (true posts)	0.36	0.63	0.01	0.07	0.06	0.01	-.16**	-.11**	-.13**	-.19**	-.18**	
11. Happy reaction (false posts)	1.02	1.04	0.02	.09*	0.01	0.01	-.18**	-.23**	-.23**	.15**	0.01	.11**
12. Angry reaction (false posts)	1.1	1.34	0.05	-.13**	-0.02	-	.44**	.50**	.52**	.14**	.38**	-.23**
						0.05						
13. Disgust reaction (false posts)	0.8	0.9	-.11**	0.08	0.02	0.04	0.01	0.01	0.01	.09*	-.08*	.24**
14. Happy reaction (plausible posts)	1.97	1.17	0.01	0.04	-0.02	0.02	.28**	.32**	.34**	.36**	.10**	0.03
15. Angry reaction (plausible posts)	0.57	0.78	-0.01	-0.02	0	0.01	.19**	.18**	.19**	.08*	.32**	-0.07
16. Disgust reaction (plausible posts)	0.44	0.71	-.08*	0.05	0.01	0.03	-.12**	-.13**	-.14**	0.02	-.09*	.33**
17. Happy reaction (implausible posts)	1.06	1.06	-0.01	.11**	0.01	0.06	-.10*	-.20**	-.19**	.16**	0.02	.10*
18. Angry reaction (implausible posts)	1.08	1.34	0.04	-.10*	-0.03	-	.42**	.48**	.50**	.14**	.40**	-.20**
						0.01						

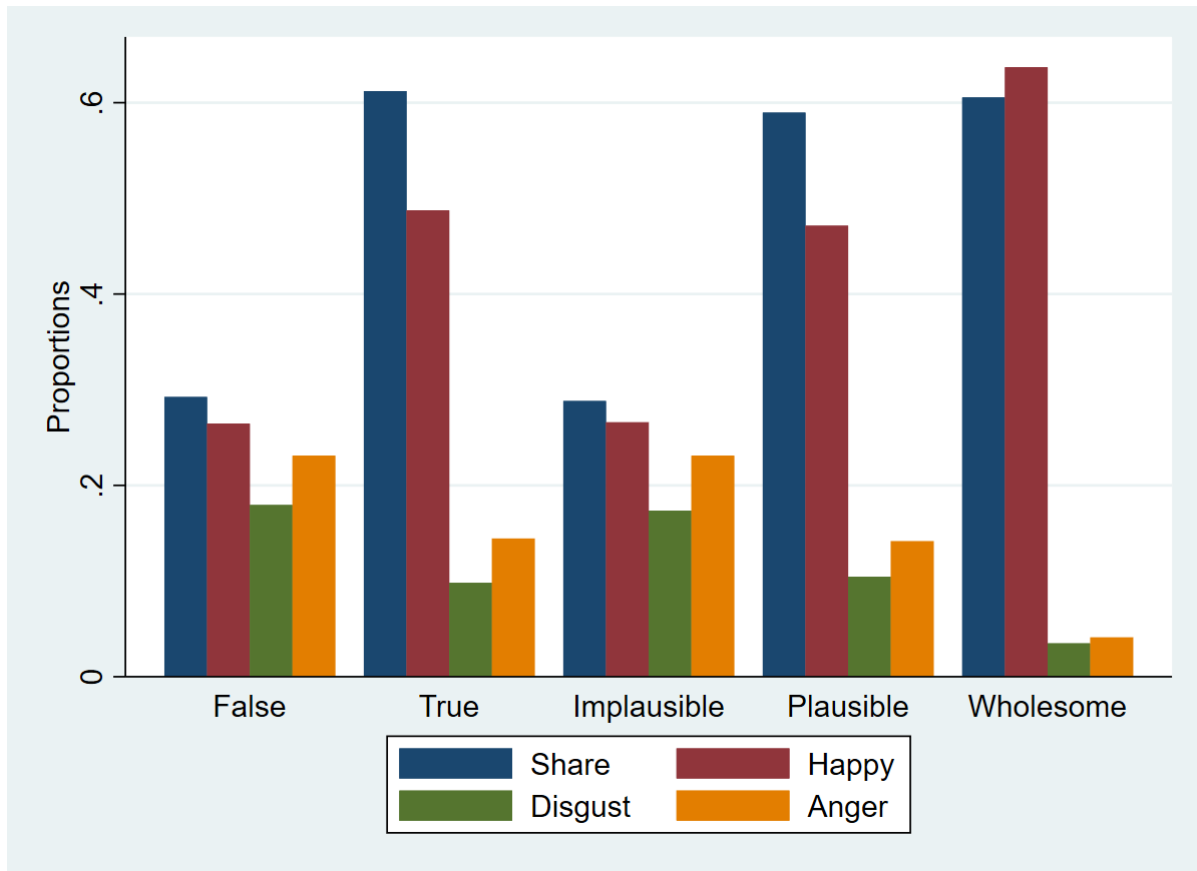
19. Disgust reaction (implausible posts)	0.77	0.92	-0.06	0.04	0.06	-	0	0.02	0.01	0.05	-.11**	.28**
						0.03						
20. Happy reaction (wholesome posts)	2.56	1.32	0	-0.07	-.11**	-	-0.06	-.16**	-.14**	.21**	0.04	0.06
						0.01						
21. Angry reaction (wholesome posts)	0.19	0.51	-0.01	0.08	0.01	0.02	0.03	0.07	0.06	-0.02	.32**	-0.03
22. Disgust reaction (wholesome posts)	0.15	0.41	-.09*	0.03	-0.01	-	-0.01	0	-0.01	-0.02	-0.02	.17**
						0.04						

Variable	11	12	13	14	15	16	17	18	19	20	21
12. Angry reaction (false posts)	-.38**										
13. Disgust reaction (false posts)	-.18**	-.32**									
14. Happy reaction (plausible posts)	.13**	.18**	.09*								
15. Angry reaction (plausible posts)	-0.02	.43**	-.10*	-.18**							
16. Disgust reaction (plausible posts)	.19**	-.20**	.18**	-.25**	-.15**						
17. Happy reaction (implausible posts)	.38**	-.27**	.16**	.23**	-0.04	0.05					
18. Angry reaction (implausible posts)	-.21**	.72**	-.14**	.19**	.45**	-.15**	-.39**				
19. Disgust reaction (implausible posts)	0.04	-.12**	.29**	0.03	-.12**	.30**	-.17**	-.24**			
20. Happy reaction (wholesome posts)	.23**	-0.02	0.03	.23**	0.03	.10**	.26**	0.02	0.03		
21. Angry reaction (wholesome posts)	.10*	.18**	-0.07	0.05	.29**	-0.03	.12**	.20**	-.13**	-.25**	
22. Disgust reaction (wholesome posts)	0.05	-.10*	.15**	0	-0.02	.16**	.09*	-.12**	.16**	-.22**	0.06

Note. *M* and *SD* are used to represent mean and standard deviation. * indicates $p < .05$. ** indicates $p < .01$. Gender was coded 1 = Man, 2 = Woman, 3 = Non binary/Trans. Education was linearly measured

Figure 1.

Proportions of shares and emoji reactions to different post content.



Appendix: Content Creation Methodology

The 250 posts were created using current or recent news after removing most, if not all, identifying elements to minimize any potential internet searches by participants during the experiment.

Consistent with Pennycook, Cannon & Rand (2018), these messages were divided into five broad categories: plausible, implausible, true, false, and wholesome. The plausible category included headlines that participants were more likely to be certain about (e.g. “There is a low chance of mobile phone use leading to cancer”) with the reverse being true for the implausible category (e.g. “There is an 87% chance that Indian women report a better quality of life than men”). True and false categories respectively contained the same information as headlines in the plausible and implausible categories but without the probability or likelihood phrasing. As a result, the phrasing of the true and false categories was more like a traditional news headline. For the aforementioned examples, the phrasing of the true headline was “Using mobile phones is not associated with an increased risk of cancer” and the false headline was phrased as “Indian women report a better quality of life than men.” The rationale for using the same basic source news but phrasing it differently was to assess if changing the certainty phrasing affects the spread of misinformation. The wholesome category was used as a control and its headlines were mainly employed to test whether participants in the social condition were more likely to share wholesome content compared to those in the financial condition. All the messages were randomized and counterbalanced, with each participant receiving all five types of messages across the three days of the experiment.

[Tattle's Khoj](#) database was used for viewing words in IFCN-certified fact-checking articles in a particular week. After considering the overall word frequency and the number of clusters a specific word belonged to, some keywords were chosen to help build an overarching theme (e.g., the

words “COVID-19”, “myths”, and “vaccines” led to the creation of the theme “health”). The [Khoj](#) database was also supplemented by manually tracking fact-checking websites like [WebQoof](#), [Alt News](#), [The Logical Indian](#), and [Boom](#), and keeping up-to-date with various online news media outlets for over a month. A breakdown of the subthemes and phrasing of headlines is provided in Tables A1-A3.

Table A1: Themes, subthemes, and headlines in each category

Themes	Subthemes	Number of Headlines
Health	COVID-19	12
	Vaccines	8
	Myths	8
	Risks	4
	NeoCov	4
Science/Tech	Aerospace	4
	Agriculture	4
	Energy	4
Tech	5G	6
	Non-gaming app	2
	Space	4
	Crypto	4
	VR	4
	Health & tech risk	4
Science	Medical discovery	4
	Planets	4
	Climate	4
	Material science	4
	Health-related advances	4

	Blackhole	4
--	-----------	---

Political	Prison	4
	Finance	16
	Flag	4
	Education	4
	Election	8
	International Relations	4
	Protests	4
	War	4
Religiopolitical	Muslims	8
Religion	Antinational	2
	Fatwa	2
	Anti-Hinduism	4
	Scripture	4
Miscellaneous	Environment	4
	Entertainment	8
	Sports	4
	Quality of life	4
	Online Dating	4
	Wildlife	4
	Plastic	4
	Toothpaste	4
TOTAL		200

Table A2: Types and frequencies of phrasing used for plausible messages

	Plausible				
Type of phrasing	probability	likely/unlikely	very likely/very unlikely	high possibility/low possibility	high chance /low chance
Number of headlines	10	12/2	5/1	10/0	10/1

Table A3: Types and frequencies of phrasing used for implausible messages

	Implausible				
Type of phrasing	probability	likely/unlikely	very likely/very unlikely	high possibility/low possibility	high chance /low chance
Number of headlines	9	12/3	5/1	9/1	9/1

Supplementary Tables

Table S1. Descriptives for Post-Task Questions

Post-Task Question	<i>M</i>	<i>SD</i>
Engagement		
How many times did you minimize the tab over the study?	3.19	2.15
How often did you contemplate the message?	5.73	1.45
How often did you lose track of time?	4.09	2.13
How emotionally engaged were you with the task?	5.7	1.49
How often did you think about the consequences to you if the message were true?	5.6	1.53
How often did you think about the consequences to your friends and family if the message were true?	5.51	1.57
How often did you think about the consequences to others in the country, if the message were true?	5.55	1.53
How often did you think about the consequences to others in the world, if the message were true?	5.64	1.46
How often did you think about the consequences to the people you dislike/hate if the message were true?	5.22	1.77
To what extent did you enjoy participating in the tasks?	6.28	1.09
To what extent did you like the messages shown to you?	5.77	1.32
To what extent were you familiar with the messages shown to you?	5.56	1.4
Would you want to participate in such a study again?	6.54	0.93
Social Condition		
To what extent did you understand how you gained likes?	5.92	1.28
To what extent did you understand how you lost likes?	5.68	1.5
To what extent were you thinking about likes while sharing the messages?	5.55	1.59
To what extent were you paying attention to your dashboard while sharing the messages?	5.8	1.42
Financial Condition		
To what extent did you understand how you gained money?	5.39	1.76
To what extent did you understand how you lost money?	5.15	1.94
To what extent were you thinking about money while sharing the messages?	4.86	2.05
To what extent were you paying attention to your dashboard while sharing the messages?	5.5	1.72

Note: Responses to engagement-related post task questions ranged from 1 = 'Not at all' to 7 = 'All the time'. Responses to incentive-related (social and financial) post task questions ranged from 1 = 'Not at all' to 7 = 'Perfectly well'

Table S2. *T* tests to Check Demographic Distribution Across Financial and Social Condition

<i>Variable</i>	<i>Social</i>	<i>Financia</i>	<i>t-stat</i>	<i>p-value</i>
	<i>M</i>	<i>SD</i>		
Age	31.14	26.4	1.115	0.265
Education	4.86	4.84	0.256	0.798
householdsize	4.4	4.37	0.39	0.697
	<i>M</i>	<i>SD</i>	<i>z-stat</i>	<i>p-value</i>
gender==Man	0.643	0.576	2.056	0.04
gender==Woman	0.344	0.399	-1.697	0.09
gender==Non-binary	0.013	0.025	-1.345	0.179
Unmarried	0.639	0.63	0.285	0.776
Unemployed	0.047	0.066	-1.212	0.225
Employed	0.659	0.58	2.429	0.015
Student	0.234	0.299	-	0.026
			2.229	
Homemaker	0.058	0.05	0.536	0.592
religion==Atheism	0.037	0.078	-	0.008
			2.666	
religion==Buddhism	0.024	0.021	0.321	0.748
Christianity	0.035	0.032	0.209	0.834
religion==Hinduism	0.798	0.79	0.277	0.782
religion==Islam	0.074	0.046	1.746	0.081
religion==Sikhism	0.017	0.014	0.429	0.668
religion==Jainism	0.015	0.018	-	0.708
			0.374	
Upper caste	0.79	0.773	0.599	0.549
Metropolitan	0.337	0.32	0.527	0.598
Urban	0.399	0.4	-	0.979
			0.026	
Semi-Urban	0.208	0.211	-0.119	0.906
Rural	0.056	0.068	-	0.439
			0.774	

Note: Caste was collapsed into upper and lower. Marital status was also collapsed into married and unmarried. M and SD are used to represent mean and standard deviation.

Table S3. Sample Descriptives and correlations for demographics, PI scales, posts shared, and read more (baseline)

<i>Variable</i>	<i>M</i>	<i>SD</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>	<i>15</i>	<i>16</i>
1. Gender	1.42	0.56																
2. Age	28.81	65.61	.21**															
3. Education	4.84	1.13	-0.01	0.02														
4. Religion	4	1.23	0	0.01	.10**													
5. Political Ideology (PI)	120.94	34.85	-.26**	0.05	0.03	0.06												
6. PI Purity Subscale	49.63	16.23	-.21**	0.06	0	0.02	.94**											
7. PI Obedience Subscale	71.31	20.4	-.27**	0.04	0.06	.08*	.96**	.80**										
8. True posts shared	0.55	0.5	0.01	0.03	-.07*	0.06	.09**	.13**	0.04									
9. False posts shared	0.29	0.45	0	0.05	-0.06	0.06	-0.01	0.06	-.07*	.26**								
10. Plausible posts shared	0.49	0.5	-0.05	0.03	-0.04	0.03	.17**	.19**	.13**	.25**	.19**							
11. Implausible posts shared	0.3	0.46	-0.02	0.05	-.11**	0	0.03	.09**	-0.03	.21**	.42**	.21**						
12. Wholesome posts shared	0.59	0.49	-0.03	0.02	-.11**	0.01	0.06	.11**	0.01	.26**	.27**	.22**	.29**					
13. True posts 'read more'	0.33	0.47	.11**	-	-.11**	-.07*	-.35**	-.31**	-.36**	0.01	0.06	-.08*	0.04	.13**				
14. False posts 'read more'	0.32	0.47	.09**	-	-0.05	-	-.31**	-.26**	-.32**	-0.01	0.05	-	0.03	.13**	.69**			
15. Plausible posts 'read more'	0.31	0.46	.09**	-	-.08*	-	-.34**	-.29**	-.34**	-0.02	0.02	-0.04	0.06	.12**	.68**	.66**		
16. Implausible posts 'read more'	0.32	0.47	.08*	-	-0.03	-	-.36**	-.31**	-.36**	0.01	0.06	-0.05	.10**	.15**	.68**	.72**	.69**	

17. Wholesome posts 'read more'	0.34	0.47	.10**	0.04	-0.02	-	-.36**	-.31**	-.37**	-	0.05	-.07*	0.05	.17**	.69**	.69**	.68**	.70**
						0.06					0.05							

*Note. M and SD are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.*

Table S4. Sample Descriptives and correlations for demographics, ME scale, and post reactions (baseline)

Variable	M	SD	1	2	3	4	5	6	7	8	9	10
1. Gender	1.34	0.51										
2. Age	26.67	5.43	-0.05									
3. Education	4.98	0.93	0.07	.30**								
4. Religion	4.01	1.07	.10*	0.04	0.05							
5. Positive Moral Emotions	15.29	4.29	-0.05	0.03	0.07	-						
						0.02						
6. Negative Moral Emotions	23.4	11.76	-0.04	0.06	.09*	0.02	.65**					
7. Total Moral Emotions	38.7	14.9	-0.05	0.06	.09*	0.01	.80**	.98**				
8. Happy reaction (true posts)	0.51	0.5	0.01	0.03	0.01	0.06	0.07	0.07	0.08			
9. Angry reaction (true posts)	0.19	0.39	0	0.01	-0.02	-	0.07	0.05	0.05	-.49**		
						0.02						
10. Disgust reaction (true posts)	0.13	0.34	-0.05	-0.04	0.02	0.01	0	-0.02	-0.02	-	-.19**	
										.40**		
11. Happy reaction (false posts)	0.32	0.47	0.06	0.03	-0.03	0.02	-0.06	-.10*	-.10*	0	0.01	0.02
12. Angry reaction (false posts)	1.1	1.34	0.05	-.13**	-0.02	-	.44**	.50**	.52**	0.02	.25**	-.10**
						0.05						
13. Disgust reaction (false posts)	0.24	0.43	-0.06	-0.03	0.03	-	0.06	0.01	0.03	-0.02	0.02	.09*
						0.04						
14. Happy reaction (plausible posts)	0.49	0.5	-0.06	0.02	-0.04	0.03	0.04	0.05	0.05	.10*	-0.02	-0.02
15. Angry reaction (plausible posts)	0.19	0.39	0.02	-0.01	0.04	-	.18**	.13**	.15**	0.02	.17**	-.09*
						0.02						
16. Disgust reaction (plausible posts)	0.12	0.33	-0.01	0.03	-0.01	-	-0.03	-0.07	-0.07	-0.03	-0.03	.12**
						0.01						
17. Happy reaction (implausible posts)	0.3	0.46	-0.04	0.04	0.01	0.03	-0.04	-.13**	-.11**	0.06	0	0.07

18. Angry reaction (implausible posts)	0.26	0.44	0	-0.04	-0.03	-	.31**	.28**	.31**	0.01	.12**	-0.07
						0.04						
19. Disgust reaction (implausible posts)	0.22	0.42	0.03	-0.03	0.01	0.02	0.01	0.02	0.02	0	-0.03	.09*
20. Happy reaction (wholesome posts)	0.73	0.45	-0.03	-0.07	-0.06	0.01	-0.05	-0.07	-0.07	0.01	.09*	-0.02
21. Angry reaction (wholesome posts)	0.06	0.24	0.01	0	-0.01	0	.16**	.11**	.13**	0.06	0.03	-0.04
22. Disgust reaction (wholesome posts)	0.04	0.19	-0.01	0.01	0	0.02	-0.01	-0.03	-0.03	0.02	-0.05	0.02

<i>Variable</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>	<i>15</i>	<i>16</i>	<i>17</i>	<i>18</i>	<i>19</i>	<i>20</i>	<i>21</i>
12. Angry reaction (false posts)	-.09*										
13. Disgust reaction (false posts)	-.38**	-0.02									
14. Happy reaction (plausible posts)	.09*	0.05	-0.07								
15. Angry reaction (plausible posts)	-0.04	.24**	0.04	-.47**							
16. Disgust reaction (plausible posts)	-0.02	-.12**	.11**	-.36**	-.18**						
17. Happy reaction (implausible posts)	.09*	-.09*	0.05	.10**	0.04	0					
18. Angry reaction (implausible posts)	-0.06	.41**	0.02	-0.04	.18**	-0.02	-.38**				
19. Disgust reaction (implausible posts)	0	-0.05	0.05	-0.01	-0.01	.09*	-.35**	-.31**			
20. Happy reaction (wholesome posts)	0.05	0.04	0.02	0.05	0.02	0	0.01	-0.03	-		
									0.02		
21. Angry reaction (wholesome posts)	-0.04	.17**	-	0.01	0.07	-0.01	-0.02	.19**	-	-.41**	
			0.06						0.05		
22. Disgust reaction (wholesome posts)	0.05	-0.06	0.03	0.01	-0.01	0.03	.09*	-0.04	-	-.32**	-
									0.02		0.05

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

Table S5. Sample Descriptives for Financial and Social incentive condition (Baseline)

<i>Variable</i>	<i>Financia</i>		<i>Socia</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1. Gender	1.46	0.58	1.46	0.58
2. Age	26.4	6.26	26.4	6.26
3. Education	4.84	1.22	4.84	1.22
4. Religion	3.94	1.32	3.94	1.32
5. True posts shared	0.54	0.5	0.54	0.5
6. Plausible posts shared	0.51	0.5	0.51	0.5
7. False posts shared	0.29	0.45	0.29	0.45
8. Implausible posts shared	0.31	0.46	0.31	0.46
12. Wholesome posts shared	0.58	0.49	0.58	0.49
13. All true posts shared (true + plausible)	1.05	0.78	1.05	0.78
14. All false posts shared (false + implausible)	0.6	0.76	0.6	0.76
15. Happy reaction (true posts)	0.47	0.5	0.47	0.5
16. Happy reaction (false posts)	0.3	0.46	0.3	0.46
17. Happy reaction (plausible posts)	0.48	0.5	0.48	0.5
18. Happy reaction (implausible posts)	0.29	0.45	0.29	0.45
19. Happy reaction (wholesome posts)	0.71	0.46	0.71	0.46
20. Disgust reaction (true posts)	0.12	0.33	0.12	0.33
21. Disgust reaction (false posts)	0.22	0.41	0.22	0.41
22. Disgust reaction (plausible posts)	0.09	0.28	0.09	0.28
23. Disgust reaction (implausible posts)	0.17	0.38	0.17	0.38
24. Disgust reaction (wholesome posts)	0.04	0.19	0.04	0.19
25. Anger reaction (true posts)	0.2	0.4	0.2	0.4
26. Anger reaction (false posts)	0.24	0.43	0.24	0.43

27. Anger reaction (plausible posts)	0.18	0.39	0.18	0.39
28. Anger reaction (implausible posts)	0.24	0.43	0.24	0.43
29. Anger reaction (wholesome posts)	0.04	0.2	0.04	0.2
30. True posts 'read more'	0.36	0.48	0.36	0.48
31. False posts 'read more'	0.34	0.48	0.34	0.48
32. Plausible 'read more'	0.35	0.48	0.35	0.48
33. Implausible 'read more'	0.35	0.48	0.35	0.48
34. Wholesome 'read more'	0.36	0.48	0.36	0.48

Note. M and SD are used to represent mean and standard deviation.

Table S6. Factorial Analysis of Variance for posts shared

<i>Source</i>	<i>df</i>	<i>Partial SS</i>	<i>MS</i>	<i>F</i>
Model	18	823	45.72	214.89**
Gender	2	6.59	3.3	15.49**
Age	1	19.77	19.77	92.92**
Education	1	5.65	5.65	26.57**
Religion	6	21.72	3.62	17.01**
Pi_purity	1	210.55	210.55	989.53**
Pi_obedience	1	67.6	67.6	317.7**
Posts 'read more'	1	13.02	13.02	61.17**
Incentive	1	0.03	0.03	0.14
Type of post	4	522.68	130.67	614.13**
Total	22,29 9	5563.79	0.25	

*Note: df indicates degrees of freedom, Partial SS indicates partial sum of squares, MS indicates mean squares. * indicates $p < .05$. ** indicates $p < .01$.*

Table S7. Factorial Analysis of Variance for post reaction (happy)

<i>Source</i>	<i>df</i>	<i>Partial SS</i>	<i>MS</i>	<i>F</i>
Model	18	508.47	28.25	127.38**
Gender	2	1.91	0.95	4.3*
Age	1	0.002	0.002	0.01
Education	1	0.0001	0.000 1	0
Religion	6	10.18	1.7	7.65**
Pi_purity	1	0.02	0.02	0.0 8
Pi_obedience	1	10.62	10.62	47.89**
Posts 'read more'	1	12.78	12.78	57.63**
Incentive	1	0.04	0.04	0.17
Type of post	4	451.28	112.82	508.73
Total	22,29	5449.64	0.24	

9

*Note: df indicates degrees of freedom, Partial SS indicates partial sum of squares, MS indicates mean squares. * indicates $p < .05$. ** indicates $p < .01$.*

Table S8. Factorial Analysis of Variance for post reaction (disgust)

<i>Source</i>	<i>df</i>	<i>Partial SS</i>	<i>MS</i>	<i>F</i>
Model	18	85.48	4.75	47.05**
Gender	2	1.11	0.55	5.48**
Age	1	0.06	0.06	0.5 5
Education	1	0.35	0.35	3.51
Religion	6	1.18	0.2	1.95
Pi_purity	1	6.11	6.11	60.57**
Pi_obedience	1	9.2	9.2	91.19**
Posts 'read more'	1	0.64	0.64	6.32*
Incentive	1	0.85	0.85	8.46**
Type of post	4	64.64	16.16	160.11**
Total	22,29 9	2334.33	0.1	

Note: *df* indicates degrees of freedom, *Partial SS* indicates partial sum of squares, *MS* indicates mean squares. * indicates $p < .05$. ** indicates $p < .01$.

Table S9. Factorial Analysis of Variance for post reaction (anger)

<i>Source</i>	<i>df</i>	<i>Partial SS</i>	<i>MS</i>	<i>F</i>
Model	18	212.7	11.82	95.85**
Gender	2	0.53	0.27	2.17
Age	1	4.07	4.07	33.01**
Education	1	0.42	0.42	3.42
Religion	6	7.25	1.21	9.81**
Pi_purity	1	27.3	27.3	221.41**
Pi_obedience	1	1.03	1.03	8.39**
Posts 'read more'	1	2.49	2.49	20.22**
Incentive	1	0.95	0.95	7.74**
Type of post	4	111.26	27.82	225.63**
Total	22,29 9	2959.58	0.12	

*Note: df indicates degrees of freedom, Partial SS indicates partial sum of squares, MS indicates mean squares. * indicates $p < .05$. ** indicates $p < .01$.*