

# Reducing WhatsApp Usage to Mitigate Misinformation Exposure During Elections: Evidence from a Multi-Country Experiment

Rajeshwari Majumdar

Tiago Ventura

Shelley Liu

Carolina Torreblanca

Joshua A. Tucker

APSA Annual Meeting  
September 6, 2024

# Motivation

*Disinformation Spreads on WhatsApp  
Ahead of Brazilian Election*

[Give this article](#)



**Zimbabwe election disinformation spreads on WhatsApp**



Tendai DUBE

Thu, August 3, 2023 at 2:41 AM EDT · 4 min read



**WhatsApp — India's leading university**

Entire revolutions can be plotted and unraveled in a WhatsApp bubble



MANDIP DODI



COMMENTS



SHARE



...

READ LATER



# Research on Social Media and Misinformation

- Existing research: focused primarily on feed-based platforms that are popular in advanced democracies (Twitter, Facebook)
- But in the Global South
  - the types of social media platforms that are predominantly used are fundamentally different (message-based; WhatsApp)
  - the harmful offline consequences of misinformation exposure may be more pronounced

# Our Goal

Identify the causal effects of WhatsApp usage on exposure to misinformation and downstream effects on users' political attitudes

# Approach

WhatsApp “deactivation” experiment in India and South Africa during the month leading up to their most recent general elections

# Classical Deactivation Studies

## The Welfare Effects of Social Media<sup>†</sup>

By HUNT ALLCOTT, LUCA BRAGHIERI, SARAH EICHMEYER,  
AND MATTHEW GENTZKOW\*

*The rise of social media has provoked both optimism about potential societal benefits and concern about harms such as addiction, depression, and political polarization. In a randomized experiment, we find that deactivating Facebook for the four weeks before the 2018 US midterm election (i) reduced online activity, while increasing offline activities such as watching TV alone and socializing with family and friends; (ii) reduced both factual news knowledge and political polarization; (iii) increased subjective well-being; and (iv) caused a large persistent reduction in post-experiment Facebook use. Deactivation reduced post-experiment valuations of Facebook, suggesting that traditional metrics may overstate consumer surplus.* (JEL D12, D72, D90, I31, L82, L86, Z13)

## Testing the effects of Facebook usage in an ethnically polarized setting

Nejla Asimovic<sup>a,b,1</sup>, Jonathan Nagler<sup>a,b</sup>, Richard Bonneau<sup>b,c,d</sup>, and Joshua A. Tucker<sup>a,b,e</sup> 

<sup>a</sup>Department of Politics, New York University, New York, NY 10012; <sup>b</sup>Center for Social Media and Politics, New York University, New York, NY 10012;

<sup>c</sup>Department of Biology, New York University, New York, NY 10012; <sup>d</sup>Courant Institute of Mathematical Sciences, New York University, New York, NY 10012;

<sup>e</sup>Department of Russian and Slavic Studies, New York University, New York, NY 10012

Edited by Christopher Andrew Ball, Duke University, Durham, NC, and accepted by Editorial Board Member Mark Granovetter May 10, 2021 (received for review November 2, 2020)

It is commonly believed that social media is altering intergroup dynamics—bringing people closer or further apart from them from one another—the impact of social media on interethnic attitudes has yet to be rigorously evaluated, especially within areas with tenuous interethnic relations. We report results from a randomized controlled trial in Bosnia and Herzegovina (BiH), exploring the effects of deactivating social media during the period of political remembrance in July 2019 on a set of interethnic attitudes of Facebook users. We find evidence that, counter to preregistered expectations, people who deactivated their Facebook profiles report lower regard for ethnic outgroups than those who remained active. Moreover, we present additional evidence suggesting that this effect is little conditioned on the level of ethnic heterogeneity of respondents' residence. We also extend the analysis to include measures of subjective well-being and knowledge of news. Hence, we find that Facebook deactivation leads to suggestive improvements in subjective well-being and a decrease in knowledge of news. These findings are relevant to recent research in the United States in a very different context, thus increasing our confidence in the generalizability of these effects.

spaces can help or hinder the goal of transforming relations. From this perspective, social media should be considered and studied as one of the forces influencing the direction of countries' postconflict paths.

Importantly, both ethnicity and partisanship are potential markers of one's identity. Given that partisanship is acquired through socialization, it is therefore a measure of one's worldview than group membership based on more immutable characteristics such as ethnicity or race (10). As such, some propose that people may assign larger blame and responsibility to others for their partisanship than for their inborn group affiliation (11). The difference between partisanship and ethnicity, however, has rarely been discussed as a potential explanation of prejudice reduction (12–14), may be less effective at mitigating outgroup hostility in the context of partisanship (15). This, however, remains an open question requiring studies that go beyond partisan affiliation, in particular discussing attitudes toward ethnic outgroups and especially in an online context.

# This Study

- **Problem:** Fully deactivating WhatsApp is neither possible nor ideal

# This Study

- **Problem:** Fully deactivating WhatsApp is neither possible nor ideal
- **Our proposed workarounds:**
  - 1 Impose a time limit on WhatsApp usage
  - 2 Cut the primary channels through which users are exposed to misinformation and polarizing content: videos, images & audio

# This Study

- **Problem:** Fully deactivating WhatsApp is neither possible nor ideal
- **Our proposed workarounds:**
  - 1 Impose a time limit on WhatsApp usage
  - 2 Cut the primary channels through which users are exposed to misinformation and polarizing content: videos, images & audio
- **Modified deactivation experiment:**
  - 1 Incentivize participants to either limit their daily WhatsApp usage or refrain from looking at multimedia content

# This Study

- **Problem:** Fully deactivating WhatsApp is neither possible nor ideal
- **Our proposed workarounds:**
  - 1 Impose a time limit on WhatsApp usage
  - 2 Cut the primary channels through which users are exposed to misinformation and polarizing content: videos, images & audio
- **Modified deactivation experiment:**
  - 1 Incentivize participants to either limit their daily WhatsApp usage or refrain from looking at multimedia content
  - 2 Test whether these restrictions reduce exposure to election (mis)information

# This Study

- **Problem:** Fully deactivating WhatsApp is neither possible nor ideal
- **Our proposed workarounds:**
  - 1 Impose a time limit on WhatsApp usage
  - 2 Cut the primary channels through which users are exposed to misinformation and polarizing content: videos, images & audio
- **Modified deactivation experiment:**
  - 1 Incentivize participants to either limit their daily WhatsApp usage or refrain from looking at multimedia content
  - 2 Test whether these restrictions reduce exposure to election (mis)information
  - 3 Test whether reduced exposure leads to changes in political attitudes

# Hypotheses

Compared to those who use WhatsApp without any restrictions, those who are randomly assigned to reduce WhatsApp usage will:

# Hypotheses

Compared to those who use WhatsApp without any restrictions, those who are randomly assigned to reduce WhatsApp usage will:

- given changes in consumption habits,

H1a: report lower levels of exposure to false information

H1b: report lower levels of exposure to true information

# Hypotheses

Compared to those who use WhatsApp without any restrictions, those who are randomly assigned to reduce WhatsApp usage will:

- given changes in consumption habits,

H1a: report lower levels of exposure to false information

H1b: report lower levels of exposure to true information

- considering theories of repeated exposure on accuracy perceptions,

H2a: be more likely to accurately identify false information as false

H2b: be less likely to accurately identify true information as true

# Hypotheses

Compared to those who use WhatsApp without any restrictions, those who are randomly assigned to reduce WhatsApp usage will:

- given changes in consumption habits,

H1a: report lower levels of exposure to false information

H1b: report lower levels of exposure to true information

- considering theories of repeated exposure on accuracy perceptions,

H2a: be more likely to accurately identify false information as false

H2b: be less likely to accurately identify true information as true

- given likely reduction in exposure to misinformation and hate speech,

H3a: exhibit lower levels of partisan polarization

H3b: exhibit lower levels of ethnic/racial prejudice

# Experimental Design

# Experimental Conditions

## Types of partial deactivation

- Time: Limit daily WhatsApp usage to 10 minutes
- Media: Do not consume multimedia content received via WhatsApp

# Experimental Conditions

## Types of partial deactivation

- Time: Limit daily WhatsApp usage to 10 minutes
- Media: Do not consume multimedia content received via WhatsApp

## Treatment and control

- Treatment: Change settings *and* behavior for **4 weeks** (one month) leading up to election
- Control: Change only behavior for **3 days** a month before election

# Experimental Conditions

## Types of partial deactivation

- Time: Limit daily WhatsApp usage to 10 minutes
- Media: Do not consume multimedia content received via WhatsApp

## Treatment and control

- Treatment: Change settings *and* behavior for **4 weeks** (one month) leading up to election
- Control: Change only behavior for **3 days** a month before election

→ **2 × 2 design**

Time Treatment	Time Control
Media Treatment	Media Control

# Treatment 1: Time Reduction

## Task

- Treatment: For 4 weeks, limit daily usage to 10 minutes per day and set a daily app timer of 10 minutes
- Control: For 3 days, limit daily usage to 10 minutes

# Treatment 1: Time Reduction

## Task

- Treatment: For 4 weeks, limit daily usage to 10 minutes per day and set a daily app timer of 10 minutes

## Checking compliance



## Treatment 2: Multimedia Deactivation

### Task

- Treatment: For 4 weeks, do not consume multimedia and turn off the automatic download of multimedia
- Control: For 3 days, do not consume multimedia

# Treatment 2: Multimedia Deactivation

## Task

- Treatment: For 4 weeks, do not consume multimedia and turn off the automatic download of multimedia

## Automatic downloads

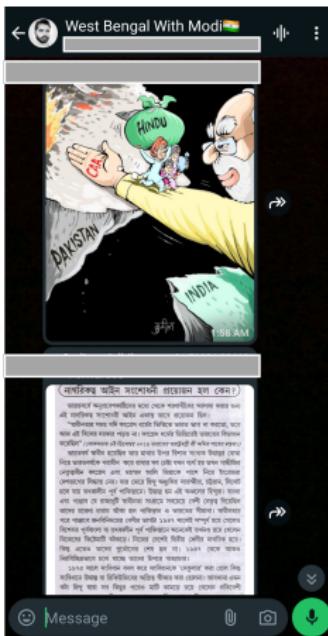


# Treatment 2: Multimedia Deactivation

## Task

- Treatment: For 4 weeks, do not consume multimedia and turn off the automatic download of multimedia

## Automatic downloads

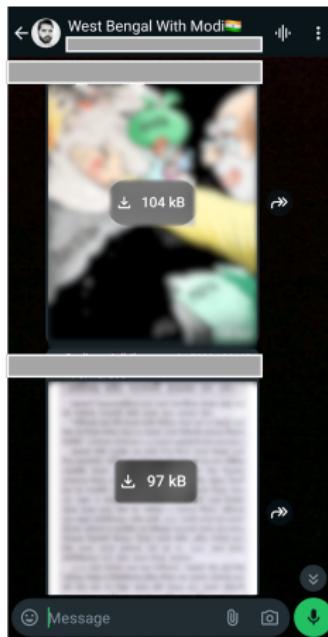


# Treatment 2: Multimedia Deactivation

## Task

- Treatment: For 4 weeks, do not consume multimedia and turn off the automatic download of multimedia

## Automatic downloads



# Treatment 2: Multimedia Deactivation

## Task

- Treatment: For 4 weeks, do not consume multimedia and turn off the automatic download of multimedia

## Checking compliance

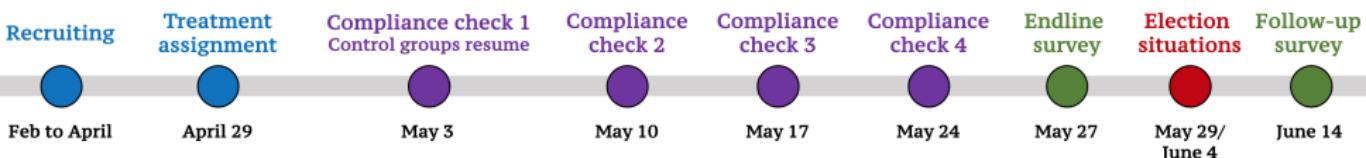


# Implementation

- Used Meta Advertisements to recruit WhatsApp users in India and South Africa
- All participants could earn a total of \$15 for completing surveys and uploading weekly screenshots
- Participants in treatment conditions who successfully reduced usage or multimedia consumption received an additional \$45

# Implementation

- Used Meta Advertisements to recruit WhatsApp users in India and South Africa
- All participants could earn a total of \$15 for completing surveys and uploading weekly screenshots
- Participants in treatment conditions who successfully reduced usage or multimedia consumption received an additional \$45
- Timeline:



# Implementation

- Used Meta Advertisements to recruit WhatsApp users in India and South Africa
- All participants could earn a total of \$15 for completing surveys and uploading weekly screenshots
- Participants in treatment conditions who successfully reduced usage or multimedia consumption received an additional \$45
- Timeline:



- Started with 1,498 respondents across the two countries; 1,394 (93%) completed the endline survey

	Treatment	Control
Time	91%	94%
Media	93%	94%

# Results

# Exposure to and Belief in News Headlines

- Participants are shown 8 news headlines; asked (a) whether they have seen this news before & (b) whether they think the news is true

# Exposure to and Belief in News Headlines

- Participants are shown 8 news headlines; asked (a) whether they have seen this news before & (b) whether they think the news is true
- Four headlines are true news; four are misinformation stories

# Exposure to and Belief in News Headlines

- Participants are shown 8 news headlines; asked (a) whether they have seen this news before & (b) whether they think the news is true
- Four headlines are true news; four are misinformation stories
- Count how many (a) recalled & (b) accurately identified as T/F

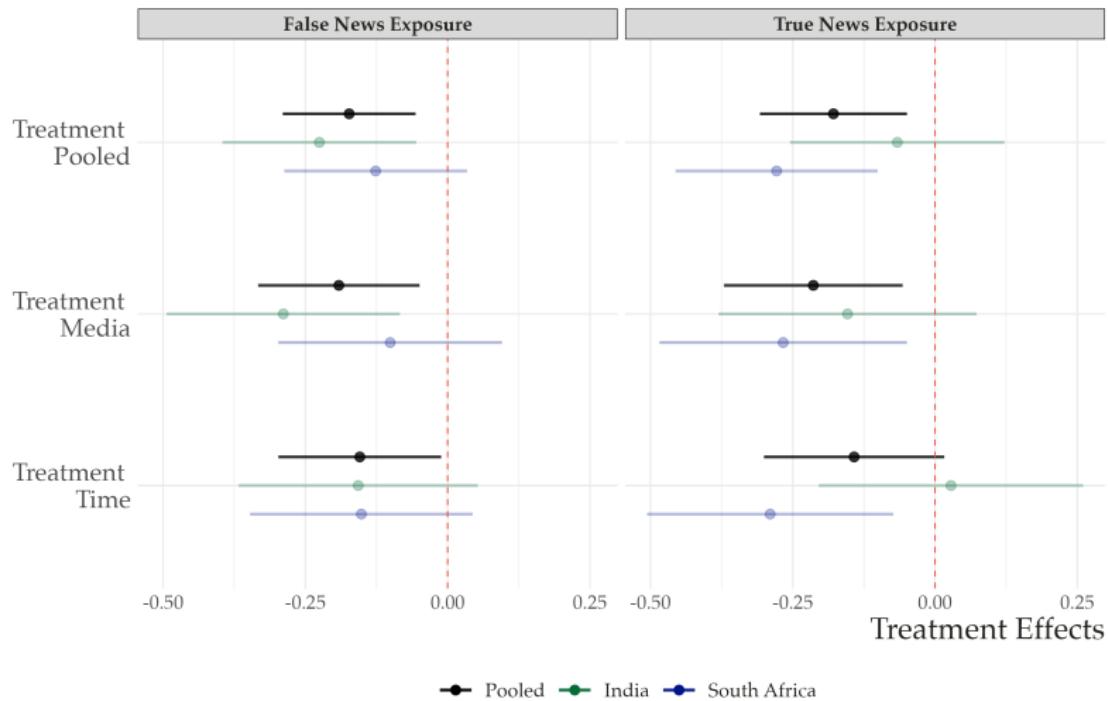
# Exposure to and Belief in News Headlines

- Participants are shown 8 news headlines; asked (a) whether they have seen this news before & (b) whether they think the news is true
- Four headlines are true news; four are misinformation stories
- Count how many (a) recalled & (b) accurately identified as T/F
- Means (sd) across the control conditions:

	True headlines	False headlines
Prior exposure	2.13 (1.30)	1.25 (1.18)
Think it is true	2.39 (1.18)	1.28 (1.15)

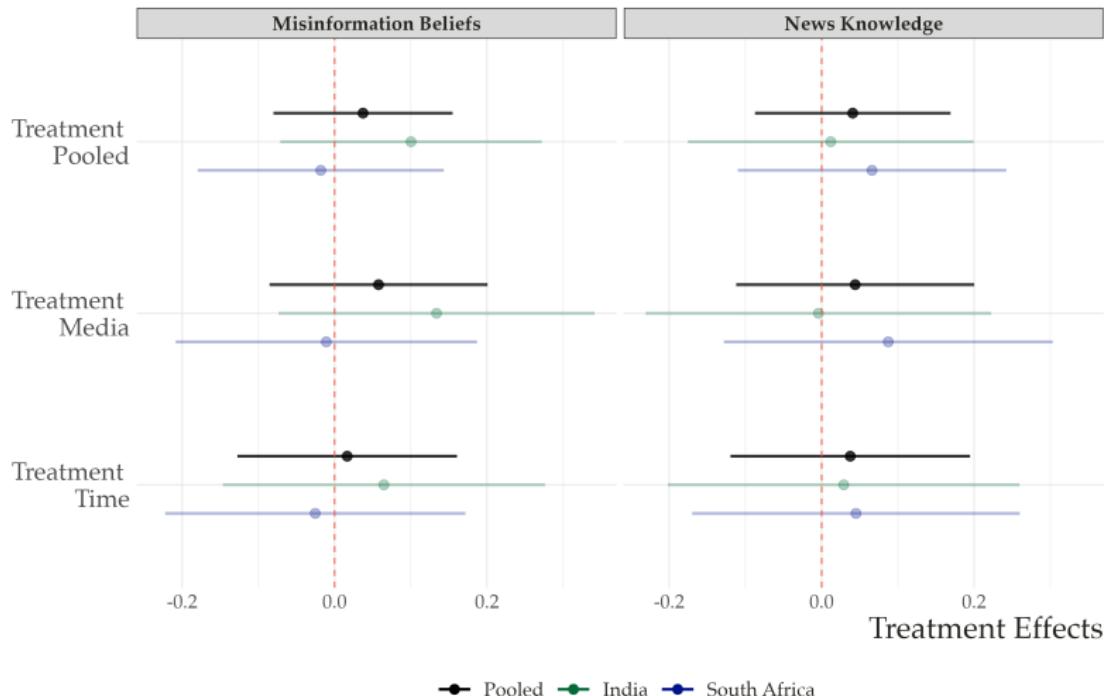
# Exposure

Limiting WhatsApp usage reduces exposure to misinformation & news



# Belief

A short-term reduction in exposure does not translate to changes in accuracy judgments



# Polarization

Would a potential reduction in exposure to misinformation, hate speech, and incivility reduce

- 1** partisan animosity?
- 2** ethnic/racial prejudice?

# Polarization

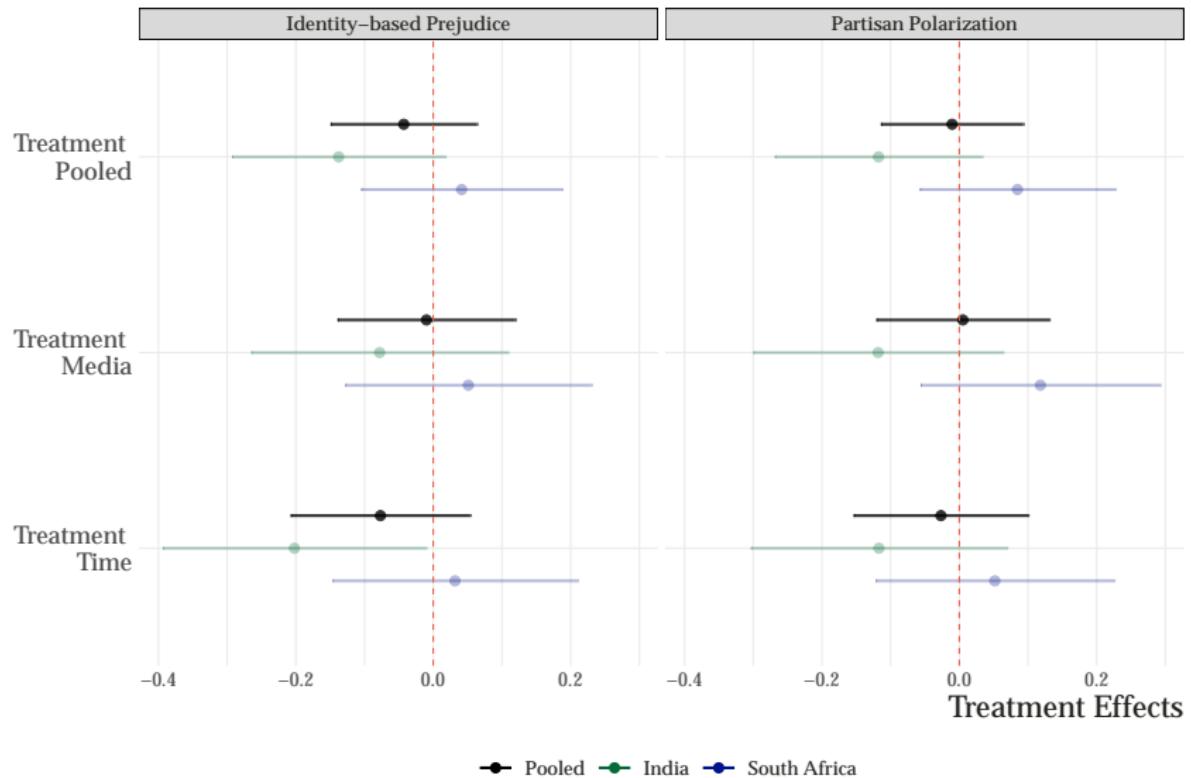
Would a potential reduction in exposure to misinformation, hate speech, and incivility reduce

- 1 partisan animosity?
- 2 ethnic/racial prejudice?

Series of outcomes related to ingroup and outgroup attitudes

- Outcomes
  - overall feelings
  - traits
  - willingness to participate in social activities
- Groups
  - in India: BJP & Congress voters; Hindu & Muslim citizens
  - in South Africa: ANC & DA voters; Black & White citizens

# Polarization



## Concluding thoughts

# Concluding thoughts

## Summary

- WhatsApp is an important channel through which voters receive misinformation and news in India and South Africa
- Reduced exposure to such information does not mechanically affect accuracy perceptions
- Reduced exposure to such information *sometimes* affects outgroup or outpartisan attitudes

# Concluding thoughts

## Summary

- WhatsApp is an important channel through which voters receive misinformation and news in India and South Africa
- Reduced exposure to such information does not mechanically affect accuracy perceptions
- Reduced exposure to such information *sometimes* affects outgroup or outpartisan attitudes

## Next steps

- More HTEs and secondary outcomes to understand divergences in results across countries
- Analyze compliance data and estimate complier effects
- 2-week follow-up survey: effects of “returning” to WhatsApp?
- Replication during local elections in Brazil: next week



What's Up with WhatsApp?

Audience is typing...



TODAY

The you!



8:30-ish ✓



Type a message





What's Up with WhatsApp?

Audience is typing...



TODAY

Thank you!



8:30-ish ✓



Type a message

