Combating Misinformation using Fact-Checking via WhatsApp in South Africa

Researchers:

Jeremy Bowles

Kevin Croke

Horacio Larreguy

Shelley Liu

John Marshall

Sector(s): Health, Political Economy and Governance

Sample: 8,947 adult South African Facebook users

Target group: Adults

Outcome of interest: Trust

Intervention type: Information Social networks Media Online learning

AEA RCT registration number: AEARCTR-0007615

Partner organization(s): Africa Check

Misinformation has spread worldwide and can be the catalyst for harmful individual and collective behaviors. In South Africa, the majority of citizens use social media platforms, such as Facebook and Whatsapp, which are used to spread misinformation. Researchers conducted a randomized evaluation to assess how fact-checks regularly sent to participants through WhatsApp can affect their ability to discern false information, as well as their beliefs and attitudes related to topics subject to viral misinformation, particularly Covid-19. Overall, the intervention improved participants' ability to detect misinformation and somewhat increased their willingness to participate in Covid-19 safety measures, particularly when the intervention was delivered in the form of a short text or a podcast with empathetic language.

Policy issue

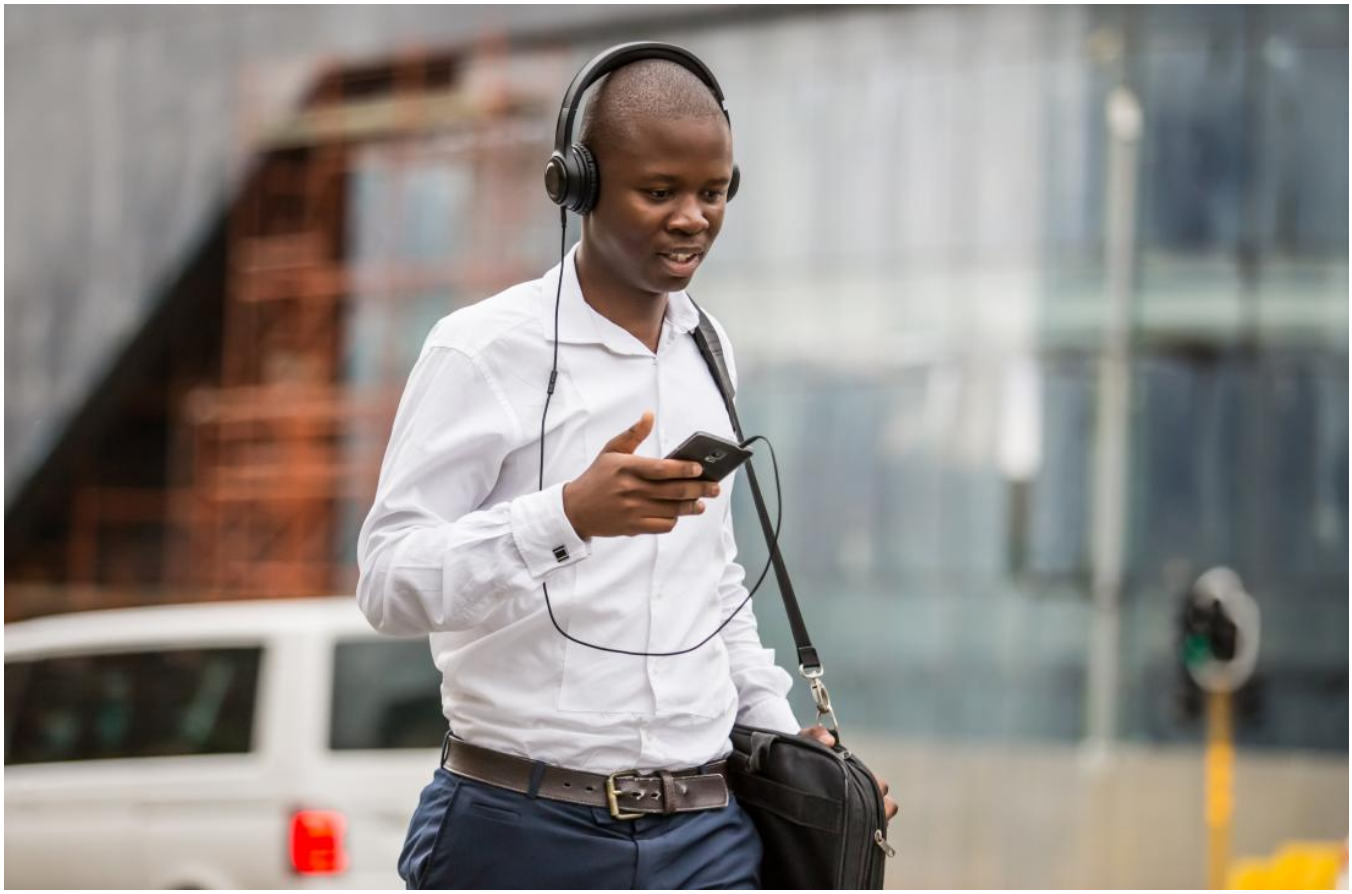
Misinformation has spread worldwide and can be the catalyst for harmful individual and collective behaviors. It has been linked to destructive actions across the globe, such as violence toward the Rohingya minority group in Myanmar[1], and mob violence against discriminated groups in India.[2], This phenomenon is particularly concerning in the Global South where citizens have limited access to independent sources of information, low digital literacy, and are increasingly reliant on social media for information.[3], [4]

Fact-checking has become a popular strategy for disarming misinformation, leading to the establishment of many fact-checking institutions across the globe. A key feature of these institutions is their ability to regularly engage citizens over a sustained period with different types of fact-checking and verification methods, equipping citizens not only with facts but with lessons to distinguish credible information. However, previous research has focused on shorter one time efforts conducted in artificial lab or online settings and has been concentrated in high-income countries. Can providing citizens in non-high-income countries with sustained exposure to fact-checking reduce the dissemination of misinformation?

Context of the evaluation

Misinformation has become a common phenomenon in South Africa, affecting citizens' understanding about social, political, and health issues.[5], [6] This is fueled by the popularity of social media platforms such as Facebook and WhatsApp that can be key vehicles for spreading misinformation, as well as the high cost of mobile data that makes it challenging to access other news sources on the internet.

Researchers partnered with Africa Check, the first independent fact-checking institution serving sub-Saharan Africa since 2012,[7], to deliver the fact-checking intervention. They recruited study participants through a Facebook advertisement, resulting in most participants to be between ages 18 to 50 years old, due eligibility constraints and the older population's low use of social media. Otherwise, the study participants were roughly representative of the South African population in age, gender, ethnicity, education, and region. The demographics of participants matched closely to the characteristics of the latest South African Afrobarometer survey, an independent pan-African research network that provides high-quality data on African society.[8]



A man walking while on his phone.

Photo Credit: Rich T Photo, Shutterstock.com

Details of the intervention

Researchers partnered with Africa Check to conduct a randomized evaluation to test how a consistent fact-checking program can affect citizens' beliefs and attitudes related to topics that are subject to viral misinformation, as well as their ability to discern false information. These messages scrutinized largely false social media stories that were trending on South African social media on topics such as Covid-19, health remedies, politics and society, and other high-profile topics. Participants were recruited bi-weekly

in batches through Facebook to receive three fact-checks through WhatsApp every two weeks over a six-month period. Researchers grouped participants by baseline characteristics to ensure each group had similar demographics, baseline social media consumption patterns, trust towards different news sources, and misinformation knowledge. Participants were then randomly assigned within each group to one of four methods of fact-check delivery or to a comparison group, ensuring that intervention groups had similar compositions. The interventions had the following components:[9]

1. Text information: A single sentence summary of the fact-check information with a link to the full article on the Africa Check website, for each of the three fact-checks in the message. About 7814 participants were assigned to this intervention and saw their assignment.
2. Long podcast: The message contained a 6–8 minute podcast featuring a lively conversation between two hosts explaining each of the three claims to be true, false, misleading, or uncertain. This form of edutainment looks to increase listeners' engagement with the informational content. 1,530 participants were assigned to this intervention and saw their assignment.
3. Short podcast: The message contained a 4–6 minute abbreviated version of the long podcast, mindful of users' limited time or shorter attention spans. 1,576 participants were assigned to this intervention and saw their assignment.
4. Empathetic podcasts: The message contained a full length podcast supplemented with empathetic content to communicate the hosts' understanding of how fear and concern for friends and family could prime citizens to be deceived by misinformation. This method builds on recent research that emphasizes the role of emotions in the spread of misinformation. 1,547 participants were assigned to this intervention.
5. Comparison: Did not receive any fact-check messages. 1,898 participants were assigned to this intervention and saw their assignment.

The researchers randomly assigned half of the participants within each treatment group to also receive text within the messages highlighting the importance of fact-checking for common good to encourage the participant to read the message or listen to the podcast. To further encourage engagement, 83 percent of participants within each treatment arm were randomly assigned to receive a quiz about the fact-check and the remaining received an unrelated quiz about popular culture. All of the participants assigned to the comparison group received the unrelated quiz on popular culture.

At the end of the intervention period, the researchers surveyed the participants to measure discernment between true and false information and trust in different media sources; information consumption, verification, and social media sharing patterns; and attitudes and behaviors relating to Covid-19 and politics. The researchers combined measures from the survey responses to create overall outcome indexes.

Results and policy lessons

Overall, some forms of the intervention improved participants' ability to detect misinformation and somewhat increased willingness to participate in Covid-19 safety measures. Simple and short text messaging, as well as podcasts with empathetic language, produced the greatest effects across the messaging varieties.

- *Discerning fact from fiction:* Participants that received text-only messaging and podcast with empathetic messaging were more likely to detect true and false information relative to the comparison group by 0.12 and 0.14 standard deviations,[10] respectively. The short and long podcast delivery methods did not have a significant impact. Similarly, participants that received the text-only messaging and empathetic podcast, as well as the long podcast, were more likely than the comparison group to identify conspiracy theories by 0.11, 0.17 and 0.11 standard deviations, respectively.

- *Trust in social media*: Researchers measure participants' trust in social media platforms other than WhatsApp (e.g., FB, Twitter, Instagram). Participants that received the text-only messaging, long podcast, and empathetic podcast interventions were less likely to report trusting in social media platforms by 0.15, 0.07, and 0.15 standard deviations. The short podcast had no significant impact on this outcome.
- *Information consumption, verification, and sharing:* No participants changed their social media consumption nor how often they sought to actively verify information. However, all participants did increase verification through Africa Check. This led to crowding out of verification through traditional media (e.g., local online news, national broadcaster), as well as crowding out of verification through online and social media for respondents who received text-only messaging. This implies that an individual's ability to verify information may not be the only reason for individuals' limited efforts to do so. While no participants changed their use of social media as a result of the messaging interventions, participants that received fact-checks through texts or empathetic podcasts were about 0.10 standard deviations less likely to report they would share information they had seen on social media. This demonstrates the potential for sustained fact-checking to limit viral spread of misinformation.
- *Attitudes and behaviors:* Participants that received the podcast fact-check did not change their beliefs and behaviors related to Covid-19. However, fact-checks delivered via short and simple text messaging increased the index by 0.14 standard deviations. This was driven by individuals' self-reported willingness to get vaccinated, increase mask-wearing, decrease indoor activity and their increased skepticism that Covid-19 is a hoax.

The study demonstrates the feasibility for using WhatsApp messaging to stimulate citizens to engage with fact-checking and approach news subject to misinformation more critically. Results suggested that direct and short forms of delivery were the most effective in driving impact, as opposed to longer-form delivery, unless empathetic language is explicitly added.

Based on the study results, Africa Check and the researchers are evaluating a similar project in Kenya and South Africa, utilizing the reach of social media influencers (high-profile journalists and social activists with large social media followings) to distribute both fact checks and digital literacy training. They aim to gather insights on the effectiveness in improving followers engagement with reliable content.

---

1. Whitten-Woodring, Jenifer, Mona S. Kleinberg, Ardeth Thawnghmung, and Myat The Thitsar. 2020. "Poison if you don't know how to use it: Facebook, democracy, and human rights in Myanmar." The International Journal of Press/Politics 25(3): 407-425.

2. Banaji, Shakuntala, Ramnath Bhat, Anushi Agarwal, Nihal Passanha, and Mukti Sadhana Pravin. 2019. "WhatsApp vigilantes: an exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India. Department of Media and Communications, London School of Economics and Political Science.

3. Arechar, Antonio A., Jennifer N. L. Allen, Adam Berinsky, Rocky Cole, Ziv Epstein, Kiran Garimella, Andrew Gully, et al. 2022. "Understanding and Combatting COVID-19 Misinformation Across 16 Countries on Six Continents." PsyArXiv. February 11. doi:10.31234/osf.io/a9frz.

4. Bowles, Jeremy, Horacio Larreguy, and Shelley Liu. 2020. "Countering misinformation via WhatsApp: Preliminary evidence from the COVID-19 pandemic in Zimbabwe." PloS one 15(10)

5. Wasserman, Herman. 2020. "Fake news from Africa: Panics, politics and paradigms." Journalism 21(1): 3-16.

6. Servick, Kelly. "Fighting scientific misinformation: A South African perspective." *Science*, February 15, 2015. https://www.science.org/content/article/fighting-scientific-misinformation-south-african-perspective

7. Africa Check. Accessed December 14, 2022. https://africacheck.org/who-we-are

8. Afrobrometer. "FAQs". Accessed December 14, 2022. https://www.afrobarometer.org/about/faqs/#:~:text=Afrobarometer%20is%20an%20independent%20pan,governance%2C%20economy%2

9. Of 8,947 participants, 1,616 never saw treatment assignments, but the remaining participants were naturally balanced across intervention groups.

10. The author combined a number of survey questions to create an index for each outcome variable (e.g. “Discernment between true and false news” and “Identification of conspiracy theories”), standardizing this measure against the control group. The results are reported in standard deviations of this index, which tells us the extent to which the participants in the intervention groups responded differently from the comparison group for these outcomes.