CUTS
International

Operation Strategy Note

Do users in India, Kenya and Ghana react differently to problematic content?

1. Introduction

1.1 Background

Social media platforms (including instant messaging applications) have revolutionised human social interactions where users can express themselves freely, fearlessly and safely, thereby making the society more inclusive and democratic. Studies have found that connecting with friends & family, exchange of information related to work/study and news are the greatest drivers behind people using social media.[1] Further, social media has fuelled the access to information in an unprecedented fashion. However, misuse of social media platforms has become a tool for spreading problematic content (like incorrect, sensational, provocative, divisive, hateful and misleading information etc.). Such misuse may cause various harms to users, such as: risk of increasing polarisation, creating trust deficits, instigating negative reactions, promoting abusive behaviour, etc. A few studies have found that many people forwarding such content are generally unaware about the fact that they may be spreading problematic content.[2] Exposure to such content may lead to different kinds and levels of harm to users depending upon various factors like their socio-economic profile, age group, educational background, among others.

1.2 Introduction to Problematic Content

Multiple studies have theoretically classified problematic content and its manifestations. Also termed as information disorder, problematic content can be classified into different categories based on the falseness and the intention of creation.[3] These are misinformation, disinformation and mal-information. The categorisation has been explained below along with a pictorial representation.

---

[1] Dodda T R and Dubbudu R (2019) Countering Misinformation Fake News in India. Factly Media & Research (Factly) and The Internet and Mobile Association of India (IAMAI). Available online at: https://factly.in/wp-content/uploads//2019/02/Countering-Misinformation-Fake-News-In-India.pdf
[2] Dodda T R and Dubbudu R (2019) Countering Misinformation Fake News in India. Factly Media & Research (Factly) and The Internet and Mobile Association of India (IAMAI). Available online at: https://factly.in/wp-content/uploads//2019/02/Countering-Misinformation-Fake-News-In-India.pdf
[3] Wardle C and Derakhshan H (2017) Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe. Available online at: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

1.2.1 **Misinformation:** Misinformation is false or incorrect information which a user shares without knowing its falseness and does not indent to cause any harm. It includes unintentional mistakes such as inaccurate photo captions, dates, statistics, and translations, etc.[4] It also includes satire or parody content which may not have been made with an intention to cause harm but has a potential to fool. Misinformation often floats as unverified information[5] which includes suspicious statistics, fake news[6], urban legend[7], spams[8], trolls[9] and unfounded opinion related to any particular event or person, or group of persons.[10]

1.2.2 **Disinformation:** Disinformation is false information which a user knows about and shares with an intention to cause harm to others. It may include fabricated or deliberately manipulated content like misrepresentation of factual information like changing date and time of genuine content where old content is often repurposed and circulated again. It also includes intentionally created conspiracy theories or rumours.[11]

1.2.3 **Mal-information:** Mal-information is genuine information which is shared with an intention to cause harm.[12] This may include sensational, hyped and provocative content pertaining to an event or person or group of persons. Hate speeches/abuses based on political, religious, caste, gender, etc. factors also fall under this category. Further, deliberately making private information public for personal gain (example: revenge porn) also falls in this category.[13]

---

[4] Derakhshan H (2020) Disinfo Wars: a taxonomy of information warfare. LSE Media Policy Project | Truth, Trust and Technology Commission. Available online at https://blogs.lse.ac.uk/medialse/2019/09/27/disinfo-wars-a-taxonomy-of-information-warfare/

[5] A piece of information can be defined as unverified information before it is verified, and those verified to be false or inaccurate obviously belong to misinformation. It may trigger similar effects as other types of misinformation, such as fear, hatred and astonishment.

[6] Fake news refers to false information in the form of news (which is not necessarily disinformation since it may be unintentionally shared by innocent users).

[7] Urban legend is intentionally-spread misinformation that is related to fictional stories about local events. The purpose can often be entertainment.

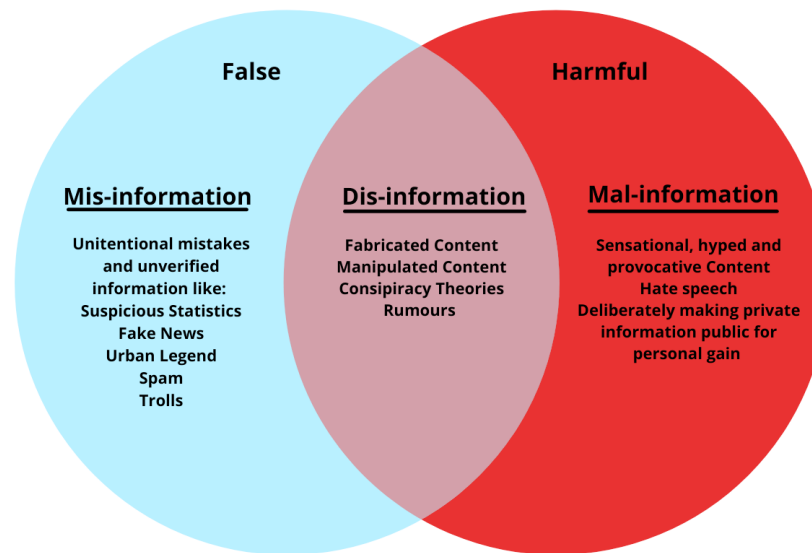[8] Spam refers to irrelevant information that is sent to a large number of users.

[9] Troll aims to cause disruption and argument towards a certain group of people. Different from other types of misinformation that try to convince its recipients, trolling aims to increase the tension between ideas and ultimately to deepen the hatred and widen the gap.

[10] Wu L, Morstatter F, Carley K M, and Liu H (2019) Misinformation in Social Media: Definition, Manipulation, and Detection. SIGKDD explorations Vol. 21 Issue 2. Available Online at: https://www.kdd.org/exploration_files/8._CR.10.Misinformation_in_social_media_-_Final.pdf

[11] Derakhshan H (2020) Disinfo Wars: a taxonomy of information warfare. LSE Media Policy Project | Truth, Trust and Technology Commission. Available online at https://blogs.lse.ac.uk/medialse/2019/09/27/disinfo-wars-a-taxonomy-of-information-warfare/

[12] Wardle C and Derakhshan H (2017) Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe. Available online at: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

[13] Karduni A (2019) Human-Misinformation interaction: Understanding the interdisciplinary approach needed to computationally combat false information. arXiv.org. Available online at: https://arxiv.org/pdf/1903.07136.pdf

CUTS®
International

False

Harmful

Mis-information

Unitentional mistakes
and unverified
information like:
Suspicious Statistics
Fake News
Urban Legend
Spam
Trolls

Dis-information

Fabricated Content
Manipulated Content
Consipiracy Theories
Rumours

Mal-information

Sensational, hyped and
provocative Content
Hate speech
Deliberately making private
information public for
personal gain

Types of Problematic Content

Source: *Wardle C and Derakhshan H (2017) in Information Disorder: Toward an interdisciplinary framework for research and policymaking.[14]*

Further, problematic content can be classified into different domains as mentioned below.[15]

1. **Hateful content based on political or religious affiliation:** Contents such as those abusing or threatening a politician or a political party, religious figure or a religion/caste. These may also include cultural messages which have religious connotations, celebrity or societal references. It may also include government related messages which are linked to the police, judiciary, and political parties and may include real or fake announcements and advisories.

---

[14] Wardle C and Derakhshan H (2017) Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe. Available online at:
https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c
[15] Akbar S and Pal J (2020) Content patterns in COVID-19 related digital misinformation in India. Microsoft Research India. Available online at:
http://joyojeet.people.si.umich.edu/an-archive-of-covid-19-related-fake-news-in-india/ and   Akbar, S., Kukreti, D., Sagarika, S. and Pal, J. (2020) Temporal patterns in COVID-19 related digital misinformation in India. Available online at: http://joyojeet.people.si.umich.edu/temporal-patterns-in-covid-19-misinformation-in-india/

2. **Unverified information on historic events, or rumours:** These may include alarming content which invoke fear like violent imagery, death or mass-killings, illness of people in the pandemic. These may also include doctored statistics messages have fake numbers which may be linked to economy, covid-19 cases, population figures etc.

3. **Sensational content:** Such content may pertain to an event or person(s) and may include cure-based messages suggest alternative or mainstream remedies which make claims of curing people of diseases. Further, it may also include hype towards any possible occurrence(s) of certain events.

4. **Scam related content:** These may be messages which include scams such as attractive yet suspicious deals/offers, such as for shopping, free gifts, job opportunities, government benefits, and availing loans. Further, it may also include panic-buying and target businesses with fake positive cases.

2. CUTS Study

2.1 Need for the Study

The dissemination and spread of problematic content generally follows a three phased chronological process. It starts with creation, is followed by publication or reproduction and finally by distribution or propagation.[16] While creation and publication may be done intentionally or unintentionally, distribution or propagation of problematic content by a user often happens once he/she is develops trust about content being authentic and truthful.[17] Studies suggest that once users are convinced about a piece of information being factual, they may share it vehemently.[18] Further, problematic content propagates more if it is based on identity and emotion which users related to, rather than on scientifically verifiable instrumental facts.[19] There are several reasons why users get convinced about such content easily. A few of them are listed below:[20]

---

[16] Wardle C and Derakhshan H (2017) Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe. Available online at: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

[17] Zhou X and Zafarani R (2020) A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. arXiv.org. Available online at: https://arxiv.org/pdf/1812.00315.pdf

[18] Zhou X and Zafarani R (2020) A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. arXiv.org. Available online at: https://arxiv.org/pdf/1812.00315.pdf

[19] Akbar S and Pal J (2020) Content patterns in COVID-19 related digital misinformation in India. Microsoft Research India. Available online at: http://joyojeet.people.si.umich.edu/an-archive-of-covid-19-related-fake-news-in-india/ and Akbar, S., Kukreti, D., Sagarika, S. and Pal, J. (2020) Temporal patterns in COVID-19 related digital misinformation in India. Available online at: http://joyojeet.people.si.umich.edu/an-archive-of-covid-19-related-fake-news-in-india/

[20] Zhou X and Zafarani R (2020) A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. arXiv.org. Available online at: https://arxiv.org/pdf/1812.00315.pdf

- They may be tempted to trust and accept information if it follows their pre-existing views and pleases them. Further, they may not accept new evidence because it might not conform to their views.
- They may also want to feel accepted or gain popularity in their social circles.
- Public figures and mainstream media are also found spreading problematic content which leads to an increase in their followers.
- Users are often inclined to believe a piece of information after repeated exposures.

Further, studies have cautioned against a likelihood risk of a significant increase in the trend of posting misinformation on popular instant messaging applications in the upcoming years and state that users who are active in posting problematic content will continue sending to do so and the receivers will continue believing them due to lack of awareness.[21] Thus, problematic content may spread adopting a snowball effect.

The impact such spread of problematic content causes is severe. While it can lead to users experiencing online harms like cyber-bullying and trolling etc., individuals may also be vulnerable to offline harm such as mental and physical trauma, among other things. The level of harm faced by the user depends upon how users react to problematic content. Though the literature gives a good understanding of the reasons behind users' belief in problematic content, it lacks in identifying how users *perceive, engage, understand, internalise, and react* to problematic content which in turn leads to them developing trust in problematic content. There may be several factors which affect the reaction patterns and thus, the level of harm users and non-users are vulnerable to may vary. These may be on account of difference in users' ethnographic, cultural and political scenarios, geographic location (urban, peri-urban and rural areas), demographic profile (age, gender, and education), economic condition (income level), experience (number of years of internet usage), psychological and cognitive variables, extent and duration of exposure to problematic content). Therefore, there is a need to study how users react to exposure to problematic content on social media platforms, and if their reaction causes any harm and if yes, the severity of such harm. Further, it is also important to understand perspective on problematic content of users so that a strategic plan can be devised which can be used to protect the users from the harms caused by these contents.

While existing literature provides an understanding of the issue in individual countries, a comparative analysis between countries in different geographies has been relatively less explored. In light of this, CUTS proposes to conduct a study which will take a detailed analysis of the users' reaction towards problematic content in three countries, namely India, Kenya and Ghana. This study is directed in the three countries to gain comparative perspective of consumers' within the global south. The three chosen countries are emerging developing economies, each of which has different cultural and political landscape. Therefore, the study will help in understanding if reaction to problematic content differs with changing ethnographic, cultural and political scenarios or not. It will further present a comparative analysis of the users' reaction towards

---

[21] Khurana P, Kumar D and Kumar S (2019) Research of Fake News Spreading Through Whatsapp. International Journal of Innovative Technology and Exploring Engineering (IJITEE). Available online at: https://www.ijitee.org/wp-content/uploads/papers/v8i6s4/F11930486S419.pdf

problematic content which is required in order to verify how different or common the reaction patterns are and to know how differently it may harm different users and non-users. In this context, a literature review of previous studies conducted on problematic content in the three chosen countries is presented in the table below.

Studies in India

Methodology	Findings	Relevance for our Study
1. University of Michigan: Content patterns in COVID-19 related digital misinformation in India[22]		
Survey of 243 unique debunked COVID-19 related messages online. Qualitative coding of messages into different categories. Annotated each message for thematic-categories, representative-tags, source of claims & attributions and medium of propagation.	Messages were classified in different categories - Alarmism[23], Culture[24], Cure[25], Nature & the Environment[26], Casualty[27], Business & Economy[28], Government[29], and Doctored statistics[30]. Amount of circulation of such messages is different for each category. Misinformation is more affective around identity and emotion, than around instrumental facts that can be scientifically verified. Mode of content relied upon to relay different kinds of misinformation varies between video, text, image and audio. Same/old news is often refabricated or re-circulated. Public figures and/or mainstream media are also noted to be spreaders of misinformation, which often adds credibility.	Classify kinds of misinformation based on subject, and type/mode. Check effects, frequency and emotions/reactions they trigger. List down factors giving messages credibility based on age, use of official seals or popular personalities etc.

---

[22] Akbar S and Pal J (2020) Content patterns in COVID-19 related digital misinformation in India. Microsoft Research India. Available online at: http://joyojeet.people.si.umich.edu/an-archive-of-covid-19-related-fake-news-in-india/ and Akbar, S., Kukreti, D., Sagarika, S. and Pal, J. (2020) Temporal patterns in COVID-19 related digital misinformation in India. Available online at: http://joyojeet.people.si.umich.edu/an-archive-of-covid-19-related-fake-news-in-india/

[23] Message is alarmist when it refers to fear-invoking messages such as those with violent imagery, death or mass-killings.

[24] Message is about culture when messages have religious connotations, celebrity or societal references.

[25] Messages comes under cure-based when messages suggest remedies – alternative or mainstream – that make claims on curing people of the virus.

[26] Message is classified as nature when messages have references to animals and the environment.

[27] Messages relating to deaths, illness of people in the pandemic, including graphic images of suffering (not including doctored statistics)

[28] Messages relating to scams, panic-buying and target businesses with fake positive cases.

[29] Messages have government announcements and advisories or refer to police, judiciary, political parties.

[30] Messages that have exaggerated numbers of positive cases or death counts and fake advisories.

Methodology	Findings	Relevance for our Study
	Different keywords are used/observed for different categories of content.	
	Popular misinformation content has been curated.	
2. Various statistics compiled by Statista.com	

The main factor which influenced Indian respondents of a survey to share information online was that it might benefit others. This was followed by a belief that mainstream media was biased and did not cover certain types of news.[31] According to the results of a survey on rural India, 760 respondents were part of family WhatsApp groups, while about 105 stated that they were part of local administration groups on WhatsApp. The highest number of respondents were part of groups with friends on the social networking app.[32] It was found that about 15 percent of respondents never believed information they received via the messaging app. While about eight percent of respondents always believed the information, they received on WhatsApp during the survey period.[33] It was seen that about 36 percent of respondents only read forwarded messages whereas 25 percent of Indians in rural areas read and forwarded messages on WhatsApp during the survey period.[34] A majority of respondents, about 43 percent trusted news channels, while about 15 percent trusted WhatsApp for news during the survey period.[35] About 45 percent of the respondents of a survey about fake news in the Indian media said they had seen stories that were completely made up for political or commercial reasons. The survey respondents were asked to state the different kinds of fake news that they came across over the preceding week of the survey period.[36] A majority of Indians under the age of 35 preferred to use online sources to access news, whereas, those above 35 years of age are more inclined towards traditional offline sources for news as of early January 2019.[37] About 51 percent of the survey respondents agreed that they are concerned when facts were spun or twisted to push a particular agenda through fake news media. On the other hand, 32 percent of the respondents seemed to	Consumer perceptions towards social media and mainstream media may be compared. Information consumption from different sources on WhatsApp may be mapped, and checked for trustworthiness, followed by re-actions. Perception towards agenda driven content being circulated online may be checked. Sources of receiving news (print, video, online etc.) may be checked and compared for

---

[31] https://www.statista.com/statistics/996489/india-motivational-factors-to-share-information-on-social-media/
[32] https://www.statista.com/statistics/961517/india-nature-of-whatsapp-groups-in-rural-areas/
[33] https://www.statista.com/statistics/961486/india-credibility-of-whatsapp-information-in-rural-areas/
[34] https://www.statista.com/statistics/961432/india-action-towards-forwarded-messages-on-whatsapp-in-rural-areas/
[35] https://www.statista.com/statistics/961504/india-most-trusted-news-platform-in-rural-areas/
[36] https://www.statista.com/statistics/1027036/india-exposure-to-fake-news/
[37] https://www.statista.com/statistics/1026298/india-main-source-of-news-by-age/

Methodology	Findings	Relevance for our Study
think satire was a concern in terms of fake news.[38] When asked about the different social networks used to access news as of 2019, a majority of the surveyed respondents in India stated Facebook and WhatsApp as their main sources at 52 percent each in the week preceding the survey period. Snapchat had the lowest use for news access at just five percent.[39] 63 percent of respondents according to a recent survey said the headline or picture was an important attribute to click-through to a news story on social media. This was in line with findings from another study that showed Indians largely rely on social cues or alternative markers like number of comments, images and the sender of the messages to access news.[40] About 29 percent of respondents trusted news they received via WhatsApp or Facebook. About 55 percent trusted the news they read in a newspaper, while this was 50 percent among TV news consumers.[41] About 41 percent of the respondents who supported BJP said they trusted most news most of the times in the week preceding the survey in early 2019, whereas, only 26 percent of those who did not support any political party and did not intend to vote said they trusted the news media that week.[42] 44.6 percent said that they had not heard of fact-checking organizations, while close to 20 percent trusted outlets that debunked fake news or fact-checked the news.[43]		trustworthiness. Factors influencing decision to believe or question the credibility of the news received online may be checked – picture, headline, source, likes etc. Awareness towards fact checking organisations, and capacity to use the same may be gauged.
3. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India[44]		
Conducted preregistered survey experiments in both the United States and India examining the	Results indicated that exposure to variants of the Facebook media literacy intervention reduces people's belief in false headlines. The media literacy treatment significantly reduced beliefs in false news	A proposed intervention to curb problematic content may be tested through the

---

[38] https://www.statista.com/statistics/1026835/india-fake-news-concern/
[39] https://www.statista.com/statistics/1026234/india-social-networks-used-to-access-news/
[40] https://www.statista.com/statistics/1026277/india-attributes-to-access-news-on-social-media/
[41] https://www.statista.com/statistics/962662/india-trust-in-news-source/
[42] https://www.statista.com/statistics/1026819/india-trust-in-news-by-political-allegiance/
[43] https://www.statista.com/statistics/996531/india-level-of-trust-in-fact-checking-organizations/
[44] Guess A.M., Lerner M, Lyons B, Montgomery JM, Nyhan B, Reifler J and Sircar N (June 2020) A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. Proceedings of the National Academy of Sciences of the United States of America (PNAS). Available online at: https://www.pnas.org/content/117/27/15536

Methodology	Findings	Relevance for our Study
effectiveness of presenting people with "tips" to help spot false news stories. Used the average treatment effect on the treated (ATT) method, using an instrumental variables approach. Exposed respondents to 2 waves of news items (pre-post literacy intervention) to check their ability to spot fake news.	stories, and respondents better distinguished between mainstream and false articles. Increased scepticism of false news headlines may come at the expense of decreased belief in mainstream news headlines—the media literacy intervention reduced the perceived accuracy of these headlines in both the US and India online surveys. Respondents' overall ability to distinguish between mainstream and untrustworthy news increases by more than 26% in the US sample and 17% in the highly educated online Indian sample. The findings provide important evidence that shortfalls in digital media literacy are an important factor in why people believe misinformation that they encounter online.	methodology.
4. Research of Fake News Spreading Through WhatsApp[45]		
Tried to estimate the spread of Fake news on Whatsapp based on the Analytic Modeling, considering the number of feasible authors those are relevant for spreading of fake news (S), the wide variety of coetaneous authors who are highly active for posting the fake information (I), the range of authors who got the right information (R) through TV channels, newspaper etc who are inactive to spread the fake	Significant likelihood of increase in the trend of posting unreal information on WhatsApp in the next upcoming years as the ones who are active in posting rumours will continue sending fake news and the receivers will continue believing them due to lack of awareness. This study, based on conclusion, suggested that all WhatsApp users required to check the truthfulness of news before sending them to other users.	If remained unchecked, there is a significant risk of rise in the spread of problematic content through WhatsApp.

[45] Khurana P, Kumar D and Kumar S (2019) Research of Fake News Spreading Through Whatsapp. International Journal of Innovative Technology and Exploring Engineering (IJITEE). Available online at: https://www.ijitee.org/wp-content/uploads/papers/v8i6s4/F11930486S419.pdf

Methodology	Findings	Relevance for our Study
information.		
	5. Countering Misinformation Fake News In India[46]	
A web survey (1296) that collected data to understand the patterns of information consumption in India and if there is a correlation between beliefs, attitudes, bias and motivation to share information. Consolidated opinions of multiple stakeholders (76) extracted from semi structured interviews - Technology & Internet Service Providers Government – Policy Law Enforcement, Media & Influencers, Fact Checkers & Verifiers, Academia, Political Parties.	The young (below 20 years) and old (above 50 years) that may be most susceptible to fake news. Those who are relatively new to use of technology/ internet & smartphones may be more susceptible to fake news than others. Newspapers still remain one of the top sources of information for people across age groups. Connecting with friends & family, exchange of information related to work/study and news are the greatest drivers behind people using social media. As age of the respondents increased, Friends or Friend Groups and groups based on political/social/cultural beliefs of the person was chosen by a greater proportion of respondents as their main source of information on social media. Background evidence and trust in organizations/persons are what make most people to believe in the information they receive on social media. Benefit to others and the lack of trust on mainstream media are what drive people to forward information on social media. is in line with global observations that most people forward/share information without knowing they are spreading fake news. People trust neutral media and fact-checking organizations more than others. A substantial percentage of people are not aware of the existence of fact	Possible solutions were recommended for different stakeholders.

---

[46] Dodda T R and Dubbudu R (2019) Countering Misinformation Fake News in India. Factly Media & Research (Factly) and The Internet and Mobile Association of India (IAMAI). Available online at: https://factly.in/wp-content/uploads//2019/02/Countering-Misinformation-Fake-News-In-India.pdf

Methodology	Findings	Relevance for our Study
	checking organizations.	
	People verify the information they receive when they are pushed to do so.	
	Fake news has a greater probability of being correctly identified when it is fact checked/debunked by multiple organizations.	
6. "I don't think that's True, Bro!" An Experiment on Fact-checking Misinformation in India[47]		
Studies if peer-to-peer corrections on the encrypted messaging app WhatsApp can reduce misinformation or not. An experimental evaluation is conducted among a large sample size of N=5104 to check the effect of different types of corrective messages against seven common rumours. Respondents are shown screenshots of WhatsApp group chat where one users posts a message carrying fake news and another user posts a corrective message as a response. The respondents are themselves not a part of the group and only view screenshot of the group chat.	Finds that peer-to-peer corrections substantially reduce belief in misinformation. Even brief, un-sourced and unsubstantiated corrective messages have an affect which is comparable to that of corrective messages which are backed by evidence from credible sources. Suggests that even signalling a doubt about a rumour, substantiated or unsubstantiated, may significantly reduce spread of misinformation.	Intervention such as corrective messages on encrypted messaging apps may lead to a great degree of awareness and spread of misinformation may be prevented.

[47] Badrinathan S and Chauchard S (2020) "I Don't Think That's True, Bro!" An Experiment on Fact-checking Misinformation in India. Working Paper. Available online at: https://sumitrabadrinathan.github.io/Assets/Paper_WhatsApp.pdf

Studies in Kenya

Methodology	Findings	Relevance for our Study
1. Audience Motivations for Sharing Dis- and Misinformation: A Comparative Study in Five Sub-Saharan African Countries[48]		
Secondary research. Conducted focus groups with undergraduate and graduate student in five countries - Kenya[49], Namibia, Nigeria, South Africa and Zimbabwe. A total of 59 participants joined the discussion, in groups ranging from 5 to 15. Discussions lasted between 50 and 65 minutes and covered the following	In established democracy such as the United Kingdom, more than half of social media users (57.7%) have reported that they have recently come across news on these platforms of which the veracity was in doubt. Nevertheless, a high percentage (42.8%) of users admit to have shared false or inaccurate news, of which 17.3% said they thought the news was false at the time of sharing it. Race is one influence on how young social media users select exposure to news on social media. Users who willingly and/or knowing shared false information on social media platforms were 'likely to be male, younger, and more interested in politics. Not only social position, but also political orientation was found to play a role in the likelihood of British social media users sharing false information. When asked to reflect on the reasons why they share news on social media, British respondents reported as the top three reasons 'To express my feelings' (65.5%), 'To inform others' (also 65.5%) and 'To find out other people's opinions' (51,1%). These reasons display an orientation towards civic participation or purpose. The social utility of sharing 'fake news' in Singapore –'cope with uncertainty, build relationships, and for self-enhancement'. The main types of news stories that are shared, are those that have a 'high informational utility – 'news you can use" 'which resonates with their own lives and that have a high emotional impact. Sharing news, is seen as	Parameters such as democracy, freedom of press may be kept in mind while data analysis and country wide comparison. Sample design may be split as per race/caste/religion, political alignment etc. for comparison. Knowingly and unknowingly sharing fake news may be checked, along with reasons for the same. These reasons may be personal, interpersonal, political, expecting response/action etc.

---

[48] Wasserman H, Madrid-Morales D, Mare A, Ndlovu K, Tully M, Umejei E and Uzuegbunam CE (2019) Audience Motivations for Sharing Dis- and Misinformation: A Comparative Study in Five Sub-Saharan African Countries. The Berkman Klein Center for Internet & Society, Harvard University. Available online at: https://cyber.harvard.edu/sites/default/files/2019-12/%20Audience%20Motivations%20for%20Sharing%20Dis-%20and%20Misinformation.pdf

[49] Kenya is an East African country with a vibrant independent media as well as a strong presence of international media, notably Chinese media who have made Nairobi a hub from where it expands into the region. Kenya also has a vocal, active community of social media users which makes it suitable for a study of dis- and misinformation online.

Methodology	Findings	Relevance for our Study
general topics: media consumption, news sharing online, sharing of political information, the prevalence of fake news, and possible solutions to existing problems with mis- and disinformation. Online survey also conducted.	contributing to social cohesion – users doing so are motivated by the emotional impact the news is seen to have, the relevance it might have for the receiver, and the sender's intention to 'provide advice or warning' (trust between sender & receiver, and inter-personal relationships). When hostile rumors are shared, the aim is to '(1) coordinate the attention and action of the audience with the goal of mobilizing against the target group and (2) signal their willingness to engage in conflict escalation (i.e., helping push the collective over the tipping point for collective action).' Studies on spread of 'fake news' in the African context, established a link between lack of trust in the news media and the sharing of false news. Second, there is a sense of civic duty that might lead social media users to share warnings of impending disasters or crises. Even if the information might turn out not to be true, the harm done by not informing others may be seen as outweighing that of not informing them. And, third, there's the sense that information is democratic and needs to be passed on. Users may take the popularity or virality of a shared piece of information as indication of its veracity. This motivation might be especially relevant in African countries where the state exercises a great deal of control or ownership over the media, which may lead to a decline in trust in mainstream media. In the African context, further cultural influences such as the long-standing importance of informal sources of information such as gossip, rumour and may play a further role in the likelihood of media users to share news found on social media.	Hostility could specifically be focussed upon, as a category of fake news, and expected result from self and receiver may be recorded. Again, indicators of checking veracity/trust on news items may be gauged with respect to similar parameters, as used in this study.

Studies in Ghana

Methodology	Findings	Relevance for our Study
	1. Media Perspectives on Fake News in Ghana[50]	

[50] Ahiabenu K, Ofosu-Peasah G, Sam J (2018) Media Perspectives on Fake News in Ghana. Penplusbytes. Available online at: https://penplusbytes.org/wp-content/uploads/2018/05/FAKE-NEWS-STUDY.pdf

Methodology	Findings	Relevance for our Study
A survey which had 154 respondents from newsrooms across the country was conducted. Majority of respondents were from radio stations. Other respondents were from converged newsrooms (a combination of two or more media types namely TV, radio, newspaper and online)	This study is important in a number of ways more so that it is ground-breaking in the Ghanaian context. The study provides unique insights into how newsrooms are evolving especially by moving away from a reliance on traditional sources of news to social media and user content-generated material, which opens them up to the possibility of peddling fake news. The study also established that fake news is mostly manifested as fabricated content and false headline without connection to content. Overall, the study revealed that fake news is a growing problem. However, the media, regulatory bodies, and government do not have a clear-cut strategy to deal with the problem. Additionally, the Ghanaian media landscape does not have systems, budget and trained personnel dedicated to combat the menace of fake news.	Highlights the roles of newsrooms, especially considering the reliance on user-generated content, in safeguarding Ghanaians from fake news. It also highlights the need for editorial systems and concerted efforts towards putting in procedures and structures to monitor the news media. Similarly, it discusses the credibility of newsrooms in the light of the prevalence of fake news and the financial responsibilities needed to combat the spread. Highlights the need to promulgate strategy options, as well as legal and regulatory policies to combat fake news.
2. Double-edged Sword: Ghanaians see pros, cons of social media, want access but not fake news[51]		
Surveys were conducted	Findings further cement social media as a common source of news for Ghanaians. Surveys showed that although radio and television remain the most dominant news sources, daily news consumption via social media and the Internet is steadily increasing. However, this study reveals that despite being a common source of news for Ghanaians, Ghanaians trust it less than traditional media.	Studies the dissemination of fake news and legal provisions for people who disseminate fake news. Discusses the perceptions of Ghanaians of political parties and politicians and their role in the dissemination of false information. Highlights the perceived role of the government in combating the spread of false information, hate speech and other problematic content.

[51] Appiah-Nyamekye Sanny J and Selormey E (2020) Double-edged sword? Ghanaians see pros, cons of social media, want access but not fake news. Afrobarometer Dispatch No. 366. Available online at:https://media.africaportal.org/documents/ad366-fake_news_and_social_media_in_ghana-7june20.pdf

2.2 Objectives of the Study:

The study will find out how users react to different kinds of problematic content and their intention behind sharing it. It will be aimed at finding the different ways in which individual users react to problematic content and the possible impact on them as well as the society as a whole. It will enable us to understand the rationale behind different kinds of users' inadvertent contribution (reposting, forwarding, believing, reacting) in the spread of problematic content, which will help in recommending practices for users on how to consciously avoid the spread of problematic content on social media platforms. Specifically, the study has the following objectives.

- Check if misuse of social media platforms (through problematic content - incorrect, sensational, provocative, divisive, hateful, and misleading information) may cause different levels and kinds of harm (in the form of increasing polarisation, creating trust deficits, instigating negative reactions, promoting abusive behaviour, etc.) to different users, particularly new as well as to non-users.

- Understand diverse perspectives of different users (difference in users' ethnographic, cultural and political scenarios, geographic location, demographic profile, economic condition experience, psychological and cognitive variables, extent and duration of exposure to the problematic content, etc.) towards problematic content.

- Devise targeted factor-specific intervention plan to help users engage more critically or consciously, to avoid problematic experiences arising due to exposure to problematic content.

2.3 Research Methodology

For meeting the aforementioned objectives, a two-fold strategy will be adopted. CUTS will conduct (a) an impersonal user perspective study for meeting the first two objectives, and (b) an in-person simulation exercise and laboratory study for meeting the last objective.

- **Impersonal user perspective study:** For this study, a user perspective survey is required. For this purpose, CUTS will design a survey questionnaire which will be aimed at understanding the reaction of the users towards problematic content and the level of harm it may cause to them. Accordingly, a Logical Framework Analysis will be prepared which will be used to prepare the questionnaire. Using the prepared questionnaire, a survey of 2000 users each in the three stated countries: India, Kenya and Ghana will be conducted.

- **In-person simulation exercises and laboratory study:** To gain a deeper insight on the perception and reaction of users to problematic content and to device a targeted factor-specific intervention plan, an in-person study will be conducted which will involve simulation exercises in a laboratory environment. Tools for this will be created using the findings from the impersonal user perspective study.

2.4 Way Forward

Various studies classify problematic content, identify the process of its spreading, its impact, and propose solutions for curbing it. However, reaction of users towards problematic content and the possibility of harm have been less explored. The Operation Strategy Note highlights the need for doing a study in India, Kenya and Ghana which can help in bridging the aforementioned research gap. To fill the research gap, three objectives have been stated and to meet the objectives, the methodology has also been identified. Going forward, CUTS will prepare a Logical Framework Analysis model which will further be used to prepare survey questionnaire. Based upon the findings from the survey, CUTS will also conducted in-person simulation exercises and laboratory study which will help in devising a targeted factor-specific intervention plan which will help users engage more critically or consciously with problematic content.