"It Matches My Worldview": Examining Perceptions and Attitudes Around Fake Videos

Farhana Shahid*
Cornell University
Ithaca, United States
fs468@cornell.edu

Srujana Kamath*
Independent Researcher
Mumbai, India
srujana.kamath@gmail.com

Annie Sidotam
Independent Researcher
Hyderabad, India
anniesowmya@gmail.com

Vivian Jiang
Cornell University
Ithaca, United States
vmj5@cornell.edu

Alexa Batino
Cornell University
Ithaca, United States
afb75@cornell.edu

Aditya Vashistha
Cornell University
Ithaca, United States
adityav@cornell.edu

ABSTRACT

We present a qualitative study with 36 diverse social media users in India to critically examine how low-resource communities engage with fake videos, including cheapfakes and AI-generated deepfakes. We find that most users are unaware of digitally manipulated fake videos and perceive videos to be fake only when they present inaccurate information. Few users who know about doctored videos expect them to be of poor quality and know nothing about sophisticated deepfakes. Moreover, most users lack the skills and willingness to spot fake videos and some were oblivious to the risks and harms of fake videos. Even when users know a video to be fake, they prefer to take no action and sometimes willingly share fake videos that favor their worldview. Drawing on our findings, we discuss design recommendations for social media platforms to curb the spread of fake videos.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Fake videos, Misinformation, India, Global South, ICTD, HCI4D

ACM Reference Format:
Farhana Shahid, Srujana Kamath, Annie Sidotam, Vivian Jiang, Alexa Batino, and Aditya Vashistha. 2022. "It Matches My Worldview": Examining Perceptions and Attitudes Around Fake Videos. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3491102.3517646

1 INTRODUCTION

Social media platforms have seen an unprecedented growth in users based in the Global South [94]. In parallel, videos have also seen the fastest growth among all formats and are the major driving force for generating more engagement and revenues on social media [28, 41]. India, the focus of our study, is one of the largest markets with steeper consumption of video ads and other video sharing services [102]. Around 97% of Indians, who are connected to the Internet, watch videos online [89]. The availability of affordable smartphones and cheap mobile data along with the emergence of high-speed 4G network have made videos more popular and accessible to the masses in India—particularly to low-literate social media users who are able to consume videos easily compared to textual content [73].

However, this has proven to be a double-edged sword as many social media users tend to trust content in videos more [93]. Fake videos often exploit the human tendency of *"Seeing is Believing"* [10, 81] to skew information and perceptions, which often lead to increased levels of social schism and political polarization. For example, doctored videos and old videos used out-of-context have led to civic unrest and lynchings in India [45, 70]. These risks and harms have been more pronounced during the COVID-19 pandemic, where fake videos spreading false health advice and anti-vaccine propaganda have led to increased hospitalizations and deaths, resulting in overburdened health systems [87].

Traditionally, people have used common video editing tools and techniques (e.g., photoshop, lookalikes, slowing down video frames, removing or stitching segments) to create cheapfake videos. One classic example of cheapfakes is the viral doctored video of the US House Speaker Nancy Pelosi, where the video was slowed down to make her speech appear slurred and drunk. Similar technique was applied on a video of a senior politician in India to make him appear intoxicated [63]. However, among the arsenal of fake videos, AI-generated deepfake videos are considered to be the most harmful, as they are becoming increasingly common and convincing. Deepfakes closely resemble the original videos and may trick untrained human eyes to consider them as real [90]. Journalists frame deepfakes as a pernicious technology that destabilizes trust in video evidence [68] and social media news [104], undermines a shared sense of sociopolitical reality [18], enables online abuse and harassment of women [108], and blurs the acceptable dichotomy between real and fake [111]. The risks emanating from fake videos generally, and deepfakes specifically, are more prominent for new users in the Global South who have limited training to assess information credibility [93] and are less-aware of emergent digital threats [24].

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI '22, April 29-May 5, 2022, New Orleans, LA, USA
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9157-3/22/04...$15.00
https://doi.org/10.1145/3491102.3517646

Despite the multifaceted threats of fake videos at large, most research to date heavily focuses on deepfake videos in the contexts of Global North. For example, scholars have examined human accuracy in spotting deepfake videos [17, 61], the harms they cause to individuals and society [34, 35, 55, 58, 68, 104], and people's concerns around the discovery and propagation of deepfakes [3, 4, 18, 30, 111]. However, despite large differences between new social media users in the Global South and mainstream users in the Global North (e.g., the US), little work has been done to understand how emerging social media users in the Global South engage with fake videos [9] and deepfakes [99].

To fill this critical gap, we conducted semi-structured interviews with 36 diverse social media users from rural and urban India, who vary in terms of literacy, digital skills, and social media experience. We sought to understand: **RQ1:** How do regular social media users perceive and interact with fake videos? **RQ2:** How do they evaluate the risks and harms of fake videos and what they consider necessary to curb its spread?

We used video probes (one real and three deepfake videos) to guide our interviews and to elicit spontaneous responses from our participants and observe if they could identify the fake videos and if so, how. Our analysis revealed several alarming findings on how social media users perceive and engage with fake videos and what their associated mental models and threat models are. Our participants had varying perceptions of what they consider as fake videos. Several participants were unaware that videos can be doctored or edited, and most knew nothing about deepfakes. A few participants, who were aware of digital manipulations of videos, expected these videos to be of poor quality and easily recognizable. In fact, most participants' understanding of fake videos was only limited to inaccurate or false information presented in video format. In addition, most participants lacked the necessary skills, awareness, and willingness to identify fake videos and assess authenticity. Even when participants knew a video to be fake, they rarely reported it and, sometimes willingly shared it. While some participants felt that fake videos could do little or no harm, many described how these videos are polarizing and divisive. The participants also expressed diverse opinions in describing the role individuals, social media platforms, and governments should play to contain the spread of fake videos.

Drawing on these findings, we synthesize key takeaways for HCI and social computing researchers interested in understanding the misinformation landscape in low-income settings in the Global South. We discuss several design recommendations to foster meaningful dialogue and civil discourse on social media, for example, by designing online educational and training programs to inform mental and threat models of new users about the emergent risks of fake videos, using automated and explainable AI-based credibility indicators to label fake videos, and making the reporting feature more transparent and accessible. In summary, our work makes the following contributions:

- A qualitative study that critically examines how diverse social media users in India identify and interact with fake videos, explicating their engagement and propagation practices and their perceptions around the risks and harms of fake videos.
- A range of design recommendations that could mitigate the spread of fake videos on social media.

2 RELATED WORK

Videos are becoming the dominant modality through which misinformation and fake news propagate on social media [28, 67]. By tapping into video's popularity and virality, false narratives and propaganda get quickly amplified on platforms like YouTube [47, 101] and TikTok [113], posing a serious threat to new social media users, most of whom are emerging in the Global South.

Despite the high risks of misinformation in the Global South, most of the existing work to date has focused on misinformation in the Global North, including its prevalence [40, 53], diffusion [7, 92], and people's interactions with it [25, 26, 59]. For example, scholars have examined factors that shape people' perceptions of information credibility, such as trust on the news source, interpersonal trust in one's social network, and perceived homogeneity of the news content [39, 103]. A recent study notes that users' perceived value of information to spark conversation and their desire for self-expression and socialization fuel the sharing of misinformation among the users based in the Global North [26]. Besides, partisan bias [91] and users' tendency to share content based on homogeneity [11] reinforce confirmation bias and lead to the propagation of misinformation and rumors.

Prior studies show that people's intention to share misinformation relates to their demographics and individual characteristics, such as age, gender, ideology, and biases [8, 20, 64]. However, most of the work so far focuses on users based in the Global North, resulting in growing concerns around the extent to which the findings from the Global North generalize to misinformation and fake news in other cultures and contexts [62, 79]. This has prompted several scholars to study socially and culturally diverse social media users in the Global South and how they interact with and propagate misinformation and fake news. We now situate our research in a body of related work examining misinformation and fake videos in the Global South.

2.1 Misinformation in the Global South

The politically charged environment in the Global South is a breeding ground for the proliferation of misinformation, fake news, and rumors [37, 86]. A recent study by Jakesch et al. [52] shows how Indian political parties orchestrated and centrally organized their supporters on WhatsApp and Twitter to engineer social media trends during elections. Recent work from Akbar et al. [5] describe how the COVID-19 pandemic led to a plethora of misinformation in India, partly owing to its existing polarized and divisive information environment. They observe that tweets from politicians tapped into communal prejudices and affective responses of the mass, and ignited hatred and coordinated attack against particular religious groups. Several scholars have noted that the fact-checking resources in the Global South are too scarce to keep up with the wild pace of misinformation. For example, recent work from Haque et al. [44] about the existing fact-checking practices in Bangladesh shows that journalists believed fact-checking to be the responsibility of third-party fact-checkers rather than that of news media. On the other hand, voluntary fact-checkers reported facing various difficulties in verifying online news due to limited resources and sparse infrastructural support. In such muddy information environment, people often adopt different informal methods to assess information credibility. Recent work from Chandra and Pal [22] shows that different stakeholders in a marketplace in India leveraged their community bonds and collective knowledge to make sense of the rumors to interpret unfamiliar and ambiguous situations. Similarly, work from Sultana and Fussell [97] describes that rural villagers in Bangladesh relied on their religious faith, local beliefs and myths, and social justice sensibilities to fact-check COVID-related misinformation. This line of work demonstrates that socio-cultural values, local beliefs, and informal offline communication not only shape people's perceptions of misinformation but also influence collective fact-checking practices that often differ from professional fact-checking methods that prioritize authenticity over anything. These studies prompted us to dig deeper and investigate how diverse social media users in India perceive and engage with fake videos considering their rapid propagation and harmful impact on the society [98]. We now present scholarly work that looks into different aspects of people's engagement with fake videos.

2.2 Engagement with Fake Videos

When it comes to how people interact with fake videos, one of the key questions is: can people detect them? Prior studies show that even experienced social media users struggle to identify fake videos [16]. As AI-generated deepfakes and cheapfakes become more common and convincing, several scholars have investigated how people and machines can detect fake videos. Research shows that most people lack the ability to spot high-quality deepfake videos and are only able to detect the low-quality ones to some extent [17, 61]. Apart from video quality, prior knowledge, familiarity with the subject, and visual cues (e.g., face, hair, eye, skin tone) help people in detecting deepfake videos [43, 56, 99]. For example, people are usually better at recognizing deepfake videos of known politicians or celebrities than the ones of unknown people [56]. In fact, the susceptibility to deepfake videos is amplified by the general lack of awareness around fake videos. A survey among students in a public southeastern U.S. university revealed that almost half of the students were not aware of deepfake technology [30]. Even though many considered this technology to be harmful, some also found it amusing [30].

Almost all the research discussed thus far investigates deepfake videos in the context of Global North. In contrast, there is little work that examines how social media users in the Global South engage and deal with different types of fake videos and how users in low-income, rural settings engage with deepfake videos specifically. The work that is closest to ours is from Tahir et al. [99] in which the researchers conducted a user study with university students in Pakistan to examine how they detect deepfakes and designed a training program to improve their skills. Our work contributes to this line of work by engaging diverse social media users in rural and urban India to examine: (1) How do they perceive and interact with fake videos, and (2) How do they evaluate the risks and harms of fake videos and what they consider necessary to curb its spread. To our knowledge, ours is one of the earliest exploratory work to critically examine the diversity in social media users' perceptions of and interactions with fake videos in low-resource settings in the Global South.

3 METHODS

We conducted an in-depth qualitative study with social media users in urban and rural India from July to August 2021. To recruit urban participants, we reached out to our personal networks and used snowball sampling to have people from different demographics ranging from low to middle income, digitally novice to literate social media users. To recruit rural participants, we partnered with a grassroots organization focusing on rural development with significant presence in several regions in India. An organization staff member approached their primary contacts in the villages, explained to them the purpose of our study, and gave us the contact information of those who were interested in speaking with us. All our interactions with the participants took place online to ensure safety of our participants due to the ongoing COVID-19 pandemic. Our study protocol comprised two phases. In the first phase, we sent four videos to the participants via WhatsApp a few hours before the interview and asked them to watch the videos before our interview. In the second phase, we conducted a semi-structured interview either via phone or Zoom calls as preferred by the participants. Following the interview, we compensated each participant with INR 200 (~USD 2.70). Our study was approved by the Institutional Review Board (IRB) at Cornell University.

Sending Video Probes. HCI researchers have advocated using cultural probes, artifacts, and technology probes to understand the needs of the underserved communities and the complexities of their everyday lives [50, 110]. These probes help collect "inspirational data" that not only stimulate the design process but also elicit responses that provide a glimpse of the workflows, processes, aspirations. For example, Okolo et al. [77] used video probes to explore the knowledge and perceptions of AI among low-income community health workers in India who had limited or no prior experience with AI. Drawing on this prior work, we carefully curated a set of one real and three deepfake videos, and used them as exploration artifacts to elicit people's understanding of and experiences with fake videos generally and deepfakes specifically. We chose to use video probes since we were unsure if our participants have knowledge of and experiences with deepfakes in particular.

All the videos were of same length (~10 seconds) and showed a person of Indian origin speaking in Hindi. Among the three deepfake videos, two were prepared using *lip-sync* and another using *face swapping* deepfake technologies. Using lip-sync, a person's lip movements are modified to match a new audio [2]. Face-swapping transfers a source face to the destination while preserving destination's facial expressions and movements [80]. We opted to create deepfake videos ourselves since we could not find good quality deepfake videos of people of Indian origin in Hindi. The three deepfake videos featured Indian cricketer Sachin Tendulkar (lip-synced), Indian actor Pankaj Tripathi (lip-synced), and a television news anchor (face swapped). To understand if participants could differentiate between fake and authentic videos, we also added a real video featuring a popular Indian actor, Shah Rukh Khan. We avoided using any politically or religiously polarizing content so as to not bias our participants. We intentionally kept the videos brief (only 10 seconds of duration) and used neutral, non-informative audio so that participants could focus more on visual cues rather than the audio content.

Conducting Semi-Structured Interviews. After the participants confirmed watching the videos, we conducted semi-structured interviews with them. The majority of the interviews were conducted in Hindi and a few in English, following the preferences of the participants. At first, we read an informed consent script to the participants and requested their verbal consent. When agreed, the participants were asked to summarize their thoughts about the videos we sent them. We then discussed their opinions about the videos before digging deeper into their existing knowledge of fake videos, specifically deepfakes and cheapfakes. Our subsequent questions sought an in-depth understanding of how they perceive fake videos, their knowledge about current technologies to create fake videos of varying qualities, their interactions with perceived fake videos, how they evaluate the benefits and harms of fake videos, and the measures they consider necessary to tackle the onslaught of fake videos. At the end of the interview, we debriefed our participants which videos were fake and which one was real. We requested them not to share the videos with others. Since we had no way to confirm whether the participants would delete the videos or not share those with others, we intentionally chose neutral content for all our videos to avoid any potential risks or harms. After each interview, we revised our questions to add new probes, stopping when participants' responses reached saturation. Each interview lasted approximately 40 minutes, and was audio-recorded with the consent of the participants.

Data Collection and Analysis. In total, we collected around 24 hours of audio recordings from the interviews. Audio recordings were translated into English and transcribed. We then used inductive thematic analysis [42] that allows key themes to emerge from the raw data through repeated examination and comparison. At first, two authors individually coded the same interview and then came together to discuss discrepancies. We avoided using any presupposed codes and instead let the codes emerge from our data. We continued to code separately and came together after each interview until we had reached intercoder reliability and codebook stabilization. Following this, the authors separately coded the remaining transcripts. Throughout the analysis, we held multiple discussions to iteratively refine the codes and reconcile disagreements through peer-debriefing [31] to ensure that our themes comprehensively represent the data. We also conducted member checks [14] to ensure that participants' subjective meanings, actions, and social contexts were authentically represented. Multiple passes through the data resulted in a total of 90 codes (e.g., out-of-context fake videos, fake videos are the most harmful, fake videos incite violence) and four high-level themes (e.g., experiences with fake videos, perceived harms of fake videos).

Participant Demographics. Table 1 lists the participants' demographics. We interviewed 15 urban and 21 rural participants aged between 18–65 years. The urban participants were from major cities in Gujarat, Karnataka, Telangana, Maharashtra, and Andhra Pradesh. Majority of the urban participants (N=14) had at least a bachelor's degree. They came from a wide range of professions, including teachers, engineers, health workers, social workers, journalists, and consultants, among others. Only four participants had a technology related background (e.g., CS student, engineer, software developer). Two urban participants were new smartphone users (with less than one year of experience) and the rest were using smartphones for over five years.

On the other hand, all rural participants were recruited from different villages in North India. About half of them had less than 12 years of formal schooling. We interviewed housewives, farmers, social workers, teachers, students, shopkeepers, private job holders, and unemployed folks. More than half of them used a smartphone for less than 5 years and about one-fourth were new to social media with less than a year of experience. All our participants used WhatsApp, YouTube, and Facebook, with an exception of urban participants who also used Twitter, Reddit, and TikTok.

Positionality. All authors except two are from countries in the Global South. Among them three are from India and have lived experiences of interacting with the kinds of fake videos many of our participants referred to. Even though the authors have several years of experience conducting fieldwork with low-income communities in India, they acknowledge that they themselves are not low-income and have not lived in rural regions for extended periods of time. Their relative privilege provides them with certain advantages that most participants in this study do not hold. However, the authors' past engagements with urban and rural communities helped elicit in-depth responses from the participants and build a nuanced understanding of the underlying socio-cultural norms that impact participants' engagement with fake videos. All authors view HCI research from an emancipatory action research mindset, which led them to examine how people from socially and culturally diverse backgrounds experience and interact with fake videos online and how this knowledge might inform design recommendations that could potentially curb the spread of fake videos in India.

4 FINDINGS

We begin by exploring participants' knowledge of and experiences with fake videos, including their encounters with deepfake and cheapfake videos (Section 4.1). We then discuss how the participants deal with fake videos (Section 4.2), before diving deeper into the potential benefits and harms of fake videos (Section 4.3), and their opinions on measures they consider important to prevent the spread of fake videos (Section 4.4). While describing our findings, we refer to our participants by pseudonyms.

4.1 Perceptions of and Experiences with Fake Videos

The probes that we shared with our participants helped us explore their opinions on fake videos. Most participants did not observe any audio-visual inconsistencies in the video probes. Only a few of them noticed some of the videos to be fake, without our probing. Even when we informed the participants that some of the shared videos were fake, they struggled to point out which ones were fake and which was real. All participants except one failed to correctly identify all the videos. About one-third of the participants believed all four videos to be real. Nearly one-fourth of the participants incorrectly classified the real video as fake.

Table 1: Demographics of the participants.

Demographics	Urban	Rural
Gender	Female: 7, Male: 8	Female: 11, Male: 10
Age (years)	Min: 19, Max: 65, Avg: 35, SD: 14	Min: 18, Max: 45, Avg: 28, SD: 9
Education level	Middle school: 1, Bachelors: 9, Masters: 5	Middle school:1, High school: 3, Bachelors: 12, Masters: 2
Smartphone use	1 yr: 2, >5 yr: 13	1 yr: 3, 2-4yr: 8, >5yr: 4

4.1.1 Perceptions of Fakeness. When we sought opinions from our participants about what they consider to be "fake" videos, they offered us a wide range of perspectives. For many of our participants, fake videos contained *"obviously fake and scripted content."* When asked for examples of fake videos, they often mentioned videos of astrological predictions, TikTok memes, pornographic videos, and prank videos, among others. Some rural participants had no idea about the existence of fake videos and thought videos could hardly be tampered with. Most of the new social media users in our sample often tended to believe almost anything they see online, had limited knowledge about fake videos, and believed videos to be a *"window of reality."* Some of them felt that they could never receive fake videos as they are only connected to their *"trustworthy family and friends"* on social media, assuming that fake videos are shared only by malicious actors and adversaries.

> *"Apart from my personal contacts, no one knows my number. Also I am not part of any external WhatsApp group. So, I never receive any fake videos."* (Krittika, Rural, Female, 40 years)

Like Krittika, many participants tended to blindly trust content shared by their personal contacts. Some participants believed the nature of WhatsApp groups (e.g., family groups, interest-based groups, public groups, work-related groups) to influence whether they can expect to be exposed to fake videos in that group. For example, Smita, a 35-year-old teacher in rural region, believed that she would never see a fake video on WhatsApp and Facebook because she only participates in education-related groups and discussions.

4.1.2 Misleading and Out-of-Context Fake Videos. About one-fourth of our participants highlighted how they considered fake videos as those that spread suspicious or unverified information and rumors. For example, Apoorva, one of the urban participants, gave the example of an anti-vaccine video as a fake video in which it was shown that *"a boy died after taking COVID vaccine."* Almost one-third of the participants perceived videos to be fake when they depicted real incidents but were presented out-of-context with misleading narrative either to exaggerate an event or provoke people. Ashish shared an example to illustrate this point:

> *"A couple of years ago, the JNU student movement stirred the whole country. There was a video where protesting JNU students were chanting, 'we want freedom.' This was turned into a propaganda saying 'they want freedom from the country', whereas, the students actually implied they wanted freedom from poverty, unemployment, and as such. We see these kinds of videos almost everyday."* (Ashish, Urban, Male, 27 years)

Some participants felt that such out-of-context fake videos which are in unfamiliar language or feature unfamiliar people are the most difficult to verify. A few participants expressed that fake videos with misleading narratives that are inline with one's political or religious beliefs are also difficult to spot, suggesting that people are susceptible to fake videos that affirm their political or communal biases. This finding is inline with prior studies that report how people are likely to consider those news sources as reliable that support their political affiliations and biases [52, 71, 72].

A few participants added that they would also consider defamatory videos as fake. They pointed that monetary gain, blackmailing, revenge, and political and business rivalry might lead to the circulation of such fake videos. Additionally, several rural participants (N=9) reported coming across a pattern of fake videos that misleadingly claimed *"If you do X, Y will happen."* Most of these videos had a religious undertone and a few even featured monetary schemes trying to financially scam low-literate rural users. Tarun explained:

> *"I came across this fake video that said, 'this is where the Goddess was born. Share this video with 15 people and you will be blessed. Otherwise, you will be the harbinger of bad luck.'"* (Tarun, Rural, Male, 22 years)

4.1.3 Digitally Modified Fake Videos. More than half of our participants (N=23) defined fake videos as those with *"digital modifications."* When we asked them how these videos are digitally modified, they reported that the audio or faces in videos are often replaced and sometimes multiple videos are added together or some words/scenes are edited out of the video to propagate a false narrative. They expected these videos to be of poor quality (e.g., having blurred or distorted facial features or asynchronous lip movements) and easily discernible.

More than three-fourths of our participants were unaware of deepfake videos. Only four participants reported having some prior knowledge about deepfake videos, out of which two were unaware of the term *"deepfake"* but could recall encountering some instances of deepfakes when we explained the term to them. Two participants understood deepfakes to be videos produced using a technique which helps replace an individual's face or other features (e.g., clothes, speech) in a way that the edited video looks similar to the original. One participant defined deepfake to be *"computer-generated"* and associated it with *"Google's Deep Dream"*—a tool that enables creation of artistic visual content through Human-AI collaboration. When we prompted the participants to give examples of deepfake videos that they might have encountered earlier, Anushka, a 27-year-old urban woman, recounted an Instagram video with AI-generated human figures that looked so close to real humans that she struggled to differentiate between real and fake figures. A few participants added that they had heard that deepfake technology is used to create pornographic content, as is widely reported in the popular press [19]. Others struggled to give a concrete example, suggesting that even those who know about deepfakes have limited ability to spot them. In contrast, many participants could recall examples of cheapfake videos that they encountered previously. For instance, one of the participants shared that his friend showed him a cheapfake video, where a cartoon with a human face was dancing around. Adarsh, a 30-year-old rural man, shared how he was able to detect a video to be fake because *"speaker's lips were not synchronizing with the audio."* However, unlike Adarsh, many rural participants trusted even such obviously inferior cheapfake videos, attributing glitches, pixelated frames, and audio mismatch to poor network connectivity.

Taken together, participants had varying perceptions of what comprises fake videos, ranging from lack of awareness about the existence of fake videos to superficial knowledge about sophisticated deepfakes. This suggests that not only users are likely to interact with fake videos differently, but also they have varying susceptibility to fake videos. For example, the participants who considered false information presented in video format as *fake videos*, did not consider the possibility that videos could be digitally manipulated, making them more susceptible to digitally modified fake videos.

4.1.4 Identifying Fake Videos. After hearing participants' perceptions of fake videos, our interviewers explained deepfake videos to them. After learning about these technologies, some participants (20%) felt that they lacked the skills, capabilities, and experience to detect sophisticated fake videos. They all pointed that due to advancing technologies they would hardly be able to detect any subtle cues from such *"well-made"* fake videos. Besides, some participants mentioned that most of the time they consume videos *"passively rather than engaging with them actively"* in the sense that they pay more attention to whether the content matches their worldview instead of considering whether the video itself is manipulated, used out-of-context, or contains misinformation. They reported that when they have some expertise or prior knowledge about the topic depicted in fake videos then they might be able to spot them better.

When we asked participants to share their thought process on how they assess the credibility of a video, about one-third of them reported never using any technological tools or fact-checking sites. Instead, they relied on their intuition. In fact, most of our participants knew nothing about fact-checking, including how fact-checkers verify information and where to find such resources. Although one-fourth of our participants mentioned that they carefully look for visual cues and re-watch the videos multiple times when they find something suspicious, many others mentioned that they do not pay attention to visual cues and are not actively thinking about *"spotting fake signs."* This is in contrast with observations from a prior research study [99], where people mentioned relying on visual cues to spot fake videos. The participants in our study reported that they focus more on the *voice*, *audio content*, and *context of the information* rather than relying solely on the visual inconsistencies. Aditi explained how she often ignore visual cues:

> *"Most of the time I don't really pay attention to visual clues. I just hear what the person is saying in the video. If they are saying something very controversial then probably I try to assess if the information is correct or not. Otherwise, I just watch."* (Aditi, Urban, Female, 34 years)

Participants reported that focusing on audio might help them detect those instances of fake videos in which a celebrity's voice is replaced with a different audio. This reaffirms that for many users like Aditi, perceptions of fake videos hardly stretch beyond audio manipulation and suspicious information presented in the video, let alone considering the possibility of AI-generated deepfakes. A more suitable approach could be to pay attention to both audio and visual cues, especially to recognize high-quality deepfakes. A couple of participants commented that watching a real video of the same subject would help them recognize corresponding deepfake videos, where the face has been swapped. This is inline with findings from Khodabakhsh et al. [57] which show that the participants who saw the actual biometric reference video of the person targeted in deepfakes were found to be better at spotting the deepfake videos.

A few participants reported that when they doubt a video to be *suspicious*, they go through the comments to see if others also considered the video to be fake. On digging deeper, some participants mentioned that this strategy often resulted in more confusion, particularly when comments call out the same video as both fake and credible, making it hard to decide what to believe. In such cases, the participants relied on the credibility of the source from where the video came to assess information credibility. Even though the majority (86%) of the urban participants were using smartphones for many years, only one-fourth of them were proactive about spotting fake videos and seeking correct information, indicating that longer experience with technology does not necessarily translate into healthy information behaviors. On the contrary, none of the rural participants relied on online resources to assess the credibility of videos they found suspicious. They felt that they lacked digital skills that could be helpful to spot fake videos.

4.2 Interaction with Fake Videos

Following participants' varying perceptions of fake videos, we queried how they tended to respond when they suspected a video to be fake. Our participants described various actions they took on such occasions.

4.2.1 Ignoring Fake Videos. Most of the participants reported that they simply ignore the videos they doubt as fake, particularly when the videos are in an unknown language, forwarded multiple times on WhatsApp, sent by a person who routinely shares fake videos, or when the topics do not align with their interests. Some commented that they don't care much about fake videos on social media and hardly pay any attention while watching them. Ashish further unpacked this indifference:

> *"When I come across a fake video, my reaction is it's just another source of wrong information. So, I prefer to do nothing about it because I have never been very active in these matters and what's the point anyways."* (Ashish, Urban, Male, 27 years)

Some participants felt that there are a lot of fake videos online and given the abundance of such videos, they considered it best to not waste any time on them *"because people keep sharing such stuff."* Aditi elaborated:

> *"Even if I do come across fake videos, I don't really spend my time looking at it, thinking about it, or even wondering why people do that. It's just a waste of my time. And I think most of the people also know it's fake. So, I just ignore it."* (Aditi, Urban, Female, 34 years)

Participants like Aditi did not take any action even after doubting videos to be fake simply because they assumed other people would know the videos were fake or thought that propagation of fake videos could not be prevented. Such passive attitudes could lead to the tragedy of the commons and result in an increased consumption of fake videos often with undesirable consequences.

4.2.2 Deleting and Blocking. A few participants (particularly WhatsApp users) deleted the videos they perceived to be fake. The participants offered different reasons for why they opted to delete videos they perceived as fake. Some of them only wanted to keep useful videos on their devices, a few worried about accidentally sharing such videos with others, and some who shared their devices with family members were concerned that their family members might be exposed to such videos. In line with the scholarship on shared phone use in the Global South [4, 106], these findings offer how shared devices among family members could lead to inadvertent sharing of fake videos. Srikant described:

> *"My kids often use my phones. So I delete anything unnecessary. Because if the kid ends up sharing something in my absence and I find it later when others comment or react, that would be embarrassing. People might wonder why I shared such thing. That's why I don't keep videos on my phone."* (Srikant, Rural, Male, 40 years)

A few participants reported that they blocked personal contacts on WhatsApp who routinely shared fake videos with them. One participant left a local WhatsApp group where fake videos were routinely circulated. This shows that when it comes to dealing with fake videos, many people simply opt to get rid of it on their end. Though these measures protect the individuals from being exposed to fake videos, they do not necessarily stop fake videos from spreading.

4.2.3 Leaving Comments and Discussion. Even though many participants mentioned that they do their best to not engage with fake videos, a few participants felt that it was their *"moral and civil duty"* to do more. They reported leaving comments on videos they perceive as fake to make others aware of it. Aditi described:

> *"If I encounter a fake video, I would leave a comment saying that it was not authentic, how other people were indiscriminately forwarding it, and just share my thoughts about it."* (Aditi, Urban, Female, 34 years)

A couple of participants found it helpful to discuss videos they find suspicious with family and friends. Three women mentioned that they often show suspicious videos to their husbands or brothers who have more *"digital and worldly experience"* to get their perspectives. This shows that male family members played a key role in forming and shaping female family members' opinions of fake videos. A few participants mentioned they had some friends, who often sent them fake videos and so, they asked these friends to share videos only after vetting the content. They added that they prefer to engage in these discussions as there are chances people might be sharing fake videos without being aware of them. Although these online and offline discussions have the potential to inform people, many participants reported how they were less inclined to do so out of the fear of hurting other's feelings and to avoid any backlash.

4.2.4 Reporting Fake Videos. When we probed our participants on whether they tended to report fake videos, we received mixed responses. More than half of our participants were reluctant to report fake videos and offered various reasons for their unwillingness. A majority of them considered reporting to be *ineffective* as some of them had prior experience where reported videos were not removed from the platform, and no actions were taken against the users who shared the fake videos. This is inline with findings from prior studies that discuss how people didn't receive the expected outcome after reporting abusive content and thus concluded that reporting doesn't work [76, 105]. A couple of participants felt that social media platforms take many days to take action on flagged videos and since these videos spread *"so fast that reporting them have zero practical impact."* A few expressed their frustration, saying that on social media people can choose to share *whatever they like* even if it is fake and hence, they preferred to not take the trouble to report anything. In addition, some participants were unaware about how to go about reporting a video and what would follow if they tried to do something. A few participants avoided reporting because they were unsure about what kinds of explanations and evidence they would be asked to provide:

> *"If somebody reports a video, I'm not sure what happens next. How is it dealt with? I think there should be more transparency around that, for example, if I reported a video, they need to tell me how they would follow up with me on what happened and what action they will take."* (Adarsh, Rural, Male, 30 years)

Only a few participants admitted reporting videos they perceived to be fake. Often such videos contained polarizing content, harassment, pornography, and vulgar language. A handful of participants also shared that they reported social media accounts and YouTube channels that they knew were spreading fake videos containing communal or political propaganda. The participants stressed the necessity to adopt a community-based approach to pointing out potential fake videos so that the platforms could block such videos and take actions against the offenders. They expected immediate removal or censoring of reported content as well as strong sanctions on users and accounts spreading fake videos, ranging from temporarily banning them to suspending them indefinitely. But such expectations were rarely met. This shows that even though social media platforms provide the option to report potentially fake content, this feature is often underutilized owing to lack of proper knowledge about its underlying processes and what ensues after reporting.

4.2.5 Sharing Fake Videos. We found that participants were split in their attitudes toward sharing videos they found to be suspicious. Many participants were comfortable sharing *obviously* fake videos, for example, cheapfakes created for entertainment. This included funny cheapfakes of politicians and actors or user-generated cheapfakes using popular applications like Adobe After Effects [1] and WOMBO [109]. Some participants expressed that they are fine sharing fake videos when the broader message in the video aligns with their beliefs and perspectives, highlighting how that was a sufficient reason for sharing. This is consistent with the findings from prior studies conducted both within the Global South [100] and the Global North [69] which report that people are willing to share fake news that aligns with their preexisting beliefs. Smita described:

> *"If my worldview matches with a video, I will definitely share it even if it's fake. Whatever they are saying, whatever views they hold doesn't matter. If I like it, then I will share it."* (Smita, Rural, Female, 35 years)

These participants were oblivious of the harm that might arise from sharing fake videos and some harbored beliefs that videos could do no harm. On the other hand, some participants claimed that they would not share fake videos with others even if the videos support their communal and political beliefs. They felt that sharing fake videos could not only harm other people but also ruin their social image. They emphasized that just because they believe in something that *"doesn't necessarily mean other people should follow that."*

Taken together, these findings show that users not only engage with videos they doubt as fake differently but also their perceptions of the potential harms of fake videos vary, which we explore next in more detail.

4.3 Harms of Propagating Fake Videos

Many participants (N=16) considered videos to be the most harmful among all modalities through which misinformation and fake news spread. They believed that the effect of visual experiences lasts longer than that of reading textual information which also demands more cognitive load. Some participants pointed that many low-literate social media users who are new to online information environments consume videos heavily and prefer to engage with audio-visual content more than textual content. They felt that these users are the most susceptible to the risks and harms of fake videos since they might lack the skills to assess the credibility of videos.

4.3.1 Societal Harm. When we probed our participants about their thoughts on how fake videos could cause harm, about one-third of them felt that fake videos lead to violence, communal tension, and chaos. They also emphasized how fake videos are weaponized to drive political vendetta and seed religious hatred. Ridhima shared:

> *"There are many fake videos on sensitive topics. These are dangerous and lead to protests and riots. Many people do not have any clear understanding of the things on the ground. They believe these fake videos, assume that they can't be tampered, and share them widely. All of these further exacerbate the situation."* (Ridhima, Urban, Female, 32 years)

Some participants labelled fake videos to be a *"danger to democracy"* and felt that they undermine freedom of speech, violate human rights, and set the ground for communal hatred. They felt that fake videos turn people against each other and thus hinder societal progress. Vikram described the harms of fake videos:

> *"People usually have strong opinions on caste, religion, nationalism, and communal issues. Whenever they watch some fake videos on the Internet spreading wrong information on these sensitive topics, it might hurt their sentiment and provoke them to hate, bash, and even cause physical harm to other groups of people having different opinions."* (Vikram, Urban, Male, 22 years)

Participants like Vikram felt that political parties reap the greatest benefit by creating fake videos and do so to increase their follower base, defame the opposition, and advance their propaganda. Around one-fourth of the participants mentioned that sharing fake videos often lead to *"misinformation cascades"* as people keep forwarding these videos without much thought, often resulting in panic and mistrust within the society. Aditi elaborated:

> *"People can be easily misled and manipulated by fake videos as they do not know how to spot them. It's dangerous for people to simply forward fake videos without much inspection and soon it results into a chain of misinformation – as one person forwards the video to another and it goes on and on."* (Aditi, Urban, Female, 34 years)

4.3.2 Harms to Individuals. A few people who knew about deepfakes considered them to be a huge blow to individual's privacy. Anushka, a 27-year-old woman living in a city, felt that her identity, face, and persona could be easily stolen to show her speaking and doing things she never did. She felt that this kind of *identity theft* without one's consent could be misused to spread misinformation and to sway the opinions of targeted groups. Several female participants like Anushka were also concerned with how fake videos are often used to harass women. They described incidents where women's images were used to create fake pornographic videos and the resultant trauma and mental harassment the women had to go through. Purva described how she learnt about deepfake pornographic videos:

> *"I saw a news on how a woman committed suicide after a video showing her intimate moments went viral. Later, it was revealed that the video was fake. Someone put woman's face to another woman's body in a pornographic content. The woman felt so humiliated that she ended her life."* (Purva, Urban, Female, 45 years)

A few participants mentioned how some people inadvertently hurt themselves while trying to imitate actions from computer-generated movie scenes. Abdul, a 19-year-old rural resident, recounted one such example where a boy in his neighborhood tried to replicate a stunt from an action movie and badly injured himself. A few participants also noted how fake videos caused health-related harms to them and their family members when they followed false health advice during the COVID-19 pandemic. They also gave examples of how anti-vaccine fake videos provoked many people to decline the vaccine for coronavirus.

4.3.3 No harm. Most of the participants who were new to social media and some rural residents struggled to articulate the harms of fake videos. Some of them felt that fake videos are harmless. These participants had different mental models than experienced users and believed only staged videos to be fake. They explained how sharing funny cheapfakes, memes, and advertisements could entertain other people and bring monetary gains to content creators. Given the appetite for entertaining cheapfake videos, two participants felt that new businesses could benefit financially by launching commercial applications for creating entertaining and funny cheapfake videos, without reflecting much on how such applications could be repurposed for nefarious means. Abdul commented:

> *"These prank and funny fake videos are all means to earn money. Because if subscribers increase, then likes and comments will increase. And you will get more money."* (Abdul, Rural, Male, 19 years)

Overall, these findings suggest that many participants, with an exception of new users in rural regions, were aware of the risks and harms of fake videos and how these videos can be misused to create a polarized and divisive society.

4.4 Mitigating the Spread of Fake Videos

As many participants pointed out the potential harms of fake videos, they listed several measures that could be adopted by individuals, governments, and social media platforms to counteract the spread of fake videos.

4.4.1 Raising Awareness. A majority of the participants suggested raising awareness as a first step to mitigate the spread of fake videos. Many of them emphasized the need of *"collective teaching"* to educate the public on the harms of fake videos and ways to assess the credibility. They recommended using social media, newspapers, TV programs, paper pamphlets, and video demonstrations to raise awareness. They emphasized the need to create infographics-based training materials that are easy to grasp than long, text-heavy educational articles. A few participants also suggested creating educational videos that explain to users various techniques to spot digital manipulations. They also advised to engage non-governmental organizations, social workers, teachers, bureaucrats, and celebrities to promote awareness. Vikram commented:

> *"If social media influencers post a two-minute video on how to spot fake videos, it would instantly spread like wildfire among their large follower base."* (Vikram, Urban, Male, 22 years)

Given the threats to civil liberties and democratic institutions, some participants strongly suggested to start discussing the harms of fake videos in schools and colleges to build *"information-aware, digitally-skilled"* future generations. They emphasized the importance of extending digital literacy programs to older adults, who often have limited technology know-how and are usually more susceptible to fake videos. A few participants drew attention to the existing digital divide between urban and rural residents and advocated for creating programs specifically for social media users in rural regions. Pratiksha elaborated:

> *"Urban residents are usually more aware. But in the villages if meetings are arranged, people are brought together, and shown videos on what fake videos are, how they look like, and what to do when one receives fake videos, then they might inform others around them and this might lead to some changes."* (Pratiksha, Rural, Female, 41 years)

Several participants acknowledged the importance of embedding these interventions and educational programs in the socio-cultural fabric of the society. Describing Indian society as strongly patriarchal, Sanjay suggested that the heads of the households in rural areas could be trained so that they can inform their family members about safe information behaviors. Participants also noted that the educational programs should respect different cultural, religious, and regional sensitivities while providing examples of fake videos and discussing their harms. They highlighted the need to have recurring training sessions and the importance of long-term engagement with such training programs:

> *"It is not the case that if you teach me today, tomorrow I will receive a fake video and immediately identify it. I don't think this can happen in one go, but only by repeating it multiple times people might retain the knowledge and understand the implications better."* (Pratiksha, Rural, Female, 41 years)

4.4.2 Role of Social Media Platforms. Our participants offered several suggestions that social media platforms can take to curb the spread of fake videos. However, not many understood the underpinnings of social media and lacked awareness of the existing features to curb the spread of fake videos. For example, a couple of participants suggested that the platforms should use community feedback to label fake videos. Some of them did know about the reporting feature, but suggested to make those more visible like the *"like and comment button."* Some participants recommended the platforms to prevent the upload of potential fake videos by detecting them in real time, not knowing the technical challenges that it presents as well as associated difficulties in labeling videos that have partly true and partly misleading information. A few participants proposed crowd-based labeling of fake videos by forcing users to answer a few questions about the authenticity of the video which they attempt to share.

A handful of participants demanded that once verified, the platforms must instantly remove all instances of fake videos in order to prevent users from watching and sharing them. A few participants added that fake videos thrive on the reactions they receive online. Thus, they proposed to disable reactions, comments, and sharing of videos that are flagged by users or platforms. They also suggested that the platforms must notify all users who may have been exposed to a fake video and remind those to not share fake videos who have done so in the past.

Not knowing how the platforms operate under the hood, a few participants demanded the platforms to be more transparent about how they deal with reported videos. Along with the number of comments and reactions to the video, they suggested to display the number of times a video is reported to increase transparency. A couple of participants believed that fake videos are often shared by unverified accounts with small follower base which social media platforms should cut loose. This suggests that users thought videos from verified accounts with large follower base are not likely to be fake. Even in several prior studies it was observed that sender's profile and follower-base largely influence how people perceive credibility of online news [23, 38, 96]

4.4.3 Role of the Government and the Public.

Many participants held the governments accountable for the propagation of fake videos and believed that they could play a greater role in controlling the spread. They suggested to create regulatory bodies and develop top-down policies to penalize those who spread fake videos. Some of them felt that social media platforms are public spaces and should come under the scrutiny of government regulatory bodies. A few urban participants cited how Twitter had to comply to new Internet regulations in India and agreed to remove or censor content based on the country's new IT law [49]. A few suggested to enforce strict laws against publishing fake videos, punishing the offenders (e.g., imprisonment, confiscating property), and banning applications that allow the creation of such content. They also advised to enlist cybercrime divisions to trace the sources that spread fake videos.

Some participants expressed that everyone should be proactive about dealing with fake videos. They advised actionable measures that individuals can take when they doubt something as fake, such as reporting potentially fake content, commenting on fake videos to inform the public, talking to friends and family about fake videos, and educating people about the necessity of consuming information from reliable sources. They also called for action to scrutinize one's own activity on social media and carefully think before sharing anything on social media.

4.4.4 A Feeling of Despair: Nothing will Stop Fake Videos.

Despite all the enthusiasm and suggestions regarding what can be done to prevent the propagation of fake videos, some participants believed that there is no way to curb the spread of fake videos. Several participants critically reflected on the merits and limitations of these measures. They expressed that creating awareness or developing knowledge about fake videos will not necessarily stop people from creating, consuming, or sharing fake videos. They described how due to advancing technologies and accessible tools, high quality fake videos could be generated so easily that training alone might not help people. They also thought that social media platforms could do very little to detect fake videos due to the massive scale of videos that are uploaded every day. Besides, they believed controlling the spread of fake videos might not directly align with the corporate interests of social media platforms. Anand explained:

> *"The people who built social media applications care more about drawing users to their platform to generate more revenues rather than stopping fake videos."* (Anand, Rural, Male, 21 years)

The participants also expressed similar worries about government initiatives to prevent fake videos and felt that the governments lack the infrastructure, knowledge, and policies to counteract the onslaught of fake videos, especially deepfakes. They highlighted how the implementation of some of these measures could result in increased surveillance and restrictions on the freedom of speech. A few participants also believed that the governments lack the will to adopt these measures, because as political parties, they themselves have *"created and spread fake videos to serve their political agendas."* They also felt that mainstream news channels are working hand in glove with the governments to polarize the public. Deepak noted how these measures can give the governments power to erode the foundations of democracy:

> *"Whatever people say against the government, it is their right. But when political parties come to power, they try to censor things according to their own interests. Be it any government or any party, the power to control people's freedom of expression should not be in their hands. I personally think the government should not be responsible for preventing fake videos."* (Deepak, Urban, Male, 65 years)

Taken together, our participants had mixed feelings on to what extent the propagation of fake videos could be controlled and what are the appropriate ways to approach that. However, they agreed that it should not be a single endeavor from a particular group. Rather, individuals, local and national governments, and social media platforms should work together to deal with the menace of fake videos.

5 DISCUSSION

Our work qualitatively examines how diverse social media users in India perceive and engage with fake videos. The increased dissemination of misinformation via fake videos in the Global South, especially in India, makes our work a timely and relevant study. Through a qualitative study of 36 social media users in rural and urban India, we explored diverse ways in which people perceive fake videos, including deepfake and cheapfake videos. We observed that people's susceptibility to fake videos is usually shaped by their varying perceptions of such content. Notably, we found that many participants did not have a deep understanding of the technological advancements and prevalence of fake videos, believing that such videos would be easier to detect due to obvious audio-visual inconsistencies or they would not come across any fake video simply because they interact only with their friends and family online. We further unpacked how people interact with fake videos and observed several alarming practices, such as passive consumption and sharing of fake videos that align with one's beliefs, and a general indifference towards reporting fake videos. We examined their perceived threat models around fake videos and found that people generally considered videos to be more harmful than other modalities of misinformation (i.e., text, image). They also labelled fake videos as a great threat to both the individuals and the society. Finally, we probed participants' views on how to curb the spread of fake videos and the roles they expect individuals, society, governments, and social media platforms to play.

Overall, our findings suggest that most social media users are ill-equipped to deal with fake videos, have varying perceptions of what qualifies as fake, and engage with such videos differently, suggesting that there is no *one-size-fits-all* approach to combat the spread of fake videos. Based on our findings, we now discuss design recommendations for social media platforms to mitigate the risks and harms of fake videos.

Capacity Building. Many participants suggested using short-form videos and infographics to educate users about what fake videos are and how to spot them. They recommended these formats to maximize user engagement with training resources. The participants' observations are in line with research findings that show how videos drive more engagement than text-based content [28, 67]. Scholars have also used similar strategies in the past, for example, to educate people about privacy risks via videos [95] that have resulted in higher adoption of security features [6]. Recently, Facebook has also designed an infographic in consultation with its fact-checking partners to educate its users on fake news [48]. However, more work is needed to bring forth the training and educational resources. For example, since many social media platforms exhibit advertisements to users, similar to sponsored content, short-form training videos and infographics could be intermixed in users' news feed. Since many participants made a point of educating people routinely, repeated exposure to such short-form videos in one's newsfeed would be helpful for users to be reminded of healthy information practices. A routine exposure might lead to behavioral change eventually as is shown in prior studies in which regular exposure to social media-based health education was deemed effective in improving users' health-related behaviors [60, 65, 83]. However, more work is needed to design training materials that appeal to a wide array of users with different skills, cultures, literacy levels, backgrounds, and prior beliefs.

Explainable-AI based Credibility Indicators. Apart from training people to spot fake videos, our participants suggested that social media platforms should proactively detect and label fake videos. Given the large volumes of data that the platforms amass, a few participants recommended that the platforms could use AI advances to identify instances of fake videos that are running rampant and use credibility indicators to label videos that are either digitally manipulated or contain misleading information. Recent advances in GAN- and CNN-based technologies have made it possible to detect deepfakes using various measures, such as visual heartbeat rhythms [85], affective cues from perceived audio-visual emotions [74], and audio-visual dissonance [29]. Scholars have also used linguistic, acoustic, and user engagement features for automated detection of misinformation in YouTube videos [46]. However, prior work shows that AI-based credibility indicator is the least effective in reducing people's trust in fake news [112].

To alleviate users' distrust from AI-based predictions of fake news, one approach could be to design *explainable* AI-based credibility indicators, which have been shown to be useful in helping people share more credible news and report more fake news [75]. However, more work is needed to make AI-based credibility indicators explainable to low-literate, digitally novice social media users, many of who access information in languages other than English. A recent Google report shows that nearly 70% of Indians access digital content in their local languages [12]. To make AI's explanations more acceptable and understandable to these users, not only the credibility indicators and underlying AI algorithms should be compatible with local language content, but they also need to accessibly convey the information about features that guided its decision to users who might have limited to no digital literacy and English-language skills.

Moreover, current AI-based fake news detection systems have high *false negative* rates as the underlying algorithms are often inept in capturing socio-cultural nuances present in the content and users' interactions [107]. In addition, a recent study from Saltz et al. [88] show that *false positives* in case of auto-labeling visual misinformation further erode users' trust in credibility indicators. Reporting uncertainty along with predicted labels might help address this issue. However, since videos can be consumed easily without requiring any added digital or literacy skills, novel approaches are needed to visually convey such information. One approach could be to use colored backgrounds for displaying different types of fake videos (e.g., fabricated, edited, or missing context) [88] or to signify the confidence level when predicting fake videos. This could be particularly useful for low-literate users who might otherwise face difficulties in understanding prediction accuracy. Work from Bhuiyan et al. [13] shows that use of such colored backgrounds for tweets from reliable, questionable, and unreliable sources improved users' recognition of content from questionable sources.

Online Training for Crowd-based Credibility Ratings. An approach to address the various issues of AI-based fact-checking and detection of digitally manipulated videos is to leverage human assistance for designing crowdsourced credibility indicators as they are shown to be as effective as professional fact-checkers [78]. Our participants also suggested asking social media users about the credibility of videos while they are watching them. However, our findings show that many social media users struggle to spot fake videos, so a crowdsourced approach would be ineffective without proper capacity building. One way to handle this is to offer voluntary online training to social media users for recognizing fake videos. Prior studies have experimented with a range of techniques to build users' capacity to spot digitally manipulated fake videos. This includes displaying actual biometric reference videos along with fake ones [56], introducing different deepfake generation techniques with examples and pointers to generated artifacts [56], highlighting inconsistent areas in fake videos with explanatory clues [99], and offering mock tests with or without feedback [66], among others. Platforms could use these tactics to train social media users to effectively spot and report a wide array of deepfake videos. In addition, to not overburden the users with identifying the fake videos themselves, the platforms could also utilize linguistic cues from user comments to assess the veracity of the content [54] and label the videos accordingly. Moreover, due to prevalent socio-cultural norms, participants' perceptions of what comprises "fake" videos may differ from what qualifies as fake videos according to platform's policy. To accommodate socially and culturally diverse viewpoints, platforms should recruit a diverse pool of fact-checkers and partner with local news agencies to evolve the understanding of fake videos.

Reflective Sharing. Our participants suggested that platforms should disable reactions, comments, and sharing options on fact-checked fake videos since many people use them as a proxy to judge the content's credibility [15, 36]. Since fact-checking videos takes time, our participants also recommended to place sharing restrictions on all videos to minimize the risks and harms of fake videos. While these measures might seem draconian and impractical, WhatsApp did adopt similar measures to curb the spread of COVID-related misinformation [84]. Although such measures might reduce the spread of misinformation, it might also hurt the spread of authentic news, helpful advice, or harmless memes. Instead, the platforms could nudge the users when they try sharing a video and ask them whether they think the video is real or fake and why. Such lightweight interventions while sharing have been found to be effective in reducing the propagation of fake news [51]. Directly querying the users about the authenticity of videos and their reasons behind sharing might discourage them from sharing videos that they find suspicious. In addition, when users try to share videos, the platforms could open a pop-up that reminds them to be conscientious and thoughtful of their online activities and to not hurt other's sentiment by sharing communally or politically charged videos. This might help reduce the sharing of fake videos as many research studies show that such reminders could motivate users to change their behaviors [27, 82].

Visible and Transparent Reporting Feature. Our participants rarely reported videos. Many participants were confused about how reporting worked, pointing to the lack of transparency in the underlying workflows and processes. To make this feature more usable and useful, platforms should clearly communicate how they process the reported posts and follow-up with what actions they have taken (or lack thereof) on the flagged videos. Merely mentioning that no action was taken because the content did not violate the platform's policy might not be helpful to most users who seldom read or could comprehend the platform's policy. Hence, the platforms should be careful in accessibly communicating their responses with proper rationale to ensure that those flagging videos do not feel disgruntled as did some of our participants. To encourage more users to report videos they perceive to be fake, platforms may find it valuable to turn to social influence. Prior work shows that simply showing people how many of their friends used security features is effective in making them adopt those measures [32, 33]. In addition to displaying how many times a video is reported, as our participants suggested, the platforms could also mention how many of their friends reported a particular video, which might motivate people to report more videos.

Although almost all platforms enable users to report privately shared messages between two users, our participants expressed hesitation in reporting possibly fake videos that were shared with them privately, thinking that the sender might get to know who reported them if a punitive action is taken against the sender. Usually the platforms keep the reporters' identity and other metadata confidential and do not reveal it to the sender [21]. That said, platforms should offer clear information on how the reporting of privately shared content is handled. In addition, some users confessed they did not know how to flag content. This might be because most platforms (e.g., Facebook, Instagram, and Twitter) hide the reporting option from the post and it might be difficult for new and digitally novice social media users to locate this option. Hence, the platforms should consider making this feature more visible and making the reporting process easier to follow.

Our participants were alarmed by the rampant sharing of fake videos on WhatsApp, attributing the absence of fact-checking features on the platform. Not only there are no credibility indicators on the posts on WhatsApp, but also the dynamics vary between publicly and privately shared information and such nuances should be taken into consideration while designing interventions. Some of our participants shared how they were forced to leave WhatsApp groups to avoid receiving fake videos. The platform may offer users the option to hide messages from specific group members and report them to the group admins who can then decide to take disciplinary action on them (e.g., temporarily or permanently suspending them, or reducing their privileges temporarily). More research is needed to design, implement, and evaluate new sets of features that enable group admins of private and public WhatsApp groups to contain the spread of fake videos. Taken together, these recommendations could enable social media platforms to be more pragmatic in curbing the spread of fake videos and creating an environment conducive to civil discourse.

6 CONCLUSION

Through interviews with 36 diverse social media users from rural and urban India, we examined how they perceive and interact with fake videos, how they evaluate the risks and harms of fake videos, and what they consider necessary to curb its spread. We found that the users had varying perceptions of what they consider as fake videos. Most of them passively consumed videos and lacked the willingness and skills to spot fake videos. Several users were unaware that videos can be doctored or edited, and most knew nothing about deepfakes. A few participants, who were aware of digital manipulations of videos, expected these videos to be easily discernible. Even when participants knew a video to be fake, they rarely reported it and sometimes willingly shared it for entertainment or when it supported their worldviews. Many participants described how fake videos are polarizing and divisive and expressed diverse opinions in how individuals, social media platforms, and governments could contend with the harms of fake videos. Based on our findings, we discussed the potential of short-form training videos, transparent reporting processes, Human-AI credibility indicators, and reflective sharing in containing the spread of fake videos.

REFERENCES

[1] Adobe. 2021. Adobe After Effects. Retrieved 2021-09-08 from https://www.adobe.com/products/aftereffects.html

[2] Shruti Agarwal, Tarek El-Gaaly, Hany Farid, and Ser-Nam Lim. 2020. Detecting Deep-Fake Videos from Appearance and Behavior. *CoRR* abs/2004.14491 (2020), 12 pages. https://arxiv.org/abs/2004.14491

[3] Saifuddin Ahmed. 2021. Navigating the maze: Deepfakes, cognitive ability, and social media news skepticism. *New Media & Society* 0, 0 (2021), 1–22.

[4] Saifuddin Ahmed. 2021. Who inadvertently shares deepfakes? Analyzing the role of political interest, cognitive ability, and social network size. *Telematics and Informatics* 57 (2021), 101508. https://doi.org/10.1016/j.tele.2020.101508

[5] Syeda Zainab Akbar, Anmol Panda, Divyanshu Kukreti, Azhagu Meena, and Joyojeet Pal. 2021. Misinformation as a Window into Prejudice: COVID-19 and the Information Environment in India. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–28.

[6] Yusuf Albayram, Mohammad Maifi Hasan Khan, and Michael Fagan. 2017. A Study on Designing Video Tutorials for Promoting Security Features: A Case Study in the Context of Two-Factor Authentication (2FA). *International Journal of Human–Computer Interaction* 33, 11 (2017), 927–942.

[7] Jussara Almeida. 2019. Misinformation Dissemination on the Web. In *Companion Proceedings of The 2019 World Wide Web Conference* (San Francisco, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 740.

[8] Ahmer Arif, John J. Robinson, Stephanie A. Stanek, Elodie S. Fichet, Paul Townsend, Zena Worku, and Kate Starbird. 2017. A Closer Look at the Self-Correcting Crowd: Examining Corrections in Online Rumors. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 155–168.

[9] Shakuntala Banaji, 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. , 62 pages.

[10] Arthur Asa Berger. 1989. *Seeing Is Believing: An Introduction to Visual Communication.* Mayfield Publishing Company, , 1240 Villa Street, Mountain View, CA 94041.

[11] Alessandro Bessi, Fabio Petroni, Michela Del Vicario, Fabiana Zollo, Aris Anagnostopoulos, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. 2015. Viral Misinformation: The Role of Homophily and Polarization. In *Proceedings of the 24th International Conference on World Wide Web* (Florence, Italy) *(WWW '15 Companion)*. Association for Computing Machinery, New York, NY, USA, 355–356.

[12] Ananya Bhattacharya. 2017. India's internet users have more faith in content that's not in English. Retrieved 2021-09-07 from https://qz.com/india/972844/indias-internet-users-have-more-faith-in-content-thats-not-in-english-study-says/

[13] Md Momen Bhuiyan, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. NudgeCred: Supporting News Credibility Assessment on Social Media Through Nudges. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 427 (oct 2021), 30 pages.

[14] Linda Birt, Suzanne Scott, Debbie Cavers, Christine Campbell, and Fiona Walter. 2016. Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? *Qualitative Health Research* 26, 13 (2016), 1802–1811.

[15] Porismita Borah and Xizhu Xiao. 2018. The Importance of 'Likes': The Interplay of Message Framing, Source, and Social Endorsement on Credibility Perceptions of Health Information on Facebook. *Journal of Health Communication* 23, 4 (2018), 399–411.

[16] Joel Breakstone, Mark Smith, Sam Wineburg, Amie Rapaport, Jill Carle, Marshall Garland, and Anna Saavedra. 2019. Students' civic online reasoning: A national portrait. , 49 pages.

[17] Joel Breakstone, Mark Smith, Sam Wineburg, Amie Rapaport, Jill Carle, Marshall Garland, and Anna Saavedra. 2021. Students' Civic Online Reasoning: A National Portrait. *Educational Researcher* 0, 0 (2021), 49.

[18] Catherine Francis Brooks. 2021. Popular discourse around deepfakes and the interdisciplinary challenge of fake video distribution. *Cyberpsychology, Behavior, and Social Networking* 24, 3 (2021), 159–163.

[19] Matt Burgess. 2020. Deepfake porn is now mainstream. And major sites are cashing in. https://www.wired.co.uk/article/deepfake-porn-websites-videos-law

[20] Joseph N Cappella, Hyun Suk Kim, and Dolores Albarracín. 2015. Selection and transmission processes for information in the emerging media environment: Psychological motives and message characteristics. *Media psychology* 18, 3 (2015), 396–424.

[21] Facebook Help Center. 2020. What happens when I report something to Facebook? Does the person I report get notified? Retrieved 2021-09-08 from https://www.facebook.com/help/1037960630447342

[22] Priyank Chandra and Joyojeet Pal. 2019. *Rumors and Collective Sensemaking: Managing Ambiguity in an Informal Marketplace.* Association for Computing Machinery, New York, NY, USA, 1–12.

[23] Apoorva Chauhan and Amanda Lee Hughes. 2020. Trustworthiness Perceptions of Social Media Resources Named after a Crisis Event. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 044 (2020), 23 pages.

[24] Jay Chen, Michael Paik, and Kelly McCabe. 2014. Exploring Internet Security Perceptions and Practices in Urban Ghana. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. USENIX Association, Menlo Park, CA, 129–142.

[25] Xinran Chen and Sei-Ching Joanna Sin. 2013. 'Misinformation? What of It?': Motivations and Individual Differences in Misinformation Sharing on Social Media. In *Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries* (Montreal, Quebec, Canada) *(ASIST '13)*. American Society for Information Science, USA, Article 104, 4 pages.

[26] Xinran Chen, Sei-Ching Joanna Sin, Yin-Leng Theng, and Chei Sian Lee. 2015. Why Do Social Media Users Share Misinformation?. In *Proceedings of the 15th ACM/IEEE-CS Joint Conference on Digital Libraries* (Knoxville, Tennessee, USA) *(JCDL '15)*. Association for Computing Machinery, New York, NY, USA, 111–114.

[27] Tom Chivers. 2019. What's next for psychology's embattled field of social priming. Retrieved 2021-09-08 from https://www.nature.com/articles/d41586-019-03755-2

[28] Amit Chowdhry. 2019. Study: Relevant Video Content Drives More Engagement And Revenue. Retrieved 2021-09-06 from https://www.forbes.com/sites/splunk/2021/09/01/why-you-should-still-care-about-malware-and-what-to-do-about-it-now/?sh=6357ab965394

[29] Komal Chugh, Parul Gupta, Abhinav Dhall, and Ramanathan Subramanian. 2020. *Not Made for Each Other- Audio-Visual Dissonance-Based Deepfake Detection and Localization.* Association for Computing Machinery, New York, NY, USA, 439–447.

[30] Justin D. Cochran and Stuart A. Napshin. 2021. Deepfakes: Awareness, Concerns, and Platform Accountability. *Cyberpsychology, Behavior, and Social Networking* 24, 3 (2021), 164–172. https://doi.org/10.1089/cyber.2020.0100 arXiv:https://doi.org/10.1089/cyber.2020.0100 PMID: 33760667.

[31] John W. Creswell and Dana L. Miller. 2000. Determining Validity in Qualitative Inquiry. *Theory Into Practice* 39, 3 (2000), 124–130.

[32] Sauvik Das, Adam D.I. Kramer, Laura A. Dabbish, and Jason I. Hong. 2014. Increasing Security Sensitivity With Social Proof: A Large-Scale Experimental Confirmation. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (Scottsdale, Arizona, USA) *(CCS '14)*. Association for Computing Machinery, New York, NY, USA, 739–749. https://doi.org/10.1145/2660267.2660271

[33] Sauvik Das, Adam D.I. Kramer, Laura A. Dabbish, and Jason I. Hong. 2015. The Role of Social Influence in Security Feature Adoption. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1416–1426. https://doi.org/10.1145/2675133.2675225

[34] Nicholas Diakopoulos and Deborah Johnson. 2021. Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society* 23, 7 (2021), 2072–2098.

[35] Tom Dobber, Nadia Metoui, Damian Trilling, Natali Helberger, and Claes de Vreese. 2021. Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics* 26, 1 (2021), 69–91. https://doi.org/10.1177/1940161220944364 arXiv:https://doi.org/10.1177/1940161220944364

[36] Jr Edson C Tandoc, Richard Ling, Oscar Westlund, Andrew Duffy, Debbie Goh, and Lim Zheng Wei. 2018. Audiences' acts of authentication in the age of fake news: A conceptual framework. *New Media & Society* 20, 8 (2018), 2745–2763.

[37] Gowhar Farooq. 2018. Politics of Fake News: How WhatsApp Became a Potent Propaganda Tool in India. *Media Watch* 9, 1 (2018), 106–117.

[38] Christine Geeng, Savanna Yee, and Franziska Roesner. 2020. Fake News on Facebook and Twitter: Investigating How People (Don't) Investigate. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–14.

[39] Natalie Gerhart and Anna Sidorova. 2017. The Effect of Network Characteristics on Online Identity Management Practices. *Journal of Computer Information Systems* 57, 3 (2017), 229–237.

[40] Amira Ghenai. 2017. Health Misinformation in Search and Social Media. In *Proceedings of the 2017 International Conference on Digital Health* (London, United Kingdom) *(DH '17)*. Association for Computing Machinery, New York, NY, USA, 235–236.

[41] Megan Graham and Jennifer Elias. 2021. How Google's $150 billion advertising business works. Retrieved 2021-09-08 from https://www.cnbc.com/2021/05/18/how-does-google-make-money-advertising-business-breakdown-.html

[42] Greg Guest, Kathleen M. MacQueen, and Emily E. Namey. 2012. *Applied thematic analysis.* SAGE, Thousand Oaks, CA.

[43] Parul Gupta, Komal Chugh, Abhinav Dhall, and Ramanathan Subramanian. 2020. The Eyes Know It: The eyes know it: FakeET- An Eye-tracking Database to Understand Deepfake Perception- An Eye-Tracking Database to Understand Deepfake Perception. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 519–527.

[44] Md Mahfuzul Haque, Mohammad Yousuf, Ahmed Shatil Alam, Pratyasha Saha, Syed Ishtiaque Ahmed, and Naeemul Hassan. 2020. Combating Misinformation in Bangladesh: Roles and Responsibilities as Perceived by Journalists, Fact-Checkers, and Users. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 130 (oct 2020), 32 pages.

[45] The Hindu. 2021. A week after election results, violence continues in Bengal. Retrieved 2021-09-08 from https://www.thehindu.com/news/national/other-states/west-bengal-violence-barrage-of-fake-videos-posts-surface-on-social-media/article34513532.ece

[46] Rui Hou, Veronica Perez-Rosas, Stacy Loeb, and Rada Mihalcea. 2019. Towards Automatic Detection of Misinformation in Online Medical Videos. In *2019 International Conference on Multimodal Interaction* (Suzhou, China) *(ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 235–243.

[47] Eslam Hussein, Prerna Juneja, and Tanushree Mitra. 2020. Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 048 (May 2020), 27 pages. https://doi.org/10.1145/3392854

[48] Andrew Hutchinson. 2020. Facebook Launches New Education Campaign to Help People Detect Fake News. Retrieved 2021-09-07 from https://www.socialmediatoday.com/news/facebook-launches-new-education-campaign-to-help-people-detect-fake-news/580835/

[49] Andrew Hutchinson. 2021. Twitter Agrees to New Indian Government Regulations, Despite Concerns About the Impacts on User Speech. Retrieved 2021-09-02 from https://www.socialmediatoday.com/news/twitter-agrees-to-new-indian-government-regulations-despite-concerns-about/603057/

[50] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology Probes: Inspiring Design for and with Families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 17–24.

[51] Farnaz Jahanbakhsh, Amy X. Zhang, Adam J. Berinsky, Gordon Pennycook, David G. Rand, and David R. Karger. 2021. Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 18 (2021), 42 pages.

[52] Maurice Jakesch, Moran Koren, Anna Evtushenko, and Mor Naaman. 2018. The Role of Source, Headline and Expressive Responding in Political News Evaluation. *SSRN Electronic Journal* 1, 1 (2018), 5 pages.

[53] Jasser Jasser. 2019. Dynamics of Misinformation Cascades. In *Companion Proceedings of The 2019 World Wide Web Conference* (San Francisco, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 33–36.

[54] Shan Jiang and Christo Wilson. 2018. Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 82 (nov 2018), 23 pages.

[55] Stamatis Karnouskos. 2020. Artificial Intelligence in Digital Media: The Era of Deepfakes. *IEEE Transactions on Technology and Society* 1, 3 (2020), 138–147. https://doi.org/10.1109/TTS.2020.3001312

[56] Ali Khodabakhsh, Raghavendra Ramachandra, and Christoph Busch. 2019. Subjective Evaluation of Media Consumer Vulnerability to Fake Audiovisual Content. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Gjøvik, Norway, 1–6. https://doi.org/10.1109/QoMEX.2019.8743316

[57] Ali Khodabakhsh, Raghavendra Ramachandra, and Christoph Busch. 2019. Subjective Evaluation of Media Consumer Vulnerability to Fake Audiovisual Content. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Manhattan, NY, 1–6. https://doi.org/10.1109/QoMEX.2019.8743316

[58] Jan Kietzmann, Linda W. Lee, Ian P. McCarthy, and Tim C. Kietzmann. 2020. Deepfakes: Trick or treat? *Business Horizons* 63, 2 (2020), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING.

[59] Jan Kirchner and Christian Reuter. 2020. Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 140 (Oct. 2020), 27 pages.

[60] Holly Korda and Zena Itani. 2013. Harnessing Social Media for Health Promotion and Behavior Change. *Health Promotion Practice* 14, 1 (2013), 15–23.

[61] Pavel Korshunov and Sébastien Marcel. 2020. Deepfake detection: humans vs. machines. *CoRR* abs/2009.03155 (2020), 6 pages. arXiv:2009.03155 https://arxiv.org/abs/2009.03155

[62] Nieman Lab. 2021. The focus of misinformation debates shifts south. https://www.niemanlab.org/2018/12/the-focus-of-misinformation-debates-shifts-south/.

[63] Karen Labelo. 2019. India Is Teeming With 'Cheapfakes', Deepfakes Could Make It Worse. Retrieved 2022-02-17 from https://www.boomlive.in/india-is-teeming-with-cheapfakes-deepfakes-could-make-it-worse/

[64] Danielle Lottridge and Frank R. Bentley. 2018. *Let's Hate Together: How People Share News in Messaging, Social, and Public Networks*. Association for Computing Machinery, New York, NY, USA, 1–13.

[65] Miaomiao Luo, Yu Hao, Ming Tang, Mengzhen Shi, Fengjuan He, Yuanmao Xie, and Weigang Chen. 2020. Application of a social media platform as a patient reminder in the treatment of Helicobacter pylori. *Helicobacter* 25, 2 (2020), e12682.

[66] Brandon Mader, Martin S. Banks, and Hany Farid. 2017. Identifying Computer-Generated Portraits: The Importance of Training and Incentives. *Perception* 46, 9 (2017), 1062–1076.

[67] Megan Mahoney. 2020. Just the Stats: Why Should You Leverage Video Marketing. Retrieved 2021-09-06 from https://www.singlegrain.com/video-marketing/just-stats-science-video-engagement/

[68] Marie-Helen Maras and Alex Alexandrou. 2019. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence & Proof* 23, 3 (2019), 255–262.

[69] Alice E Marwick. 2018. Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technology Review* 2, 2 (2018), 474–512.

[70] Shruti Menon. 2020. Coronavirus: The human cost of fake news in India. Retrieved January 8, 2021 from https://www.bbc.com/news/world-asia-india-53165436

[71] Miriam J. Metzger, Andrew J. Flanagin, and Ryan B. Medders. 2010. Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication* 60, 3 (08 2010), 413–439.

[72] Robert B. Michael and Brooke O. Breaux. 2021. The relationship between political affiliation and beliefs about sources of "fake news". *Cognitive Research: Principles and Implications* 6 (2021), 15 pages. Issue 1.

[73] Ananda Mitra. 2020. Short-Form Video Dominates Social Media In India. Retrieved 2021-09-08 from https://www.forbes.com/sites/anandamitra/2020/06/23/tiktok-short-form-video-dominates-social-media-in-india/?sh=30b9e49c6803

[74] Trisha Mittal, Uttaran Bhattacharya, Rohan Chandra, Aniket Bera, and Dinesh Manocha. 2020. *Emotions Don't Lie: An Audio-Visual Deepfake Detection Method Using Affective Cues*. Association for Computing Machinery, New York, NY, USA, 2823–2832.

[75] Sina Mohseni, Fan Yang, Shiva Pentyala, Mengnan Du, Yi Liu, Nic Lupfer, Xia Hu, Shuiwang Ji, and Eric Ragan. 2020. Machine Learning Explanations to Prevent Overtrust in Fake News Detection. arXiv:2007.12358 [cs.IR]

[76] UCL News. 2021. Young peoples' rates of reporting online harassment and abuse are 'shockingly low'. Retrieved January 2, 2022 from https://www.ucl.ac.uk/news/2021/dec/young-peoples-rates-reporting-online-harassment-and-abuse-are-shockingly-low

[77] Chinasa T. Okolo, Srujana Kamath, Nicola Dell, and Aditya Vashistha. 2021. "It Cannot Do All of My Work": Community Health Worker Perceptions of AI-Enabled Mobile Health Applications in Rural India. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 701, 20 pages.

[78] Laura Hazard Owen. 2020. Crowds of regular people are as good at moderating fake news on Facebook as professional fact-checkers. Retrieved 2021-09-07 from https://www.niemanlab.org/2020/10/crowds-of-regular-people-are-as-good-at-moderating-fake-news-on-facebook-as-professional-fact-checkers/

[79] Gordon Pennycook and David G. Rand. 2021. The Psychology of Fake News. *Trends in Cognitive Sciences* 25, 5 (May 2021), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

[80] Ivan Perov, Daiheng Gao, Nikolay Chervoniy, Kunlin Liu, Sugasa Marangonda, Chris Umé, Mr. Dpfks, Carl Shift Facenheim, Luis RP, Jian Jiang, Sheng Zhang, Pingyu Wu, Bo Zhou, and Weiming Zhang. 2020. DeepFaceLab: A simple, flexible and extensible face swapping framework. *CoRR* abs/2005.05535 (2020), 10 pages.

[81] Greg Philo. 2002. *Seeing and Believing: The Influence of Television*. Routledge, London. https://doi.org/10.4324/9780203136560

[82] Peter Pirolli, Shiwali Mohan, Anusha Venkatakrishnan, Les Nelson, Michael Silva, and Aaron Springer. 2017. Implementation Intention and Reminder Effects on Behavior Change in a Mobile Health System: A Predictive Cognitive Model. *Journal of Medical Internet Research* 19, 11 (2017), e397.

[83] Zachary Pope, Daheia Barr-Anderson, Beth Lewis, Mark Pereira, and Zan Gao. 2019. Use of Wearable Technology and Social Media to Improve Physical Activity and Dietary Behaviors among College Students: A 12-Week Randomized Pilot Study. *International Journal of Environmental Research and Public Health* 16, 19 (Sep 2019), 3579.

[84] Jon Porter. 2020. WhatsApp says its forwarding limits have cut the spread of viral messages by 70 percent. Retrieved 2021-09-06 from https://www.theverge.com/2020/4/27/21238082/whatsapp-forward-message-limits-viral-misinformation-decline

[85] Hua Qi, Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, Wei Feng, Yang Liu, and Jianjun Zhao. 2020. *DeepRhythm: Exposing DeepFakes with Attentional Visual Heartbeat Rhythms*. Association for Computing Machinery, New York, NY, USA, 4318–4327.

[86] Julio C. S. Reis, Philipe Melo, Kiran Garimella, Jussara M. Almeida, Dean Eckles, and Fabrício Benevenuto. 2020. A Dataset of Fact-Checked Images Shared on WhatsApp During the Brazilian and Indian Elections. *Proceedings of the International AAAI Conference on Web and Social Media* 14, 1 (2020), 903–908.

[87] Niranjan Sahoo. 2020. How fake news is complicating India's war against COVID-19. Retrieved 2021-09-08 from https://www.orfonline.org/expert-speak/how-fake-news-complicating-india-war-against-covid19-66052/

[88] Emily Saltz, Claire R Leibowicz, and Claire Wardle. 2021. Encounters with Visual Misinformation and Labels Across Platforms: An Interview and Diary Study to Inform Ecosystem Approaches to Misinformation Interventions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 340, 6 pages.

[89] Sanam. 2020. Digital and Social Media Landscape in India. Retrieved 2021-09-08 from https://sannams4.com/digital-and-social-media-landscape-in-india/

[90] Oscar Schwartz. 2018. You thought fake news was bad? Deep fakes are where truth goes to die. Retrieved 2021-09-08 from https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth

[91] Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar. 2017. Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction. *New Media & Society* 19, 8 (2017), 1214–1235.

[92] Anu Shrestha and Francesca Spezzano. 2019. Online Misinformation: From the Deceiver to the Victim. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (Vancouver, British Columbia, Canada) *(ASONAM '19)*. Association for Computing Machinery, New York, NY, USA, 847–850.

[93] Manish Singh. 2019. Reliance Jio partners with Facebook to launch literacy program for first time internet users in India. Retrieved January 7, 2021 from https://social.techcrunch.com/2019/07/03/reliance-jio-facebook-digital-literacy-udaan-india/

[94] Statista. 2021. Social media users in India. Retrieved 2021-09-09 from https://www.statista.com/forecasts/1145422/social-media-users-in-india

[95] Alexa Stein, Norman Makoto Su, and Xinru Page. 2020. Learning through Videos: Uncovering Approaches to Educating People about Facebook Privacy. , 5 pages.

[96] David Sterrett, Dan Malato, Jennifer Benz, Liz Kantor, Trevor Tompson, Tom Rosenstiel, Jeff Sonderman, and Kevin Loker. 2019. Who Shared It?: Deciding

What News to Trust on Social Media. *Digital Journalism* 7, 6 (July 2019), 783–801.

[97] Sharifa Sultana and Susan R. Fussell. 2021. Dissemination, Situated Fact-Checking, and Social Effects of Misinformation among Rural Bangladeshi Villagers During the COVID-19 Pandemic. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 436 (oct 2021), 34 pages.

[98] S Shyam Sundar, Maria D Molina, and Eugene Cho. 2021. Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps? , 29 pages.

[99] Rashid Tahir, Brishna Batool, Hira Jamshed, Mahnoor Jameel, Mubashir Anwar, Faizan Ahmed, Muhammad Adeel Zaffar, and Muhammad Fareed Zaffar. 2021. Seeing is Believing: Exploring Perceptual Differences in DeepFake Videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 174, 16 pages.

[100] Shalini Talwar, Amandeep Dhir, Puneet Kaur, Nida Zafar, and Melfi Alrasheedy. 2019. Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services* 51 (2019), 72–82.

[101] Lu Tang, Kayo Fujimoto, Muhammad Tuan Amith, Rachel Cunningham, Rebecca A Costantini, Felicia York, Grace Xiong, Julie A Boom, and Cui Tao. 2021. "Down the Rabbit Hole" of Vaccine Misinformation on YouTube: Network Exposure Study. *Journal of Medical Internet Research* 23, 1 (2021), e23262.

[102] Saumya Tewari. 2020. India is sixth largest market in consumption of video ads: Report. Retrieved 2021-09-08 from https://www.livemint.com/industry/advertising/india-is-sixth-largest-market-in-consumption-of-video-ads-report-11582024457022.html

[103] Russell Torres, Natalie Gerhart, and Arash Negahban. 2018. Epistemology in the Era of Fake News: An Exploration of Information Verification Behaviors among Social Networking Site Users. *SIGMIS Database* 49, 3 (jul 2018), 78–97.

[104] Cristian Vaccari and Andrew Chadwick. 2020. Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society* 6, 1 (2020), 13 pages. https://doi.org/10.1177/2056305120903408 arXiv:https://doi.org/10.1177/2056305120903408

[105] Aditya Vashistha, Abhinav Garg, Richard Anderson, and Agha Ali Raza. 2019. Threats, Abuses, Flirting, and Blackmail: Gender Inequity in Social Media Voice Forums. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[106] Fie Velghe. 2012. Deprivation, distance and connectivity: The adaptation of mobile phone use to a life in Wesbank, a post-apartheid township in South Africa. *Discourse, Context & Media* 1, 4 (2012), 203–216. https://doi.org/10.1016/j.dcm.2012.09.004

[107] James Vincent. 2018. Why AI isn't going to solve Facebook's fake news problem. Retrieved 2021-09-08 from https://www.theverge.com/2018/4/5/17202886/facebook-fake-news-moderation-ai-challenges

[108] Travis L. Wagner and Ashley Blewer. 2019. "The Word Real Is No Longer Real": Deepfakes, Gender, and the Challenges of AI-Altered Video:. *Open Information Science* 3, 1 (2019), 32–46. https://doi.org/10.1515/opis-2019-0003

[109] wombo. 2021. Ai powered lip sync app. Retrieved 2021-09-08 from https://www.wombo.ai/

[110] Susan Wyche. 2019. Using Cultural Probes In New Contexts: Exploring the Benefits of Probes in HCI4D/ICTD. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (Austin, TX, USA) *(CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 423–427.

[111] Aya Yadlin-Segal and Yael Oppenheim. 2021. Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence* 27, 1 (2021), 36–51.

[112] Waheeb Yaqub, Otari Kakhidze, Morgan L. Brockman, Nasir Memon, and Sameer Patil. 2020. *Effects of Credibility Indicators on Social Media News Sharing Intent.* Association for Computing Machinery, New York, NY, USA, 1–14.

[113] Brandy Zadrozny. 2021. On TikTok, audio gives new virality to misinformation. Retrieved 2021-09-09 from https://www.nbcnews.com/tech/tech-news/tiktok-audio-gives-new-virality-misinformation-rcna1393