K4D | Knowledge, evidence and learning for development

Countering online misinformation, hate speech or extremist narratives in the Global South

Huma Haider
Independent consultant
4 March 2022

Question

What can we learn from previous and ongoing programmes in the Global South and specifically in Pakistan aiming to counter online mis/disinformation, hate speech or extremist narratives?

Contents

1. Summary

The widespread expansion of social media outlets has enabled the spread of mis/disinformation, hate speech and extremist narratives online. Internet-based technologies can also be used to confront these types of communication. **This report focuses on counter-messaging efforts (also referred to as strategic communications) through online platforms.** This can comprise **counter-narratives** that challenge false information or existing narratives (e.g. undermining the credibility of an extremist group); or **alternative narratives** that seek to replace existing narratives, rather than directly confront them (e.g. introducing messages of coexistence).

There is a significant knowledge gap regarding the effectiveness of counter-messaging, with few studies that look at their impact and effectiveness (Chan, 2020; Zeiger & Gyte, 2020; Whittaker & Elsayad, 2019; Reed et al., 2017). Studies that exist are largely descriptive; or are based on campaigns originating in the Global North (Zeiger & Gyte, 2020). In addition, studies often discuss effectiveness based on number of views and level and degree of user engagement (likes, shares, comments), with less known about the ability to achieve attitudinal and behavioural change and/or deter radicalisation (Chan, 2020; Whittaker & Elsayad, 2019).

This rapid literature review draws on a wide range of academic, civil society and donor literature, including reports that provide general guidance on counter-messaging programming. It surveys a range of initiatives from the Global South, with a focus on Pakistan.[1] Some of the profiles of initiatives in this report are solely descriptive, based on the availability of information. Countering violent extremism (CVE) programming is most common in the literature, but is encompassing to include not only counter-messaging to extremist narratives and hate speech, but also digital literacy to address mis/disinformation espoused by extremist actors. The following is a brief summary of lessons learned and guidance for programming, along with select examples:

Effective strategic communications require a clear identification of the target audience— and tailored messaging for different target audiences and local contexts and cultures (Reed, 2017; Barzegar et al., 2016). This can be achieved in part by empowering communities with the knowledge, skills and tools to counter messages in their networks (Zeiger & Gyte, 2020). Kenya-based *YADEN's #insolidarity* campaign provides capacity-building, tools, and platforms for youth to develop their own messages and share their stories about how terrorism has influenced their lives (Zeiger & Gyte, 2020). The *Dare to be peace #beranidami* campaign, targeted at youth in urban Indonesia, drew from a wide range of local community groups. Localising online messages of tolerance and moderation is considered to be an effective tactic to reach the target audience (Zeiger, 2016).

Catering to different audiences includes giving minority groups who have been targeted in online spaces, through hate speech and harassment, the space to engage online safely. Hamara internet, in Pakistan aims to provide a secure digital environment for women, teaching feminist activists how to utilise digital tools for secure communications (e.g. encryption), enabling them to safely engage in online challenges to patriarchal narratives (Rehman, 2016).

A key consideration for message designers is whether messages are offensive (sets the agenda) or defensive (reactive) – corresponding with alternative and counter-narratives, respectively (Whittaker & Elsayed, 2019). A survey of campaigns in the Middle East and North Africa (MENA) find that the vast majority of messages in campaigns are offensive, rather than

---

[1] Literature on online counter-messaging initiatives in South America was not readily available during this rapid literature review, thus examples from the region are absent.

defensive. When designing CVE messages, it is suggested that offensive messages should outweigh defensive messages in order to avoid the perceived danger of allowing violent extremists to control the debate (Whittaker & Elsayed, 2019).

An important aspect of alternative messaging is to provide a non-violent action that still addresses grievances that communities might have, which can be underlying drivers of radicalisation. This can include messages that address structural or personal grievances, such as building positive identities and enhancing social cohesion (Zeiger & Gyte, 2020).

Alternative narratives should make use of positive messages centred on a community's history and roots (Udupa, 2021). Media productions in Pakistan, representing girls (Gogi and Burka Avenger) provide alternative narratives—engaging in complex discussions of gender, adolescence, religion, and urban reality, with attention to positive aspects; and promoting peace and tolerance in society (Pirzada, 2017; Barzegar et al., 2016).

Appealing to emotions can be an effective strategy in achieving greater engagement with counter-messaging. In Pakistan, the military's has adopted a "motivational frame", focusing on heroism and resilience, in its campaigns to gain support for its war against terrorism. This frame has triggered the highest level of user engagement (comments) than other frames, such as an informational frame (Khan & Pratt, 2020). A Kuwaiti video commercial, which sought to counter Daesh's narrative of divided communities and justification for violent acts, aimed to have an emotional appeal, showing clips of victims of terrorism to highlight the human consequences of such violence. The video received over 10 million views, with active online engagement in the form of over 10 thousand comments on YouTube and Twitter (Elsayed et al., 2017).

The credibility and trust worthiness of the messenger—the individual, group or institution delivering the message or narrative—is as important as the message itself (RAN, 2016). Different messengers should be utilised depending on the target audience. While local cultural influencers may appeal to a general audience, they may be less credible to potential violent extremists (Bilazarian, 2020; RAN, 2016). Research on Burka Avenger finds that the success of the show can be linked in large part to its use of popular singers, and the status of the show's creator as a pop star, who was able to draw a large audience (Zeiger & Gyte, 2020). In contrast, former violent extremists may be more credible and effective in providing counter-narratives to potential violent extremists (Bilazarian, 2020; RAN, 2016). An online video of an interview with a former Daesh member from Indonesia, undermines the credibility of Daesh by highlighting the corruption within the group (Zeiger, 2016).

Female preachers and/or experts are considered to be credible messengers to deliver gender-based counter and alternative narratives (Elsayed, 2019). Muslim women within MENA are more likely to engage with and communicate with other women, which makes the role of female preachers and/or experts particularly important (Elsayed, 2019).

Target audiences of counter-messaging should not be told what to think, but rather taught how to engage in critical thinking—for example, through digital media literacy programmes (Braddock, 2019; Rosand & Winterbotham, 2019; Sher & Sturm, 2018). A recent evaluation of two educational interventions for countering misinformation in urban Pakistan finds that provision of personalised feedback about participants' own responses to fake news stories, which also highlight the features of each fake news item, is more effective than general informational messages (Ali & Qazi, 2021).

Online CVE efforts need to institutionalize systems not only to dispel extremist narratives but also to replace the immersive social experience provided by extremist groups (Bilazarian, 2020). This can include various forms of engagement, such as: sharing messages, posts and comments—which can promote dialogue and community; and the formation of online networks (Bilazarian, 2020; Merchandani et al. 2018). Against violent extremist (AVE)'s mobile app brings together various actors, including former violent extremists and survivors of violent extremist—enabling them to share experiences and lessons; and work together to combat extremist propaganda or narratives (Amit et al., 2021).

In order to be effective, mobile apps designed to provide counter-messaging and/or to facilitate the formation of networks should be available in local languages and downloadable for free. The Young Power in Social Action initiative is a mobile app, launched in Bangladesh, in Bangla and free to download. It includes sessions on family and society, the negative sides of extremism, and the victims of extremism—which can serve as counter-narratives that undercut extremist messaging (Amit et al., 2021).

The integration of online and offline programming may achieve greater results. Online counter-messaging might trigger doubt in particular beliefs and views. They may not interrupt the radicalisation process on their own, however, requiring offline counter-measures, alongside –e.g. educating young people at school about the consequences of violent extremism (RAN, 2016). The integration of online and offline programming may also be required out of pure necessity in geographic areas where online forms of communication are less common, such as in rural areas.

2. Background

The widespread expansion of social media outlets has allowed individuals espousing mis/disinformation, hate speech or extremist narratives to interact not only with others who are like-minded, but also with those vulnerable to persuasion and potential violent radicalisation (Braddock, 2019). Extremists exploit social media platforms, and the Internet more generally, for various reasons—ranging from spreading hateful narratives and propaganda to financing, recruitment, and sharing operational information (Ganesh & Bright, 2020; Braddock). At the same time, Internet-based technologies can be used to confront the threats posed by extremist groups online (Braddock, 2019). The UN Secretary-General has urged all member states to engage with strategic communications, the Internet and social media in their CVE strategies (Elsayed, 2019). Many governments, civil society actors and private companies have begun to develop programmes to counter mis/disinformation, hate speech and extremist narratives online. They fall into two key categories:

▪ **Disruption**: This involves preventing and prohibiting extremists the ability to spread false news, terrorist content and propaganda online. Specific mechanisms include: legislative and policy measures; blocking content and access to social media platforms; and filtering and removal of extremist content from platforms. Many extremists end up barred from social media (Zeiger & Gyte, 2020; Braddock, 2019).

▪ **Counter-messaging**: This entails an engagement approach and is related to strategic communications that counter hate speech and the narratives of extremism and terrorism online and/or provides alternative messaging (Zeiger & Gyte, 2020; Braddock, 2019). The aim is to affect the beliefs and attitudes of individuals exposed to mis/disinformation, propaganda and extremist messages and to undermine the appeal of those messages (Braddock, 2019; van Eerten et al., 2019). Programming is in a nascent phase, with various

approaches, including: government-led counter messaging through dedicated websites and presence on social media platforms; civil society curated videos and online content; mobile apps to encourage collaboration in counter-messaging; and digital literacy (Bilazarian, 2020; Elsayed, 2019; van Eerten et al., 2019).

Both disruption and counter-messaging are essential: it is important to restrict the online space for extremist narratives to spread; while also constructing a compelling alternative that resonates with local people (Rosand & Winterbotham, 2019). This report focuses on counter-messaging, centred on online initiatives. Long-term efforts to combat hate speech and extremist narratives are also outside of the scope of this report. These include: legal protections for minority populations; the fostering of tolerant and diverse societies, ensuring appropriate redress and protection mechanisms are in place; and addressing past legacies of violence (GAAMAC, 2021).

Concepts

For definitions of misinformation, disinformation, hate speech and extremism, see the companion report: *Kelly, L. (2022). Online hate speech and mis/disinformation in fostering extremism.*

Counter-narratives comprise content that challenges the themes that appear in another type of narrative (Braddock, 2019). In the field of CVE, it is seen as an attempt "to challenge extremist and violent extremist messages, whether directly or indirectly through a range of online and offline means" (Zeiger & Gyte, 2020). Common themes include: deconstructing extremists' arguments; emphasis of suffering at the hands of the extremist group in question; undermining the credibility of the group; or the re-characterisation of those portrayed as enemies of the group (Zeiger & Gyte, 2020; Braddock, 2019).

Alternative narratives seek to replace narratives that advocate hate and (violent) extremism rather than react to them and challenge them head-on (Braddock, 2019). They often comprise positive messages, such as peace, unity and coexistence (O'Connor & Hlaing, 2017); and messages that can address structural or personal grievances such as building positive identities and enhancing social cohesion (Zeiger & Gyte, 2020).

(Digital) media literacy entails teaching people to be more judicious in consuming information, including having the natural inclination to access, evaluate, and fact-check the materials they read (Haunschild et al., 2022; Vasu et al., 2018). Instilling critical thinking skills in national education systems specifically with the aim of countering fake news is a new concept, with very few extant case studies (Vasu et al., 2018). There may be lessons from the CVE experience, where critical thinking skills, which are useful in steering youth away from radicalisation, can be applied to fake news (Vasu et al., 2018).

Online mediums for counter-messaging includes (see Windisch et al., 2021; Rasheed, 2020):

- Videos: short films or animations
- Texts: slogans, hashtags or open letters
- Images: photos or memes
- Online literature, including newspapers and e-magazines
- Interactive social media messaging (e.g. Facebook, Instagram, TikTok, WhatsApp, YouTube)
- Audio recordings: podcasts or short audio-clips
- Comics: manga, short panels or graphic novels
- Online education-based interventions (e.g. providing people with digital media literacy skills)

The state of evidence

There is a significant knowledge gap regarding the effectiveness of counter-messaging: research and evaluations on strategic communications are in the early stages (Whittaker & Elsayad, 2019; Reed et al., 2017). While there has been a rise in online counter radicalisation programmes, the efficacy of such programmes remains largely unknown (Agbedejobi, 2017; Ferguson, 2016). There are few studies that look at the impact and effectiveness of counter-narratives (Chan, 2020; Zeiger & Gyte, 2020). Further, studies that do exist are largely descriptive: these studies collect and organise message data into descriptive categories, with the provision of some guidance on good practice (Zeiger & Gyte, 2020; Whittaker & Elsayed, 2019). The few studies that have attempted to evaluate effects are usually based on campaigns conducted in the Global North (Zeiger & Gyte, 2020). Due to the dearth of research that engages empirically with counter-messaging, their dissemination and the outcomes they generate, the effectiveness of this strategy in CVE remains contested (Chan, 2020; Whittaker & Elsayad, 2019).

3. Guidance on good practice

There is a literature base that provides general guidance on good practice for counter-messaging interventions (including key steps in producing narratives – see Chart 1).

Target audience

Effective strategic communications require a clear identification of the target audiences of a messaging campaign and a nuanced behavioural and attitudinal understanding of that audience (Reed et al., 2017). Different target audiences require different messages (Barzegar et al., 2016). While the overall audience is likely to be diverse, priority must be given to those who may be susceptible to violent extremist propaganda (Reed et al., 2017).

It is important to tailor messaging to local contexts and cultures: this can be achieved in part by empowering communities with the knowledge, skills and tools to contest and counter messages of violent extremism and terrorism in their networks (Zeiger & Gyte, 2020). Some organisations have provided capacity-building to "local voices" as part of their broader communications campaign against terrorism (Zeiger & Gyte, 2020). This is the case, for example, with Kenya-based YADEN's #insolidarity campaign, which provides capacity-building, tools, and platforms for youth to develop their own messages and share their stories about how terrorism has influenced their lives[2] (Zeiger & Gyte, 2020). UNHCR has developed a comprehensive community training package on Using Social Media in Community-based Protection that covers various aspects of recognising, protecting and countering misinformation, rumours and hate speech against refugees and internally displaced persons.

---

[2] In this video, two young women from Kwale County share their stories and encounters with terrorism, talking about how youth have been recruited for the terrorist groups and the dangers their community faces with respect to violent extremism (Zeiger, 2018).

Chart 1: Key steps in developing a counter- or alternative narrative

Source: Zeiger, 2016: 5. This chart has been removed for copyright reasons. The chart can be viewed at https://hedayah.com/app/uploads/2021/09/File-3182016115528.pdf

Message design and content

A key consideration for message designers is whether messages are offensive (sets the agenda) or defensive (reactive—aimed at countering an existing narrative) – overlapping in large part with alternative and counter-narratives, respectively (Whittaker & Elsayed, 2019). In the case of the latter, message content is crafted dependent on the adversary's messaging; as such, the adversary tends to not only initiate, but also shape the pace and nature of the information contest (Reed et al., 2017). When designing CVE messages, it is suggested that offensive messages should outweigh defensive messages, due to perceived danger in allowing violent extremists to control the debate (Whittaker & Elsayed, 2019). A survey of civil society organisations (CSOs) and private-sector actors finds that there is an overwhelming preference for the development of alternative narratives, and on resilience, than on counter-narratives, in addressing violent extremism (Barzegar et al., 2016). A recent analysis of over two hundred messages from ten multi message campaigns drawn from the MENA and the Daesh Defector data collections, hosted online by the Hedayah Centre finds that (Whittaker & Elsayed, 2019):

▪ Most messages in campaigns are offensive (85.8%), rather than defensive (14.2%).
▪ There is an equal distribution of positive and negative messages[3]; but messages become more negative as they reach a more extreme target audience (tacit or active supporter vs. antis, curious or engaged), in order to trigger a behavioural change away from support.
▪ CVE campaigns have a tendency to highlight atrocities rather than the incompetency or hypocrisy of groups.

An important component of alternative messaging is to provide a non-violent action that still addresses grievances that communities might have that can be underlying drivers of radicalisation (Zeiger & Gyte, 2020). The Anataban ("I am tired") movement recognizes that violence and corruption are normalized in South Sudan and citizens are encouraged to take proactive steps towards peace while not tolerating violence and corruption in their communities. The campaign reaches people through music, art and videos; and encourages social media users to use hashtags promoting peace in their posts (Zeiger & Gyte, 2020).

Appealing to human emotion is important as evidence alone can be countered (RAN, 2016). A growing body of research on the spread of online messages finds that messages that arouse positive emotions or perceptions of novelty are more likely to be disseminated through online social networks (see Braddock, 2019). The "Destroyed Lives" campaign in Somalia sought to reach its audience with an emotional appeal, featuring a video that highlights the stories of mothers of victims of Al-Shabaab attacks and the effects on them (Zeiger, 2018).

Humour, satire or sarcasm can also be a tool to delegitimise extremist narratives and discredit terrorist organisations – working through distortion and exaggeration (parody), critique and mockery (satire) of serious issues (Haunschild et al., 2022; Zeiger & Gyte, 2020; RAN, 2016). However, given the subject matter, these tactics should be used carefully and in a sensitive manner (RAN, 2016).

---

[3] Negative messages attack the linkages between violent extremists and their proposed solutions; while positive messages emphasise possible solutions to crises beyond violent extremists (Whittaker & Elsayed, 2019).

Counter-messaging campaigns can also involve local popular cultures, where appropriate, such as hip-hop, rap, photography, cinema, music and art). Several ongoing experiments have demonstrated the value of such creative interventions (Udupa, 2021). These tactics can also involve the dissemination of artwork, memes and GIFs. In Myanmar, in the midst of violent attacks against the Rohingya Muslims online and offline, Facebook stickers featuring a flower in an animated character's mouth were created as a symbol of peace with the message, "End hate speech with flower speech" (Udupa, 2021).

Effective counter-messaging needs to be prominent on diverse webpages, reaching diverse communities. A study on counter-speech, analysing online posts and comments on prominent public pages in India finds that counter-speech is not limited to a single narrative (Merchandani et al. 2018). It suggests that in order to be effective, counter-speech must take place on pages that serve as a common meeting ground, rather than stay restricted to niche pages that serve solely to reinforce the views of like-minded individuals (Merchandani et al. 2018). The study identifies the following additional lessons for the effective design of counter-messaging (Merchandani et al. 2018):

▪ Counter-speech and counter-narratives must focus on current affairs and directly address present concerns. Otherwise, they risk being irrelevant.

▪ Narratives that can smooth tensions between religious communities are more successful than those that address political differences.

▪ Calls to reduce the pitch of rhetoric, without necessarily offering a counterpoint, are prevalent but not effective.

▪ Counter-speech is effective when it addresses ideas, not specific content.

▪ Administrators of mainstream pages with large followings must be encouraged to identify, highlight and promote counter-speech.

▪ Counter-speech tends to take place more in comments than in posts; as such, users of Facebook and other social media platforms can be encouraged to become a part of an ecosystem that promotes dialogue and community.

Messenger

The credibility and trust worthiness of the individual, group or institution delivering the message or narrative is as important as the message itself (RAN, 2016). Different messengers should be utilised for different types of alternative narratives—distinguishing, for example, between individuals with established appeal to broader audiences and those more credible to potential violent extremists (Bilazarian, 2020; RAN, 2016). A survey of CSOs and private sector actors finds that alternative narratives are most effective when they emerge organically and are disseminated by trusted community leaders who already have credibility and an understanding of the kinds of messaging that will connect with at-risk audience (Barzegar et al., 2016). Recruiting local cultural influencers to be part of community WhatsApp groups to counter hateful speech, can also be effective. Influencers could include poets, musicians, cinema celebrities, online meme creators, digital influence service providers, and online game developers (Udupa, 2021). Involving these actors will help to ensure that positive narratives are culturally resonant in local contexts; and adopt the formats and logics of how discourses actually circulate in online networks (Udupa, 2021).

Online and offline counter narratives often fail to recognize that **civil society actors, thought leaders, and social media influencers, who are credible to general audiences, may be less convincing to the small group that sympathizes with violent extremism** (Bilazarian, 2020). Instead, former members of terrorist groups, members of the community being targeted by violent extremist propaganda, or individuals who agree with message targets' beliefs in principle (but not with violent practices to defend them) could be particularly useful for distributing messages (Braddock, 2019). Former violent extremists can be particularly adept at presenting counter narratives utilizing their pre-existing network connections and knowledge, and pointing out hypocrisies of a group, such as highlighting corruption or other injustices (Bilazarian, 2020; Zeiger & Gyte, 2020). A publicised online video of an interview with a former Daesh member from Indonesia, for example, undermines the credibility of Daesh by highlighting the corruption within the group (Zeiger, 2016). Synchronizing various counter-narrative efforts online by different credible messengers across multiple platforms, in real time, creates greater complexity, however, in developing and implementing digital campaigns (Bilazarian, 2020).

Critical thinking

Target audiences should not be told what to think, but rather taught how to think (Braddock, 2019). There is growing recognition of the potential effectiveness of incorporating the development of critical thinking skills in counter-narrative activities (Rosand & Winterbotham, 2019). CVE programming aimed at promoting alternative interpretations of the ideology in question could gain more traction by not contending that the target audiences are "wrong" about their beliefs, but rather that there may be multiple ways of looking at and interpreting an issue (Braddock, 2019).

(Digital) media literacy programmes and social media education campaigns can play an important role, raising students' critical thinking and awareness of the tactics of online ideological propagation and recruitment; and contributing to more constructive interactions (Sher & Sturm, 2018). People often do not learn proper online search techniques, basic critical thinking skills, and the ways to identify bias in an argument (Sher & Sturm, 2018). Critical thinking and media literacy involve teaching people to be more judicious in consuming information; and to develop a natural inclination to question and fact-check what they read (Vasu et al., 2018).

Integration into new networks

Online CVE efforts need to institutionalize systems not only to dispel extremist narratives but to replace the immersive social experience provided by extremist groups (Bilazarian, 2020). While extremist groups attempt to foster strong intergroup bonds, counter-narrative campaigns have tended to provide weak or non-existent relational alternatives (Bilazarian, 2020). Studies of social media usage find that interpersonal connection in the form of messaging or comments is linked to higher rates of well-being and belonging. This includes various forms of engagement, such as: sharing messages, posts and comments (Bilazarian, 2020). Online counter narratives need to adopt interactive counter-narrative techniques that can reduce alienation and promote belonging – providing new social connections outside extremist networks alongside a positive alternative narrative; however, this can be challenging (Bilazarian, 2020).

Integration of online and offline—and longer-term—programming

Counter-messaging on the Internet and social media might place "seeds of doubt", but may not interrupt the radicalisation process on its own. Rather, they should go hand in

hand with offline counter-measures, for example, educating young people at school about the consequences of violent extremism (RAN, 2016). While online counter narratives are receiving increased attention, online and offline countering violent extremism approaches are rarely well coordinated (Bilazarian, 2020). As the fields develop, both areas would benefit from mutual learning that builds on areas of past success and failure, and that works together to address common challenges (Bilazarian, 2020).

The integration of online and offline programming may also be required out of pure necessity in geographic areas where online forms of communication are less common, such as in East Africa (Zeiger, 2018). Similarly, in Pakistan, a large section of population in, especially in rural areas, has no or limited access to the Internet (Jan, 2016). Facebook is very popular among the urban youth, but its proliferation is still limited in much of the country and among women (Sahotay et al., 2020). As such, online social media campaigns cannot replace offline activities, especially when targeting rural communities and women (Sahotay et al., 2020).

Similarly, in Southeast Asia, even though internet use is generally high, there are large rural populations in many areas of Indonesia, the Philippines and Malaysia that do not have consistent access to the internet (Zeigar, 2016). In such situations, lessons can be drawn from what has been called the "Sneakernet", whereby **the Internet still plays a role in influencing networks, but face-to-face contact and interpersonal relationships reinforce and interact with messages originating online** (Zeiger, 2016). It is essential for counter-narratives to reoccur across several platforms so as to be reinforcing rather than "one-off" messages (Zeiger, 2016).

There is also a growing recognition among researchers and practitioners of the potential effectiveness of combining counter-narrative activities with other longer-term preventive processes (Rosand & Winterbotham, 2019). In particular, policymakers must acknowledge unaddressed socio-economic and political factors that gave rise to violent extremism in the first place (Rosand & Winterbotham, 2019). In the case of Pakistan, for example, it is argued that if counter-narrative efforts fail to pay attention to the social conditions which provide a pool of potential terrorists, it will have limited success in dislodging terrorist ideology (Iqbal et al., 2019).

4. Survey of programming in the Global South

Against violent extremist (AVE)

AVE, launched in 2012, is a network created by the Institute for Strategic Dialogue (ISD) to tackle violent extremist globally. The key focus is preventing the recruitment of "at-risk" youth (Amit et al., 2021). AVE's mobile app, launched in 2017, along with its website, brings together former violent extremists, from different histories and backgrounds, and survivors of violent extremist—enabling them to share experiences, leverage lessons, and work together to combat extremist propaganda or narratives. Specifically, the app enables downloaders to connect and communicate with the members of the network, discover AVE events and initiatives near them and collaborate with them (Amit et al., 2021). It has reached 2641 connections, including 310 formers and 164 survivors, and together, they have been engaged in about 81 projects (Amit et al., 2021). ISD has reported that the AVE network has played an integral role in the success of their counter-narrative programming and other initiatives (Amit et al., 2021).

Asia

The "Challenging hate narratives and violations of freedom of religion and expression online in Asia" (Challenge) project, funded by the European Instrument for Democracy and Human Rights (EIDHR) is a three year project focused on Bangladesh, India, Indonesia, Myanmar and Pakistan. These countries are characterised by the presence of majority religious communities and challenges for secularists and minorities (APC, 2021). One of the key focuses of the project is: understanding and countering hate speech online, by generating narratives and discourse that defend diverse opinions (APC, 2021). Within this focus, two of the objectives are:

▪ To build the capacity of individuals and civil society organisations in South and Southeast Asia to counter hate speech online.
▪ To generate creative, artistic and critical content that promotes discourses and narratives that are secular, diverse, inclusive and rights-respecting on issues relating to religion.

APC and its partners will work with a wide range of individuals, such as activists; artists (writers, novelists, poets, painters, musicians, actors, cartoonists and comedians); editors and journalists from selected media outlets; internet users who are influential over various social media platforms; and users who are targeted for their comments and expression online (APC, 2021). They will also work with legislators, internet service providers and intermediaries, human rights lawyers, and regional and international human rights bodies and mechanisms (APC, 2021).

Pakistan

Internet penetration in Pakistan is relatively low at around 35%, with 17% of the population holding a social media account in 2019, compared to the global average of 49% (see Sahotay et al., 2020). Nonetheless, 17% of the population amounts to around 36 million people—most of whom are urban males between the ages of 18 and 34. Facebook is the most popular social media platform, and WhatsApp the most popular messaging platform among this demographic. While it can be effective to target this demographic with online counter-messaging, campaigns that target rural youth or women should rely more on offline initiatives (Sahotay et al., 2020).

Government strategic communication: Pakistan's military has engaged in strategic use of social media to encourage and sustain public support for its ongoing war against terrorism (Khan & Pratt, 2020). The Inter-Services Public Relations (ISPR) is the strategic military communication wing: it publishes Hilal, a monthly magazine published in English and Urdu; produces films and dramas for electronic media; and since 2013, has been using social media as a key tool in communication. This has involved the appointment of social-media experts and design of an anti-terrorism campaign on Facebook and Twitter (Khan & Pratt, 2020).

A study on the ISPR's strategic use of Facebook finds that, between 2012–2016, it engaged 5.5 million people on Facebook in Pakistan (15.6% penetration) with diverse ethnic, educational, and socioeconomic backgrounds (Khan & Pratt, 2020). Nonetheless, the vast majority of Facebook users are between the ages of 18 and 44 (86.2%) and male (78.5%) (see Khan & Pratt, 2020). During the study period, ISPR's Facebook page received posts from 45 countries, most prominently from Pakistan, Saudi Arabia, United Arab Emirates, India, Bangladesh, Afghanistan, the U.K. and the U.S. (Khan & Pratt, 2020). Overall, ISPR's critical debates on Islamic extremism and terrorism in Pakistan, and its strategic communication on social media in the war against terrorism reached nearly 19 million people in three years (Khan & Pratt, 2020).

The study also finds that the ISPR used four different types of strategic frames in its posts, in response to a fast-changing, evolving security situation in the country (Khan & Pratt, 2020):

▪ **Diagnostic** frames were posts that identified the causes of and problems from terrorism and assigned blame (e.g. Indian involvement in terrorism in Pakistan; Afghanistan factor);
▪ **Prognostic** frames were posts that proposed solutions, strategies, and tactics to respond to terrorism in Pakistan (e.g. security components; democratic issues; and education)
▪ **Motivational** frames provided a rationale (logical and emotional) to act against terrorism or terrorists (e.g. heroism and resilience);
▪ **Informational** frames provided general information about Pakistan's armed forces and their readiness to confront conventional and nonconventional threats.

Framing was used strategically to facilitate public–military and people-to-people engagements. Motivational frames were the most dominant forms of communication (44.7%), used to generate dialogue between the military and the public on virtual platforms in the war against terrorism (Khan & Pratt, 2020). The military also focused on prognostic frames (42.7%), with a focus on security, which can generate significant people-to-people dialogue that might help in building a consensus on taking measures against Islamic extremism in Pakistan (Khan & Pratt, 2020).

The study operationalised public engagements with ISPR posts at three levels:

▪ low level of engagement (likes),
▪ moderate level of engagement (shares); and
▪ high level of engagement (comments).

Strategic communications were dominated overall by a low level of engagement (86.33%), followed by moderate (11.48%), and high level of engagement (2.19%) (Khan & Pratt, 2020). Motivational frames received the highest public engagement (e.g. comments), followed by the prognostic frames, likely due to the use information-rich and powerful photos and videos embedded with highly emotional and logical arguments (Khan & Pratt, 2020). Public engagement with informational frames and diagnostic frames were significantly low (Khan & Pratt, 2020). The study also finds persistent increases in the high level of public engagement (comments) with people-to-people engagement overtaking public-military engagement over time. This suggests that public–military engagement facilitated people-to-people dialogues (Khan & Pratt, 2020).

Counter-narratives: Research on youth in Pakistan finds that they are prolific social media users, with some using social media for up to 10 hours a day (Sahotay et al., 2020). Civil society counter-messaging campaigns directed at youth should thus focus on social media platforms, particularly Facebook and WhatsApp. Educating youth about misinformation and disinformation, with a focus on improving digital literacy, could help to counter the proliferation of 'fake news'; and help young people to identify and counter discriminatory speech and propaganda spread by extremist groups (Sahotay et al., 2020). The research also suggests that incorporating pop culture trends in the design of messages is an effective way to reach audiences, as it embeds critical narratives and deconstruction of stereotypes in content that different groups of young people seek out. The approach is more subtle, rather than an explicit attempt to exert strategic influence (Sahotay et al., 2020). While pop culture in Pakistan has often solidified stereotypes and discriminatory norms against women and other minority groups, more inclusive perspectives and narratives have been emerging in recent years (Sahotay et al., 2020).

Harakut-ut-Taleem: Research on an online counter-narrative campaign, including Harakut-ut-Taleem, which targeted Taliban extremist narratives promoted in Pakistan, finds that sharing

counter-narratives online can foster online conversations. It can also provide a better understanding of the target audiences' reactions towards the counter-messages (Silverman et al., 2016). The Harakut-ut-Taleem campaign consisted of six videos (three videos in Urdu and three with English subtitles), which highlighted the negative aspects of the Taliban's actions, including one in which the protagonist details the crimes he witnessed as a member of the Taliban (Silverman, et al., 2016). The videos were shared on Twitter, YouTube and Facebook, with the vast majority of engagement (e.g. how many people liked, disliked, shared, retweeted, or commented) occurring on Twitter.[4] Comments were supportive and negative in equal share. While this engagement suggests that such campaigns can be effective in driving conversations among vulnerable communities, it is uncertain whether such counter-narratives can actually deter radicalisation (Silverman et al., 2016). If conversations are sustained, however, already radicalised people and at-risk individuals could get ongoing exposure to alternative viewpoints that can produce a "seed of doubt" in their minds (Silverman et al., 2016).

Addressing online gender-based hate speech and extremism: Women and other minorities in Pakistan have gone online in search of safe spaces, which are not readily available offline; however, harassment and hate speech has often followed them online (George, 2016). The Digital Rights Foundation in Pakistan provides digital safety trainings for women; a helpline (Ab Aur Nahin/not anymore) that provides free legal counsel and referrals in cases of online harassment and abuse; and campaigns such as "Hamara internet/our internet" to map and raise awareness about online harassment using interactive online formats like quizzes (Udupa, 2021; Rehman, 2017). Hamara internet aims to provide a secure digital environment for women, teaching feminist activists how to utilise digital tools for secure communications (e.g. privacy and encryption features), thus enabling online challenges to patriarchy and dissent to continue in a safe manner (Rehman, 2016). There are concerns, however, that such online spaces are inaccessible, 'western' and elitist. They need to cater to all Pakistanis, in different languages, ensuring that social hierarchies are not replicated online (Rehman, 2017).

Peer support groups: Given that sex, harassment and abuse are taboo subjects in Pakistan's patriarchal society, women often refrain from reporting their experiences with rape and marital abuse (Younas et al., 2020). Studies reveal that an important factor in the decision to disclose is the presence of supportive, non-judgmental peers (Younas et al., 2020). Technologies, particularly Facebook, have been leveraged in the Global South to find such connections and anonymous support (Younas et al., 2020). In Pakistan, there has been a recent rise in closed women's Facebook groups as an alternate to the lack of safe, non-judgmental spaces offline. A study of one such group with over 15,000 female members reveals the vital importance of such digital safe spaces (Younas et al., 2020). These private groups with only female members are spaces where the rules of patriarchy are suspended; and women can be vulnerable and share their narratives, without fear of being harassed or stalked (Younas et al., 2020). Analysis of the Facebook group in Pakistan shows that women will only share their narratives anonymously; and often only in spaces with little overlap with their real world social circles (Younas et al., 2020).

Alternative narratives: Online video channels have emerged in Pakistan in response to harassment and to promote feminist politics. For instance, The "Girls at Dhabas" project in Pakistan invites women in South Asia to "reclaim public spaces on their own terms" and

---

[4] There were 5,463 engagements in total (Twitter, Facebook and YouTube), with the vast majority from Twitter (4,814) during the time the videos were posted. Harakat ut-Taleem, created as a new organisation for this project, gained 116 page likes on Facebook, 62 followers on Twitter, and six subscribers to their YouTube channel, suggesting the value of targeted Twitter campaigns to build the online presence of new organisations (Silverman, 2016).

publishes the stories of women taking a stroll on the streets in the night, or hanging out at "male-only" venues, and celebrating these moments of "transgression" and "occupation" through visual and textual narratives online (Udupa, 2021).

Burka Avenger and Gogi: Graphic narratives like Gogi (1970–present) and its newly hosted website, by Nigar Nazar, and Burka Avenger (2013–present), by Haroon Rashid, are popular productions in Pakistan. Gogi appears in several different media such as newspapers, printed books, television, and online resources to reach its desired audiences. Burka Avenger is an animated television show, broadcast on national television in Pakistan, Afghanistan and India; and garnering millions of views online through its free YouTube channel (with English subtitles) (Pirzada, 2017; Barzegar et al., 2016). Both adopt alternative narratives, deviating from the binary us-versus-them way of thinking—engaging in complex discussions of gender, adolescence, religion, and urban reality; and promoting peace and tolerance in society (Pirzada, 2017; Barzegar et al., 2016).

In Gogi comics, Gogi is a powerful social activist, trendsetter, and conscientious citizen. She regularly voices her opinions on issues that plague Pakistani society like honour killings, street harassment, violence against women, and restrictive views on female education (Pirzada, 2017). In Burka Avenger, Jiya is a school teacher by day and Burka Avenger at night. By celebrating Jiya's role as a superhero and teacher, the show implicitly honours the pursuit of leadership roles and professional careers by young women (Pirzada, 2017). Both girls also visit social spaces, an act generally considered taboo for young girls. Burka Avenger has also addressed issues of terrorism, while drawing out uniting factors in society, such as music (Pirzada, 2017).

Research on Burka Avenger finds that it has achieved much success: it is the most watched children's TV show in Pakistan and has received support from wide-sections of Pakistani, Afghani and Indian society, with a notable absence of pushback from conservative and extremist groups (Barzegar et al., 2016). The show is also internationally acclaimed, having receiving many awards (Barzegar et al., 2016). The success of the show can be linked in large part to its culturally sensitive content and use of local talent (popular singers, producers and other artists) (Barzegar et al., 2016). The show's creator is also a pop star, whose celebrity status made her an ideal messenger, capable of drawing large audiences for Burka Avenger to deliver its messages of peace, tolerance, acceptance, cooperation, gender-equality, and other critical values (Zeiger & Gyte, 2020).

Digital literacy: There is little evidence on the effectiveness of educational interventions that focus on teaching social media users how to recognize fake news (Ali & Qazi, 2021). A recent evaluation of two educational interventions for countering misinformation in urban Pakistan finds that provision of personalised feedback is more effective than general informational messages (Ali & Qazi, 2021). In the first intervention, individuals are shown an informational video in the national language that educates them about common features of misinformation; in the second intervention, individuals are shown the video, followed by personalized feedback about their own responses to fake news stories. The feedback highlights the features of each fake news item that can enable the individuals to identify the news as fake. While the first intervention did not have any significant impact, participants in the second intervention were more likely to correctly identify fake news relative than the control group (Ali & Qazi, 2021). Males benefited more from the intervention, which could be due to the general lower level of digital literacy among women (Ali & Qazi, 2021). Such educational interventions have the potential to be delivered at scale on social media platforms, in participants' own languages (Ali & Qazi, 2021).

For further discussion, please see K4D's companion report on online spaces in Pakistan targeting women, religious minorities, activists and voices of dissent.

Bangladesh

Knowledge sharing and counter-messaging: Civil society and public officials have launched mobile apps in Bangladesh aimed at sharing information and knowledge about extremism and crime. The *YPSA CVE Initiative* is a mobile app, launched by the non-governmental organisation Young Power in Social Action (YPSA). The app is entirely in Bangla and free to download and use, focusing on countering violent extremism in Cox's Bazar (Amit et al., 2021). It includes sessions on family and society, extremism and violence, the negative sides of extremism, and the victims of extremism. *Hello CT* is a mobile app, launched by the Counter-Terrorism and Transnational Crime (CTTC) unit of Dhaka Metropolitan Police (DMP). Users of the app can share information, including through audio and video clips, regarding aspects of extremism and terrorism. The app is in Bangla and English and free to download and use. The app also has a news feed section, where local news, slogans, best practices, and campaigns are shared for awareness building among users (Amit et al., 2021).

Addressing misinformation: The Bangladesh Peace Observatory (BPO), based at the University of Dhaka, is part of the UNDP's project on Partnership for a Tolerant, Inclusive Bangladesh. It seeks to advance peace and tolerance by creating a virtual platform that shares knowledge about the state of political, ethnic, communal, criminal, gender-based, and extremist violence (Amit et al., 2021). Its website contains different sources of publicly available data and visualises those in a useful and interactive way, which makes information easily digestible, contributing to countering rumours and fear by providing facts (Amit et al., 2021).

India

(Digital) media literacy: A study on in-person media literacy training to counter misinformation, based on two tools—reverse image searching and navigating a fact-checking website—finds that there was no significant effect on the respondents' ability to identify fake news on average (Badrinathan, 2020). There was a difference in outcomes, however, based on political partisanship: non-BJP respondents were able to successfully learn from the treatment, identifying fake news at a higher level; whereas the presence of partisan motivated reasoning subdued any positive effects for BJP respondents, who became significantly less able to identify fake news, particularly pro-attitudinal fake news (Badrinathan, 2020)

Indonesia

Government strategic communication: The national counterterrorism agency (BNPT) engages in counter-radicalisation messaging offline (seminars, workshops and training activities for BNPT partners, government agencies, NGOs, and community leaders) and online (media literacy programmes and other activities through website, social media, media electronics, and or print media) (Al Tahaj & Logahan, 2019). The website is managed by a third party in collaboration with the BNPT. The programme has yet to demonstrate success and still needs time to achieve optimal results (Al Tahaj & Logahan, 2019). Online and offline radicalisation processes and radical terrorist ideologies persist in the country. Research suggests that the BNPT should prioritise a religious approach by presenting a counter-narrative of a peaceful Islam and

subsequently an alternative narrative emphasising love for the country (Al Tahaj & Logahan, 2019). They should also rely more on third party messengers and involve the community and related government agencies to support the programme (Al Tahaj & Logahan, 2019).

Counter-messaging: The Dare to be peace #beranidami campaign was targeted at youth in urban areas in Indonesia who are active on social media such as Facebook, Twitter and YouTube. They are considered particularly vulnerable due to limited understanding of religion (particularly in Jakarta) (Zeiger, 2016). The campaign coordinated the efforts of 130 community groups in Indonesia, including theatre groups, religious groups, women's groups, human rights activists, business groups, youth groups, ex-jihadis, network of victims of religious and terrorist attacks and labour unions. Key messengers included Indonesian celebrities, such as political figures, sports stars and musicians (Zeiger, 2016). The programme included a large rally, TV broadcasts of celebrities promoting the programme, and YouTube videos—all of which had the aim of promoting messages of moderation and tolerance to counter messages of intolerance reinforced by violent extremists (Zeiger, 2016). Drawing on aspects of local cultures to localise the message of tolerance and moderation is considered to be an effective tactic (Zeiger, 2016).

Timor Leste

Digital media literacy: A part of a larger programme empowering youth to engage positively online and prevent violent extremism in their communities, UN Women engaged 44 young (age 18-26) active social media users (primarily women) to develop digital literacy capacities to counter hate speech and promote social cohesion (UN Women, 2021). This included ensuring a gender-sensitive approach in countering gender stereotypes and promoting social cohesion. It is part of UN Women's wider efforts in empowering youth to engage positively online and prevent violent extremism in their communities (UN Women, 2021). A post survey conducted in 2020 finds that all the participants' gained knowledge on misinformation, fake news, fact checking, social cohesion and content creation after the 16 hours training. In particular, they were able to identify, protect and report against online hate speech and misinformation; and create and upload positive-themed videos and posters on their social media channels (UN Women, 2021).

Middle East and North Africa

Gender-based counter-narratives using messengers of female preachers and/or experts can play a central role as credible messengers in providing alternative and counter-narratives (Elsayed, 2019). This is a strategy that has been adopted by various international organisations and governments in their counter-messaging programming in MENA. The Organisation of Islamic Cooperation, for example, has launched the Centre for Dialogue, Peace and Understanding (CDPU) or Sawt Al-Hikma (The Voice of Wisdom) as a counter-messaging platform dedicated to delegitimising extremist discourse on social media and the Internet (Elsayed, 2019). The initiative relies on female Muslim preachers and/or experts in a number of multimedia, graphics and animation clips that aim at correcting the concepts used by terrorist groups in justifying their actions (Elsayed, 2019). Muslim women within MENA are more likely to engage with and communicate with other women, which makes the role of female preachers and/or experts particularly important (Elsayed, 2019). While the effectiveness of these narratives at reducing violent extremism has yet to be seen, they have managed to raise the profile of women, presenting female role models for other women (Elsayed, 2019).

Egypt

Counter-messaging: The Al Azhar Observatory for Combating Extremism provides counter-arguments on some of the misconceptions claimed by terrorist groups; and suggestions for providing replies to religious arguments used by terrorist groups that promote violence (Elsayed, 2019). It amplifies the voices of women, through its awareness-raising campaigns and hashtags addressing women. They comprise animated sketches telling the stories of women that aim at not only debunking campaigns against the Islamic culture, particularly with regard to the allegedly inferior status of women vis-à-vis men in the Islamic societies; but also promoting a culture of respect and appreciation of a religious character, using sacred texts and taking viewers back to the time of prophet Mohammed (Elsayed, 2019). These campaigns include:

- "Anti Malikah" (You are a Queen): highlights the role of women and reaffirms their high standing in Islam.
- "Law kona fi zaman al nubowah" (What if were living in the Prophet's time): features stories of women that played a prominent role in the Islamic history.
- "Hunna" (Feminine "They") and "Enjazatuha tulhemoni" (Her achievements inspire me): provide inspirational alternative narratives through featuring profiles of successful women in a broad range of fields.

Kingdom of Saudi Arabia

Counter-messaging: Etidal (Moderation), a centre based in Riyadh, works in collaboration with governments and organisations to counter the exclusionary discourse promoted by groups like Daesh; and to draw attention instead to concepts of moderation in coexistence in MENA societies. It carries out its alternative and counter-narrative campaigns Facebook, Twitter and YouTube, including through posting videos (Elsayed, 2019).

Assakina (Tranquility) is an online website, offered in Arabic, English, Russian and Turkish. It collaborates with Islamic scholars to address issues such as jihad, political violence and radicalisation; and to expose and refute extremist ideologies. Assakina has an online multi-media database (with over 40,000 digital, audio and video materials) of religious guidance to answer questions on Islamic belief (Elsayed, 2019). Its "Creative Minds for Social Good' campaign seeks to promote tolerance, multiculturalism, inter-religious dialogue and inclusiveness. The campaign targeted cities that were sources of terrorist groups' recruitment through short video campaigns (Elsayed, 2019). It draws in particular on the experiences of Saudi females, targeting parents and mothers. One of the videos portrayed a mother who was inadvertently instilling discriminatory principles on her son; as the video progressed, she changed her views by promoting the values of respect for others regardless of ethnicity, ancestor or family/tribe (Elsayed, 2019).

Kuwait

Counter-messaging: A video commercial, produced by Zain (a Kuwaiti communications company) sought to counter Daesh's narrative of divided communities; and of justifying acts of violence as religiously and morally acceptable. It also sought to promote peace and tolerance and to publicise that there is a "way out" for terrorism (Elsayed et al., 2017). The target audience was the general audience in the MENA region. The messenger for the narrative was a popular singer, Hussain Al Jassmi, from the UAE, who sang in Arabic. The video aims to have an emotional appeal, showing clips of victims of terrorism to highlight the human face of the

consequences of violence and terrorism. It also shows the lead singer extending his hand to a potential suicide bomber, with the line "confront your enemies with love, not war"— emphasizing that love, peace, tolerance, patience and mercy are core values of Islam and Middle Eastern society, not violence or terrorism (Elsayed et al., 2017). The commercial was disseminated on Arabic TV channels and various social media platforms, including on YouTube. It received over 10 million views as of October 2017, with active online engagement in the form of over 10 thousand comments on YouTube and Twitter (some pro and some against) (Elsayed et al., 2017)

Qatar

Countering mis/disinformation: Governments and independent groups have set up websites to debunk fake news that constitute disinformation (Vasu et al., 2018). In Qatar, "Lift the Blockade" is a government website, set up in 2017, to counter what Qatar regarded as fake news distributed by geopolitical rivals to justify the imposition of economic sanctions amid the gulf crisis (Vasu et al., 2018). Illustrated with images of the country and its people, the site profiled Qatar's perspective on the dispute; detailed the impact the blockade had on citizens; and included recent international news articles that show Qatar in a positive light (Scott, 2017). While these sites have the potential to dispel 'fake news', the audience is likely to be limited to those predisposed to fact-checking due to their cognitive biases or to digital illiteracy (Vasu et al., 2018). It also assumes that the reader will trust the findings of the fact checkers (Vasu et al., 2018). Given these challenges, it may be beneficial to run such websites in tandem with wider strategic communications efforts (Vasu et al., 2018).

Africa

Counter-messaging: The *UNDP Africa Toolkit: Tackling Extremist Narratives* mobile app has been designed as a toolkit for CSOs and groups to develop their own counter narratives campaigns, targeting violent extremism. The developers of the app suggest that alternative narratives are more effective that counter-narratives, as the former can define "what we are", rather than "what we are against". The app provides a step-by-step guideline on strategic and communication elements, starting from primary concepts of narratives to planning a campaign— with an emphasis on positive and inclusive messages rooted in a community's history and traditions (Amit et al., 2021). It also shares case studies and other knowledge. The app is available in English and French; and can be downloaded for free – after which, it does not require Internet connect (Amit et al., 2021).

Kenya

Countering mis/disinformation and promoting alternative narratives: During the 2007-2008 election crisis, social media was considered to be a catalyst for violence, spreading hate speech and violent narratives; and a tool for peacebuilding—sharing (or re-tweeting in the case of Twitter) traditional media news items that advocated peace (Sher & Sturm, 2018). Peacebuilding efforts also included the establishment of the website called Ushahidi (Testimony), designed to allow people to report instances of violence via email, SMS, or directly through the site (Sher & Sturm, 2018). Ushahidi was used for violence prevention prior to the 2013 elections, compiling and circulating peaceful messages posted on Facebook, Twitter, and SMS-delivered web postings (Sher & Sturm, 2018). In addition, peace advocates teamed up with a popular comedy series on television during the 2013 elections to embed awareness raising information about online propaganda and hateful speech in the comedy narrative. An impact assessment study

showed that viewers who watched these episodes had a keener grasp of how leaders manipulate their followers with polarising language (Udupa, 2021).

In advance of the 2017 elections, the Sentinel Project and iHub Research launched Una Hakika ("Are you sure?") – a mobile-phone based information service that monitors and checks the spread of rumours and acts to contain them (Sher & Sturm, 2018; IDS, 2018). It was launched in the Tana Delta region, which is especially susceptible to ethnic conflict due to misinformation, particularly with regard to neighbouring villages. Community leaders play a pivotal role is disseminating verified stories (Prabhakar & George, 2019). It is difficult to isolate the effects that the Kenyan social media campaigns had on the election process, but overall the 2017 elections had a small fraction of the violence that afflicted the 2007 elections (Sher & Sturm, 2018).

Zimbabwe

Countering misinformation: There have been multi-pronged social media-induced threats to national security in Zimbabwe, chief among them being social media-instigated violent protests and the spread of 'fake news'. Social media platforms have also been used to enhance Zimbabwe's national security (Mugari & Chisuvi, 2021). A study conducted in Harare finds that social media plays a significant role in information dissemination and as a news platform for citizens. There have been several occasions that government ministers have taken to social media to dispel 'fake news', also disseminated through social media, that could have produced tensions (Mugar & Chisuvi, 2021). Respondents of the study note, however, that there is very limited interaction between the government and citizens through social media channels: there is a one way flow of information from the government to citizens, whereas citizens preferred two-way engagement over social media networks (Mugari & Chisuvi, 2021).

5. References

Agbedejobi, P. (2017). Challenging online radicalisation: A refutation of counter-radicalisation strategies, counter-narratives and a new approach. *Counter-Narratives and a New Approach (March 10, 2017).* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2931179

Al Tahaj, R., & Logahan, J. M. (2019). Counter Radicalization Analysis in Efforts to Prevent the Spread of Radical Ideology in Indonesia. *Journal of Strategic and Global Studies*, *2*(1), 5. https://scholarhub.ui.ac.id/jsgs/vol2/iss1/5/

Ali, A., & Qazi, I. A. (2021). Countering Misinformation on Social Media Through Educational Interventions: Evidence from a Randomized Experiment in Pakistan. *arXiv preprint arXiv:2107.02775.* https://arxiv.org/pdf/2107.02775.pdf

Amit, S., Barua, L., & Kafy, A. A. (2021). Countering violent extremism using social media and preventing implementable strategies for Bangladesh. *Heliyon*, *7*(5), e07121. https://doi.org/10.1016/j.heliyon.2021.e07121

APC. (2021) Challenging hate narratives and violations of freedom of religion and expression online in Asia. Association for Progressive Communications (APC). https://www.apc.org/en/project/challenge

Badrinathan, S. (2020). Educative interventions to combat misinformation: Evidence from a field experiment in India. Working Paper. https://pdri.upenn.edu/wp-content/uploads/2020/09/Fake_News_Paper.pdf

Barzegar, A. P., Karhili, S., & El, N. (2016). Civic approaches to confronting violent extremism. https://www.academia.edu/download/48700293/Civic_Approaches_Sept_8_2016_Digital_Release.pdf

Bilazarian, T. (2020). Countering violent extremist narratives online: Lessons from offline countering violent extremism. *Policy & Internet*, *12*(1), 46-65. https://doi.org/10.1002/poi3.204

Braddock, K. (2019). Communicatively countering violent extremism online. In *The Handbook of Communication and Security* (pp. 247-261). Routledge.

Chan, N. (2020). In moderation of 'strangers': terrorism, ontological (In) security and counter-narratives in Malaysia. *Critical Studies on Security*, *8*(1), 28-45. https://doi.org/10.1080/21624887.2020.1734906

Elsayed, L. (2019). Counter-Messaging for Countering Violent Extremism in the Middle East and North Africa: Using a Gender Lens. In *Enhancing Women's Roles in Preventing and Countering Violent Extremism (P/CVE)* (pp. 95-104). IOS Press.

Elsayed, L., Talal F., and Zeiger, S. (2017) 'Undermining Violent Extremist Narratives in the Middle East and North Africa: A How-To Guide,' Abu Dhabi: Hedayah. https://www.hedayahcenter.org/wp-content/uploads/2019/11/File-18122018114055.pdf

Ganesh, B., & Bright, J. (2020). Countering extremists on social media: challenges for strategic communication and content moderation. *Policy & Internet*, *12*(1), 6-19. https://doi.org/10.1002/poi3.236

George, C. (2015). Managing the dangers of online hate speech in South Asia. *Media Asia*, *42*(3-4), 144-156. Roundtable discussion. https://doi.org/10.1080/01296612.2016.1150585

Haunschild, J., Kaufhold, M. A., & Reuter, C. (2022). Cultural violence and fragmentation on social media: Interventions and countermeasures by humans and social bots. In *Cyber Security Politics* (pp. 48-63). Routledge. https://library.oapen.org/bitstream/handle/20.500.12657/52574/9781000567113.pdf?sequence=1#page=63

IDS. (2018). *Using digital media and social media to monitor and reduce violence in Kenya's elections.* Policy briefing, no. 144. Falmer: IDS. https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/13002/PB144.pdf?sequence=3&isAllowed=y

Iqbal, K., Zafar, S. K., & Mehmood, Z. (2019). Critical evaluation of Pakistan's counter-narrative efforts. *Journal of Policing, Intelligence and Counter Terrorism*, *14*(2), 147-163. https://doi.org/10.1080/18335330.2019.1574020

Jan, F. (2016). Social Media and Cyber-Jihad In Pakistan. *Al-Idah*, *32*(1), 34-47. http://www.al-idah.szic.pk/index.php/al-idah/article/view/154

Khan, M. K., & Pratt, C. B. (2020). Strategic communications: The Pakistan military's use of social media against terrorism. *Media, War & Conflict*, 1750635220972127. https://doi.org/10.1177/1750635220972127

Mirchandani, M., Sahai, D., & Goel, O. (2018). Encouraging counter-speech by mapping the contours of hate speech on Facebook in India. Observer Research Foundation. https://www.academia.edu/download/57492780/Encouraging_Counter_Speech_by_Mapping_the_Contours_of_Hate_Speech_on_Facebook_in_India_.pdf

Mugari, I., & Chisuvi, R. (2021). Social media and national security in Zimbabwe: Embracing social media for national security and addressing social media threats. *African Security Review*, *30*(1), 86-101. https://doi.org/10.1080/10246029.2020.1857806

O'Connor, T., & Hlaing, K. H. (2017). Peacebuilding strategies to combat extremist narratives. *Peacebuilding and violent extremism*, 45. https://www.peacedirect.org/us/publications/full-report-peacebuilding-violent-extremism/

Pirzada, T. (2017). Narrating Muslim girlhood in the Pakistani cityscape of graphic narratives. *Girlhood Studies*, *10*(3), 88-104. https://www.researchgate.net/publication/321656606_Narrating_Muslim_Girlhood_in_the_Pakistani_Cityscape_of_Graphic_Narratives

Prabhakar, T., & George, D. (2019). Considerations in Archiving Misinformation from Encrypted Messaging Apps. https://cyber.harvard.edu/sites/default/files/2019-12/Tattle_Disinformation_Workshop_revised.pdf

RAN. (2016). Preventing radicalisation to terrorism and violent extremism. Amsterdam: Radicalisation Awareness Network. https://ec.europa.eu/home-affairs/system/files_en?file=2021-05/ran_collection-approaches_and_practices_en.pdf

Rasheed, A. (2020). *Countering the radical narrative.* New Dehli: Manohar Parrikar Institute for Defence Studies and Analyses. https://www.idsa.in/system/files/book/book_countering-the-radical-narrative.pdf

Reed, A. G., Ingram, H. J., & Whittaker, J. (2017). Countering terrorist narratives. *Countering Terrorist Narratives.* https://scholarlypublications.universiteitleiden.nl/access/item%3A2951675/view

Rehman, Z. (2017). Online feminist resistance in Pakistan. *SUR-Int'l J. on Hum Rts.*, *26*, 151. https://heinonline.org/HOL/Page?handle=hein.journals/surij26&div=17&id=&page=&collection=journals

Rosand, E. & Winterbotham, E. (2019). *Do counter-narratives actually reduce violent extremism?* Washington, DC: Brookings Institute. https://www.brookings.edu/blog/order-from-chaos/2019/03/20/do-counter-narratives-actually-reduce-violent-extremism/#cancel

Sahotay, J., Ashraf, S., & Hakim, Z. (2020). Voices of Pakistani youth: Lessons for civil society in the development of effective counter-narrative campaigns. ISD. https://www.isdglobal.org/wp-content/uploads/2020/11/2020_11_20-YC-PAKISTAN-V8.pdf

Scott, V. (2017). *Qatar launches new website to counter 'fake news'.* Doha News. 18 September. https://medium.com/dohanews/qatar-launches-new-website-to-counter-fake-news-2f7ef9d0a625

Sher, G., & Sturm, E. (2018) Social Media and Peacebuilding: Could Mindsets be Positively Affected? *Strategic Assessment 21*(3). https://www.inss.org.il/wp-content/uploads/2018/11/Sher-Strum.pdf

Silverman, T., Stewart, C. J., Amanullah, Z., & Birdwell, J. (2016). The impact of counter-narratives. *Institute for Strategic Dialogue*, *54*. https://www.jugendundmedien.ch/fileadmin/PDFs/anderes/schwerpunkt_Radikalisierung/Impact-of-Counter-Narratives_ONLINE_1.pdf

Udupa, S. (2021). Digital Technology and Extreme Speech: Approaches to Counter Online Hate. Commissioned Research Paper for the United Nations Peacekeeping Technology Strategy https://epub.ub.uni-muenchen.de/77473/1/Digital_technology_and_extreme_speech_Udupa.pdf

UN Women (2021) Choosing to Change the Narrative, 44 Youth Make Use of Digital Tools to Counter Hate Speech and Promote Social Cohesion. 8 February. https://asiapacific.unwomen.org/en/news-and-events/stories/2021/02/choosing-to-change-the-narrative

Vasu, N., Ang, B., Teo, T. A., Jayakumar, S., Raizal, M., & Ahuja, J. (2018). *Fake news: National security in the post-truth era*. S. Rajaratnam School of International Studies. http://journres1.pbworks.com/w/file/fetch/135129183/Fake_News_National_Security_in_the_Post-.pdf

Whittaker, J., & Elsayed, L. (2019). Linkages as a Lens: An Exploration of Strategic Communications in P/CVE. *Journal for Deradicalization*, (20), 1-46. https://journals.sfu.ca/jd/index.php/jd/article/view/243

Windisch, S., Wiedlitzka, S., & Olaghere, A. (2021). Protocol: Online interventions for reducing hate speech and cyberhate: A systematic review. *Campbell Systematic Reviews*, *17*(1). https://hope-radproject.org/wp-content/uploads/2021/12/Windisch-et-al.-2020-Online-interventions-for-reducing-hate-speech-and-cyber-hate-A-systematic-review.pdf

Younas, F., Naseem, M., & Mustafa, M. (2020, June). Patriarchy and social media: Women only Facebook groups as safe spaces for support seeking in Pakistan. In *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development* (pp. 1-11). https://dl.acm.org/doi/abs/10.1145/3392561.3394639

Zeiger, S., & Gyte, J. (2020). *Prevention of Radicalization on Social Media and the Internet*. Chapter 12 in Handbook of Terrorism Prevention and Preparedness. International Centre for Counter-Terrorism (ICCT).

Zeiger, S. (2018). 'Undermining Violent Extremist Narratives in East Africa: A How-To Guide,' Abu Dhabi: Hedayah. https://www.hedayahcenter.org/resources/reports_and_publications/undermining-violent-extremist-narratives-in-east-africa-a-how-to-guide-2/

Zeiger, S. (2016). 'Undermining Violent Extremist Narratives in South East Asia: A How-To Guide,' Abu Dhabi: Hedayah. https://hedayahcenter.org/app/uploads/2021/09/File-3182016115528.pdf

Acknowledgements

Suggested citation

About this report