COUNTERSPEECH

IN SRI LANKA

TOOLKIT

GENERATION
පරපුර
தலைமுறை

Co-funded by
the European Union

UN
DP

GENERATION
ஹாஷ்டேக்
தலைமுறை

Hashtag Generation is a youth-led organization run by a group of young tech-savvy Sri Lankans working towards building a society where everyone has the skills, information and tools to be active participants in making the decisions that affect their communities, technologies and bodies. The group adopts a non-partisan approach and works with the strong conviction that decision-making at all levels should remain transparent and inclusive in order to remain sustainable and build lasting peace in Sri Lanka. We mobilize social and new media tools to raise awareness and catalyse dialogue on important social issues that affect youth. We work on three main components, namely - trainings and outreach, communications and advocacy, and social media trend analysis.

The organization which started as a completely volunteer-run model in 2015 has now transformed into a full-time operation. Our work consists of initiatives such as advocating for youth participation in decision-making, strengthening women's civic and political engagement, advocacy for the rights of ethnic, religious and sexual minorities and raising awareness on the importance of cyber security and countering misinformation and online hate speech.

Developed by Sachini Perera for Hashtag Generation

TABLE OF CONTENTS

INTRODUCTION

I see no evil, I hear no evil, I speak no evil,
rather, my credo is,
I see all evil and speak of it,
I hear all evil, and give ear to it,
I speak of evil, my utterance heard aloud,
Hope remains, everlasting, that vanquishes all evil that has
escaped from Pandora's mythic box.

Pandora's Box by Jean Arasanayagam

In 2019 about 34% of the population in Sri Lanka had access to the internet and it is reported that by January 2021, this had increased to 50%[1]. As internet penetration gradually and steadily increases in Sri Lanka – including due to Covid-19 responses that sent education, information, work and other aspects of life online to varying levels of success – there is also an increase in various kinds of expression and speech on the internet. Almost all internet users in Sri Lanka are social media users[2] and much of this expression and speech occurs on social media. Research by Hashtag Generation and other civil society groups[3] indicate that this includes a prevalence and proliferation of online hate speech.

Such speech is not new nor does not occur in a vacuum. "Since its independence, Sri Lanka has seen the rise of various ethno-nationalist factions, often backed by governments, which used ethno-nationalism for political gain",[4] and online hate speech is situated in this online-onground continuum of ethno-religious nationalism. And all of this is in keeping with a groundswell of racism, ethno-nationalism and intolerance in South Asia and worldwide.

---

[1] Annex 3

[2] Helani Galpaya, Ayesha Zainudeen and Tharaka Amarasinghe, AfterAccess: ICT access and use in Sri Lanka and the Global South (LIRNEasia 2019) https://lirneasia.net/2019/05/afteraccess-ict-access-and-use-in-sri-lanka-and-the-global-south-presentation/ accessed 2 April 2021.

[3] Annex 1

[4] Democracy Reporting International, Social Media Analysis - What Facebook Tells us about Social Cohesion in Sri Lanka (Democracy Reporting International 2019)

As noted by the United Nations Special Advisor on Genocide Adama Dieng, **"we all have to remember that hate crimes is preceded by hate speech......we need therefore to make every effort to address hate speech"**.[5] There is a number of tactics that attempt to address hate speech and counterspeech – **"any direct response to hateful or harmful speech which seeks to undermine it"**[6] – is an important one.

This trilingual toolkit aims to strengthen the capacities of civil society actors and anyone else who wants to address hate speech through counterspeech. It offers definitions and contours of hate speech, introduces types of counterspeech, shares some good practices and strategies for creating counter narratives, and briefly discusses the ecosystem of tactics in addressing hate speech that counterspeech is also a part of. This toolkit was developed by collecting good practices of countering hate speech in Sri Lanka[7] as well as a desk review of existing knowledge resources and toolkits on counterspeech[8], critically engaging with the practices and resources in the context of the current internet and social media landscape of Sri Lanka[9] and the legal frameworks applicable to freedom of expression in the country[10], and coming up with strategies and recommendations that are specific to the geographical, political and cultural specificities of Sri Lanka.

We hope you find this to be a useful resource in our collective struggle to free our communities – both online and onground – from hate speech and nurture narratives that celebrate our diversity.

---

[5] Adama Dieng, Stopping Hate Speech (United Nations 2019) https://www.youtube.com/watch?v=rnbcQT-b8ak&t=142s accessed 3 May 2021
[6] Dangerous Speech Project https://dangerousspeech.org/counterspeech/
[7] Annex 4
[8] Annex 1
[9] Annex 3
[10] Annex 2

IDENTIFYING
HATE SPEECH

There is no single or singular definition of hate speech and what constitutes "hate" can be both universal and extremely context specific. With this in mind, we share a few frameworks and definitions that explore the parameters and challenges of hate speech.

THE CONCENTRIC CIRCLES OF HARMFUL SPEECH

Developed by Gehan Gunatilleke, a lawyer and scholar from Sri Lanka, the concentric circles of harmful speech[11] is a model that is helpful in identifying where hate speech is located within various kinds of harmful speech. Below is an adaptation of it for the purposes of this toolkit.

EXAMPLES

POSSIBLE RESPONSES BY THE STATE AND OTHER ACTORS

Speech that defames, stereotypes, or insults a person

Speech that conveys hatred towards a particular group

Speech that incites others to commit acts of hatred against a particular group

Speech that specifically incites others to commit acts of violence

HARMFUL SPEECH

HATE SPEECH

INCITEMENT

INCITEMENT TO VIOLENCE

'Providing remedies' E.g. civil damages for defamation, reporting mechanisms on platforms

'Countering' E.g. creating and sharing counter - messaging

'Regulating' E.g. impose civil sanctions (dismissal, revoking licenses)

'Prohibiting' E.g. the state may investigate prosecute and punish

As shown in the above figure, the types of harmful speech can be differentiated according to the intensity of harm caused by the speech. Similarly, the intensity of the responses should correspond to the intensity of the harm. This model identifies hate speech as "speech that conveys hatred towards a particular group" and it is important to unpack what "hatred" may entail. Below are some definitions that could be helpful.

[11] Gehan Gunatilleke, Countering Harmful Speech: Why Trust the State? (Groundviews 2020) https://groundviews.org/2020/12/17/countering-harmful-speech-why-trust-the-state/ accessed 25 March 2021

DEFINITIONS

★ "The terms 'hatred' and 'hostility' refer to intense and irrational emotions of opprobrium, enmity and detestation towards the target group" [12]

★ "There is one common thread among definitions of hate speech, which is that it vilifies a person or group of people because they belong to a group or share an identity of some kind. Therefore it's not hate speech to say 'I hate you' without referring to a group." [13]

★ "Any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor." [14]
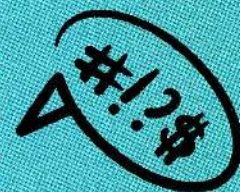
CHALLENGES

As noted by scholar Susan Benesch and the Dangerous Speech Project, hate speech can be "vague, broad, and in practice, everyone defines it differently" [15]. A few questions to keep in mind when trying to identify hate speech are:

★ How intense and sustained must speech be to be identified as hatred?

★ Who feels the hate in hate speech? The speaker, the target of speech, others exposed to the speech, all of them?

★ What are the social norms and contexts (whether social, political, cultural) that inform how hate is expressed and perceived?

---

[12] Article 19, Principle 12(1)9i) of The Camden Principles on Freedom of Expression and Equality (Article 19 2009)

[13] Dangerous Speech Project, Dangerous Speech: A Practical Guide (Dangerous Speech Project 2020) https://dangerousspeech.org/guide/ accessed 25 March 2021

[14] United Nations, United Nations Strategy and Plan of Action on Hate Speech (United Nations 2019) https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml accessed 25 March 2021

[15] Dangerous Speech Project, FAQ (Dangerous Speech Project 2020) https://dangerousspeech.org/faq/ accessed 25 March 2021

The Rabat Plan of Action[16] introduces a six-part threshold test which is also useful in making a distinction between freedom of expression and incitement to hatred which is as follows:

1. The social and political context
2. Status of the speaker
3. Intent to incite the audience against a target group
4. Content and form of the speech
5. Extent of its dissemination
6. Likelihood of harm, including imminence

It is particularly important to be reflexive in how we identify and address hate speech when a country is in a democratic deficit.[17] As illustrated by recent examples, including in Sri Lanka[18], the legal framework that is applicable to hate speech[19] can often be used to criminalize legitimate speech and to silence and shut down dissent including by journalists, activists, cartoonists, bloggers, writers, artists, members of ethnic minorities, etc.

freedom of expression or incitement to hatred

LOW

HIGH

1 The social and political context

2 Status of the speaker

3 Intent to incite the audience against a target group

4 Content and form of the speech

5 Extent of its dissemination

6 Likelihood of harm, including imminence
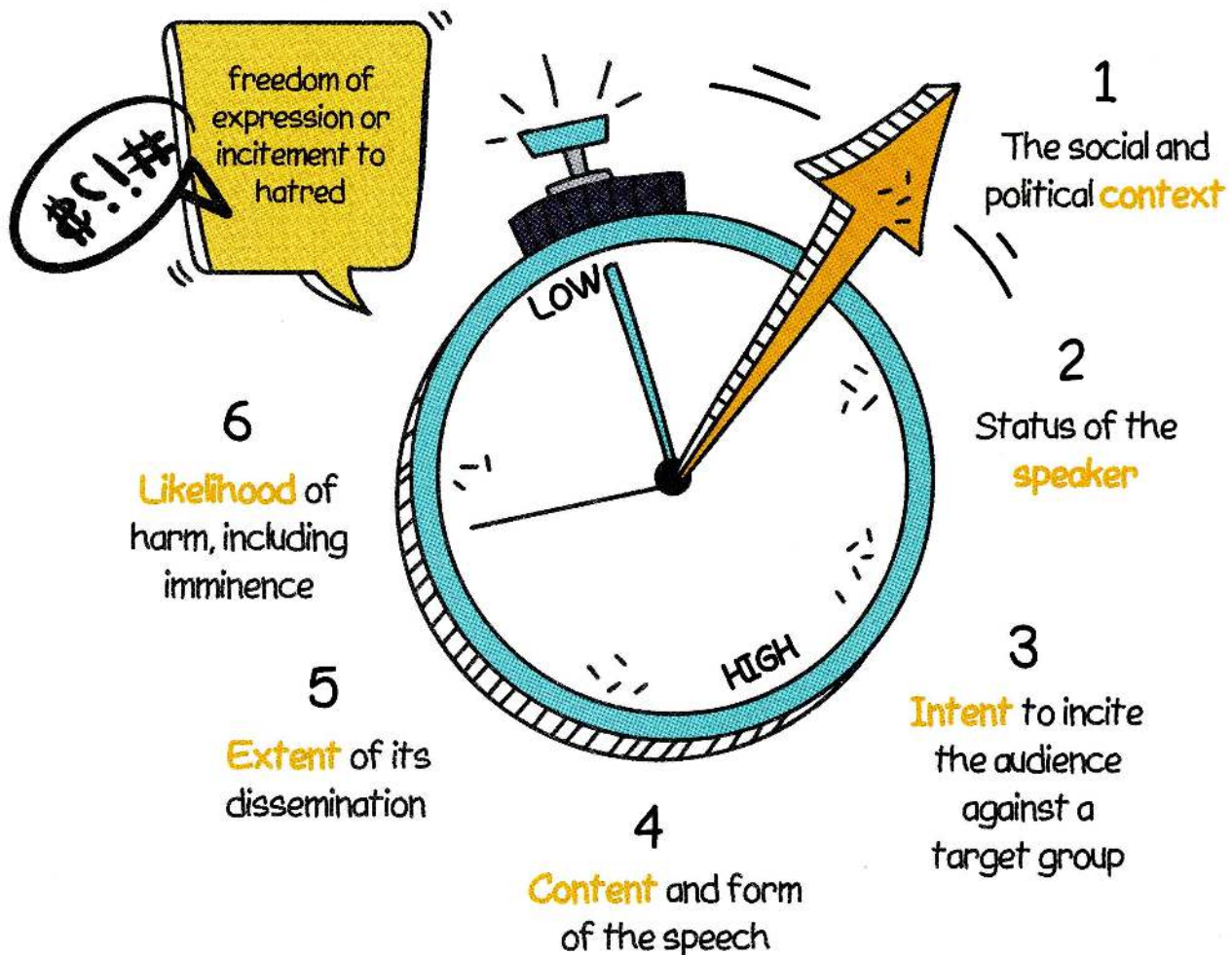
[16] Office of the United Nations High Commissioner for Human Rights, Freedom of expression vs incitement to hatred: OHCHR and the Rabat Plan of Action (OHCHR 2013) https://www.ohchr.org/en/issues/freedomopinion/articles19-20/pages/index.aspx accessed 3 May 2021
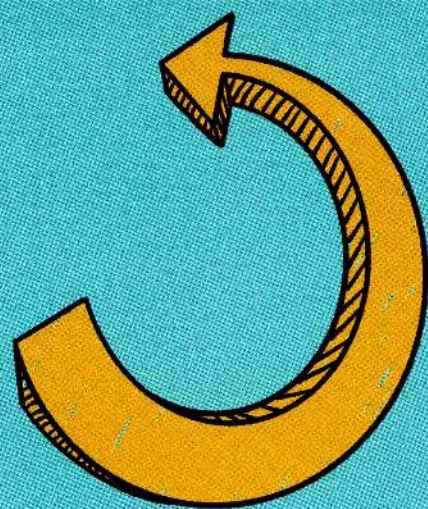[17] Gehan Gunatilleke (n 11)
[18] Freedom House, Freedom on the Net 2020 - Sri Lanka (Freedom House 2020) https://www.ecoi.net/en/document/2039119.html accessed 10 April 2021
[19] Annex 2

WHY?
COUNTERSPEECH

Many researchers have documented the prevalence and proliferation of hate speech in Sri Lanka[20] and a key recommendation emerging from much of this research is to promote counterspeech as a key strategy to address hate speech[21]. The Dangerous Speech Project defines counterspeech as "any direct response to hateful or harmful speech which seeks to undermine it"[22] and recognizes two main types of counterspeech: "organized counter-messaging campaigns and spontaneous, organic responses".[23] This toolkit considers both of these and provides tips and strategies for both proactive and reactive responses to hate speech.

Counterspeech is a long term strategy that needs to be nurtured and sustained in order for it to be effective. While it can be difficult to ascertain if counterspeech has been successful or effective, generally there are some broad parameters that can be considered.

* The person/people who made the original (hate) speech have constructively changed their beliefs and/or behavior, or
* Others who witnessed or participated in the exchange between the hate speech and counterspeech have had their beliefs and/or behavior positively influenced by it, or
* Audiences of counterspeech begin to exercise more skepticism, diligence and reflexivity when they see or experience hate speech, or
* Audiences of counterspeech are motivated to further disseminate them.

These are not easy parameters to monitor or track, whether manually or by automatic including through methods like neuro-linguistic programming (NLP). However, the following infographic[24] provides some examples of the ways in which counterspeech can be done and detected.

---

[20] Annex 1
[21] ibid
[22] Dangerous Speech Project, Counterspeech (Dangerous Speech Project 2020) https://dangerousspeech.org/counterspeech/ accessed 25 March 2021
[23] ibid
[24] Dangerous Speech Project (n20)

8

SUCCESSFUL COUNTERSPEECH ON TWITTER

Lucas Wright, Derek Ruths, Kelly P Dillon, Haji Mohammad Saleem, and Susan Benesch
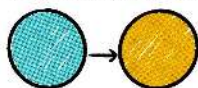
Berkman Klein Center for Internet & Society, Harvard University, School of Computer Science, McGill University

Department of Communication, Wittenberg University Dangerous Speech Project

FOUR VECTORS OF COUNTERSPEECH

One-to-one

Exchanges, often lengthy, between two accounts. Remarkable shifts in tone were observed.

In honour of MLK day today, I'm taking a vow to use the word "n****" as many times as possible in the most inappropriate times.

"Let no man pull you so low as to hate him" – Martin Luther King Jr.

"I hope I get shot soon" - Martin Luther King Jr.

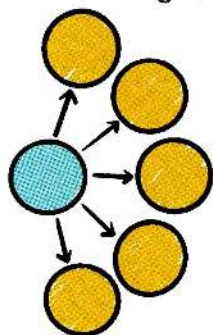[…] Does your mum know how you spend your time on the internet trying to hurt people?

I doubt it. She's been dead for a year and a half now so…

I'm sorry for your loss. And I hope you find a better way to honour her.

You are so nice and I'm so sorry

Many-to-one

Viral counterspeech overwhelms its target. Does it convince, or silence?

I am literally soo mad right now a ARAB won #MissAmerica

Once day I hope you realize how shameful this tweet is. I hope you realize it tomorrow.

Your hatred made it onto sky news! Congrats!

Ignorant/illiterate/racist/idiot

@MissAmerica sorry for being rude and "racist" and calling you a Arab please tweet back so everyone will know its real

One -to- many

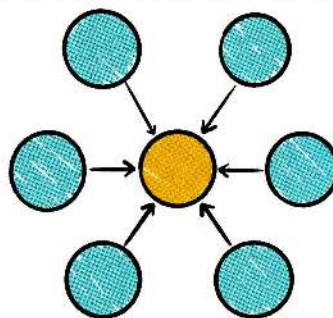A single user responds to many others. Can also be described as multiple one-to-one exchanges.

I was prompted to write this after a recent Facebook discussion […] when the term "abeed" (slaves) was used in a thread, in reference to a news story about an African American woman, who flashed an Arab American businessman in Detroit, during a verbal dispute. What was disturbing […] was not simply the racist comments that were used about the unruly woman, but some showed a profound lack of empathy when I mentioned the term "abeed" is a hurtful word.

The basis of Islam is equality before God. So how can we tolerate derogatory, racial slurs and still call ourselves Muslim? #DropTheAWord

Many-to-Many

Competing Hashtags, or struggle over control of/opinions expressed on a single hashtag.
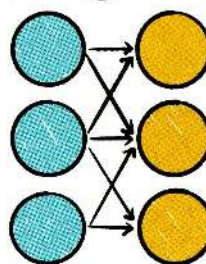
It's going to be a -27 without wind chill tomorrow morning and I have class at 8 #FuckPhyllis #Cunt #Bitch #Whore

In a room with Phyllis Wise, Adolf Hitler, and a gun with one bullet. Who do I shoot? #FuckPhyllis

As an alum, I am completely embarrassed with the behaviour of some of you assclowns. #fuckPhyllis, If I were UIUC I'd publicly shame all of u.

Oh and good luck, bitching students, your asinine and ignorant tweets are viewed by EVERYONE NOW #FuckPhyllis

So happy to the #fuckPhyllis tag has been taken over by people calling out the racism that was in it.

CREATING COUNTERSPEECH CONTENT
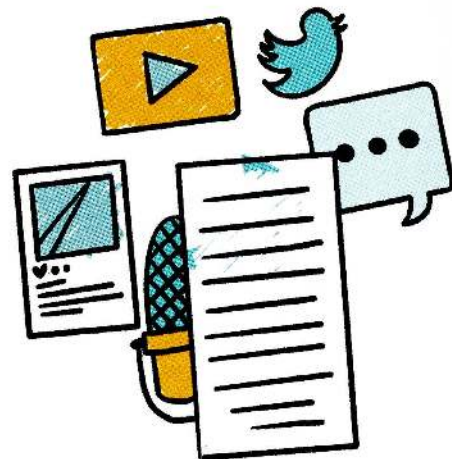
Many civil society organizations and community groups in Sri Lanka already create counterspeech content and campaigns. We collected examples of such content and campaigns[25] to identify and distill good practices and strategies embedded in them, and put them together as a toolkit for anyone interested in countering hate speech.

THE FOUR P'S: PLANNING, PRODUCING, PUBLISHING AND POST-PUBLISHING

PLANNING

PRODUCING

PUBLISHING
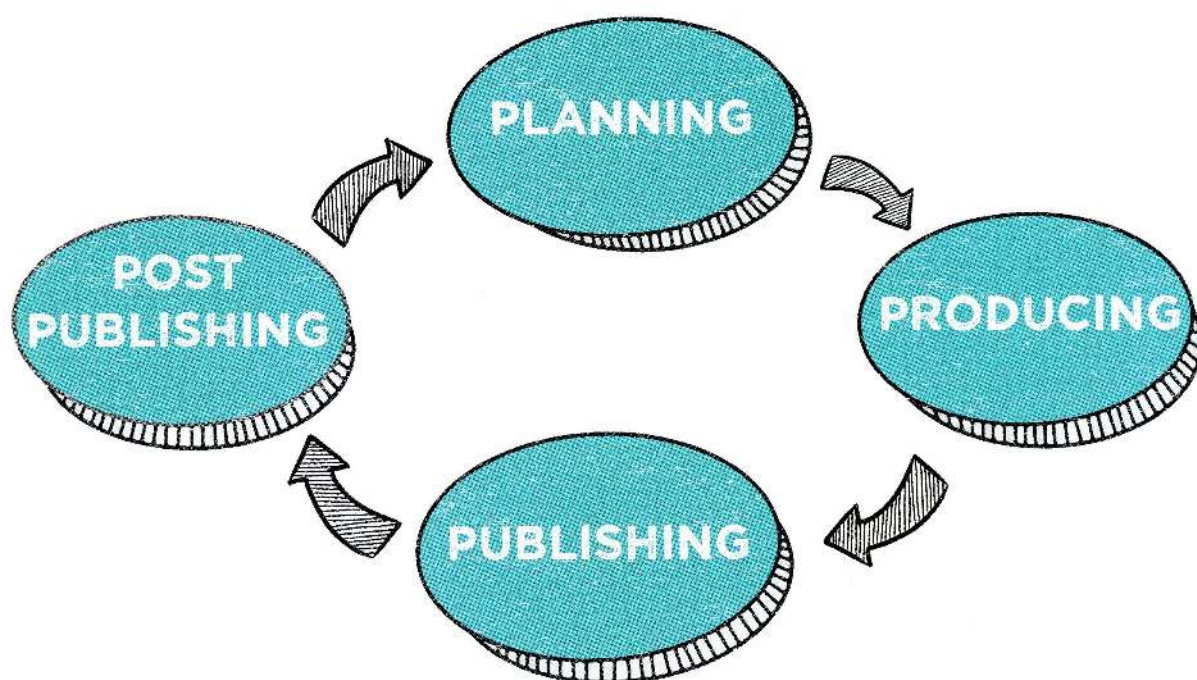
POST PUBLISHING

This toolkit explores four phases of creating counterspeech content; planning, producing, publishing and post-publishing. Each phase is detailed below and it's important to remember that while they are sequential, they are also cyclical with the lessons learnt from each project or campaign informing the next.

_____

[25] Annex 4

PLANNING

PRODUCING

PUBLISHING

POST PUBLISHING
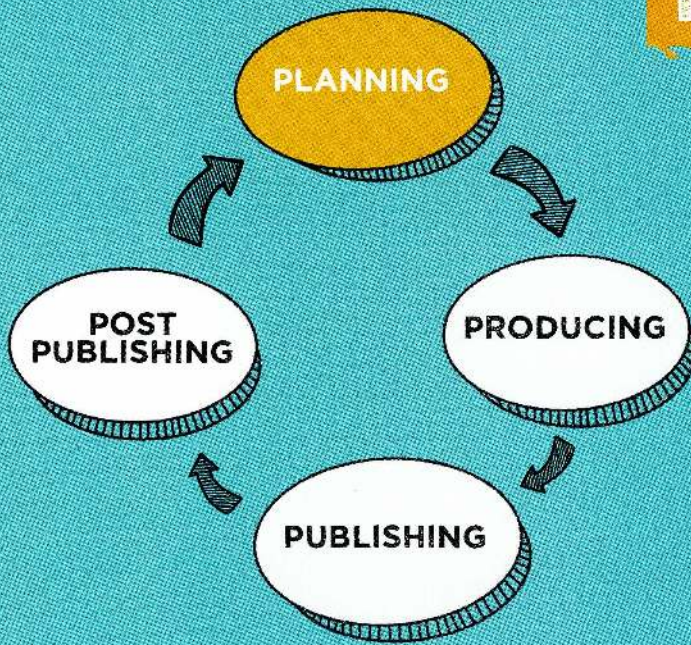
1. IDENTIFY THE ISSUE

You can use the earlier sections on identifying hate speech and understanding counterspeech to identify the issue(s) you want to address. You can try to determine whether they constitute hate speech and in instances where it might have escalated to incitement to violence, consider additional strategies to address it like reporting to platforms or flagging to law enforcement where there's imminent threat of violence.

SO FAR YOU HAVE

☑ Identified the hate speech you want to counter

2. PROACTIVE OR REACTIVE

Once you have identified hate speech that you want to counter, you need to decide which approach you will take. As noted earlier, there are two main types of counterspeech; "organized counter-messaging campaigns and spontaneous, organic responses".[26] In the planning phase, it is helpful to be clear if you're responding instantaneously to a real time instance of hate speech (more reactive) or if you're going to plan a long term and sustained response to a pervasive or anticipated occurrence of hate speech (more proactive). This clarity will influence all the other decisions you make and strategies you use in the four phases.

✳ **Example of a more reactive counterspeech:**
#StopForcedCremations video by Hashtag Generation
https://twitter.com/generation_sl/status/1354304500341858309

✳ **Example of a more proactive counterspeech:**
පුංචි දුපතේ කතාව by Blok & Dino
https://www.youtube.com/watch?v=KwNVqTVKUP8

FURTHER READING

For more proactive interventions
Defusing Hate: A Strategic Communication Guide to Counteract Dangerous Speech by Rachel Brown https://www.ushmm.org/m/pdfs/20160229-Defusing-Hate-Guide.pdf

For more reactive interventions
Considerations for Successful Counterspeech by Susan Benesch , Derek Ruths , Kelly P Dillon , Haji Mohammad Saleem , and Lucas Wright
https://dangerousspeech.org/wp-content/uploads/2016/10/Considerations-for-Successful-Counterspeech.pdf

---

[26] Dangerous Speech Project (n 20)

3. CONTEXT ANALYSIS

A context analysis is essential before creating counter messages. This includes the **origins and histories** of the hate speech you've identified, the **dominant, current and emerging narratives** around it, characteristics and behaviors of the groups of people spreading it, characteristics and concerns of the people being targeted by them, and the languages, mediums and platforms on which the hate speech you've identified is spreading or emerging. Some methods of doing this analysis would be interviews or focus group discussions, desk research, monitoring social media or referring to monitoring reports by others, monitoring mainstream media narratives, case studies of specific incidents, etc.

An analysis of the context your organization or group is embedded in can also be useful in creating effective counterspeech. There are a number of frameworks[27] you can use such as SWOT, SOAR or PEST but the best way to go about it is to decide which questions are most relevant for you and try to answer them. Your internal strengths and weaknesses and your external threats and political-social-economic-technological contexts can help you understand many things like whether you have **internal capacity** for all the skills and tools needed to create your content and campaigns, whether there are **security and safety concerns** in you being associated with the content and campaign, etc. While it is not necessary or practical to consult a lawyer each time you plan and produce a counterspeech campaign, you can be reflexive and seek counsel if you feel certain content or campaigns will be particularly provocative.

SO FAR YOU HAVE

- ☑ Identified the hate speech you want to counter
- ☑ Identified if your campaign is proactive or reactive
- ☑ Done your context analysis

[27] Annex 5

4. AUDIENCE AND PLATFORMS

Identifying audiences and platforms during the planning phase is essential because this would influence your objectives, messages, formats and languages of the messages, etc.

AUDIENCE

It is helpful to segment your audience in relation to the hate speech you want to counter and where possible, identify specific actors within those segments. Some of these segments could include those who begin or lead the instances of hate speech (ex-extremist groups, politicians), those who are most engaged in propagating it (ex- some media institutions, admins of ethno-national Facebook pages), those who are most engaged in challenging it (ex- some media institutions, human rights activists), and those who are reluctant or undecided participants/spectators (usually difficult to pinpoint but an important segment ex- Sinhala Buddhist university undergraduates with moderate views). Once you have identified the audience segments, you can map them according to their interest or investment in the issue and the power and influence they hold over it (Annex 6). This will help you determine which audience segments your counterspeech content may have a positive effect on as well as those who don't need to be reached either because they are already convinced of your messages or because they are extremely opposed.

PLATFORMS

It is important to ensure that your counterspeech content is published and shared **where the audience segments you want to reach are usually active**. But these are not the only platforms to use and you can also publish and share across other platforms. The current social media landscape of Sri Lanka[28] points to Facebook having the highest number of users so it may make sense to always publish and share to Facebook even if the main incidents you're trying to address didn't occur there.
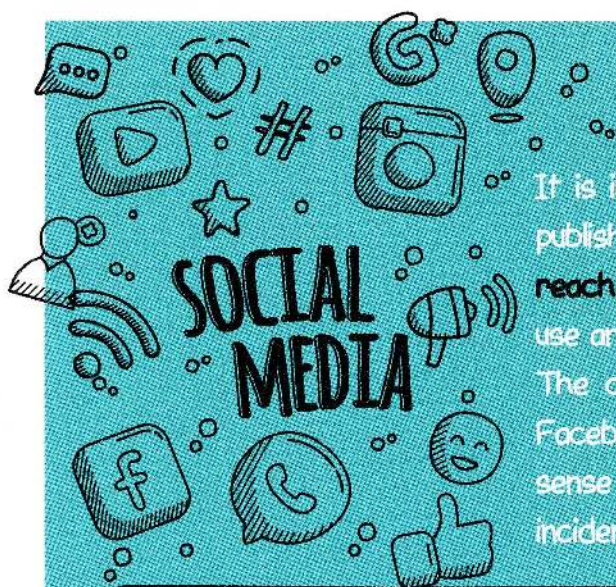
28 Annex 3

☑ Identified the hate speech you want to counter
☑ Identified if your campaign is proactive or reactive
☑ Done your context analysis
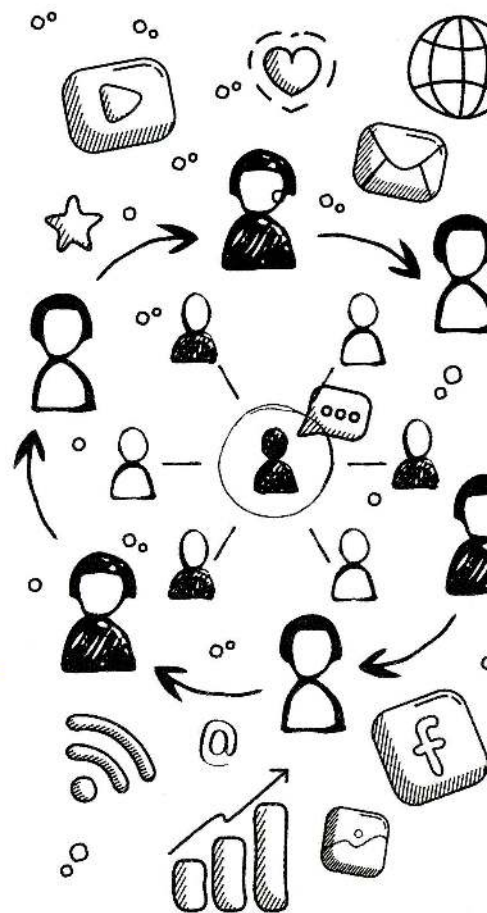☑ Identified audience segments
☑ Mapped the audience segments
☑ Identified your target audience segments
☑ Selected your platforms

5. SETTING GOALS

Next it is important to clarify your **counterspeech goals**. All the above steps would have given you enough information to set these. To refresh your memory, these were some of the possible outcomes of counterspeech that were mentioned earlier in this toolkit.

✳ The person/people who made the original (hate) speech have constructively changed their beliefs and/or behavior, or

✳ Others who witnessed or participated in the exchange between the hate speech and counterspeech have had their beliefs and/or behavior positively influenced by it, or

✳ Audiences of counterspeech begin to exercise more skepticism, diligence and reflexivity when they see or experience hate speech, or

✳ Audiences of counterspeech are motivated to further disseminate them.

These outcomes are all related to **behavior and behavior change**. There are numerous and varied theories of behavior change stemming from many disciplines[29] and rather than adopting them as they are, it is more useful to learn from them and then create your own theories that respond to your geographical, social and cultural contexts. Some things to keep in mind when determining the kind of behavior change you want to bring about are; the **social and cultural norms** at the root of the issues you want to address, the **implicit and explicit attitudes** of your target audience segments, and the **emotions and beliefs** that inform and drive their behavior.

---

[29] Rachel Davis and others, Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review (Health Psychology Review 2015) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4566873/ accessed on 2 April 2021

The following guidelines from page 122 of *Defusing Hate: A Strategic Communication Guide to Counteract Dangerous Speech*[30] by Rachel Brown could be useful in thinking through your counterspeech goals.

"**Overall Content Goal-setting:** First, you can develop overarching goals for your messaging strategy. You can create these goals based on your audience journey timelines. For example, "we want some messages that will cause people generally to evaluate information more critically and be less susceptible to misinformation/rumors about X; we want other messages that will make people understand the negative consequences if they are thinking about spreading or acting on specific messages." These types of content categories will help you design a content strategy and prioritize time and resources for different types of content. You can develop general guidelines for content in addition to specific pieces of content."

"**Micro-Goal Setting:** Once you have your overall content goals, you can develop specific micro-goals for each piece of message content. For example, "this radio spot/Facebook post/text message/public statement/sticker/meme should influence people to go home and wait for more information." Each message or piece of content should have a clear goal (what you want to influence the audience to do/feel/think) and target a specific audience group or groups. Your goals can be very specific about the exact behavior you want to influence. For example, a message could seek to: stop audience members from participating directly in violence, harassing members of a targeted group, or providing assistance to violent groups; encourage audience members to speak or act against violence; or prevent audience members from cheering on violent action in a crowd."

6. TIMELINE AND CALENDAR

The timeline and duration of your counterspeech interventions will determine the volume and intensity of content, allocation of work time, budget allocations and other factors. If the behavior change you're seeking is more urgent and reactive, then you can decide on a strictly timebound campaign with intensive messaging. If the goals are more longterm and sustained then you can plan a longer campaign with intervals marked for stocktaking and learning. For either of these, you can also take into account national and international days that are significant and relevant as well as any events and other milestones.

---

[30] Rachel Brown, Defusing Hate: A Strategic Communication Guide to Counteract Dangerous Speech (United States Holocaust Memorial Museum 2016) https://www.ushmm.org/m/pdfs/20160229-Defusing-Hate-Guide.pdf accessed 25 March 2021

7. BUDGET AND FUNDRAISING

While most counterspeech content and campaigns are more effective when they are carried out over a longer period of time, often the funding raised and received by civil society organizations tends to be project-based rather than core support. This can be especially tricky when funders want to see immediate results at the end of a project rather than understanding that each project is a building block in getting closer to our goals. We hope that this toolkit can help you with donor education/ advocacy to **understand not just the costs of an effective counterspeech campaign but also the importance of longevity.** However, you can also plan your budget so that the content created for specific projects have a longer shelf life.

Some of the things to budget for can include: copywriting, translations, proofreading, costs of content creation (illustrations, infographics, videos, comics, etc.), advertising, salary of communications staff, purchasing devices, tools and software, and monitoring and evaluation tools or personnel.

You should also remember that some of the most effective counterspeech content are made without any budget. So a **small budget or no budget** at all should not discourage you from creating counterspeech, and practices like **mutual aid, crowdfunding and collaborations** can be helpful.

PLANNING RECAP

- Identified audience segments
- Mapped the audience segments
- Done your context analysis
- Identified your target audience segments
- Identified if your campaign is proactive or reactive
- Selected your platforms
- Identified the hate speech you want to counter
- Decided on your counterspeech objectives
- Planned your timeline and budget

PLANNING

POST PUBLISHING
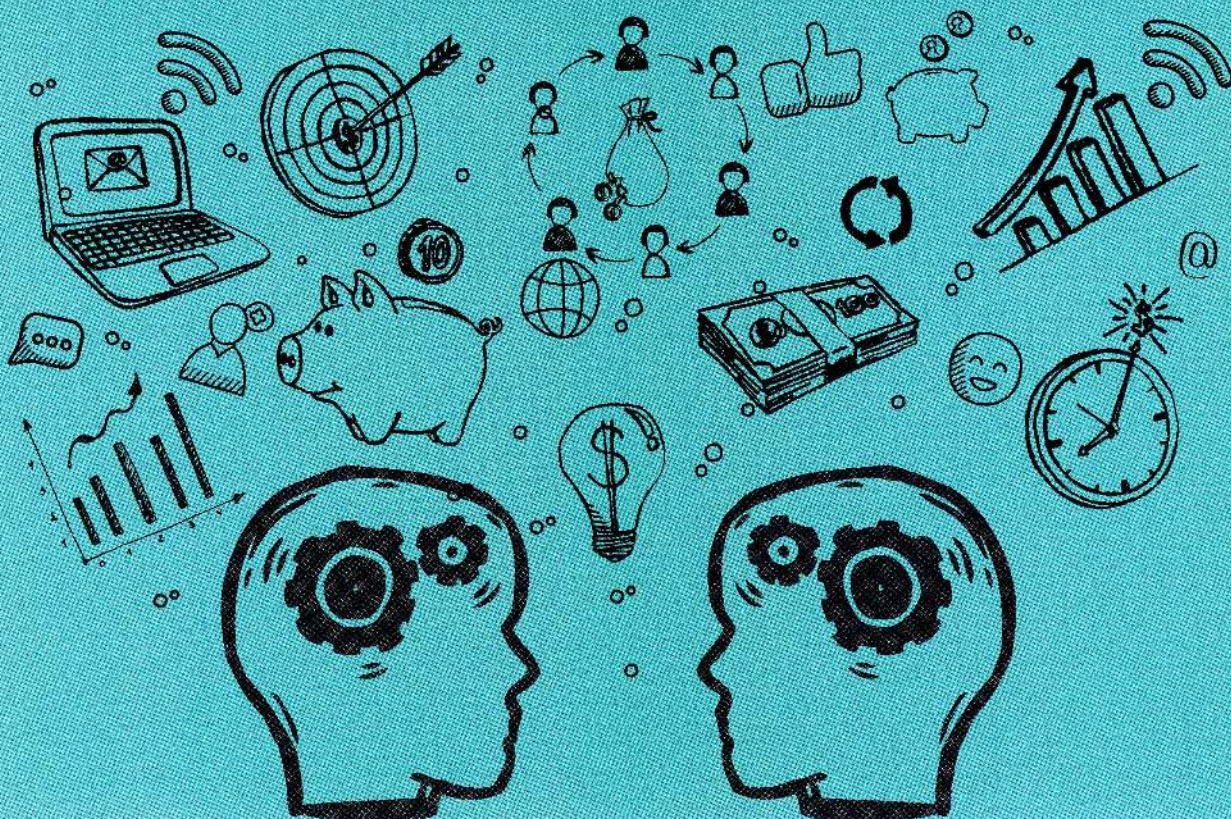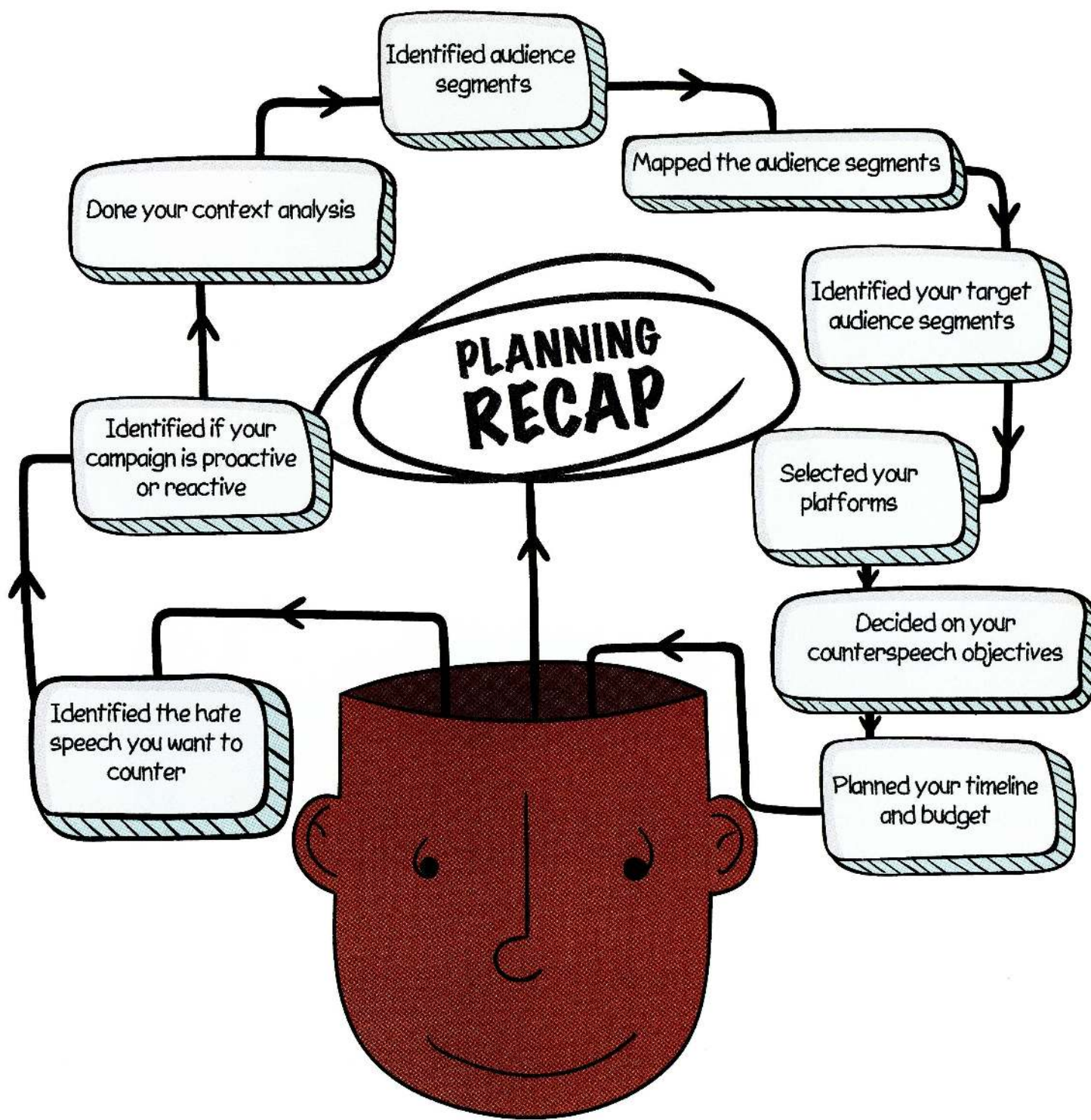
PRODUCING

PUBLISHING
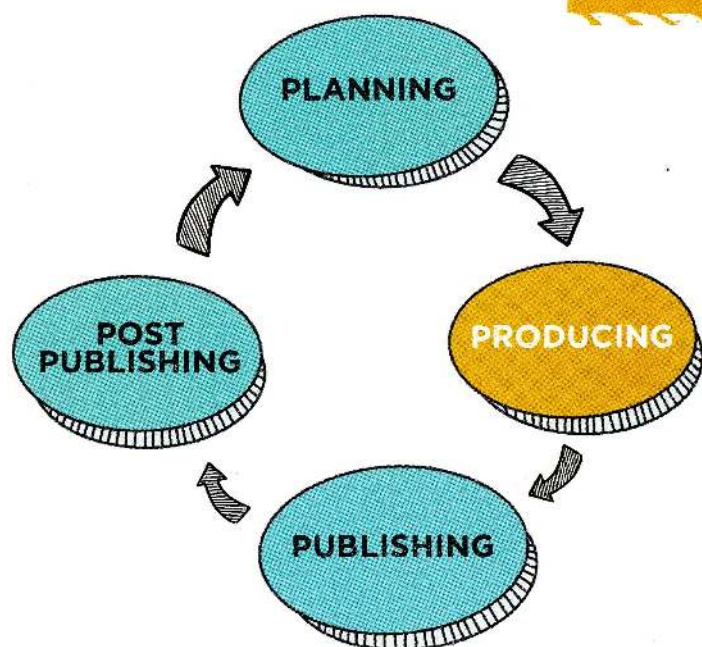
1. CRAFTING MESSAGES

There is no prescribed way of creating messages. Some effective strategies learnt from the repository of good practices[31] include humor, storytelling, testimonies and actionable messages. However as the planning phase illustrated, you are best positioned to make informed decisions on the messages that will work best for your target audience. So this section will focus on what to avoid when crafting messages.
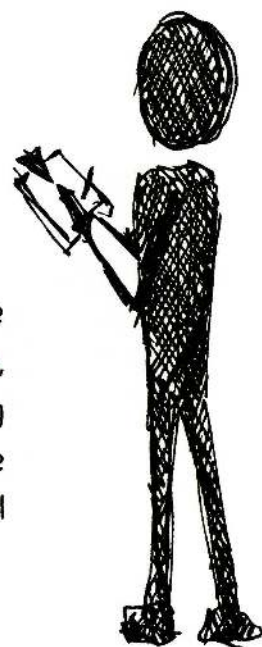
Avoid Motivated Reasoning

"People tend to accept information that confirms their existing beliefs and feelings, and reject information that contradicts them.[32]" This is known as motivated reasoning. So counterspeech that "corrects" or "preaches" can make people defensive because they feel that their core beliefs or values – as harmful as they may be sometimes – are being attacked. So you can use strategies like storytelling and testimonies to deliver your messages and also choose to target audience segments whose beliefs are not as entrenched.

---

[31] Annex 4

[32] Rachel Brown (n 28)

It can be tempting to retort in the same ways hate speech is expressed and delivered but this is usually not effective. We have to always keep in mind that our audience comes with all the complexities and **intersecting identities of being human** and try to appeal to those. For example, countering hateful speech against an ethnic minority with counterspeech that is classist or elitist may not have the intended effect.

Avoid Hostile and Dehumanizing Approaches

There are conflicting positions on the effectiveness of fact-checking as a strategy of counterspeech. However, fact-check messages can work on the undecided/intermediate segments of your audience rather than those who are extremely invested in the beliefs and ideologies behind hate speech.

2. SOURCES, CONSENT AND PRIVACY

CREDIT

It is important that your counterspeech content is backed by evidence, research, case studies, lived experiences, etc. that ground them. You may not always need to cite source material, especially when based on your own research, but crediting your sources can increase the credibility of your counterspeech as well as acknowledge their roles.

CONSENT AND PRIVACY

If you're using case studies, testimonies or content created by others (ex- tweets, TikTok videos, memes), it is imperative that you get consent to do so. And accommodate requests like anonymizing, hiding the real identity and other things that may make them feel safer about sharing their stories or content. Even when it is something that was published publicly, it is possible that the author did not intend it to be used for any other purpose than what they wanted. Consent is also ongoing and dynamic so you should make it clear that they can take their consent back if they feel unsafe or uncomfortable at any point.

If your organization does not already have a responsible data policy, it would help create one. DatNav by The Engine Room is a useful resource to learn how to do this.
https://www.theengineroom.org/datnav-digital-data-in-human-rights-research/

3. FORMAT OR MEDIUM

The planning phase would have helped you determine your audience and their preferred platforms. If you're already on those platforms, you can check your analytics to see what kind of content is most engaging. You can also monitor content by others that gain traction and make note of formats. Below is a non-exhaustive list of formats to inspire you, in no particular order.

FORMAT	EXAMPLE
✳ Short form videos	Countering Hate Speech Campaign by Hashtag Generation https://www.instagram.com/p/CGFL2PoBkEH/
✳ Long form videos	Patta History https://www.youtube.com/channel/UCHbyuXOgrziVQ1bJYNsFEFQ
✳ Memes	https://pbs.twimg.com/media/EodCAbIX1AAC5p-?format=jpg&name=small
✳ Newsletters	On Living and Dying as a Muslim in Sri Lanka by Adilah Ismail https://alifeofsaturdays.substack.com/p/on-living-and-dying-as-a-muslim-in
✳ Infographics	Sisterhood Initiative https://www.instagram.com/sisterhood.initiative/
✳ TikTok videos	Respect every religion https://vt.tiktok.com/ZSJAT1yt4/
✳ Cartoons	Political cartoons of Sri Lanka by various cartoonists https://twitter.com/cartoonlka
✳ Comics or graphic	Juwairiya Mohideen's story by Isuri
✳ Novels	https://www.frontlinedefenders.org/sites/default/files/cypher04.pdf
✳ Photo essays	Home by by Amalini De Sayrah, Oliver Friedmann and Natasha Van-Hoff http://momac.lk/home/
✳ Twitter bots	Queer Affirmations Bot by Jackie Henke https://twitter.com/queer_affirm
✳ Live broadcasts or Webinars	Youth Press Conference on Forced Cremations https://www.facebook.com/1455075774809484/videos/717978558855743
✳ Podcasts and Talk Shows	Shhh Talkshow by Shanuki de Alwis and Shifani Reffai https://www.youtube.com/c/ShhhTalkshow/featured
✳ Listicles	අසත්‍ය ප්‍රකාශ මැඬලීම සඳහා ඔබට ගත හැකි ක්‍රියාමාර්ග 10 ක් by Hashtag Generation https://www.facebook.com/1455075774809484/videos/1162355434118850
✳ Community groups	Sri Lankans Against Hate Speech https://www.facebook.com/groups/284832152529707
✳ Zines	Tales from the Crypt by the Women and Media Collective https://womenandmedia.org/tales-from-the-crypt/
✳ Slogans	#butterfliesfordemocracy https://www.facebook.com/hashtag/butterfliesfordemocracy
✳ Instagram stories	Easter Attack Survivors Project https://www.instagram.com/easterattacksurvivorsproject/
✳ Audience participation through competitions, exhibitions, etc.	Open call for storytellers https://www.instagram.com/p/CM_2nenhoxx/

It is argued that having a Visual Identity for your organization and the content you produce can increase the credibility of the content and the trust your audience has in your messages. However, if your PEST context analysis shows an increasing hostile environment towards civil society organizations, foreign funding, social justice activism and the messages you're trying to deliver, then your organizational affiliation may be used by detractors to discredit you.

SO FAR YOU HAVE

☑ Completed the PLANNING phase
☑ Crafted your messages
☑ Considered citations and consent
☑ Decided on the format(s) of your content

4. LANGUAGE AND ACCESS

Language and access are separate but interconnected considerations when producing your counterspeech content. Following are a few tips to keep in mind.

LANGUAGE

✳ The primary language(s) of your content should ideally correspond to the primary language(s) of the hate speech you're trying to address and the audiences you're trying to reach. However a secondary language can be included in subtitles or closed captions.

✳ While trilingual content is commendable if you have the resources to manage it, it should be noted that translating for the sake of translating may not always convey the nuances and cultural contexts of the messages (ex- A series of memes that are based on a Sinhala cultural references like song titles may not evoke the same effect or message Tamil or English).

✳ Creative content that requires little or no language in order to be understood is also effective and helps you overcome some of the challenges around language.

✳ If your team is not trilingual, make adequate budget allocations for translation and interpretation services. Or set up mutual aid skills exchanges with other groups.

ACCESS

* Language is a key part of your content becoming more accessible. Ensuring accessibility for people with disabilities is equally important. Descriptive alternative text, descriptive URLs, font sizes, colors, audio descriptions, subtitles, captions, transcripts, layout are some of the factors to be considerate of when producing your counterspeech content (or any other content for that matter).

* Access is also about connectivity, internet speeds and devices. According to reports, 88.5% of Facebook users in Sri Lanka access Facebook through mobile phones only. This means the content you produce should be optimized for mobile phone screens, easy to (down)load on to a wide spectrum of smartphone devices and easily shareable.

> **More resources on improving accessibility**
> W3C Accessibility Standards https://www.w3.org/WAI/standards-guidelines/
> The A11Y Project https://www.a11yproject.com/
> accessibility = https://twitter.com/a11y
> #a11y on Twitter https://twitter.com/hashtag/A11y?src=hashtag_click

5. CREATING CONTENT

The context analysis would have helped you identify which skills and tools you have within your organization to create counterspeech content and what you need externally.

If you're able to collaborate (paid or unpaid) with popular content creators who are not antithetical to your organizational values, this could be a useful strategy to produce high quality content that is informed by your research, evidence and point of view. However this strategy could backfire if your collaborators change their positions and views. Some civil society organizations also partner with advertising agencies to create campaigns and if you choose to do this, it's important to sensitize them to the nuances of the messages you want to deliver.

Creating content in-house is an option many civil society organizations choose. An organizational communications strategy that is implemented by dedicated personnel is the ideal scenario but depending on your resources, you can also use online tools to create content and attend training workshops to build your skills.

23

This is a non-exhaustive list of free and freemium tools in no particular order.

Creating graphics, editing photos, creating a logo

* Canva https://www.canva.com/
* GIMP https://www.gimp.org/
* Pixlr https://pixlr.com/
* Wondershare https://ps.wondershare.com/home

Stock photos

* Canva https://www.canva.com/
* Pexels https://www.pexels.com/
* Unsplash https://unsplash.com/
* Pixabay https://pixabay.com/
* Freepik https://www.freepik.com/home
* Envato https://elements.envato.com/

Social media management

* Hootsuite https://hootsuite.com/en-gb/
* Buffer https://buffer.com/
* MediaCMS https://mediacms.io

Bots

* Cheap Bots, Done Quick! https://cheapbotsdonequick.com/
* Artbot http://artbot.combinatorium.com/#!/tracer
* y The bot ideas generator https://matteomenapace.github.io/random-generator-generator/examples/bot-idea-generator/

Video and audio editing

* Avidemux http://avidemux.sourceforge.net/
* Audacity https://www.audacityteam.org/
* Wavosaur https://www.wavosaur.com/
* Wevideo https://www.wevideo.com/
* Kate's video toolkit https://kate-s-video-toolkit.informer.com/
* Animaker https://www.animaker.com/

This is a list of toolkits and tutorials, again in no particular order

* Take Back The Tech campaign kit https://www.takebackthetech.net/campaign-kit
* Feminist Forwards: Creating content to make the internet work for you https://www.takebackthetech.net/blog/whatsapp-forwards-good-morning-feminism
* Even machines dream: Feminist robots for Twitter https://github.com/stepaola/feminist-robots-for-twitter/blob/master/guide.md
* Digital marketing and fundraising tips for nonprofits https://www.nptechforgood.com/
* Best practices for online subtitling https://engagemedia.org/help/best-practices-for-online-subtitling/
* Video and audio compression https://engagemedia.org/help/video-compression-guide/
* Starting your podcast https://www.npr.org/2018/11/15/662070097/starting-your-podcast-a-guide-for-students

FURTHER READING

Counter-Narrative Handbook (page 53 onwards), by Henry Tuck and Tanya Silverman
https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf

6. TESTING AND FINALIZING

It is important to leave ample time in your plan for testing, proofreading, revising and finalizing your content. Testing includes audience responses, language and accessibility and technical aspects of your content. Once you have finalized your content, you're ready for publishing.

PRODUCING RECAP

Decided on the format(s) of your content

Considered language and accessibility

Considered citations and consent

Learnt about content creation tools and skills

Crafted your messages

Created content

Completed the PLANNING phase

Tested and finalized content
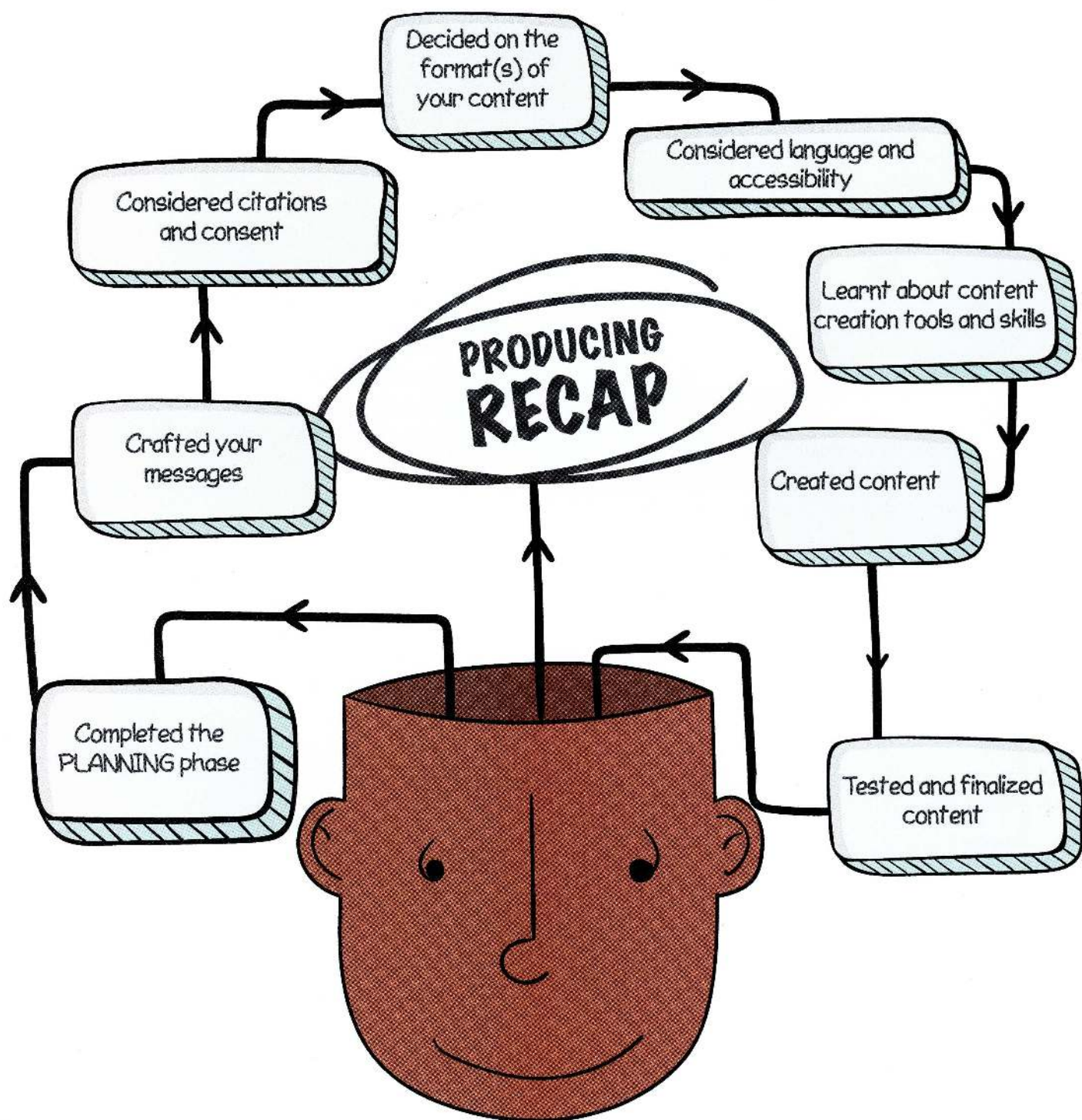
PUBLISHING
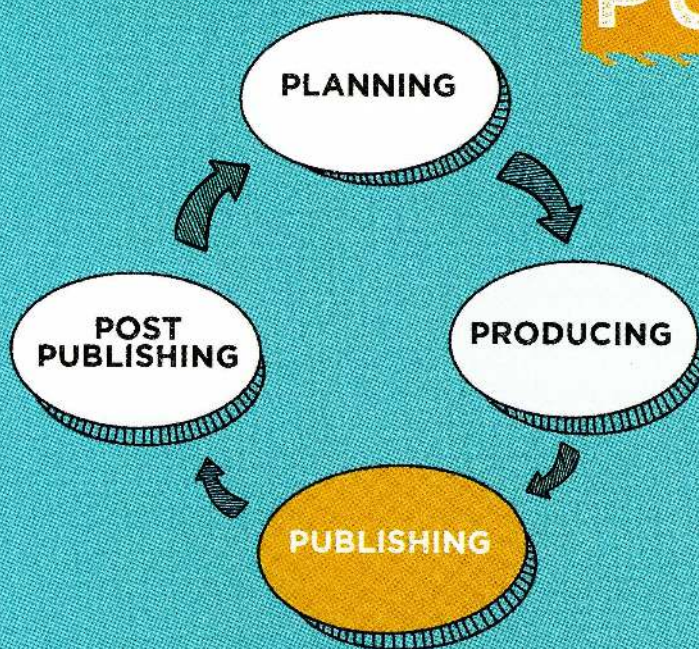
A clear dissemination strategy should be developed, including the following issues.

1. PLATFORM ECOSYSTEM

You can take stock of the platforms you identified during the PLANNING phase and make any changes if needed. It's also important to remember that it's not at all necessary to post on every platform you're on or every platform that is popular. Focusing on the audience you want to reach and disseminating to the platforms they are active on remains key.
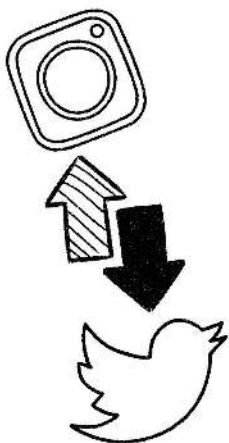
2. PLATFORMS VS CONTENT

Cross-posting the same piece of content across a number of platforms is convenient but often not effective. For example, an infographic that was created for Instagram won't display or read the same way on Twitter. Tools like Canva offer templates that help you customize content for the platform of your choice. It's also important to remember that all platforms are not equally optimized for all formats/ types of content so being selective can save you time and resources.[33]

---

[33] Alfred Lua, What to Post on Each Social Media Platform: The Complete Guide to Optimizing Your Social Content
https://buffer.com/library/what-to-post-on-each-social-media-platform/#what-to-post-on-each-social-media-platform accessed 26 April 2021

3. WHEN TO POST AND HOW OFTEN

Most often, it's you who has the answer to these two questions. The built-in analytics/insights on your social media pages, the analytics for your website and newsletter all offer insights on which days and times your audience is most engaged. It can also be useful to observe and even consult other local groups with a similar profile and mission to yours and compare notes. If you're just starting your communications work, you can use the testing period to experiment with different combinations of content, dates and times to find out what works best for you.

How often you post content will depend on the urgency and timeframe of the campaign and the platform you use. If reach and engagement is driven by algorithms as with Facebook and Instagram, it makes sense to post less frequently so that you don't dilute them. But if they are driven by chronology as with Twitter, then you can post more frequently.[34]

4. BOOSTING AND ADVERTISING

Here we come to a tricky point. There are certain financial, ethical and political considerations to make around paid advertisements to boost reach and engagement on social media. Financially it can eat into your already limited budgets and there are various reports of platforms like Facebook being opaque about paid boosts and exaggerating the reach of their advertising.[35] Such lack of transparency also brings up ethical dilemmas on whether it is a good use of funds, including public money.

There are political contradictions too, to paid advertising. Data in the form of what we give to and do on social media is what drives their platform business model, sometimes known as platform capitalism[36].

[34] Daria Marmer, How Frequently Should I Publish on Social Media? A HubSpot Experiment
https://blog.hubspot.com/marketing/how-frequently-should-i-publish-on-social-media accessed 26 April 2021.
[35] Lara O'Reilly, Facebook Admits It Exaggerated Ad Metrics (Inc 2016)
https://www.inc.com/business-insider/facebook-ad-metrics-exaggeration-update.html accessed 22 April 2021.
[36] Nick Srnicek, The challenges of platform capitalism: understanding the logic of a new business model (The Progressive Policy Think Tank 2017)
https://www.ippr.org/juncture-item/the-challenges-of-platform-capitalism accessed 2 April 2021.

Researchers have noted how social media platforms also benefit from the proliferation of hate speech[37]. And platforms like Facebook have notoriously neglected to address hate speech and dangerous speech that incites violence[38]. So while it's inevitable that our counterspeech campaigns on those platforms will benefit their business model, directly remunerating them to promote counterspeech can be a greyer area to navigate. There are practical reasons sometimes to pay for such promotions but we should consider accompanying strategies of **lobbying platforms to boost counterspeech campaigns for free.**

There are other, and **unpaid, strategies to boost your content**. One would be to tap into your networks, affinity groups and any people with large online (and offline) following you may have access to, and asking them to share your content. Another would be to have the content and campaign covered by popular blogs and media (ex- Roar Media[39]). Using relevant and popular hashtags can also be a useful tactic.

FURTHER READING

Counter-Narrative Handbook (page 31 onwards) by Henry Tuck and Tanya Silverman
https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf

5. IMPLEMENTING THE DISSEMINATION STRATEGY

Once you have laid out your dissemination plan, you'll be able to gauge how much time needs to be committed by you for its full implementation, whether this is a practical possibility, and adjust plans accordingly.

---
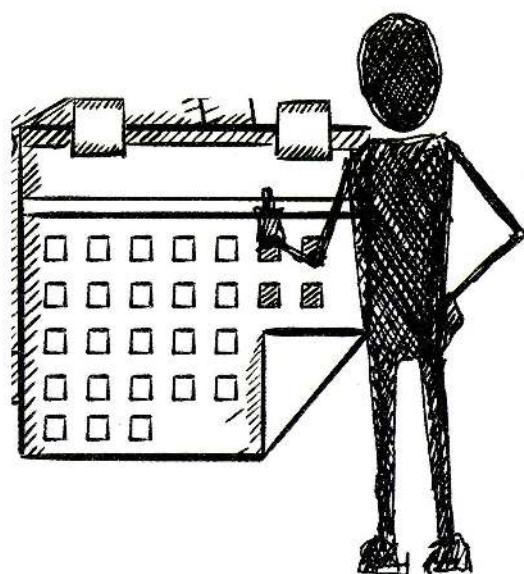
[37] Annex 1

[38] Joshua Brustein, Facebook Apologizes for Role in Sri Lankan Violence (Bloomberg 2020)
https://www.bloomberg.com/news/articles/2020-05-12/facebook-apologizes-for-role-in-sri-lankan-violence accessed 2 April 2021

[39] https://roar.media/english
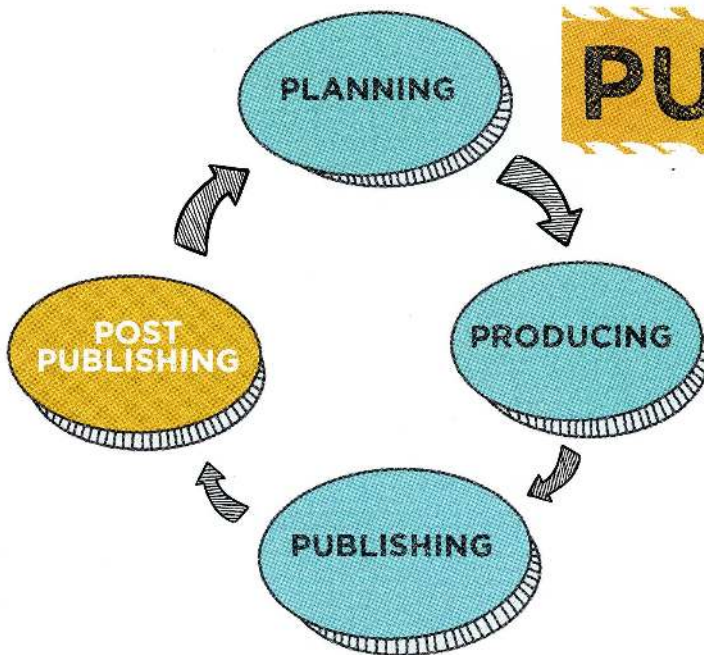
POST-PUBLISHING

Your counterspeech intervention doesn't end at the point of publishing. In fact this final phase of creating counterspeech content can often be the most crucial in ensuring sustained and longterm counterspeech work.

1. AUDIENCE ENGAGEMENT

"Talk to rather than at your audience; if you are responsive and communicative, then they will be too[40]."

Just as we attempt to find connections rather than being condescending or combative when we craft our counterspeech, the language and tone we use in dissemination should also be a way to invite engagement and conversation. The hallmark of successful or effective counterspeech would be a high level of audience engagement – both positive and negative – and this is something to be prepared for. For positive or neutral engagements, you can decide how often you want to engage and the kinds of tactics that might make your audience respond (ex- getting onboard a popular meme or hashtag, an exciting and not-too-complicated call to action).
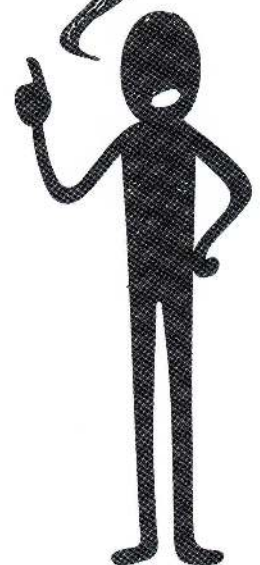
[40] Henry Tuck and Tanya Silverman, The Counter-Narrative Handbook (Institute for Strategic Dialogue 2016) https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf accessed 2 April 2021

In the case of negative feedback, in most instances the best practice is to ignore it rather than fuel it. And report it if it's harmful speech, attacking the people featured in your content, attacking other individuals, etc. However, there can also be occasions in which you can engage the negative feedback to further your counterspeech narrative[41]. In addition to all of these scenarios, there is also the possibility that other people on a platform take it upon themselves to counter or engage the negative feedback.

IGNORE
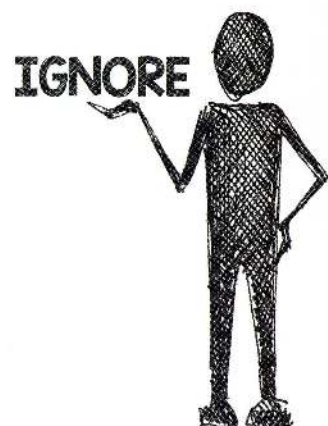
2. SAFETY AND SECURITY

We noted earlier in this toolkit that the legal framework that is applicable to hate speech in Sri Lanka[42] is more often used to criminalize legitimate speech and to silence and shut down dissent including by journalists, activists, cartoonists, bloggers, writers, artists, members of ethnic minorities, etc. In a democratic deficit, it is quite possible for your counterspeech campaign – especially if it becomes well known – to be framed as hate speech towards the dominant ethnicity, religion or both. And for those associated with the campaign to be criminalized.

While you may have considered this in the PEST during the context analysis (Annex 5), it is worth revisiting this discussion once the campaign has been published and keeping it as an ongoing conversation. And seek counsel from a lawyer if you're worried about how your counterspeech is being received and shared.

Another aspect of safety and security is that any content and campaigns that are published and hosted on third party platforms like social media platforms are never fully yours. Platforms could remove them indiscriminately or block access to your account or in some instances temporary or longterm internet shutdowns or platform blocking by the state[43] could make you lose access. While we don't always have the financial and technological capacities to maintain autonomous infrastructure[44], at a minimum we should ensure we don't treat social media platforms, Google drive and other third party infrastructure as our only storage spaces.

---

[41] Ibid (p. 40)

[42] Annex 2

[43] Keep It On https://www.accessnow.org/keepiton/

[44] Luisa Bagope and Sonaksha Iyengar, *Networks Woven with Care* (GenderIT 2021) https://genderit.org/feminist-talk/networks-woven-care accessed 2 April 2021

3. MONITORING AND EVALUATION

The final part of the final phase is monitoring and evaluation (M&E). This ensures a feedback loop between the four phases with the goals identified in the planning phase informing the quantitative and qualitative indicators that should be monitored and evaluated and the M&E findings informing the planning phase of your next counterspeech intervention. The M&E process should happen both while the campaign is running and at various intervals after it has concluded.

You can use quantitative and qualitative indicators and methods to find out if the goals you set at the start of your campaign have been accomplished. Quantitative indicators like the number of people reached, number of new followers, number of shares, etc. can be useful in understanding if your counterspeech reached people. Each platform has its own analytics system that you can use to keep track of these though as with advertising, there isn't always a guarantee if the numbers are overstated or understated. And even if accurate, quantitative indicators have inherent limitations and don't tell you a full story. Therefore it is important that your M&E also looks at qualitative indicators like feedback you received for your campaign, ways in which it was described by people who shared it, any shifts in online conversations about the issue your counterspeech addressed, and use qualitative methods like discourse analysis, and interviews or focus group discussions with some of your target audience members.

FURTHER READING

Counter-Narrative Handbook (page 43 onwards) by Henry Tuck and Tanya Silverman
https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf

Here concludes the four phases of creating counterspeech content online. We hope you have found useful tips, strategies and frameworks for your work.

Good luck!

CREATING COUNTERSPEECH CONTENT RECAP

2

Producing
* Crafting messages
* Sources, consent and privacy
* Format or medium
* Language and Access
* Creating content
* Testing and finalizing

3

Publishing
* Platform ecosystem
* Platforms vs Content
* When to post and how often
* Boosting and advertising
* Implementing the dissemination strategy

1

Planning
* Identify the issue
* Proactive or Reactive
* Context analysis
* Audience and Platforms
* Setting goals
* Timeline and calendar
* Budget and fundraising

4

Post-publishing
* Audience engagement
* Safety and Security
* Monitoring and Evaluation

OTHER WAYS
OF ADDRESSING
ONLINE HATE SPEECH

There are also other strategies to address online hate speech and usually our counterspeech interventions need to work in tandem with these to be effective and comprehensive.

Reporting and redress mechanisms on platforms can be used to flag harmful speech that violate human rights and platform community guidelines.
https://onlineharassmentfieldmanual.pen.org/reporting-online-harassment-to-platforms/

Continue to **document and research** the prevalence and proliferation of online hate speech in Sri Lanka.
https://hashtaggeneration.org/publications/ and https://deletenothing.org/

Prioritize **mental health support** for victims and survivors of hate speech and dangerous speech.
https://www.facebook.com/ShanthiMaargamSL/

Capacity strengthening and awareness raising on recognizing hate speech, understanding the consequences of hate speech, staying safe on the internet, etc.
https://sites.google.com/site/digisecsl/home

ANNEXURES

ANNEX 1

NEEDS ASSESSMENT BASED ON EXISTING LITERATURE

Prevalence and proliferation of online hate speech in Sri Lanka are documented in:

Countering Hate: An Analysis of Online Counter & Alternative Narratives to Harmful Speech in Sri Lanka by Hashtag Generation[45]

"One strategy adopted by both institutions and informal collectives working towards strengthening democracy, human rights, social cohesion and pluralism has been to develop counter and alternative narratives as well as fact-checks. While previous analyses, including by Hashtag Generation has analysed the circulation and dissemination of harmful speech such as hate speech, disinformation and online harassment, there has been little focus on the efficacy of such positive narratives. As such, this report analyses the performance of Facebook pages that share counter and alternative narratives to harmful speech in Sri Lanka based on the methodology mentioned outlined below."

Social Media Analysis - What Facebook Tells us about Social Cohesion in Sri Lanka by Democracy Reporting International, in consultation with Hashtag Generation[46]

"With the debate increasingly taking place in the virtual space, the spread of misleading information and hate speech has become a significant phenomenon that helps nurture anti-democratic tendencies."

"The combination of misleading information and hate speech within ethno-nationalist pages and their wide sharing on Facebook deepens social divisions according to political, ethnic and religious fault lines and therefore potentially endangers social cohesion in Sri Lanka."

---

[45] Hashtag Generation, Countering Hate: An Analysis of Online Counter & Alternative Narratives to Harmful Speech in Sri Lanka (Hashtag Generation 2021)

[46] Democracy Reporting International, Social Media Analysis - What Facebook Tells us about Social Cohesion in Sri Lanka (Democracy Reporting International 2019) <https://democracy-reporting.org/wp-content/uploads/2019/02/Social-Media-Analysis-draft-8-1-1.pdf> accessed 5 March 2021

Opinions, B*tch: Technology Based Violence Against Women in Sri Lanka by Ghosha, Centre for Policy Alternatives and Hashtag Generation[47]

"What emerged was a clear pattern of speech that was sexist, or objectified, harassed or otherwise targeted women and members of the LBT community. The non-consensual dissemination of intimate photos and videos was another disturbing trend found in the lead-up to this study, with entire pages dedicated to such content, or alternatively linking to such content on third-party websites. The findings of this report indicate the normalisation of sexist commentary, escalating to and including violence against women and LBT communities, both online and offline."

Saving Sunil: A study of dangerous speech around a Facebook page dedicated to Sgt. Sunil Rathnayake by the Centre for Policy Alternatives[48]

"Social media such as Facebook groups create virtual communities who 'gather', bond and vent over shared interests or causes. This like-mindedness is the glue that binds group members together. Such groups become echo chambers where an inkling of hate is quickly amplified into something larger by feeding group members information which simply reinforces that initial thought. Intolerance of 'outsiders' or those who disagree or challenge the ethos of the group and vicious hate speech targeting them as exemplified in the Saving Sunil group, is common."

Liking violence: A study of hate speech on Facebook in Sri Lanka by the Centre for Policy Alternatives[49]

"A considerable amount of social media hate speech in Sri Lanka occurs on Facebook. The ability to like, share and comment on posts allows forums for supporters to engage, to plan rallies and other events and keep all similar posts in one place. It also allows admins of pages to remove and ban dissenting voices, allowing a greater degree of control than platforms such as Twitter."

[47] Ghosha, Centre for Policy Alternatives and Hashtag Generation, Opinions, B*tch: Technology Based Violence Against Women in Sri Lanka (Centre for Policy Alternatives 2019) https://drive.google.com/file/d/16ohQ7P2K08qz-kgiQUfSrOJCGvxDLHNR/view accessed 5 March 2021
[48] Centre for Policy Alternatives, Saving Sunil: A study of dangerous speech around a Facebook page dedicated to Sgt. Sunil Rathnayake (Centre for Policy Alternatives 2015) <https://www.cpalanka.org/saving-sunil-a-study-of-dangerous-speech-around-a-facebook-page-dedicated-to-sgt-sunil-rathnayake/> accessed 5 March 2021
[49] Centre for Policy Alternatives, Liking violence: A study of hate speech on Facebook in Sri Lanka (Centre for Policy Alternatives 2014) <https://www.cpalanka.org/liking-violence-a-study-of-hate-speech-on-facebook-in-sri-lanka/> accessed 5 March 2021

Opinions, B*tch: Technology Based Violence Against Women in Sri Lanka by Ghosha, Centre for Policy Alternatives and Hashtag Generation[50]

"CSOs must support and act as advocates for long term public awareness campaigns on technology-related violence, as well as on campaigns around digital literacy and comprehensive sexuality education. This can include counter-messaging campaigns to advocate for respectful and inclusive speech online. This should ideally be supported by relevant line Ministries where possible. CSOs must lobby the State to include digital literacy and comprehensive sexuality education in the national curriculums."

Liking violence: A study of hate speech on Facebook in Sri Lanka by the Centre for Policy Alternatives[51]

"Though there is no easy or prescribed solution, progressive thinking, proactive content production, strategic interventions and careful monitoring can identify and neutralise the wider harm online hate speech can, if unchecked and allowed to grow, sustain and strengthen."

Countering Harmful Speech: Why Trust the State? by Gehan Gunatilleke[52]

"The type and intensity of interference may depend on the classification of the harmful speech. The intensity of the interference can be placed on a spectrum, and should correspond to the intensity of the harm that the speech causes. At one end of the spectrum, there can be interference through private plaints and counter-messaging. Next, there can be civil sanctions of some kind, including boycotts, dismissals from employment, and the revocation of licences. At the other end of the spectrum, there can be criminal sanctions such as fines and imprisonment. It is natural for the state to play a role in how society curbs harmful speech. Yet there may be good reasons for minimising this role – particularly when it comes to imposing criminal sanctions. One of the main reasons – and this reason is particularly relevant to the Sri Lankan context – is the loss of trust in the state."

---

[50] Ghosha, Centre for Policy Alternatives and Hashtag Generation (n 47)
[51] Centre for Policy Alternatives (n 49)
[52] Gehan Gunatilleke, (n 11)

ANNEX 2

INTERNATIONAL LEGAL FRAMEWORK

Universal Declaration on Human Rights

Article 2

Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, jurisdictional or international status of the country or territory to which a person belongs, whether it be independent, trust, non-self-governing or under any other limitation of sovereignty.

Article 19

Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.

International Covenant on Civil and Political Rights (ICCPR)

Article 19

1. Everyone shall have the right to hold opinions without interference.
2. Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.
3. The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:
   * For respect of the rights or reputations of others;
   * For the protection of national security or of public order (ordre public), or of public health or morals.

Article 20

1. Any propaganda for war shall be prohibited by law.
2. Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.

The Constitution of the Democratic Socialist Republic of Sri Lanka 1978

Article 14(1)(a)

Every citizen is entitled to the freedom of speech and expression including publication

Article 15(2)

The exercise and operation of the fundamental right declared and recognized by Article 14(1)(a) shall be subject to such restrictions as may be prescribed by law in the interests of racial and religious harmony or in relation to parliamentary privilege, contempt of court, defamation or incitement to an offence.

Article 15(7)

The exercise and operation of all the fundamental rights declared and recognized by Articles 12, 13(1), 13(2) and 14 shall be subject to such restrictions as may be prescribed by law in the interests of national security, public order and the protection of public health or morality, or for the purpose of securing due recognition and respect for the rights and freedoms of others, or of meeting the just requirements of the general welfare of a democratic society. For the purposes of this paragraph "law" includes regulations made under the law for the time being relating to public security.

International Covenant on Civil and Political Rights (ICCPR) Act, No. 56 of 2007

Article 3(1)

No person shall propagate war or advocate national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.

Article 3(2)

Every person who—

* attempts to commit;
* aids or abets in the commission of; or
* threatens to commit,

an offence referred to in subsection (1), shall be guilty of an offence under this Act.

Penal Code, No. 2 of 1883

Chapter XV on offences relating to religion, specifically Section 291A and 291B Section 120 on "exciting or attempting to excite disaffection"

Prevention of Terrorism (Temporary Provisions) Act (PTA), No. 48 of 1979

Section 2(1)(h)

Any person who by words either spoken or intended to be read or by signs or by visible representations or otherwise causes or intends to cause commission of acts of violence or religious, racial or communal disharmony or feelings of ill-will or hostility between different communities or racial or religious groups shall be guilty of an offence under this Act.

Section 14(2)(a)(ii)

No person shall, without the approval in writing of a competent authority, print or publish in any newspaper any matter relating to incitement to violence, or which is likely to cause religious, racial or communal disharmony or feeling of ill-will or hostility between different communities or racial or religious groups.

Police Ordinance, No. 16 of 1865

Section 79(2)

Any person who in any public place or at any public meeting uses threatening, abusive or insulting words or behaviour which is intended to provoke a breach of the peace or whereby a breach of the peace is likely to be occasioned, shall be guilty of an offence under this section.

Section 98

Any person who shall spread false reports with the view to alarm the inhabitants of any place within Sri Lanka and create a panic shall be guilty of an offence, and be liable to a fine not exceeding two hundred rupees, or to imprisonment, with or without hard labour, for any period not exceeding twelve months ; and if he shall be convicted a second time, or shall persist in the offence after warning to desist, he shall be liable to corporal punishment not exceeding twenty lashes.

Public Security Ordinance, No. 25 of 1947

Section 5 on "Power of President to make emergency regulations"
Section 7 on "Emergency regulations to prevail over other law."

Bail Act , No. 30 of 1997

Section 3(1)

Nothing in this Act shall apply to any person accused or inspected of having commuted, or convicted of, an offence under, the Prevention of Terrorism (Temporary Provisions) Act. No 48 of 1979, Regulations made under the Public Security Ordinance or any other written law which makes express provision in respect of the release on bail of persons accused or suspected of having committed, or convicted of, offences under such other written law.

Sri Lanka Telecommunications Act (SLTA), No. 25 of 1991

Section 5(f)

For the purpose of achieving the objects referred to in section 4, the Authority shall have the following powers, and duties to take such regulatory measures as may be prescribed to comply with any general or special directions that may be given to him from time to time by the Government of Sri Lanka in the interest of national security, public order and the defence of the country

Section 58

Every person who by himself or by his agent or servant tenders for transmission at any telecommunication office any message of an indecent, obscene, seditious, scurrilous, threatening or grossly offensive character, shall be guilty of an offence and shall be liable on conviction to a fine not exceeding five thousand rupees or to imprisonment of either description for a term not exceeding six months, or to both such fine and such imprisonment

Computer Crimes Act, No. 24 of 2007

Section 6

Offences committed against national security.

Any person who intentionally causes a computer to perform any function, knowing or having reason to believe that such function will result in danger or imminent danger to—

✳ national security;

✳ the national economy; or

✳ public order,

shall be guilty of an offence and shall on conviction be punishable with imprisonment of either description for a term not exceeding five years.

Sections 11-21

ANNEX 3

INTERNET AND SOCIAL MEDIA LANDSCAPE OF SRI LANKA

Source: Digital 2021 Sri Lanka Report by DataReportal

Internet users in Sri Lanka

10.90 MILLION

There were 10.90 million internet users in Sri Lanka in January 2021.

7.9%

The number of internet users in Sri Lanka increased by 800 thousand (+7.9%) between 2020 and 2021.

50.8%

Internet penetration in Sri Lanka stood at 50.8% in January 2021.

Social media statistics for Sri Lanka

7.90 MILLION

There were 7.90 million social media users in Sri Lanka in January 2021.

23%

The number of social media users in Sri Lanka increased by 1.5 million (+23%) between 2020 and 2021.

36.8%

The number of social media users in Sri Lanka was equivalent to 36.8% of the total population in January 2021.

7.00 MILLION

Potential number of users on Facebook - 7 million

88.5%

Percentage of users accessing Facebook through mobile phones only - 88.5%

SONGS TAMIL MOVIE SONG MOVIES SRI LANKA

Top 5 YouTube search queries - Song, Songs, Sri Lanka, Tamil movie, Movies

1.30 MILLION

Potential number of users on Instagram - 1.3 million

3.80 MILLION

Potential number of users on Facebook Messenger - 3.8 million

187,200

Potential number of users on Twitter - 187,200

Internet speed, devices and browsers statistics for Sri Lanka

16.9 MBPS

Average download speed of mobile internet connections - 16.9 MBPS

26.86 MBPS

Average download speed of fixed internet connections - 26.86 MBPS

75%

Highest share of web traffic by browser - 75% on Chrome

91.5%

Share of web traffic originating from Android devices - 91.5%

ANNEX 4

REPOSITORY OF GOOD PRACTICES IN COUNTERING HATE SPEECH FROM SRI LANKA AND BEYOND

https://docs.google.com/spreadsheets/d/1KeZdzfsGq24_JQtT6_hb--NWLqyC-vZadns-IqcFPdI/edit?usp=sharing

ANNEX 5

CONTEXT ANALYSIS FRAMEWORKS

PEST Analysis (Political, Economic, Social and Technological)

What is it?

It is best to begin with this tool as it maps the external trends or forces that may have a positive or negative impact on the organisation. This exercise can be framed around three things; identify the key political, economic, social and technological factors in the external environment of the organization, analyse the factors in terms of opportunities, risks and threats that they would form for communications of the organization and determine what the next steps for the organization would be (ex- can any threats be turned into opportunities).

Shortcomings

Some possible shortcomings of this tool are that we could miss out on key factors just because we've already got some under a category and decide to move on, while carrying out the exercise we may forget why external factors are strategically important for the organization and merely go through the motions for the sake of it, and we can sometimes get stuck in only listing what we already know so we need to proactively think of new/emerging factors we may not be very familiar

Sample template

POLITICAL
* Legislation – domestic, international, future, etc.
* Regulatory bodies and processes
* Treaty bodies
* Wars and conflicts
* Space for civil society

ECONOMIC
* Economic growth
* Interest rates
* Exchange rates
* Economic trends
* Funding, grants

SOCIAL
* Lifestyle trends
* Demographics
* Media landscape
* Media consumption behavior
* Celebrities/influencers
* Major events
* Influence of religion, culture
* Advertising and publicity
* Ethical issues

TECHNOLOGICAL
* Technology landscape
* Information and communications
* technology
* Social media
* Platform capitalism
* Technology legislation
* Intellectual property issues

What is it?

SWOT helps us consider the internal strengths and weaknesses, and the external opportunities and threats as we look at the organization's future, including while operationalizing a communications strategy. Results of the PEST analysis can be used when identifying external opportunities and threats. After identifying strengths, weaknesses, opportunities and threats, these need to be analysed in relation to communications. For an example, how can the strengths be used to their full capacity through effective communications?

Shortcomings

A common criticism of SWOT is that the framing of the tool results in an organization's thinking remaining at the same place rather than generating new ideas and helping the organization grow. Many times a SWOT analysis could leave an organization feeling overwhelmed about weaknesses and threats without constructive future steps. So it is important to keep reminding ourselves during the course of the exercise to give equal time to discuss next steps, including anticipated results.

One option is to carry out a SOAR exercise instead of a SWOT. "The SOAR approach to strategy starts with a strategic inquiry. During this inquiry an organization's greatest **S**trengths and **O**pportunities are discovered and explored among the participants. The participants are invited to share their **A**spirations and co-construct their most preferred future. Then, recognition and reward programs are design to inspire employees to achieve measurable **R**esults."[53]

Strenghts
Weaknesses
Opportunities

SWOT
Strenghts

SOAR

Opportunities
Threats
Aspirations
Results

ANALYSIS

ANALYSIS

---
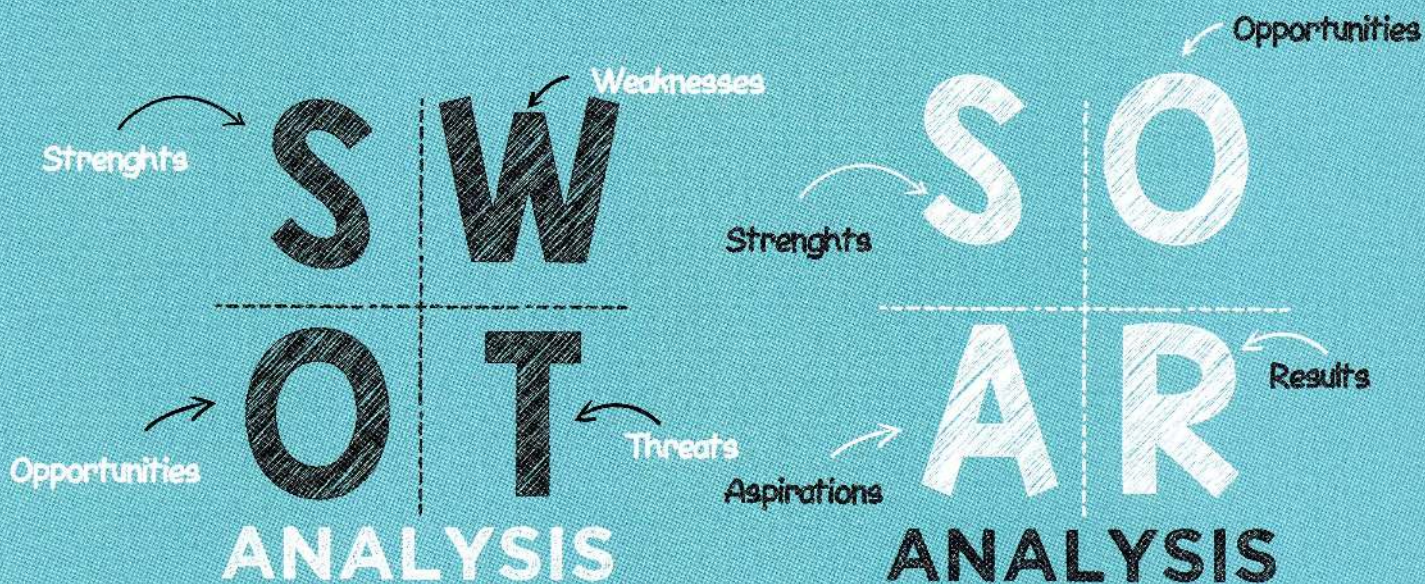
[53] Jacqueline Stavros , David Cooperrider , D. Lynn Kelley, Strategic Inquiry > Appreciative Intent: Inspiration to SOAR, 2003

SWOT

STRENGTHS
* Capacities of the organization (resources, assets, people, experience, knowledge, data, etc.)
* USPs (unique selling points)
* Network
* Accreditations
* Processes, systems

WEAKNESSES
* Gaps in capacities
* Reputation, reach, visibility
* Deadlines, pressure
* Processes, systems

OPPORTUNITIES
* Innovation
* Technology developments
* Attitudinal changes
* Legal reform

THREATS
* Political
* Legislative
* Democratic deficit

SOAR

S STRENGTHS
What are our greatest assets and strengths?

O OPPORTUNITIES
What are the opportunities to create change?

A ASPIRATIONS
What is the future and change we dream of?

R RESULTS
What are the outcomes we want? Measurable or otherwise.

ANNEX 6

1. Brainstorm to identify all possible audience segments related to your issue
2. Place them in one or both of the frameworks below to determine the target audiences of your counterspeech content and campaign.

Audience power and interest mapping

		QUADRANT 4	QUADRANT 3
POWER	HIGH (veto power)	Keep satisfied	Engage closely
		QUADRANT 1	QUADRANT 2
	LOW (little influence)	Monitor (not necessary to communicate with)	Keep informed

INTEREST

LOW (Indifferent) —————————————— HIGH (strongly support or oppose)

Audience attitude and involvement mapping[54]

Involvement — HIGH / LOW

HIGH INVOLVEMENT, NEGATIVE ATTITUDE

Most Likely Changes: Move to high involvement, low Attitude (do less harm); Move to low involvement, negative Attitude (do less harm)

Examples of Audience Types: DS Speakers, People Likely to Participate in Group-Targeted Harm Willingly; People Who Encourage Others to Participate in Ds and/or Group-Targeted Harm, Influential Leaders, Information Spreaders

HIGH INVOLVEMENT, NEUTRAL ATTITUDE

Most Likely Changes: Move to high Involvement, high Attitude (make a positive impact); Prevent from becoming negative Attitude (prevent future harm); move to low Involvement, low Attitude (reduce potential for harm)
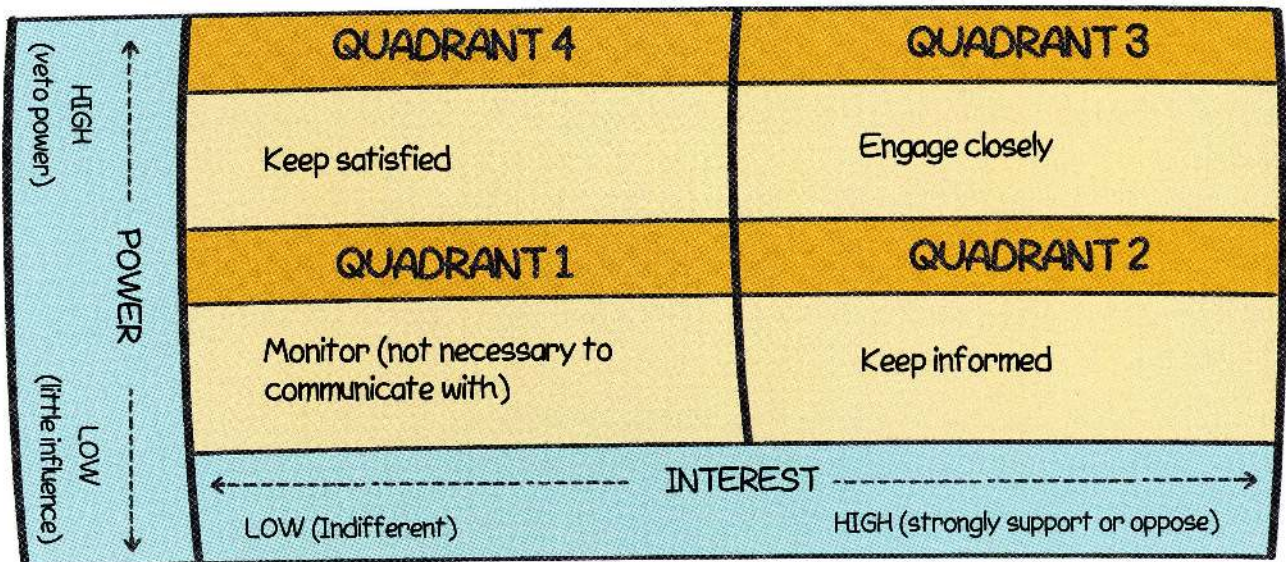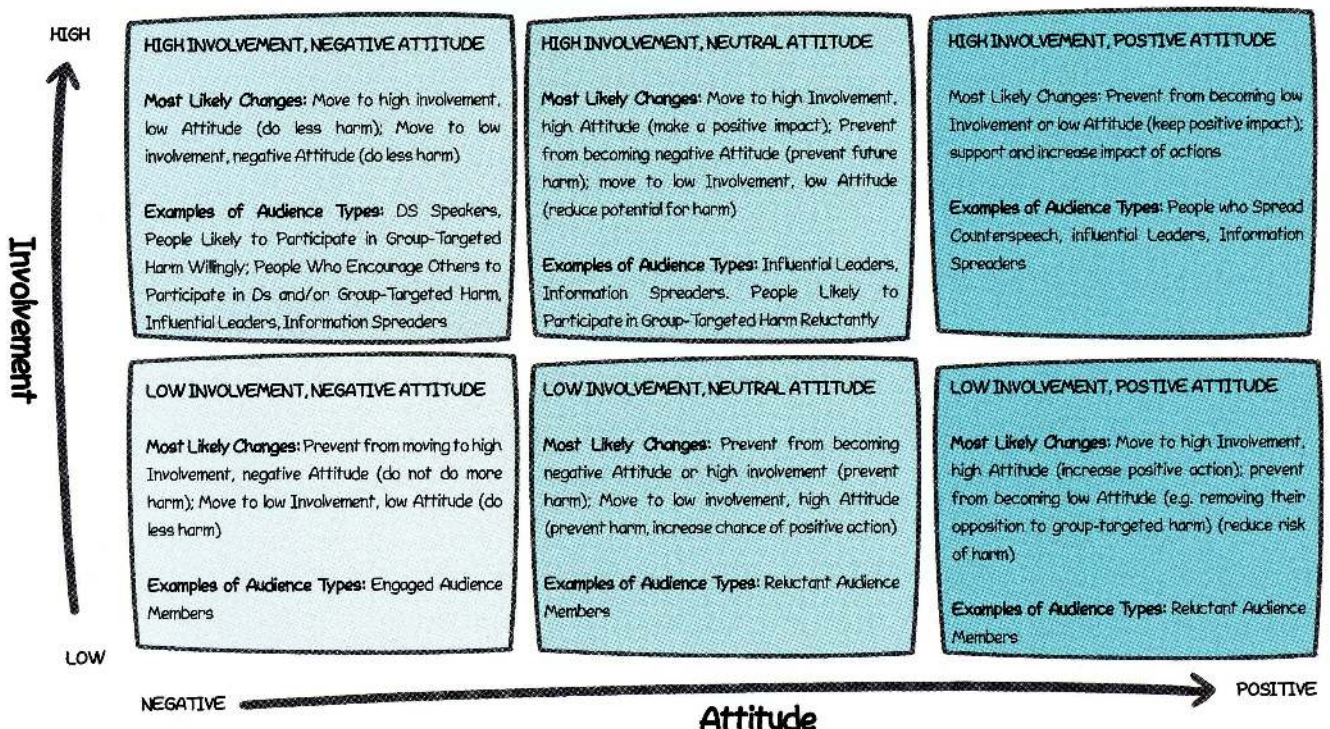
Examples of Audience Types: Influential Leaders, Information Spreaders. People Likely to Participate in Group-Targeted Harm Reluctantly

HIGH INVOLVEMENT, POSITIVE ATTITUDE

Most Likely Changes: Prevent from becoming low Involvement or low Attitude (keep positive impact); support and increase impact of actions

Examples of Audience Types: People who Spread Counterspeech, Influential Leaders, Information Spreaders

LOW INVOLVEMENT, NEGATIVE ATTITUDE

Most Likely Changes: Prevent from moving to high Involvement, negative Attitude (do not do more harm); Move to low Involvement, low Attitude (do less harm)

Examples of Audience Types: Engaged Audience Members

LOW INVOLVEMENT, NEUTRAL ATTITUDE

Most Likely Changes: Prevent from becoming negative Attitude or high involvement (prevent harm); Move to low involvement, high Attitude (prevent harm, increase chance of positive action)

Examples of Audience Types: Reluctant Audience Members

LOW INVOLVEMENT, POSITIVE ATTITUDE

Most Likely Changes: Move to high Involvement, high Attitude (increase positive action); prevent from becoming low Attitude (e.g. removing their opposition to group-targeted harm) (reduce risk of harm)

Examples of Audience Types: Reluctant Audience Members

NEGATIVE ———————————————→ POSITIVE

Attitude

54 Rachel Brown (n 28)

BIBLIOGRAPHY

Article 19, Principle 12(1)9i) of The Camden Principles on Freedom of Expression and Equality (Article 19 2009)

Bagope L and Iyengar S, Networks Woven with Care (GenderIT 2021)
<https://genderit.org/feminist-talk/networks-woven-care> accessed 2 April 2021

Brown R, Defusing Hate: A Strategic Communication Guide to Counteract Dangerous Speech (United States Holocaust Memorial Museum 2016)
<https://www.ushmm.org/m/pdfs/20160229-Defusing-Hate-Guide.pdf> accessed 25 March 2021

Brustein J, Facebook Apologizes for Role in Sri Lankan Violence (Bloomberg 2020)
<https://www.bloomberg.com/news/articles/2020-05-12/facebook-apologizes-for-role-in-sri-lankan-violence> accessed 2 April 2021

Centre for Policy Alternatives, Liking violence: A study of hate speech on Facebook in Sri Lanka (Centre for Policy Alternatives 2014)
<https://www.cpalanka.org/liking-violence-a-study-of-hate-speech-on-facebook-in-sri-lanka/> accessed 5 March 2021

Centre for Policy Alternatives, Saving Sunil: A study of dangerous speech around a Facebook page dedicated to Sgt. Sunil Rathnayake (Centre for Policy Alternatives 2015)
<https://www.cpalanka.org/saving-sunil-a-study-of-dangerous-speech-around-a-facebook-page-dedicated-to-sgt-sunil-rathnayake/> accessed 5 March 2021

Dangerous Speech Project <https://dangerousspeech.org>

Davis R and others, Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review (Health Psychology Review 2015)
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4566873/> accessed on 2 April 2021

Democracy Reporting International, Social Media Analysis - What Facebook Tells us about Social Cohesion in Sri Lanka (Democracy Reporting International 2019)
<https://democracy-reporting.org/wp-content/uploads/2019/02/Social-Media-Analysis-draft-8-1-1.pdf> accessed 5 March 2021

Dieng A, Stopping Hate Speech (United Nations 2019)
https://www.youtube.com/watch?v=rnbcQT-b8ak&t=142s accessed 3 May 2021

Freedom House, Freedom on the Net 2020 - Sri Lanka (Freedom House 2020)
<https://www.ecoi.net/en/document/2039119.html> accessed 10 April 2021

Galpaya H and others, AfterAccess: ICT access and use in Sri Lanka and the Global South
(LIRNEasia 2019)
<https://lirneasia.net/2019/05/afteraccess-ict-access-and-use-in-sri-lanka-and-the-global-sou
th-presentation/> accessed 2 April 2021

Ghosha, Centre for Policy Alternatives and Hashtag Generation, Opinions, B*tch: Technology
Based Violence Against Women in Sri Lanka (Centre for Policy Alternatives 2019)
https://drive.google.com/file/d/16ohQ7P2K08qz-kgiQUfSr0JCGvxDLHNR/view accessed 5
March 2021

Gunatilleke G, Countering Harmful Speech: Why Trust the State? (Groundviews 2020)
<https://groundviews.org/2020/12/17/countering-harmful-speech-why-trust-the-state/>
accessed 25 March 2021

Hashtag Generation, Countering Hate: An Analysis of Online Counter & Alternative Narratives to
Harmful Speech in Sri Lanka (Hashtag Generation 2021)

Lua A, What to Post on Each Social Media Platform: The Complete Guide to Optimizing Your
Social Content
<https://buffer.com/library/what-to-post-on-each-social-media-platform/#what-to-post-on-e
ach-social-media-platform> accessed 26 April 2021

Marmer D, How Frequently Should I Publish on Social Media? A HubSpot Experiment
<https://blog.hubspot.com/marketing/how-frequently-should-i-publish-on-social-media>
accessed 26 April 2021

Office of the United Nations High Commissioner for Human Rights, Freedom of expression vs
incitement to hatred: OHCHR and the Rabat Plan of Action (OHCHR 2013)
https://www.ohchr.org/en/issues/freedomopinion/articles19-20/pages/index.aspx accessed
3 May 2021

O'Reilly L, Facebook Admits It Exaggerated Ad Metrics (Inc 2016)
<https://www.inc.com/business-insider/facebook-ad-metrics-exaggeration-update.html>
accessed 22 April 2021

Srnicek N, The challenges of platform capitalism: understanding the logic of a new business model
(The Progressive Policy Think Tank 2017)
<https://www.ippr.org/juncture-item/the-challenges-of-platform-capitalism> accessed 2 April
2021

Tuck H and Silverman T, The Counter-Narrative Handbook (Institute for Strategic Dialogue 2016)
<https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf>
accessed 2 April 2021

United Nations, United Nations Strategy and Plan of Action on Hate Speech (United Nations
2019) <https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml> accessed 20
March 2021