Defusing Hate

A Strategic Communication Guide to Counteract Dangerous Speech

By Rachel Hilary Brown

"Dangerous speech" is a term for hate speech that, under the right conditions, can influence people to accept, condone and commit violence against members of a group. How can those seeking to prevent violence strategically use communication to preempt and counter the influence of dangerous speech?

To be successful, peace actors must be able to identify audiences for their interventions, understand these audiences, choose speakers and mediums that can reach and influence them, craft effective messages, and avoid risks.

The concepts, approaches, tools, and examples included in this Guide can help peace actors think through opportunities and risks and use strategic communication to prevent dangerous speech from facilitating group-targeted violence. This Guide is not a "How-To" manual, but a tool that peace actors can use and adapt according to their knowledge, judgment, and goals.

@

If you are using this resource, please let us know!

We are building a community of practice, and we would like to improve this resource for you. You can share your comments, questions, feedback, or experiences by emailing us at: **dangerousspeech@ushmm.org**

How can you use this Guide to meet your needs? You can use the Guide to:

1. **Become familiar with concepts** about how to create effective communications to influence audiences in the face of dangerous speech.

2. **Learn skills** that will help you understand the environment for your communications, select and analyze the audiences you wish to influence, and choose effective mediums, speakers, and messages.

3. **Design interventions** that avoid unintended consequences and effectively influence intended audiences.

4. **Adapt and respond** to changing contexts and audiences

Table of Contents

Introduction

1 | DANGEROUS SPEECH

"Dangerous speech" is speech that increases the risk for violence targeting certain people because of their membership in a group, such as an ethnic, religious, or racial group. It includes both speech that qualifies as incitement and speech that makes incitement possible by conditioning its audience to accept, condone, and commit violence against people who belong to a targeted group.[1] For example, Hutu extremists were able to incite genocide in Rwanda in part because years of propaganda had influenced Hutus to view Tutsis as less than human and so dangerous that they must be eliminated from the country. The propagandists' goal may not have been genocide, but their work prepared Hutus to understand and answer the call to act when extremist leaders launched the genocide.

Dangerous speech can take a variety of forms, such as an actual speech, a pamphlet, an online post, a video, an image or message on a T-shirt, or even a song. Its message may call for violence against a target group or may portray the target group in a way that makes violence against it seem reasonable, justified, and necessary. Dangerous speech often dehumanizes the group it targets (e.g., by calling its members rats, dogs, or lice), accuses the target group of planning to harm the audience, and presents the target group's existence as a dire threat to the audience. Speech may be dangerous even if it isn't intended to cause violence: for example, a false rumor that a rival group is planning to attack could make violence against the group's members seem like justified self-defense.[2]

The message by itself cannot make speech dangerous; the other factors that give speech the power to provoke violence include: [3]

- A speaker who is influential or popular with the audience;
- A medium (the means used to communicate a message) that makes the audience more likely to access, believe, or spread the speech;
- A context that increases the risk that the speech will provoke violence toward a group;
- An audience that is receptive to speech that promotes violence, fear, or hatred toward a group.

Measures that prevent dangerous speech from conditioning and inciting audiences to commit group-targeted harm (such as discrimination, persecution or violent attacks) can enhance efforts to prevent genocide, mass atrocities, and other forms of collective violence that target victims based on their group identity. Interventions that strategically use communication to prevent and reduce the impact of dangerous speech on its intended audience provide a means to counteract dangerous speech without restricting the right to free expression. The goal of this Guide is to acquaint users with information, skills and tools they can use to design safe and effective interventions that use strategic communication to counteract the impact of dangerous speech (hereafter referred to as Dangerous Speech Interventions or DSIs).

Dangerous Speech Interventions should aim to:

- Reduce the likelihood that audiences will accept and spread dangerous speech;
- Reduce the likelihood that audiences will condone or participate in group-targeted harm;
- Increase willingness among audience members to speak out against efforts to foment group-targeted hate.

To be successful, DSIs must reach and influence audiences that may be influenced by dangerous speech. In designing DSIs, it is important to identify specific audiences for the interventions, understand why they may be receptive to the message of dangerous speech, choose speakers and mediums that can reach and influence them, craft effective messages, and recognize and avoid risks.

3 | GUIDING PRINCIPLES

There are three Guiding Principles for DSIs:

1. Local Approach: Working with local partners is key. Community-level partners understand their specific context and have knowledge and social capital that is needed to make interventions successful. While intergroup hostility and conflict occurring within a country may be based on national-level grievances, local factors and influences help determine when and how violence happens.[4] Working with local partners will be crucial in order to understand, monitor, and respond to dynamics at the local level and to create approaches that are tailored for specific demographic and geographic audiences.

2. Goal-Oriented & Strategic: Clear and specific goals can guarantee that each DSI is designed based on intended outcomes and that each piece of DSI strategy and content has a clear and measureable aim. General goals, such as reducing the risk that audience members will condone or participate in group-targeted harm, can guide an overall intervention. Smaller, "micro-goals" (such as prompting audience members to question a rumor or refrain from repeating a message) should determine the intended impact of each piece of the strategy and message content, and should clearly connect to the intervention's overarching goals.

3. Do No Harm & Manage Risks: Because they are likely to be implemented in the context of ongoing intergroup hostility or conflict, DSIs face significant risks and challenges. Poorly designed interventions can result in simply "preaching to the converted," or worse, can increase an audience's exposure and receptivity to dangerous speech. This Guide provides insights and risk analysis tools to help practitioners analyze risks and ensure that, above all else, they do not cause additional harm. While each practitioner must ultimately make decisions based on his or her knowledge and best judgment as to whether the intervention could cause unintended harm, the concepts and tools included here can help facilitate this evaluation. Beyond this Guide, there are several resources available to assist with Do No Harm assessments.[5]

The Guide includes relevant concepts, approaches, tools, and examples that can be used to design DSIs. The Guide is broken into three phases of intervention design: Understand Context and Conflict; Select and Analyze Audiences; Select and Design Mediums, Speakers, and Message Content.

A workbook with exercises for conducting each of these design phases accompanies this guide. Consider, choose, adapt, and implement the approaches and the exercises (found in the workbooks) based on your context, priorities, and needs. The topics and design components covered under each phase are as follows.

Phase 1
Understand Context and Conflict

Phase 2
Select and Analyze Audiences

Phase 3
Select and Design Mediums, Speakers & Message Content

As a first step in designing a DSI, it is critical to review and consider a range of contextual factors that relate to existing dangerous speech and efforts to counteract it. In addition to providing an overview of contextual factors to consider, this section of the guide outlines how contextual factors relate to the design process, gives guidance on predicting conflict trajectories and setting broad goals for the intervention, and provides descriptions of context research tools.

Phase 1
Understand Context and Conflict

Phase 2
Select and Analyze Audiences

Phase 3
Select and Design Mediums,
Speakers & Message Content

Phase 1
Understand Context and Conflict

Phase 2
Select and Analyze Audiences

Phase 3
Select and Design Mediums,
Speakers & Message Content

PHASE 2: SELECT AND ANALYZE AUDIENCES

A clear understanding of the audience is crucial for setting clear, targeted goals and identifying whom to influence and how. This section provides frameworks for breaking the audience into different target groups, understanding the audience, and setting audience-specific goals. The section also suggests tools for audience research.

PHASE 3: DESIGN MEDIUMS, SPEAKERS, AND MESSAGE CONTENT

1: Mediums: This section outlines an approach for selecting mediums that can reach and influence the targeted audience, and it provides a summary of relevant concepts and strategies. It also explains medium-related risks and provides suggestions for how to mitigate these risks.

2: Speakers: In addition to outlining strategies and concepts for selecting speakers, this section includes a set of tools for developing and leveraging a brand. It also describes different types of risk and risk mitigation strategies.

3: Message Content: This section outlines an approach for setting message goals and developing message content, then explains relevant theories and concepts that can be employed and considered for message development. It also includes an overview of message risk analysis and message development research tools.

Phase 3
Select and Design Mediums,
Speakers & Message Content

6 | EXAMPLES FROM THE FIELD

This guide draws on experiences and examples of actual interventions that helped reduce the impact of dangerous speech on its intended audience. The following interventions are referenced throughout the Guide:

Interfaith Mediation Center in Northern Nigeria

The Interfaith Mediation Center (IMC) recruits pastors, imams, and other religious leaders to counter negative messages and events. They use a variety of mediums, including gathering people in their communities to talk face-to-face and using news media such as radio to make joint statements and discuss issues. They use verses from the Bible and the Koran to counter negative messages. They respond to incitement and events in real-time through a system of trained community members who monitor events at the local level.[6]

Sisi ni Amani Kenya

Sisi ni Amani Kenya used a text messaging platform combined with face-to-face outreach and grassroots programming to reach more than 20 communities and 65,000 subscribers with messages that aimed to change behavior around group-targeted harm during the 2013 election cycle. The messages focused on civic education (to decrease vulnerability to rumors), civic engagement (using the messages to invite people to public events that created unity), and violence prevention (sending messages to mitigate the impact of rumors and incitement, and remind people about the cost of violence).[7]

Tuzla in Bosnia (1990s)

In the early 1990s, the people of Tuzla, a city in Bosnia, were able to reject the violence that spread throughout the country based on ethnic and religious divides. The city's leadership predicted the violence that was to come and, together with the city's residents, conducted a strong and successful campaign to promote a unified Tuzlan identity through strong communications and actions. Not only did the city's inhabitants reject violence, people stayed in the city and continued to work and mix together across ethnic and religious lines.[8]

"I am Karachi" in Pakistan

As part of the "I am Karachi" initiative, artists and designers have come together to collaborate and replace the negative graffiti (called wall chalkings, which have historically been used as a means of public communication and in the recent past have turned into a space for divisive propaganda) that cover the city's walls with positive images and messages that unite people with a common identity of living in Karachi. The negative wall chalkings often target groups and build and deepen divides and conflict. In addition to reclaiming public spaces, this initiative helps build a group of artists dedicated to challenging hatred and divisiveness and promoting unity and a positive place-based identity.[9]

Muslim Community in Rwanda (1994)

Rwanda's Muslim community is an example of a group (a full community rather than isolated individuals) that resisted the appeal of dangerous speech and other pressures to participate in the genocide. The Muslim community, which had both Hutu and Tutsi members, not only refused to participate in the genocide but actively opposed it. Its actions during the genocide included rescuing, hiding, and taking care of Muslim and non-Muslim Tutsis, and providing safe haven in mosques. Muslims also rejected commands to kill or reveal Tutsis hidden in their communities, on several occasions going so far as to fight back and be killed themselves. The Muslim religious leadership played a key role in inspiring and organizing the community's resistance to the genocide.[10] While many factors contributed to this resistance to the genocide, there are several key strategies that relate to DSIs.

Sawa Shabab Radio Show in South Sudan (USIP)

Sawa Shabaab is a radio drama designed to "increase [listeners'] knowledge and change their attitudes regarding conflict." It was created in response to broadcasts of "war songs and radio stories that glorify conflict, combined with social media filled with comments that instigate violence" in South Sudan.[11] Sawa Shabaab is an example of edutainment, using an education curriculum and presenting it through a radio drama. It aims to get citizens to take an active role in peacebuilding.

Phase 1
Understand Context and Conflict

Phase 2
Select and Analyze Audiences

Phase 3
Select and Design Mediums,
Speakers & Message Content

The first step towards designing an intervention is gathering insights about the environment you're working in, analyzing the contextual factors that affect the impact of dangerous speech, and collecting information that will help you design your intervention.

Understanding How Context Affects Intervention Design

Contextual factors impact how audience members receive and react to mediums, speakers, and message content.

For example, the historical context – not only what happened but how people interpret it – influences how people view current events. It also feeds into the narrative context – such as how people in a group talk about their own identity and their relations with other groups. If a commonly accepted narrative promotes dangerous speech themes (e.g., it portrays all members of another group as inherently bad or dangerous), you will need to develop a strategy to compete with this narrative. If, on the other hand, such a narrative is not yet commonly accepted, you can attempt to prevent it from becoming mainstreamed.

Understanding the context in which your DSI will operate will enable you to adopt strategies and gain insights that are important for the success of your intervention.

By understanding the context, you can:
- Avoid Motivated Reasoning
- Use Geography-based Planning
- Assess Outside involvement
- Identify Speakers, Partners, and Resources
- Learn from the Negative
- Build on What Exists
- Anticipate Risk
- Understand Pressures Audience Groups Face
- Understand and Identify Proof Points for Different Narratives

Avoid Motivated Reasoning: People tend to accept information that confirms their existing beliefs and feelings, and reject information that contradicts them. This is called "motivated reasoning," and it means that providing people with corrective information often does not work and may even strengthen their original beliefs.[12] This also means that when people receive new information, their existing beliefs and feelings may have more influence over whether they believe or reject this information than rational reasoning.

Confronting information that directly challenges existing beliefs can be psychologically threatening to people, especially if the information challenges their sense of identity.[13] This means that while it is tempting to directly counter speech that promotes group-targeted harm with more accurate information, this strategy is often ineffective and may even backfire.

By understanding relevant history and how people understand and interpret events, the narratives people tell, the beliefs they hold true and the values that are important to their identities, you can develop messaging content that will avoid motivated reasoning. You can also strategize about how to address long-standing beliefs and values that connect to group-targeted harm. Understanding group values is important because framing new information using values a group prioritizes may make individuals in that group more open to changing their minds. You should also avoid framing new information or positive goals in terms of values the group rejects. If you know how a group views itself, you can ensure that your messaging does not go against the group's self-image and can reframe aspects of the group's self-image for peace (e.g., it is actually more heroic to promote an alternative way to resolve grievance than to engage in violence).

Use Geography-based Planning: Understanding the geographic spread and risk of group-targeted harm will help you define the scope of your intervention. You can identify key places to prioritize and strategize about where to intervene and when for the highest impact (and in order to distribute your resources effectively). You can identify locations where your intervention is most likely to be able to operate with stability over the longest time period possible and consider ways to support these locations' resistance to violence.

Insight from the Field: Interfaith Mediation Center (IMC) in Nigeria

When IMC decides where to go to identify and train new speakers, they select communities where conflict has happened or is likely to happen. They prioritize at-risk areas for in-depth workshops and training with religious and traditional leaders.[14] This demonstrates how geographic analysis and planning can help interventions target at-risk locations and ensure that localized programming has an impact.

Assess Outside Involvement: If you are an outsider to the context or receiving outside support, understanding how you are perceived can help you identify the best ways to frame, make public, or not make public, outside involvement. It is important to be careful and thoughtful about international involvement, and it may be best not to include any international branding. Foreign support and perceptions of foreign donors and their motives can create a reputational and security risk for DSIs. In many situations, the target group of dangerous speech is perceived and described to be engaged with outside actors, or to be "foreigners," and conspiracy theories about outside actors (and negative motives for their engagement) may easily gain traction and discredit an initiative, even if the situation is currently friendly. In addition, publicized foreign support can create high financial expectations for partners and participants, jeopardizing your ability to get local buy-in without financial incentives.

Identify Speakers, Partners and Resources: Your contextual understanding can help you identify speakers, partners, and resources. Potential speakers can be drawn from your analysis of different identities and groups and their leaders. By identifying people who are already countering dangerous speech and group-targeted harm, you can develop a list of initial speakers and potential partners and assess their strengths, challenges, and the resources that exist in support of countering dangerous speech. Understanding what motivates these people and groups can help you understand how to motivate new speakers.

Learn from the Negative: If you can figure out how dangerous speech is being successfully spread (through mediums, types of speakers, and types of message content) with specific audiences, you can gain insights into potential strategies (e.g., the use of a specific medium) for your intervention.

Build on What Exists: Your understanding of the general context will let you build on and learn from what already exists. For example, you can use your historical analysis of when groups were more moderate or extreme to identify narratives that accompanied moderation. You can also identify non-conflict related identities and relationships that that transcend conflict-related identities (such as an identity related to a particular town or profession) and build on them to create spaces for inclusiveness and non-violence. By understanding how the groups interact with each other, you can identify examples of peaceful coexistence and interdependencies between groups that can be used to show the value of peaceful coexistence (Note: This may not work if there are grievances related to these dependencies).

You can also build on existing narratives and norms that restrain dangerous speech and group-targeted harm (or promote peace and inclusiveness), and learn from them (for example, how successful they are and with which audiences). You can use your knowledge of the reaction to those narratives and norms (e.g., any backlash they have received) to predict and analyze risk, and to anticipate and prepare responses to rebuttals against narratives that counteract dangerous speech.

Anticipate Risk: Understanding the risks and backlash (if any) that people speaking against dangerous speech or group-targeted harm already face can help you anticipate risks for your intervention. In addition, your analysis of the actors that promote and support dangerous speech and/or group-targeted harm will help you identify spoilers and people who could pose a threat to your intervention. Finally, understanding how norms that promote dangerous speech and/or group-targeted harm are enforced and who promotes them is an important practical consideration that can help you predict and manage risk for your intervention.

Understand Pressures Audience Groups Face: Fixed, sticky (hard to change), and visible identities (e.g., marked by physical features) can increase pressures for members of audience groups to accept, condone or commit violence against a target group, since they can be easily identified and cannot change their affiliation. Since people need to feel acceptance and belonging in a group, if group identities are rigid and homogenous, it will be more difficult for people to keep their group identity while opposing group-wide norms or beliefs. As tensions or conflict escalate, groups tend to become (or seem) more homogenous and extreme. You can try to build on existing diversity and subgroups within audience groups to prevent the full group identity from being defined by violent actors. You can also support sub-groups or people within the group who do not support group-targeted harm. Once a group's identity has become dominated by support for dangerous speech and/or group-targeted harm, those who advocate moderation or a different view of identity may be targeted for harm.

Group norms have a strong influence on group members' actions. Norms that promote prejudice or group-targeted harm can create strong social pressure to conform to such views and even lead to punishment (e.g., social ostracism) for those who speak out against them.[15] Understanding the types of rewards and punishments that your audience groups face will help you understand how they are making decisions. You can also figure out whether broader societal norms can counter or compete with group norm pressures.

Consider Proof Points: Evidence for different claims is important. People act based on the information they have, and based on whether what they see and experience confirms or contradicts that information. When, for example, dangerous speech asserts that the target group poses an existential threat, audience groups will look for confirming evidence (which dangerous speech speakers will often provide).[16] History is a common source of proof points, and a history of intergroup conflict is often cited to support dangerous speech claims.[17] Understanding history can help you identify proof points that can be used to create doubt about dangerous speech claims or to create a competing narrative.

Part 2:

Gathering Contextual Understanding

Understanding the context from a variety of perspectives will enable you to capture a breadth of information on the factors that influence how people think and act. You can consider nine types of context:

- Historical
- Current
- Narrative
- Dangerous Speech
- Identity
- Beliefs and Values
- Actors
- Social Norms
- Geographical

On the following pages are series of questions that can help you understand each type of context.

1. Historical Context

What is the relevant history, and how is it discussed and interpreted by different groups?

❓ What are the main historical events that are relevant to the current dangerous speech and/or conflict?

❓ How are these historical events told and taught? What stories surround them? Do different identity groups tell different versions of the same events? How is blame assigned? What evidence is used? How is the motivation of different historical groups and actors described? Are these historical events and stories discussed in current dangerous speech? What conclusions, if any, are drawn from them?

❓ Is there a history of intergroup violence? If so: What was the level of harm/violence? What stories does each group tell about victims, heroes, and villains? What triggered the violence? What justification did perpetrators give for using violence?

❓ How have the groups historically interacted over time? When have they been more moderate or extreme? Who has led them at each of these times?

❓ What are historical group grievances? How does each group talk about these grievances? What stories are told about how the grievances came to be? What proof is used?

2. Current Context

What is the current situation of concern?

❓ Are there specific events/actors of concern?

❓ Is there currently prejudice? To what degree? Are any groups currently being targeted for economic harm, social harm, harassment, or physical violence? By whom? What form does it take? When does it happen? Who leads it?

❓ What role do rumors play in people's lives? In how people think about targeted groups?

❓ Are there existing structures (such as gangs, community policing groups, militias) that could be mobilized for violence?

3. Narrative Context

What narratives ("collective stories that frame individuals' understanding of the events in the world around them" and that people use to interpret situations and take action)[18] do people use to explain and understand new events?

❓ How are new events interpreted, and what stories are told about them? Which narratives (that are commonly used in society) promote group-targeted prejudice or harm? Do they frame the targeted group as an existential threat? What reasoning or proof is given? Do they use dehumanizing language and/or stereotypes? Do they use stories, myths, or metaphors? How do these narratives justify group-targeted harm? Do they target moderates within the group that is speaking? Who spreads these narratives? Are they effective or ineffective? For which audiences?

❓ Which, if any, existing narratives promote inclusiveness, tolerance, peace and/or non-violence? Who is spreading them? What arguments, examples, metaphors, stories, myths, or facts are used? Are these narratives effective or ineffective? For which audiences? What arguments are being used to rebut or discredit these narratives? By whom? How successful are they? What is their content?

4. Dangerous Speech Context

What do you know about speech that is or may become dangerous?

❓ Is there speech that has the tendency to influence a group of people to accept, condone, or commit harm to members of another group ("dangerous speech")? How frequent or common is it? Is it concentrated or widespread?

❓ Who are the main speakers?

❓ Are the speakers trusted and influential with the audience they are addressing?

❓ Who is the intended audience? How is the audience reacting to the dangerous speech? Are people repeating the dangerous speech that they hear? If so, how?

❓ How is dangerous speech being disseminated (through what mediums)?

❓ What are the main narratives? What examples, stereotypes, or dehumanizing language is used? Are moderates being targeted? Is the targeted group blamed for any problems? Is there a moral implication of who is good or bad? How is group-targeted harm justified? What action, if any, is proposed?

❓ Are there any efforts to stop or counter dangerous speech? Describe these efforts. Who is leading them? Who are the speakers? What mediums are used? What strategies are used?

❓ How has the audience reacted to these counter efforts? Have there been any negative consequences? (e.g., censorship, punishment, propaganda against the speakers)?

❓ What is the role of state authorities? Are they spreading, condoning, or facilitating dangerous speech? Punishing or speaking out against dangerous speech?

5. Identity Context

How does the group that dangerous speech addresses (the audience group) and those that it targets for harm (the target group) define themselves and each other? What other identities exist?

❓ What are the main identities of the audience and target groups? Are they visible (easy for others to identity)? Are they "sticky" (they cannot be easily changed)? How did they come to be? How are members of the different groups identified (e.g. through physical markers, political affiliation)? How rigid are the group identities (e.g., How easily and often do aspects of the group identity change? How much variation is there within each group? Are differences within the group accepted or discouraged?)?

❓ How does the audience to dangerous speech view the targeted group? Are there stereotypes?

❓ Does either group view the other as a threat (e.g., to the physical survival of its members, or to its continued existence as a separate group)? What information is used to support this argument?

❓ What other identities do members of each group have (e.g., members of an ethnic group may also have identities based on factors such as their religion, livelihood, region of residence, etc.)?

❓ What are current interactions like between the groups? How often do they interact and in what situations? Are they physically separated? Do they depend on each other? If so, how do they feel about that? Are there ties (business, social, religious) between groups? How strong and frequent are such ties? Are they positive or negative? Do they have any shared interests? Are there tense or conflict-related interactions? What is the level of trust or distrust between the groups?

❓ Within each group, are there people who are more moderate or extreme? What is the range of viewpoints about the current dangerous speech situation?

6. Beliefs and Values Context

What are the main values, beliefs, and belief systems that the audience and target groups of dangerous speech claim to hold? How do each of these relate to narratives that promote or counter dangerous speech?

❓ Which values does each group prioritize or find important? What reasons are given for why these values are the most important?

❓ What pressures are there to act in line with these values, if any? Are people judged based on these values (i.e., people are seen as good/bad based on whether they embody the values)?

❓ Which values does each group reject or look down on? What reasons are given for why these values are not good? Who do they think holds these values?

❓ Does either group have a set belief system (e.g., a religion, political ideology)? Who are leaders within each group's belief system? Describe.

❓ Which beliefs or values, if any, is the dangerous speech using to justify prejudice and group-targeted harm?

❓ How is violence viewed within each group's value and belief system? What beliefs and values are associated with violence, if any (e.g., honor, shame, vengeance, justice)?

❓ Which beliefs or values, if any, are used to call for inclusiveness, tolerance, or peace/nonviolence? To speak against prejudice and group-targeted harm? How are inclusiveness, tolerance, and peace/non-violence viewed within each group's value and belief system? What beliefs and values are associated with inclusiveness, reconciliation, and peace/non-violence, if any (e.g., forgiveness)?

7. Actors Context

Who are the relevant actors?

❓ Who are leaders in promoting dangerous speech or organizing discrimination or group-targeted harm? What capacities do they have (in terms of influence, access to political power, capacity to mobilize violence)? Who has more or less power within this group of actors?

❓ Are you able to identify which of these actors is more moderate or more extreme?

❓ Who offers support for group-targeted harm and/or dangerous speech? What type of support are they offering? What motivates these people to offer support?

❓ What is the leadership structure of the audience for dangerous speech? Who plays leadership roles at the local, regional, and national levels? Are these leaders promoting, neutral, or against group-targeted harm, discrimination, and dangerous speech?

❓ Are there people who have already begun to speak or act out against dangerous speech and discrimination/group-targeted harm? What capacities do they have? Are they facing any backlash, risks, or consequences for their actions?

8. Social Norms Context

What existing social norms promote or counter dangerous speech and/or group-targeted harm?

❓ Are there social norms that promote, encourage, or normalize dangerous speech, prejudice and/or group-targeted harm? What are they? Which groups do they apply to? Who, if anyone, promotes or enforces these norms?

❓ Are there social norms that promote inclusiveness and tolerance or discourage dangerous speech, prejudice and/or group-targeted harm? What are they? Which groups do they apply to? Who, if anyone, promotes or enforces these norms?

9. Geographical Context

What is the geographic spread of tensions and conflict that fuel dangerous speech?

❓ What is the overall geographic scope of potential group-targeted harm?

❓ Are there areas that are potential conflict flashpoints, where group-targeted harm is likely to start? What are these places and what are their characteristics?

❓ Are there areas that are likely to be the first places group-targeted harm spreads after the initial flashpoints? What are these places and what are their characteristics? What are likely to be the second, third, fourth places group-targeted harm spreads after the initial flashpoints? What do you predict the overall spread of conflict and group-targeted harm might look like?

❓ Are there places that are likely to be hot spots for group-targeted harm (where there is likely to be a high level of intensity to the violence)?

❓ Are there areas that are less prone to violence and may remain stable and/or resist violence? What are these places and what are their characteristics?

❓ Are there places that are likely to become bases for violent actors or for counter-violence or defensive tactics (e.g., by the targeted group)? Where the targeted group might relocate in the event of persecution?

Outside Actors

If you are an outside actor or are receiving outside support, it will be important to analyze how outside actors are perceived and discussed.

Part 3:

Predicting Conflict Trajectories

Because you are designing an intervention to reduce the risk of group-targeted violence, it is important to predict how you believe current dangerous speech could connect to such violence. This means mapping scenarios and conflict trajectories – thinking about the series of events that might occur in the location of your intervention that could lead to an escalation of dangerous speech or to group-targeted violence. For example, if there is a rumor that some members of the group targeted by dangerous speech are heavily armed, that one of them raped a woman from "our group," or that they are raising market prices unfairly, what are the specific potential responses that could lead to group-targeted violence? What are the events and actions that led to this rumor, and to its ability to spread? You can map out multiple trajectories and pick the trajectories you wish to focus on based on their likelihood and degree of harm.
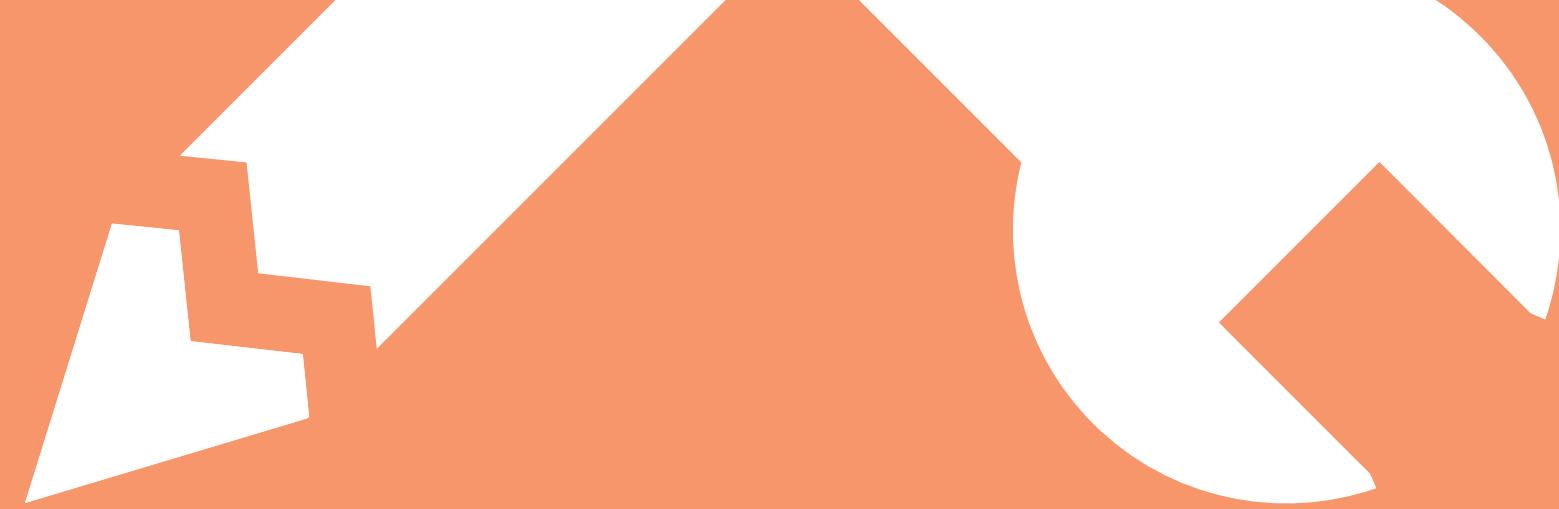
Your conflict trajectory is your explanation of the situation you are trying to prevent. Using this conflict trajectory, you can map out overall intervention goals based on how you would like to change the trajectory.

Based on your predicted conflict trajectories, you can identify potential triggers and warning signs, and use these to inform any ongoing monitoring that you will conduct.

Toolbox

You can use the following research tools and techniques to gather contextual information. As you do so, keep in mind that all individuals come with a set of biases. Try not to depend too heavily on a single person, group, or network to identify participants, partners, or interviewees. Instead, identify a wide variety of people, groups, and networks to consult and study in order to develop a full understanding of the context as well as the social networks and relationships between different groups.

Participatory Analysis: You can often get the best information by learning directly from people on the ground who bring different perspectives and biases. You can generate a wealth of knowledge and a baseline understanding of contextual factors by bringing together individuals with diverse backgrounds and information about groups and places and conducting a participatory analysis (e.g., through a workshop or series of workshops). People from relevant communities (for example from specific villages or demographic groups)have information outsiders cannot obtain. Even if you are working in your own country or area, you can use this approach to involve multiple perspectives (from different identity groups, different local communities/areas, genders or professions) and get buy-in from partners in your planning and research process.

If you're in a hurry, a participatory analysis can be a quick way to produce a foundation of information and analysis that can be supplemented with other research techniques. You can ask the participants to identify areas for further research, and work with them to identify the most promising follow-up research tools and methods. You can also ask participants to come up with questions they think are missing but are relevant.

If you are unable to conduct a participatory analysis, you can do a series of interviews with individuals and partners. While they might not have thought about this specific set of questions, it is likely that a group of participants or interviewees will have opinions about the questions asked by each type of context analysis, and that they have relevant knowledge from their own experience.

Historical reading and literature reviews: Many types of written sources can help you get an initial understanding of a context. In addition to books, academic articles, and reports by relevant organizations, you can identify influential writers and news sources (including those with strong biases) and monitor what they are saying. This can help you understand how groups talk about events and issues. Op-ed and comments sections of news articles are promising places to find information about how people are discussing events. You can also review textbooks to understand how different groups are teaching history, especially with regard to events you've identified as relevant.

Key informant interviews: Key informant interviews include interviews with experts and with people on both sides of the conflict. This may include local

leaders (e.g. religious and traditional leaders); local community members; and hardliners on both sides. These interviews can help you understand how people are making decisions, especially the reasoning behind their conflict-related decision-making and how they view their group and other groups. You can also observe how people talk about current and historical events. The interviews will require being non-judgmental and making people feel comfortable so that they can share their thoughts, feelings, opinions, and stories and you can learn as much as possible about their perspectives. The point of these interviews is not to convince, change, or teach any of the participants.

Observational Research: You can observe people's interactions and conversations to learn about: how people interpret and make sense of events; where they get their information; which mediums they access throughout the day; who the key community influencers are; how dangerous speech spreads throughout the community; whether there are attempts to counter dangerous speech; and how such attempts are received, among other things. You may naturally be observing these things, but you can structure your (and your partners') observations by coming up with some hypothesis

or specific questions for observation. For example, you could ask "how are people talking about X event if at all" and observe conversation content, location, who is involved, what sources are mentioned, and whether people talk openly or whisper. You can observe different spaces, such as specific local communities or conflict hot-spots, and online spaces such as Facebook, YouTube, or Twitter. This type of research can turn into monitoring of local events, rumors, and information.[19]

For online observational research, you can join, follow, and watch different groups, individuals, hashtags, etc. For in-person observational research, there are two approaches you can take:

1. Work with a team of people from the community (or communities) you are observing, and plan where, how, and what (e.g., questions or events) the team will observe. The team could observe specific people, places (e.g., a bar, the marketplace), or events (e.g., church services). People from the communities they are in are more likely to pick up on nuances and to blend in and make people feel comfortable. If someone is a known peace activist, community members may not feel comfortable expressing

discrimination or conflict preferences in front of him or her, so finding someone who is not associated directly with violence prevention could be valuable. Someone who clearly belongs to a particular group will likely get more candid responses from members of that same group than from members of a different group. This applies for various characteristics, such as gender, occupation, age. You can build a team of researchers based on their ability to gain access to specific demographic groups.

2. You can identify specific people from the communities you are observing and ask to shadow them. You could explain your intervention or give another reason why you'd like to shadow them based on what will make them feel free to go about their daily life. If you are not from the community, people may feel comfortable sharing certain things with you as an outsider that they would not share with someone they perceive as an insider.

You can use different tools (e.g., a recording device, note-taking, video, photography) to record information if you are able to explain the purpose of your observation and get informed consent. You may also wait to take notes in private after the observation. Your choice of tools should aim to make the people you're observing comfortable.

Analysis of Frames: You can pay attention to how people frame an event or issue -- that is, how they choose to emphasize or ignore certain facts in order to promote a specific interpretation of the problem, cause, and solution.[20] For example, whether or not to allow a hate group to hold a rally could be framed as a free speech issue or a security issue.[21] Frames define the problem, propose a reason why the problem exists, make a judgment about what is causing the problem and those being harmed, and provide recommendations for how to treat the problem. For example, an economic downturn may be blamed upon: a) weather that is considered unusual; b) leadership that is considered to be corrupt; c) a group that is viewed as undermining the economy. These frames would lead to different conclusions about what action should be taken. In most societies, certain frames are used repeatedly and taken as truth. These frames can impact what people think should be done and what they do (e.g., someone who supports freedom of speech might oppose a rally by a group that is seen as threatening security).[22] Frames can provide clues about how people understand problems, their causes, and potential solutions. Because

framing impacts how people understand situations and form opinions,[23] changes in frames that support or counter dangerous speech and group-targeted harm may be important to watch.

Social and Traditional Media Analysis: You can analyze and monitor social and traditional media, paying attention to dangerous speech, efforts to counter dangerous speech, how people react to both types of speech, narratives around events, and how information spreads. One way to think about media analysis and monitoring is as 'social listening': paying attention to which narratives are being told, specific words, tone, and metaphors that are used, how different groups are talking about themselves and each other, and which, if any, calls for action are being used.[24] You can monitor statistics, specific speakers (e.g., Facebook pages, Twitter handles, public statements, etc.), or follow keywords.[25] If you can, identify if and how online discussions are interacting with offline interactions to see how these conversations impact behavior.

Analysis of Group Documents and Public Statements: You can read and analyze relevant documents (e.g., manifestos, papers, pamphlets) that are produced by specific groups. You can also observe, document, and analyze public statements about the targeted group.

Investigative Research: You can use investigative research for things that you can't examine through open sources (e.g., how dangerous speech promoters are coordinating, who if anyone they report to, etc.). This can be risky, and it is best to partner with people who are already doing these types of investigations to access relevant information. If you have trusted "inside" sources, you can also work with them to get additional information.

Creating Feedback Loops: The context (especially the identity, narrative, actors, and dangerous speech contexts), will change over time. If you create feedback loops to learn about changes in the context (e.g., through monitoring and observation), you can use the information to adapt your intervention with the changing context. If monitoring efforts (for example, for conflict early warning and response) already exist, you can consider the value in collaborating with these efforts. If you are working with a team or network of people, you can develop a set of variables (e.g., conversations about specific types of events, rumors, etc.) to monitor, and create a strategy for how you

will track and observe each variable over time. You can figure out which type of monitoring best fits your capacity. For example, your team can observe and report on the situation in their communities or take on tasks such as calling people to find out what is happening, watching Facebook or Twitter feeds, or listening to radio stations. If you have limited resources, you can identify a small number of locations that are representative of different characteristics (e.g., hot spots/flash points) and monitor them.

Phase 1
Understand Context and Conflict
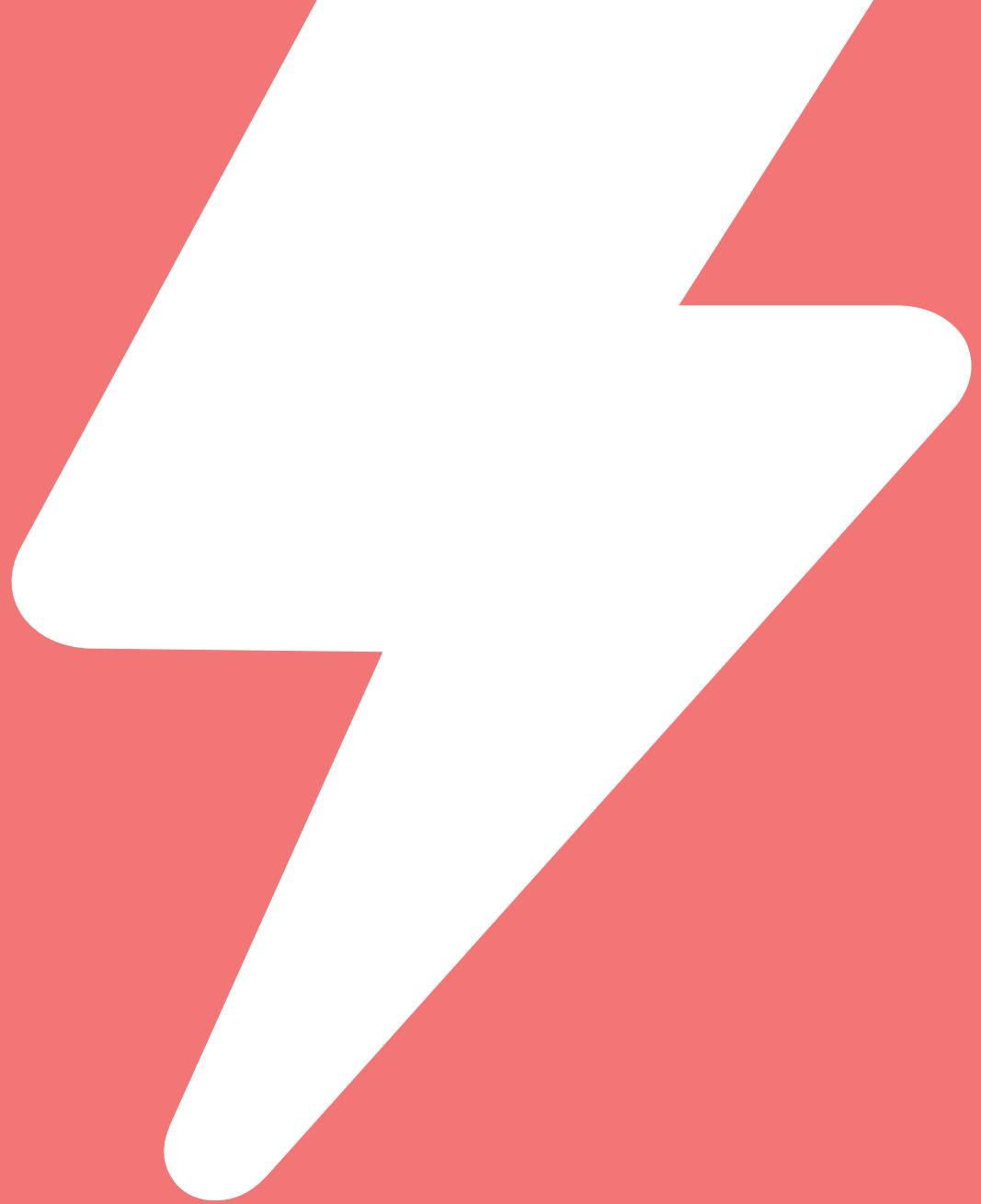
Phase 2
Select and Analyze Audiences

Phase 3
Select and Design Mediums,
Speakers & Message Content

The contextual understanding provides critical information for selecting and understanding the audience that your intervention seeks to influence. The audience analysis enables you to identify distinct groups within the audience and to assess how to influence each group. Then, you can identify which groups you can reach with the greatest impact, develop an understanding of the people in each of these groups, and set audience-specific goals for your intervention.

Part 1:

Selecting and Understanding the Audience

Within each type of audience your intervention seeks to influence (e.g., people who live in a particular area, youth, members of a religious group), people play distinct roles with respect to dangerous speech, DSIs, and group-targeted harm. Identifying these distinct groups within your audience based upon their behavior with respect to dangerous speech will help you set specific goals for influencing them. You can do this by segmenting the audience into groups and charting each group's attitudes and involvement related to dangerous speech.

Once you know the groups within your audience that you want to target and how you want to change their behavior and engagement with dangerous speech, it is important to develop a detailed understanding of the members of each group. This should be an audience-centered and empathy-based understanding: how do the people in that group think, feel, act, reason? You can do this by creating audience profiles, or characters, that represent each target group.

Step 1

Segmenting the Audience

You can break the dangerous speech audience into groups based on roles they play in dangerous speech, analyze each group based on its characteristics, and target each group with specific behavioral goals (goals for how you want to change their behavior). Key relevant categories are listed below.

Note: Women are often overlooked in violence prevention efforts. Remember that each of these categories may include women, who often play critical roles in speech that leads to violence.[26]

Influential Individuals have a high level of influence over an audience (a specific group of people), whether at the local, national, or identity group level. They may have varying relationships to dangerous speech, but play an important role even if they do not take a stand (they have an impact even by disengaging). Types of people who may fall in this category include local opinion leaders, religious leaders, national or local celebrities, and politicians. Influential Individuals can be effective DSI speakers, can endorse and lend credibility to DSIs, and can encourage those within their influence to engage. If they are currently having a negative impact, an intervention can aim to reduce their level of engagement with dangerous speech.

Information Spreaders play important roles in the spread of information. They may have varying relationships to dangerous speech, but play an important role regardless by the information they choose to pass (or not pass) on. Examples include vendors at a marketplace or people in the transport industry who become key sources of gossip and information in the community, and members of the media. DSIs can aim to influence the types of information that information spreaders do and do not pass on. For example, one goal could be to get information spreaders to question the motives of violent propaganda or rumors. Another goal could be to engage information spreaders in spreading positive messages.

Reluctant Audience Members listen to or witness dangerous speech and/or group-targeted harm, but do so reluctantly. This could be any member of society and this group can be targeted with a variety of goals. For example, DSIs can seek to influence reluctant audience members to engage in positive behavior; speak up in uncomfortable situations; leave situations when they are witness to dangerous speech and/or group-targeted harm; or document and report dangerous speech/group-targeted harm.

Engaged Audience Members are receptive to the messages of dangerous speech and to condoning group-targeted harm, but are not hardliners. For example, they may be easily influenced by charismatic leaders who promise to resolve their grievances, or be receptive to blame narratives. This could be any member of society, but certain types of people (based on demographic or other characteristics) may be disproportionately engaged. DSIs might attempt to prevent this group from spreading dangerous speech or participating in group-targeted harm or prompt them to leave situations when dangerous speech or group-targeted harm begins to take place.

People Who Encourage Participation in Dangerous Speech/Group-Targeted Harm spread dangerous speech and encourage action. This could be any member of society, though certain types of people may be disproportionately engaged. DSI's can attempt to reduce participation (e.g., aim to get them to go home when a situation escalates), encourage this group to speak up against group-targeted harm in escalating situations, or provide them with support to raise questions about the utility or risk of participating in group-targeted harm.

People Likely to Participate in Group-Targeted Harm (Reluctantly) participate or are likely to participate in the future, but do so reluctantly or with reservations. This could be any member of society, and may disproportionately include young men. DSIs can attempt to reduce participation (e.g., aim to get them to go home when a situation escalates) or to encourage this group to speak up against group-targeted harm in escalating situations.

People Likely to Participate in Group-Targeted Harm (Willingly) are likely to participate willingly or very willingly, with few if any reservations. This could be any member of society and may include unemployed young men, criminal gang members, and political hardliners. This group may be hard to reach, and DSIs may try to get them to engage in alternative ways of resolving their grievances or prevent them from participating in violence at specific times. These goals might change over time depending on actual levels of collective violence.

Note: Individuals can move between categories over time. When you think about setting goals, you can think about the steps it would take to move someone from their category to the category where you would like them to be. This will help you set realistic goals.

People (Audience Members) Who Spread Dangerous Speech can be any member of society, though some types of people may be disproportionately represented. DSI goals may include getting them to stop repeating dangerous speech or to do so less often. This could mean prompting them to talk about something else instead, or to discuss grievances without drawing a conclusion that group-targeted harm is the answer.

People (Audience Members) Who Counter Dangerous Speech could be any member of society, and some types of people may be disproportionately represented. The goal for this group is almost always to support them in their efforts and ensure that they continue this positive behavior. Additional goals may center on increasing their visibility and audience, for example by getting them to use new mediums, talk to additional people, or recruit other speakers.

Step 2
Charting Attitudes & Involvement

One helpful way to set realistic goals is to think about attitudes and involvement. What is each group's attitude towards DSIs (positive, negative, or neutral)? How involved or impactful are they in spreading speech and information in general? Low involvement means they aren't very involved with spreading information and communication; high involvement means that they are very involved.[27]
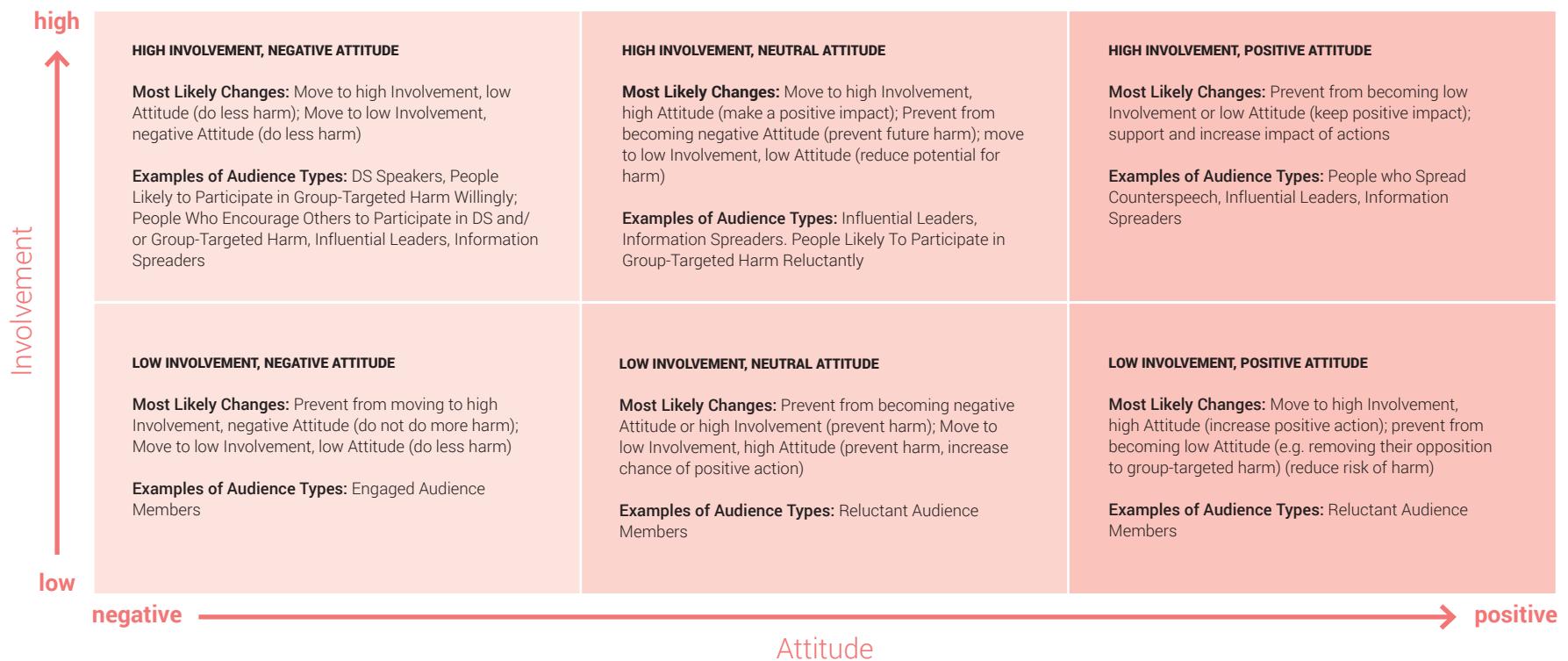
Helpful Hint: One way to think about attitudes and involvement is using the metaphor of Coca Cola and Pepsi Cola competing for the largest share of the soft drink market. We can say that people who buy a lot of soda are high involvement and people who don't buy much soda are low involvement. If we are Pepsi, we can say that people who like Pepsi have a positive attitude, people who don't have a strong preference for Coke or Pepsi have a neutral attitude, and people who like Coke have a negative attitude (towards Pepsi). To get the majority of the market share, we want to target each of these types of consumers differently. We might target consumers who are high involvement but have a negative attitude (buy a lot of soda but love Coke) to get them to drink less soda (become low involvement). For consumers with positive attitudes but low involvement (they like Pepsi but don't drink much soda), we might try to get them to drink more soda (become high involvement). For consumers who drink a lot of soda but don't really like Coke or Pepsi, we could try to: get them to like Pepsi (have a positive attitude); drink less soda (become low involvement), or prevent them from drinking Coke (getting a negative attitude). If they love Pepsi and drink a lot of soda (they are high involvement and have a positive attitude), we could try to get them to drink even more Pepsi or get their friends to drink Pepsi.
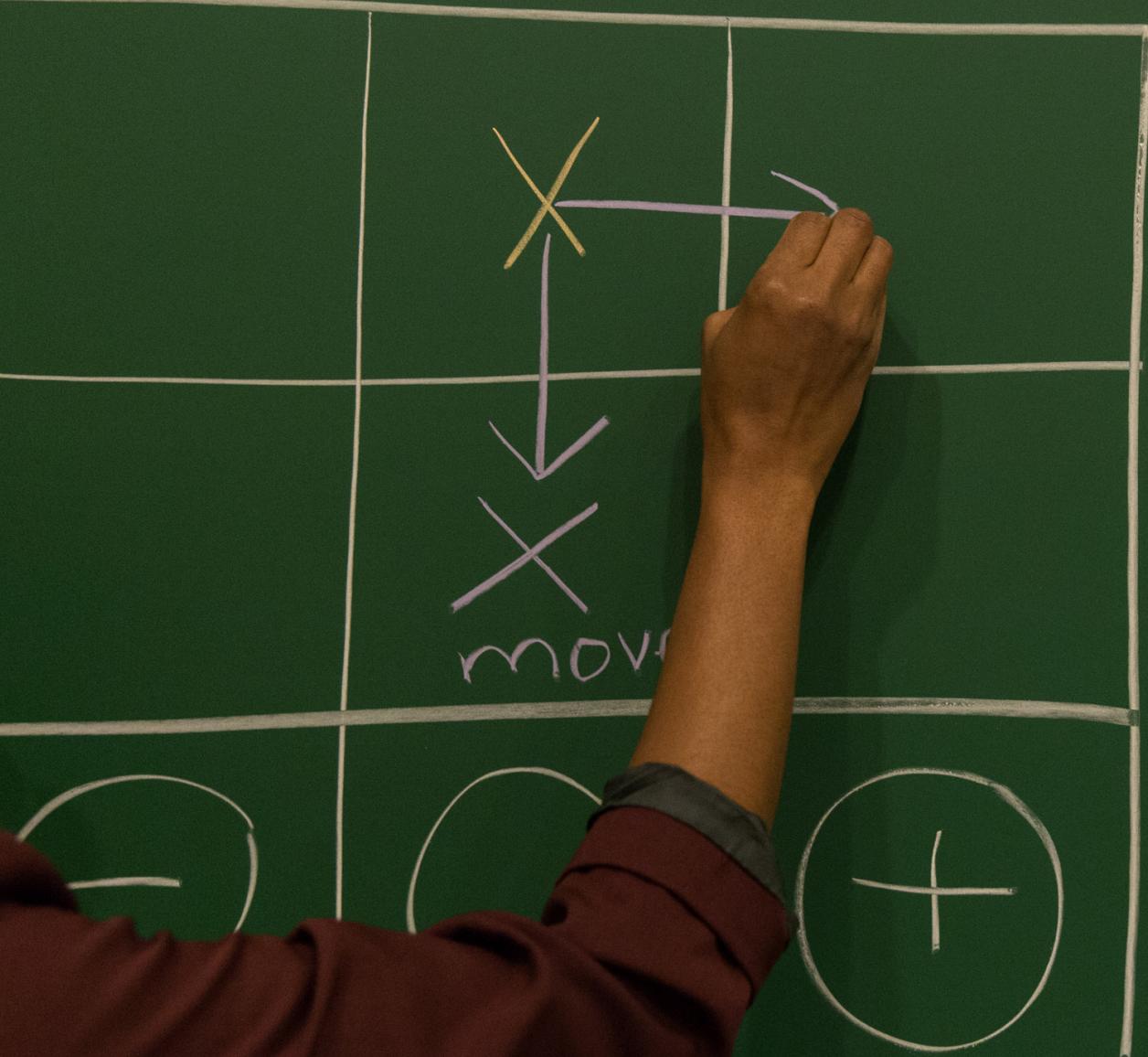
The following chart illustrates the spectrum of attitudes and involvement. This chart can help you see who has the biggest impact on dangerous speech and countering dangerous speech, which groups you can most realistically target, and for which goals. You can more easily move someone to a box next to the one they are in than to a box on the other side of the chart, and you can use this chart to plan a series of realistic goals to change people's attitudes and involvement.

Remember: People's attitudes and involvement aren't static: they may take on different roles and move throughout the chart over the course of the conflict trajectory. Your intervention can aim to prevent people from becoming more negatively involved and/or aim to increase their positive involvement.

Attitude & Involvement Chart

high ↑

Involvement

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

low

negative →→→→→→→ positive

Attitude

Attitudes and Involvement Chart in action

Step 3

Creating Audience Profiles

⚡ See Workbook 2, pp.13-21

To influence audience groups, it is important to understand the world from their perspective. One helpful way to do this is by developing a Profile (or composite identity) for each group you want to target.[28]

These Profiles should be developed from an audience perspective and should become characters that represent each audience group you aim to influence. You may decide to create multiple Profiles for each group -- you want to understand it fully, but keep the number manageable. You can give each Profile a name, and reference this person throughout the design process.

There are several types of information that a Profile should cover. First, it should give basic demographic information about the person, and give some general information about the person's life. Information about where and how each Profile accesses information can help you identify mediums that will reach the audience group, while information about whom they interact with, trust, and look to for guidance can give you clues about which speakers are likely to be credible, relevant and influential. Information from your contextual understanding about how they view their group and other identity groups can be important to identify speakers and for message content development. Finally, information about how they experience the current dangerous speech situation (their thoughts, feelings, and attitudes), as well as about what motivates or constrains their behavior can help you develop high-impact message content.

Part 2:

Setting Clear,
Audience-Specific Goals

Building on your context and audience analysis and your predicted conflict trajectories, you can set clear goals for how you want to change the behavior of each of your audience groups over time.

To do this, you can predict how each audience group will behave at different points on a trajectory and then set a specific goal for how you would like each group to behave instead at that point.

For example, your overall trajectory may predict that currently occurring dangerous speech causes people to believe that Group X is a threat, and take negative rumors about Group X seriously. Thus, any rumor of a wrong-doing by a member of Group X will be spread quickly and cause people to mobilize against Group X. If you are targeting information spreaders, you might predict that they will believe the rumor, spread the rumor, and urge people to take action. You might prefer that they question the rumor, do not spread it, and urge people to wait to find out more information before taking action. Your specific goals would be that: (1) information spreaders question rumors when they hear them; (2) refrain from spreading these rumors; and (3) urge people not to take action based on rumors. You can get even more specific by adding types of rumors, locations, etc. These goals describe how you want your audience member to move from the predicted behavior (on your conflict trajectory) to a preferred behavior.

Your intervention will be most effective if it is based on a clear understanding of how your audience interacts with each of these goals. What might motivate or prevent them from behaving in line with the conflict trajectory or your specific? Here are two approaches for analyzing how your audience groups interact with your goals:

1: Behavioral Drivers & Barriers: At each point in the conflict trajectory, you can think about what barriers prevent people from moving from the conflict status to the preferred status.[29] These behavioral barriers may be things like, "I'm afraid that if I do that, people won't like me," in which case the barrier would be social ostracism. You can then think about behavioral drivers, or things that would drive the person to move towards the preferred status. Finally, you can think about barriers that would prevent people from moving towards the conflict status and drivers that are moving people towards the conflict status. Based on this understanding, you can develop strategies that: (1) reduce barriers that are preventing people from moving towards the preferred status; (2) increase drivers (things that motivate people towards a behavior) for the preferred status; (3) increase barriers to prevent people from moving towards the conflict status and (4) decrease any drivers that are motivating people towards the conflict status.

2: Profile of the Moment: In addition to analyzing behavioral drivers and barriers, you can develop an understanding of the audience group's experience of each moment or point in time on the conflict trajectory. For example, consider what they are doing, who they are with, who is influencing their decision, what mediums and information they have access to, how they feel, and who else (e.g., a child) they are considering at the moment.

Concepts and Approaches for Influencing Behavior

For each audience group, you should ultimately aim to influence their behavior, both general behaviors (whether or not they condone or engage in group-targeted harm), and more specific behaviors based on the specific goals you have created (e.g., wearing a badge, attending or not attending a rally, not spreading a rumor, or protecting a neighbor from harm). There are many theories about how to influence people's behavior and what drives behavior. Do people behave based on what they believe? What they think everyone else is doing? How they feel? What they know? Their attitudes? The following pages outline some important factors to consider when you develop your own theory and approach for influencing behavior.

1: Social Norms: People may choose how to behave based more on social norms than on their own beliefs.[30] Social norms are informal rules that regulate how most people in a social group think and behave. They can be understood as what people within a social group think is acceptable and unacceptable, or normal and abnormal, in terms of behaviors and attitudes.[31] Social norms can describe what people actually do (e.g., "most people get married before 30") or what people approve/disapprove of (e.g., "most people approve of marriage outside of your identity group"), but perceived social norms – what people think that most people in their group do and do or don't approve of, may be just as powerful an influence on people's behavior. Social norms exert peer pressure, in that group members must obey (or think they must obey) these unwritten rules in order to fit in or be accepted by their group.

Social norms can have a significant impact on how people behave, either because they actually accept and internalize the norms, or because they fear the consequences if they do not conform to the norms (e.g., social pressure, ostracism, and even violence).[32] People may change their behavior to be in line with the norm even if it contradicts their beliefs, values, or how they want to behave. Interventions may be able to change dangerous speech-related social norms and behaviors without actually changing people's beliefs.[33]

Because of the power of social norms, purveyors of dangerous speech seek to promote and enforce social norms that are conducive to group-targeted harm. For example, in Rwanda in 1994, group norms about what it meant to be a Hutu, along with the threat of punishment for not acting within those norms, were used to pressure Hutus to engage in violence.[34]

Similarly, DSIs can use the power of social norms to positively influence people's behavior. Interventions that have used social norms to change behavior have been successful in areas ranging from recycling and energy use, to drinking alcohol and driving, to tax payment, to bullying. Learning from these approaches, DSIs can attempt to do several things:

> • **Influence perceived norms (how people think that others within their social group think or behave)**. For example, "I think that most of my peers drink alcohol" is a perceived social norm, but may not describe what other people are actually doing.[35] People tend to overestimate how much other people are engaging in negative behaviors, so interventions can try to correct these misperceptions. In areas at risk for mass atrocity, purveyors of dangerous speech may promote or reinforce social norms that are conducive to group-targeted harm and violence by creating the impression that most people support such norms, even if they don't. DSIs can help combat this risk by correcting and changing perceived social norms. For example, when group-targeted harm is being promoted, a perceived norm could be "I think everyone else believes that Group A is a threat and we should get rid of them, and plans to/is spreading the message." By demonstrating that in fact 80% of people do not want to get rid of Group A and are not spreading the message, a DSI can undermine the perceived norm.

> • **Influence injunctive norms (how much people think other people approve or disapprove of a behavior).**[36] Dangerous speech can be used to convey that certain behaviors and beliefs are or are not acceptable for a specific group (e.g., "all real Group A's think it's okay to discriminate against Group B"). DSIs can aim to change people's perceptions of what others approve or disapprove of (e.g., "most people in Group A think it's not okay to discriminate against Group B").

Limitations: Changing perceived norms may alter people's behavior, such as by giving someone the confidence to speak peacefully or by pressuring someone not to speak dangerously even if they want to; it is unlikely, however, to alter underlying attitudes or beliefs (such as prejudice). These underlying attitudes or beliefs may continue to make people receptive to dangerous speech and to embracing norms that are conducive to group-targeted harm.[37]

2: Attitudes: An intervention can target people's attitudes towards their own group and towards other groups (for example, towards the group being targeted for harm). This type of attitude can be explicit (people express their attitude openly and intentionally) or implicit (people may not even be aware of their attitude towards another group, for example they may subconsciously
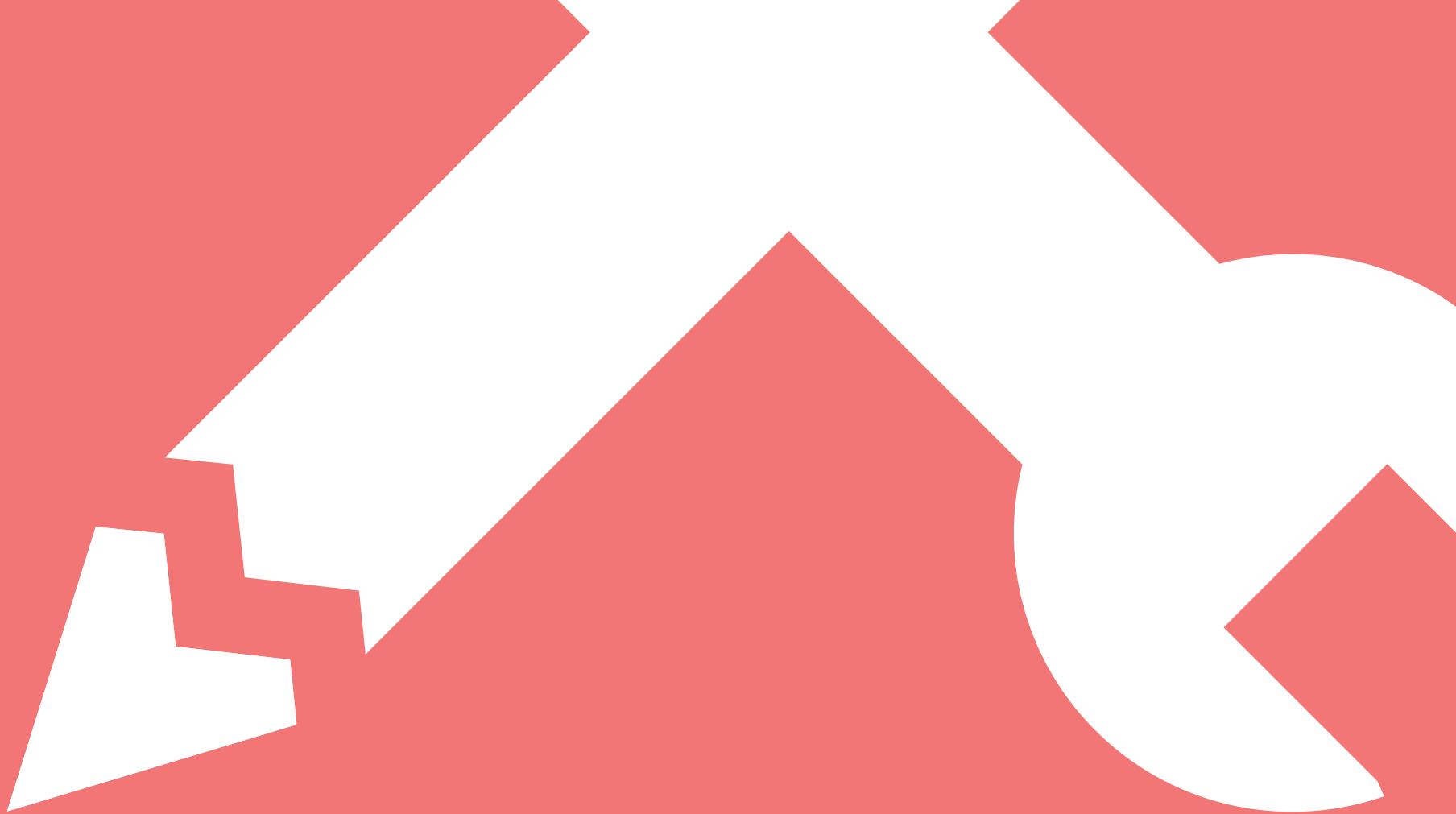
be prejudiced without ever intentionally expressing that attitude). Changing social norms can help change whether or not people explicitly express negative attitudes about a targeted group, but it's more likely that longer-term changes in people's experiences and beliefs can change their implicit attitudes.

You can also target people's attitudes towards specific behaviors (such as spreading a rumor or speaking out publicly) by targeting how people feel about the behavior (for example, if they are nervous to speak publicly or have fun spreading rumors), and what they think will happen if they do the behavior.[38] For example, you could try to move someone from thinking: "If I speak publicly everyone will make fun of me" to: "If I speak publicly everyone will respect me" to make their attitude towards speaking out publicly more positive, or you could try to move someone from thinking "If I spread rumors, everyone will see me as important" to "If I spread rumors, people will stop trusting me," to make their attitude towards spreading rumors less positive.

3: Beliefs: There is a lot of evidence that it's very hard to influence and change people's beliefs. People's experiences have more power to shape their beliefs than information they hear; in marketing, it is understood that experience "beats the message."[39] Beliefs that people hold strongly and that relate to their identity are particularly hard to change. These beliefs are subject to "motivated reasoning," when people decide what new information to accept or reject based on whether it supports what they already feel or believe to be true.[40] This makes challenging beliefs particularly difficult in situations with inter-group tension. In addition, beliefs themselves may or may not influence behavior. Beliefs may sometimes influence behavior (for example, stereotypes and beliefs about another group), but a person may also choose how to behave based something other than a belief, such as a social norm, a religious tenet, or a law.[41]

4: Emotions: People often react to situations based on how they feel, not based on rational reasoning. People often choose what to believe based on how they feel about a situation or about information. For example, fear of another group may influence how someone decides to act during a tense or violent event, or emotions towards a speaker may influence how someone

interprets what he or she says.[42] In addition, emotional responses to conflict triggers may make it difficult for people to use rational thinking when they are deciding how to act. Creating a break between people's initial emotional reaction and their decision about how to act can give them space to "cool down" and use reasoning rather than only emotions in their decision.[43]

Toolbox

You can draw on your context analysis for initial audience insights, and you can use many of the tools from the previous section for your Audience Analysis research.

Participatory Analysis: Just like for the context and conflict research, you can bring together partners or relevant participants to engage in the audience analysis process. If you are working with diverse partners, they are likely to be familiar with a variety of audience groups. They can help you fill in initial information and identify areas for further research.

Observational Research: Using the same observational research techniques, you can focus on questions and hypotheses from the audience analysis and profiles. This can be particularly helpful for observing what mediums people access, whom they interact with often, their daily routines, and their overall information network.

Key Informant Interviews: You can conduct interviews with individual audience group members (for each group the intervention aims to influence) about their lives and thoughts to compile the information needed to develop Profiles. A variety of interviews with different group members can guarantee that you capture different viewpoints and information. When you do these interviews, setting the interviewees at ease (e.g., starting by talking about things that interest them) will help you get honest feedback: the more you can get them to

let their guard down the better. This may mean simply getting to know group members and talking to them rather than setting up more formal interviews, or it may mean a combination of both. You can also pay attention to differences in how people respond to questions when different people are around – this can provide important clues about how the person wants to be seen by others.

Social and Traditional Media Analysis: Observing specific profiles and people online (and through general "social listening") can help you learn about how people discuss values and current events, and how they judge each other. This can also help you identify social networks and groups.

Active Monitoring and Observation: You can develop plans to observe how audience group members change over time and adjust your strategy accordingly.

Focus Groups: Focus groups are a way to bring people (usually 4-8 people)[44] together for a group discussion. Focus groups generally ask participants to give feedback or discuss specific topics, questions, or materials. These topics, questions, and materials should prompt people to think and talk about the

questions and topics outlined in the persona template, so that you can observe group conversations instead of using the traditional question-answer interview format.

You can recruit many participants from an audience group that your intervention seeks to influence, or recruit a spread of participants from different audience groups. You can select the types of people you want (e.g., 5 major information spreaders), and then identify people who fit the description and recruit them. When you recruit participants, you may not want to explicitly state your purpose (e.g., you can say you want to discuss the economy or learn more about the community). If people are difficult to recruit (e.g., people likely to participate in violence), you can ask them to talk about topics that you know interest them, then add some questions, materials, or exercises that will push the conversation towards your areas of interest.

Friendship Groups: Friendship groups are like focus groups, but instead of identifying a range of participants, you bring together small groups of friends and family (2-5 people) and observe them discussing relevant topics.[45] Just like a focus group, you facilitate the conversation by asking questions or providing prompts for the group to discuss. For example, you could ask them to discuss specific beliefs and values; their beliefs about their group or the target group; who is influential, trustworthy and credible, and which mediums they use to access information. Friendship groups can be created for specific audience groups by identifying a member of that audience group (for example, a reluctant audience member), and asking him/her to invite some friends or family members. You can interview each participant separately then bring them together for a joint discussion. Participants should not feel judged by any of the people facilitating the process. The goal is to make people feel safe and comfortable so that they can share their interactions and stories. As with focus groups, one challenge with developing friendship groups is that you may need to mask the topics (and questions) in order to convince people to participate and to manage risk for the participants and the facilitators.

Phase 1

Understand Context and Conflict

Phase 2

Select and Analyze Audiences

Phase 3

Select and Design Mediums,
Speakers & Message Content

1: Mediums

The mediums (channels of communication) you use will determine whether your message content reaches your audience, how often, and in which situations. The goal is to find a medium or combination of mediums that can reach as many of your audience members as possible.

Identifying MediumsThat Will Reach the Audience

When you choose mediums, you should aim to identify a medium or combination of mediums that can reach each audience group you want to influence. You can also consider how frequently and when the mediums need to reach your audience in order to influence them and accomplish the goals you have set. You can begin by building on your audience analysis (specifically, the Profiles) to create an initial list of all of the mediums that you think can reach your audience groups. You can supplement this list with any additional mediums you can think of (for example, a medium that is not widely used yet but that you predict will become important). You can do further research (e.g., observational research or statistical research about how many people listen to particular radio stations, are on Facebook, attend a specific church or mosque) to identify additional mediums.

Then, analyze the characteristics, strengths, weaknesses, and opportunities of each medium (see Medium Analysis Reference Chart). Consider relevant concepts and approaches and identify the medium or mediums that can best reach your audience and accomplish your goals.

Finally, it is critical to do a risk analysis of each medium in order to identify risks to the intervention and possible unintended negative consequences. Once you have completed these levels of analysis, you can narrow your selection down to a final combination of mediums.

Medium Analysis Reference Chart

?	Relevance
Whom does this medium reach (which audience groups)?	Communication through the medium should be targeted (in terms of speaker and message) at the groups it can reach.
Whom doesn't this medium reach (which type of audience member)? Why?	This will help you identify groups that need to be reached through another medium. If groups will be excluded from this medium, you can assess if this will create risks (for example, if one side of the conflict is excluded).
How often do members of the audience group access this medium?	You can design the timing of your messages based on how frequently and in what situations you will be able to reach your audience groups.
Where and when do they access the medium (e.g., alone or in a group? Are they fully focused while they are accessing it or are they performing multiple tasks)?	If people see/hear content while they are with a group, their reactions can be shaped by other people's reactions and social pressures. This can be positive or negative depending on how people influence each other. Consider how it will impact people's reactions to the message and how they behave.[46]

?	Relevance
What prompts the audience to access this medium and how much control do you have over when they access it (e.g., if they turn on the TV when they're in the mood vs. if they respond to prompts such as their phone ringing)?	This will help you assess how much control you have over when and how people access the medium. If you can control when you reach people (e.g., with a loudspeaker or a text message), you can reach them in response to specific events. You can target mediums that people turn to for information about events (e.g., radio) to change the type of information people have access to. You can also target mediums for consistent long-term programming (e.g., TV soap operas).
What type of interaction does this medium allow? For example, does it enable a one-way or two-way conversation? Does it enable multiple people to participate? Is it moderated? Does it enable a depth of discussion or a minimal level of discussion?	This helps you gauge how much control you have over the interactions you initiate. For example, on Facebook, you have control over an initial post but may not have control over what people post afterwards, so being prepared for potentially negative commenting will be important. This also helps you assess limitations. In a one-way medium, it will be harder to get feedback from the audience, but this could be combined with two-way messaging (such as a Facebook page or SMS platform) to enable audience engagement.
How much do the audience groups trust this medium?	You want people to trust the information you are spreading. For example, in some countries, Facebook is seen as a reliable source; in others it is not.
Do people share information they get through this medium with other people? If so, how? And how often?	Information that is passed through social networks can become credible, trusted, and can help build perceived social norms. If people use the medium to share information, you can build on this behavior so that people spread your messages to their social networks.
What other mediums do people who access this medium use? How frequently, and how much do they trust them?	These questions can help you assess where to look for content you are competing with. It can also help you assess the value of this medium in relation to other mediums, and consider which mediums could be integrated for the highest impact.

Part 2:

Concepts and Approaches for Choosing Mediums

To build a strong Medium Strategy:
- Consider Traditional Mediums
- Make the Audience a Speaker
- Build on Existing Behaviors
- Use an Integrated Medium Strategy
- Claim and Reclaim Contested Spaces
- Consider Breaking Into Echo Chambers
- Use Information and Communications Technology (ICT) Carefully

Consider Traditional Mediums: Marketers have found that "word of mouth is the most powerful way of communicating a message," because the message generally comes from people within a trusted social network.[47] While new media make it possible to spread messages to much larger and broader audiences, ultimately, messages that people hear from their friends and social networks may have the biggest impact.[48]

Make the Audience a Speaker: Your choice of medium can take into account how easily a medium enables an audience member to become a speaker, and through which mediums. For example, does the medium prompt audience members to share or like something on Facebook, to talk about it with a friend, to engage with it in a public meeting?

getting information through printed fliers and discussing them at tea shops, or watching TV every night, how can you tap into these behaviors and use them for your intervention?

Use an Integrated Medium Strategy: Mediums lend themselves to different types of interaction and reach different segments of the audience. For the greatest impact, you can use multiple mediums that interact with each other and reinforce each other.[50] For example, if you are using an ICT-based medium, you can integrate it with on-the-ground programs or a more traditional medium (e.g., using Facebook or SMS together with radio or face-to-face interactions).

Sisi ni Amani Kenya (SNA-K) in Kenya: Subscribers to SNA-K's SMS (text) messaging platform reported forwarding the messages they received to others. Text messages are commonly used and forwarded in Kenya. Some subscribers said that the messages gave them the courage to spread peace, that they forwarded messages to people they knew were in tense areas, and that they discussed the messages with others around them.[49]

Muslim Community in Rwanda: In Rwanda, Muslim leaders used the many mediums available to them to reach a wide audience and spread their message. They "spoke out publicly in ways that could reach Muslims as well as non-Muslims." They "issued a 'pastoral letter,' posted in mosques around the country, calling upon Muslims to avoid becoming involved in any political parties that involved ideologies or actions counter to the teachings of the Quran."[51] They also used broadcast radio and spoke at mosques and schools.[52]

Build on Existing Behaviors: Try to build on people's existing behaviors. Getting people to do new behaviors (e.g., use a new medium) can be difficult, and the more you can tap into what people are already doing the better. If people are accustomed to spreading rumors by word of mouth at the marketplace,

"I am Karachi" in Pakistan: The I am Karachi wall paintings have led to coverage of the new walls in traditional media, and photographs of the walls have been posted to online news sites and social media, raising the awareness and visibility of support for positive speech.[53]

Interfaith Mediation Center (IMC) in Nigeria: To spread its message of interfaith cooperation and inclusiveness, IMC uses the media (e.g., radio, television, and joint press statements), and its representatives speak at local churches and mosques. Mass media (radio and television) enable it to reach audiences it cannot communicate with directly, but they are expensive. IMC's in-person communication programming in target hot spots, where it trains local leaders to speak directly in their communities (at churches and mosques) enables it to reach audiences on an ongoing basis. This strategy builds on the strengths and limitations of different mediums: churches/mosques are accessible, build on existing behavior, and do not cost anything for the imams/pastors, but they are limited in reach. Mass media enable leaders to reach many people but less frequently, because of the cost. The combination of mediums lets IMC spread its message widely and engage more substantively and consistently with the people who are most at risk.[54]

Sawa Shabab Radio Show in South Sudan: From the beginning, the Sawa Shabab program has incorporated SMS (text messaging) and call-in. After each episode, the main character asks listeners to text in to share their thoughts on his/her dilemma, helping young listeners directly engage with the stories and practice their own critical thinking and decision-making skills.[55] This is an example of complementary mediums: the radio drama, which is a nuanced one-way communication; and the phone-based components that create opportunities for two-way discussion, feedback, and audience engagement in problem-solving and critical thinking.

Claim and Reclaim Contested Spaces: Your medium strategy can either claim or reclaim space that is, or is likely to become, dominated by hateful rhetoric. You can do this by predicting new mediums and claiming these new spaces, or by reclaiming spaces that have become dominated by dangerous and hateful speech.

Claim Contested Spaces (Predict New Mediums): Communication, especially through ICTs, is changing rapidly in many countries around the world. You can stay ahead of the curve by using contextual information about how people communicate to predict which mediums are likely to be used and how.

Example from Myanmar: Following dramatic increases in mobile phone availability, a local organization in Myanmar predicted how Facebook will be used in the future. It considered how Facebook will interact with the ways people already communicate, and concluded that Facebook is likely to become a main source of information because it builds on existing communication habits. The ability to passively scroll through Facebook and have plausible deniability about interacting with different content makes it an appealing choice in a country that is coming out of authoritarian rule and has a history of repression of freedom of speech. The risk of posting "potentially illicit information" is limited to a few people willing to take the risk.[56] In addition, people have historically depended on rumor (information shared between friends, family, and social networks) for information. Facebook mimics and enhances this type of information sharing, and is often seen as credible. The ability to post multiple sources also adds to perceived reliability.[57] This type of prediction can enable DSIs to proactively claim spaces that could be used to promote conflict

Example from Pakistan: The "I am Karachi" campaign in Pakistan has reclaimed public walls, changing the norms of public space in Karachi. Wall chalkings have traditionally been used as a way to publicly communicate in Karachi, and over time became dominated by negative and divisive propaganda, targeting groups and building and deepening divides and conflict. By painting over the negative chalkings with positive images uniting people based on a shared identity (Karachi), these artists are physically reclaiming spaces and resetting the norms that are promoted in these public spaces.[58]

Sisi ni Amani Kenya (SNA-K) in Kenya: As Kenya's 2013 election approached, SNA-K's local partners in one town noticed that butcher shops, where men from different tribes usually sat and talked together, were becoming divided: people were sitting only with members of their own tribe, and conversations became defensive (e.g., about what the other group was planning). They persuaded some shop owners to put SNA-K stickers on their walls to make it known that people should not use that space to create divisions. They were attempting to reset the norm for that space.

Reclaim Contested Spaces: Communication norms may already be set, or may have recently shifted, such that certain spaces (e.g., physical spaces such as walls or communal meeting places, and virtual spaces such as social media) are already dominated by dangerous speech.

Consider Breaking Into Echo Chambers: Social networks can create bubbles in which people only get content and information from other people like them, who are likely to have the same opinions and views. This is exacerbated by new ICTs (such as Facebook), and these bubbles are called "echo chambers."[59] Echo chambers increase gaps and misunderstandings between groups, who each get completely different information from their networks.[60] You can try to break into echo chambers by using ICTs to create interactions between people who haven't had the chance to interact in real life.

that you like the same basket ball team, same movie, you do the same job, you both hate your boss…..then maybe you start talking, maybe you really become friends."[61]

The Peace Factory in the Middle East: The Peace Factory, a Facebook-based initiative with pages all over the Middle East, has created opportunities for people who support peace, but come from different countries and backgrounds, to become part of each other's Facebook social networks. The Peace Factory created "Friend Me 4 Peace," where individuals can post "Friend me for peace" on a Peace Factory page, and The Peace Factory will make a poster with a link to their page, then publicize that they want friends from different sides of the conflict. People can see the poster and friend them. Through this initiative, The Peace Factory has helped expand people's Facebook networks. The Peace Factory writes that: "most of us have friends just from 'our side…imagine having an Israeli friend, a Palestinian friend, an Iranian, a Syrian….anyone from 'the other side'. Just one, on your fb [sic] list you don't even have to talk to them. Looking at their birthday pictures, reading their status…makes you realize that you are just the same. You realize

Use ICTs Carefully: New ICTs have enabled people to communicate more quickly and broadly with their social networks. They enable people to coordinate and communicate more easily, and they can be used to facilitate collective action and promote peace or violence. If you are using ICTs, remember to pay attention to how your audience accesses, processes, and spreads information, and how people decide how to act. It is important to identify context and audience-appropriate technologies, and to consider which audiences they are reaching, which audiences they are excluding, and whether they are promoting offline action.[62]

Part 3:

Risk Analysis & Mitigation

There are several categories of risks for medium strategies, which include "First Do No Harm" risks, risks of individual harm, and risks to the intervention's success. For each category, you should identify the risks, analyze them, then decide whether and how to mitigate them.

Instructions for how to identify and mitigate specific risks are included below. All risks should be analyzed based on the following two questions: What is the likelihood of this risk (high, medium, or low)? If this risk happened, how much negative impact would it have (high, medium, or low)? You can assess the likelihood and potential impact of each risk using a chart like the one on the following page. You can decide how to invest resources or adapt programming for risk mitigation based on the likelihood and potential harm caused by each risk. For example, unless it's an easy fix, it may not make sense to devote limited resources to a low impact/low likelihood risk, but it may make sense to spend time and resources on a high impact/low likelihood risk because of the potential damage that could be caused.

Risk Analysis Chart

Likelihood:	Impact		
	High	**Medium**	**Low**
High	High Likelihood/ High impact	High Likelihood/ Medium Impact	High Likelihood/ Low Impact
Medium	Medium Likelihood/ High Impact	Medium Likelihood/ Medium Impact	Medium Likelihood/ Low Impact
Low	Low Likelihood/ High Impact	Low Likelihood/ Medium Impact	Low Likelihood/ Low Impact

Do No Harm Risks

Risk of Increasing the Power of Dividers:[63]

Identification: At a broad level, consider which mediums are currently the biggest dividers (being used to increase divisions along conflict lines). Is there any chance that your intervention could increase the use or power of these dividers (e.g. by driving more people to use them and be exposed to divisive content)?

Mitigation: Mitigation strategies will vary depending on the mechanisms by which your medium(s) could contribute to dividers.

Risk of Decreasing the Power of Connectors:[64]

Identification: At a broad level, which mediums are currently the biggest connectors (are able to bring people together across conflict lines, or combat Dangerous speech)? Is there any chance that your intervention could reduce the use or power of these mediums?

Mitigation: Mitigation strategies will vary, but should seek to ensure that your intervention does not reduce the positive impact of existing mediums.

Misuse of the Medium:

Identification: Can the medium be misused by people with negative intentions (e.g., could a radio station or technology platform be taken over by people engaged in group-targeted harm)? What type of information could they spread through the medium? ICTs in particular risk being hacked or falling into the wrong hands. ICTs can also create risks for misinformation to spread more quickly, so risk management of content becomes even more important.[65]

Mitigation: Risk mitigation strategies may include security measures and contingency planning.[66]

Risks of Exclusion:

Identification: Your medium or combination of mediums may exclude some people. Are any audience groups for your intervention not reached through your medium strategy? Is any group disproportionately excluded, and could this exclusion have a negative impact on the conflict? Does the medium in any way disempower a portion of the audience, or individuals outside the audience? Does it contribute to conflict dynamics by favoring one group over another? ICT-based interventions in particular can create a "Bias of Connectivity," when ICTs aimed at reaching more people end up reaching only those who have access to a specific technology. This can isolate members of the audience who do not have access.[67]

Mitigation: One way to mitigate this risk is by using a variety of mediums.

Individual Harm:

Risks to Individuals:

Identification: Consider the risks to people participating in the DSI. Does the chosen medium put anyone at risk, and are they are able to give their consent? A radio station presenter may agree to participate in a show, but the show might put others (e.g., presenters who have not agreed to participate) at risk.

Mitigation: People ultimately have to decide what risks they are willing to take. It is your responsibility to ensure that everyone who is put at risk is aware and can make informed decisions. This is particularly important if you are using ICTs, since people may not fully understand the risk associated with different technologies.[68] You must therefore educate those affected so that they can make informed decisions.

Risks to the Intervention's Success:

Risks of Exclusion:

Identification: In addition to the Do No Harm risk of disproportionately excluding a group, you can ask whether the choice of mediums misses any audience group that your intervention aims to reach. If so, this could limit the impact of your intervention.

Mitigation: The best way to mitigate this risk is generally by using a variety of mediums in order to reach the full target audience.

Encouraging Passivity:

Identification: This risk category focuses on whether the medium can impact the audience's behavior, and specifically on whether it in any way disempowers the audience from meeting the intervention's behavioral goals. This is particularly important for ICTs, which can create a risk of "clicktivism," meaning the audience only passively views or interacts minimally with the message online without feeling compelled to take action.[69] This can become a Do No Harm risk if your intervention could make people who are currently taking action offline inclined to only take action online.

Mitigation: You can develop mitigation strategies for this risk by integrating mediums and developing clear strategies and message content that prompt action.

Part 4:

Iteration

The purpose of your mediums is to reach the target audience groups. How, on an ongoing basis, will you know whether your medium strategy is working? What feedback loops will you create so that you can consistently check if any of your assumptions are wrong? For example, are people actually accessing the message content you are providing? Are they sharing it with friends? By creating feedback loops (regular touchpoints with local partners who are in the field) to learn the answers to these questions, you can consistently update your understanding of the available mediums and improve your strategy based on what you learn.

2: Speakers

The speaker or speakers who deliver your message content will determine whether your audiences see your message content as reliable, relevant, and appealing.

Identifying Speakers Who Will Influence Your Audience

A relevant, credible and influential speaker will increase the likelihood that your audience will react positively to your message. How people feel about a speaker or source of information can influence whether people believe your message (even more than the message itself).[70] In other words, the speaker will determine how people feel about your message (e.g., whether they are open to the message or pre-decide that it is irrelevant). The wrong speaker can discredit a message, and the right speaker can make a message more influential. Speakers can deliver content directly and/or encourage people to participate in the intervention. They can also lend credibility to the intervention through simple actions such as liking a Facebook page, attending an event, or subscribing for alerts from a text messaging platform or social media page.

To identify your combination of speakers, first use your audience profiles and analysis to determine which speakers are credible (people believe they are trustworthy and knowledgeable), influential, and relevant (people believe the speaker shares values or a worldview with them) to your audience groups. Supplement this list with any potential speakers you might have missed. Consider if your audience includes groups (such as people involved in a particular occupation or activity, like a women's group) that could become speakers.

Next, analyze the characteristics, strengths, weaknesses, and opportunities for each speaker. Consider which audience group each speaker can reach and whether the audience groups see each speaker as credible and relevant. You can also evaluate the level of influence that the speaker has over the audience (e.g., whether people talk about what he/she says and look to him/her for advice or as a leader). Finally, assess which mediums the speaker is already using. Think about whether you can tap into their existing behaviors to get them to spread DSI messages, and what it will take to get them to use any new mediums that are part of your intervention strategy.

Consider the Do No Harm risks, risks that each speaker poses to the intervention, and the risks that he/she might face by being part of your intervention before finalizing your speaker strategy.

Note: If you can, involve partners in the speaker identification process. This can increase partners' buy-in, and can increase the network of speakers you have access to (partners may be willing to help recruit speakers you can't access). If you are targeting a variety of locations, local partners will be necessary to help you identify locally influential speakers.

IMC in Nigeria: IMC was founded by two credible, well-known, and well-liked imams and two pastors/reverends. These four leaders identify likeminded imams and pastors, then recruit and mentor them. They also recruit and train traditional leaders, which in this case includes leaders of ethnic groups or communities.[71] This is an example of using influential leaders to recruit and support a network of speakers who can influence audiences in target geographic locations.

Part 2:

Concepts and Approaches for Choosing Speakers

To build a strong speaker strategy:
- Tap into existing social networks to make the audience a speaker
- Find speakers who have things in common with the audience groups you aim to influence
- Think multi-partisan, not non-partisan or neutral
- Consider using a surprise speaker
- Use speakers who can model desired behaviors and attitudes
- Recognize the challenges
- Identify, support, protect, and build the voice of potential speakers early on
- Think about the implications of international and institutional support for the initiative
- Consider speaking "in unison"
- Consider fictional speakers
- Be creative (find creative ways to limit risk)

Tap into Existing Social Networks to Make the Audience a Speaker: People are highly influenced by their own social networks, and care about how others in their immediate circles view them. The audience often becomes a speaker for dangerous speech; for example, during the genocide in Rwanda, many individuals were recruited to participate in violence through close ties to relatives and friends who were influenced by anti-Tutsi propaganda.[72] The audience to dangerous speech can also become a powerful DSI speaker. Audience members can echo and spread positive messages, and make them relevant to the people in their social networks and local context. This will give messages more credibility, since people will hear them from people they already trust.

You can identify and recruit locally relevant, credible, and influential speakers in your target geographic areas or demographic groups. These can be people in leadership positions (e.g. they are highly influential for a particular group), or people who have diverse social networks and are practiced in spreading information and adapting it for different audiences.[73] These local speakers can model positive behavior on an ongoing basis and encourage others to spread DSI messages. Think about how your speakers will get the audience to buy into your message and spread it to their social networks, maximizing your impact. This approach is known in marketing as brand advocacy.

Find Speakers Who Have Things in Common with the Audience Groups You Aim to Influence: People are more likely to believe speakers they think share their identity, value system, or worldview. People may disbelieve a speaker simply because they think he/she doesn't have similar values, find a speaker more credible just because he/she does have shared values,[74] or judge whether someone is an expert based on whether they have shared values.[75] It is important for your audience members to see that they have things in common with a speaker (e.g., identity, traits, values, or social group membership).[76] To ensure your speakers are reaching the full audience, consider whether each audience group has a speaker they think shares a common identity, worldview, or values. This may require thinking beyond the usual speakers. People who promote messages of peace or directly contradict dangerous speech may not be able to appeal to people engaged in dangerous speech and group-targeted harm. Targeting these groups may require recruiting speakers who are not traditional peace actors.

Think Multi-Partisan, not Non-Partisan or Neutral: Multi-partisan speaker groups (speaker groups that represent all sides of the conflict), rather than groups of purely neutral speakers who express no conflict-related opinions (speaker groups made up solely of known peace actors) are likely to have a stronger impact and reach a wider audience. Each audience group you seek to influence should be able to identify strongly with one of your speakers. Strong speakers who represent relevant group identities and grievances can appeal to different members of the audience. If you aim to reach people who identify strongly with dangerous speech, group grievances, and violence, this may mean recruiting people who are vocal about grievances but are against group-targeted harm, or people who were formerly involved in violence. In addition, it can be hard to find speakers who are seen as neutral, especially if the current dangerous speech revolves around fixed identities. People may be seen as part of one group or the other just because they belong to a specific religion, ethnic background, etc., even if they have explicitly declared themselves neutral. If you want your intervention to be seen as neutral, consider using multiple speakers who represent different sides of existing grievances or conflict.

Consider Using a Surprise Speaker: A "surprise speaker" can grab the audience's attention and create discussion. This is a speaker the audience wouldn't expect to deliver a DSI message. An example from traditional advertising is Dove, which is part of the beauty industry, building a campaign about how the beauty industry makes women feel bad about themselves. Even people who had heard that message before were surprised to hear it coming from Dove.[77] In the context of DSIs, surprise speakers might include people who staunchly support one group or the other, celebrities with a reputation for being tough, or former participants in violence. For example, when IMC pastors and imams speak together, the combination of speakers is surprising. (Note: This approach may create risks if someone is seen as "traitorous" by appearing publicly with someone from "the other side.")

Use Speakers Who Can Model Desired Behaviors and Attitudes: DSI speakers can take actions – such as publicly speaking out against dangerous speech, speaking about the benefits of cooperation and diversity, or advocating non-violent resolution of grievances – that set an example for people in their

community. People may feel more comfortable speaking out or refusing to spread dangerous speech if they have seen relevant influential figures do so.[78]

Recognize the Challenges: It can be difficult to identify speakers when dangerous speech has already begun to become the norm and when group identities have become rigid. Dangerous speech often targets moderates first, portraying them as traitors to their group because of their opposition to group-targeted harm and discrimination. Moderates can be targeted with physical violence, restraint, or attacks to their credibility (e.g., claiming that they are agents of the group targeted for harm). Once moderates have been targeted, it may be difficult to identify moderate speakers who still have credibility with their groups, and anyone speaking against dangerous speech may automatically lose credibility or face consequences (for example, a speaker may be automatically categorized as a traitor for straying from dominant dangerous speech narratives). You can address this challenge by supporting moderates early on, understanding implications of external support for your intervention, speaking in unison, and finding creative solutions.

Identify, protect, and build the voice of potential speakers early on. As dangerous speech progresses, it can become increasingly difficult to protect speakers from risks of harm and from being discredited. Your intervention can build on any existing efforts. You can attempt to match speakers with messages that will allow them to stay credible. You can also think about ways speakers can assert their identity strongly (even by expressing group-related grievances) while speaking out against group-targeted harm.

Think about the implications of international and institutional support for your initiative and about whether these types of support will help or hurt the speakers' credibility. Conspiracies about international intervention are easily spread and often used to target moderate speakers. If your intervention receives international support, it may be risky to make this public, and you should think about how to frame this support and how to respond if dangerous speech actors use it to discredit your work.

Speak "in unison": One mitigation strategy when speakers are vulnerable to attack is to speak "in unison," by having a group of people speak out together to reduce individual risk.[79] You can facilitate this through social media or arrange for speakers to make joint statements and appearances. When dangerous speech becomes louder and more visible than opposition to group-targeted harm, people may think that everyone supports violence. If groups speak out together, it can help overcome the perception that most people support violence, and may make others feel more comfortable supporting DSIs.

Use Fictional Speaker(s): You can consider using a fictional speaker or speakers (e.g. by creating a radio drama, story, comic, play, pamphlet marked with a brand or symbol) to illustrate relevant concepts subtly and without putting individual speakers at risk. For example, you can use a brand as a speaker to reduce individual risk and create the perception that there is support for speaking up against dangerous speech.

Be creative and think about any other ways you can limit risk, for example by anonymizing speakers and creating the perception of mass participation.

Example from Myanmar: During authoritarian rule in Myanmar, "students distributed pamphlets by placing them in stacks on top of buses, to be scattered as the vehicles pulled away. Sharing pamphlets is the crime of the distributor, not the passive user who stoops to pick up an unknown piece of paper…"[80] While this was a form of resistance to authoritarianism, this type of creative approach can convey mass support for a piece of communication, distribute it widely, and open opportunities for people to begin difficult conversations (they could pick up a pamhlet and use it to start a taboo conversation or access information they might be uncomfortable looking for on their own).

Part 3:

Risk Analysis & Mitigation

There are three types of risk for your speaker strategy: (1) "First Do No Harm" risks (2) intervention impact risks; (3) risks to individual speakers. For each type of risk, consider the likelihood that it will happen and the potential negative impact, as well as any mitigation strategies that can reduce or prevent the risk. Remember, it is ultimately up to you to decide if the risks of a particular strategy are too high.

Instructions for identifying and mitigating specific risks are below. You can analyze risks, and decide how to invest resources or make programmatic changes, based on the likelihood that the risk will happen, and the potential negative impact if it does. As with medium risks, you can analyze risk using the following chart:

Risk Analysis Chart

Likelihood:	Impact		
	High	Medium	Low
High	High Likelihood/ High impact	High Likelihood/ Medium Impact	High Likelihood/ Low Impact
Medium	Medium Likelihood/ High Impact	Medium Likelihood/ Medium Impact	Medium Likelihood/ Low Impact
Low	Low Likelihood/ High Impact	Low Likelihood/ Medium Impact	Low Likelihood/ Low Impact

First Do No Harm Risks

Risk of Increasing the Power of Dividers:[81]

Identification: At a broad level, which speakers are the biggest dividers (successfully promoting dangerous speech and group-targeted harm or other societal divisions)? Is there any chance that your intervention could increase the power of these dividers (e.g., by making more people aware of them, increasing their audience, giving them credibility, providing recognition or legitimacy)?

Mitigation: Your mitigation strategies will depend on the ways your intervention could increase the power of divisive speakers.

Risk of Decreasing the Power of Connectors:[82]

Identification: At a broad level, which speakers are currently the biggest connectors (are able to bring people together across conflict lines or other divisions, or combat dangerous speech)? Is there any chance that your intervention could disempower or reduce the credibility, influence, or reputation of these speakers?

Mitigation: Your mitigation strategy should ensure that your intervention does not reduce the impact of speakers who are currently having a positive influence.

The Speaker May Cause Future Harm:

Identification: Being part of your intervention will potentially legitimize and support speakers' other activities. Consider whether any of the speakers you are supporting have the potential to use this legitimization and support to cause future harm (e.g., by recruiting supporters for violent causes).

Mitigation: Based on the likelihood and impact of this risk, you may decide not to work with these speakers.

Intervention Impact Risks:

The Speaker Changes Sides:

Identification: People can change their behavior during a conflict: an influential leader might support peace for a period of time then choose to participate in dangerous speech or violence. This could discredit your intervention. To predict the likelihood that each speaker will switch sides, consider the speaker's history, what he/she has to lose, his/her social ties to various networks, and who is likely to apply pressure on him/her.

Mitigation: You can engage a variety of speakers and use social pressure (and a brand) to hold the speakers accountable. You can also develop action plans and narratives to use if a speaker switches sides.

The Speaker is Discredited:

Identification: People promoting group-targeted harm often discredit moderates in their own group early on, making it hard to find credible speakers. DSI speakers can become targets of negative propaganda, and if they are discredited, it can discredit their message and the overall intervention. Most speakers can be discredited, even through false information or conspiracy theories.

Mitigation: You can predict narratives that could be used to discredit each speaker. You can review how others have been discredited and analyze real factors, such as past scandals, to anticipate rumors and conspiracy theories, then proactively discredit these narratives.

The Speaker Makes the Intervention Look Biased:

Identification: If you aim to appeal to multiple sides of a conflict or multiple groups on one side, you can think about whether each speaker will make your intervention seem biased. Consider which groups the speaker is considered part of; whether the speaker has spoken out on behalf of any

of these groups; and whether the speaker has ever spoken negatively about other groups.

Mitigation: Biased speakers can be the most effective speakers for specific groups, and they can even be surprise speakers (e.g., if they are seen as unlikely to oppose group-targeted harm). If you do not want your intervention to appear biased, the best way to mitigate this risk is by recruiting speakers from multiple groups. Multiple biased speakers can create a balanced initiative and appeal to multiple groups.

The Speaker is Very Disliked by Some:

Identification: If some people strongly dislike a speaker, this could damage the intervention's overall reputation.

Mitigation: This risk should be considered based on whether the reputational damage is worth the value the speaker brings, and whether it can be mitigated (e.g., by balancing this speaker with another speaker who is very liked by that group). Consider the pros and cons of working with this type of speaker: he or she may have a lot of credibility and influence with an audience that's hard to reach, even if he or she isolates another group. If you decide to work with a disliked speaker, using multiple speakers or content that can reduce negative associations with the speaker can mitigate risk. You may also decide that being disliked by one group is okay because it gives the speaker added credibility with another group. Consider the long-term implications of this strategy, and be sure to have a plan about how this will change over time.

Not All Members of Audience Groups are Reached:

Identification: The speakers chosen for each audience group should have something in common with the members of the audience group, such as a fixed identity, explicit set of values, or membership in a social group, and should be influential, credible, and relevant.

Mitigation: If audience group members are not represented by your speakers, you can identify additional speakers to reach these groups.

Risks to Individual Speakers

Speakers Face Personal Risks:

Identification: Speakers may face personal risks by joining a DSI and speaking out publicly. If so, you can predict the types of risk each speaker may face (e.g., job loss, arrest, physical violence).

Mitigation: You should assess ways to protect the speaker from any identified risks. At minimum, it is your responsibility to ensure that each speaker is fully aware of the risks and makes an informed decision when he/she consents to participate.
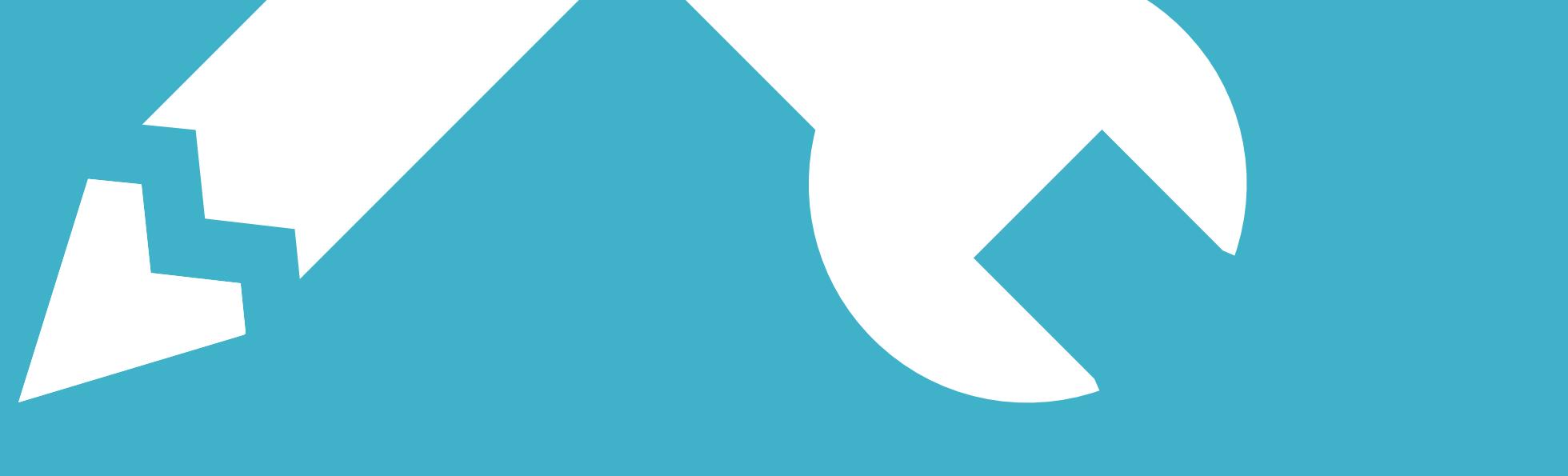
Part 4:
Iteration

The purpose of your speakers is to get target audience groups to listen to your message: the speaker is what will make your message credible, trustworthy, relevant, and influential for the audience groups you seek to influence. As you implement your intervention, how will you know whether your speaker strategy is working? How will you find out about its impact? How will you create feedback loops so you can find out if any of the assumptions you've made about your speakers are wrong? For example, how will you know: if a speaker you thought was influential has lost credibility; if another speaker is isolating a certain audience; or if a speaker is engaging in activities that could threaten the credibility of your intervention? Local partners and team members will likely be able to provide these types of feedback loops simply by watching the situation on the ground. How can you give them incentives to do so, and to report accurately? Ensure that your feedback loops will enable you to hear from a diverse group of stakeholders, and be sure to verify information.

Toolbox

Building a Brand

A brand can create a common unifying identity. By creating a strong brand around your DSI, you can begin to create an alternative to the identities that are being targeted for and by dangerous speech. A strong brand will define your intervention's values, beliefs, and personality.[83] It can create a clear set of expectations for anyone who is part of your intervention, and help hold speakers accountable. Speakers and audience members can show off their allegiance to the brand, and show that they endorse its values, beliefs, and behaviors. This can change social norm perceptions by creating visible support for non-violence, and can create a positive identity.

For these things to happen, you need to create a strong brand identity. Why does your brand exist? What promise are you making to anyone who participates in the brand? What are your values and beliefs? You can answer these questions by going through a three-part brand building process:

1: Brand Purpose: this will help audience members and speakers understand the vision or goal that they are supporting. Your brand purpose should be a clear, simple, and easy to repeat sentence that explains why your brand exists: it's your tagline, a simple way people can understand what it means to support your brand. It should explain what you are trying to accomplish, why people should care (even if they don't care about peace, conflict, or politics), and what you are bringing to the table that is new or different.

2: Brand Promise: this will create clear expectations that hold speakers accountable and give audiences clarity about what they can consistently expect from the brand. Your brand promise should be a clear, simple, and easy to repeat sentence that explains what people can expect from the brand. To build your brand promise, consider what promises you can make (and keep), what people can expect from your brand (consistently over time), and the end goals that the brand will achieve over time.

3: Brand Personality: a brand personality creates behavioral, linguistic, and tonal expectations for how people represent the brand. It can be used as an internal set of guidelines to ensure that everyone behaves consistently with the brand values when they represent the brand. The brand can also create a positive alternative identity to conflict-related identities, so the personality should appeal to your audience groups. This is important for a couple of reasons. While multiple speakers may be able to reach different audience

groups through their unique personalities and experiences, the consistency of the brand can help unify these speakers with a set of brand-related behaviors. The brand personality shouldn't take away from people's own personalities, but should let different speakers work together in a way that appears and is consistent. This creates opportunities for audience members to buy into the brand and express their support: they can identify strongly with one speaker, and that speaker can connect them to the brand even if the brand includes other speakers they don't identify with.

To build a brand personality, identify values and beliefs that the brand represents, then connect these values and beliefs to behaviors. For example: What tone (e.g., calming or aggressive) will your brand use? What does your brand say about someone's identity? What are your beliefs and values and how are they expressed? What other linguistic guidelines do you need (for example, if there are specific words that should or shouldn't be used)?

Brand Personality

Promoting Peace through Unity
- be positive, calming, hopeful and optimistic
- avoid aggressive tonality

Empowering and Inclusive
- promote an inclusive identity
- use "we"
- listen to others and react to their opinions
- show empathy...

Credible
- fact-che...

Creating a Brand Advocacy Strategy

If you choose to create a brand, you can develop a brand advocacy campaign that recruits audience members to advocate for the brand and pass messages through their own social networks. This is a way to get people to spread your message beyond your direct communications, and to buy in to a unifying identity (your brand). One way to do this is by recruiting speakers who have different types of interactions with audience groups. Here are three types of speakers you can combine to reach a breadth of people and tap into people's social networks:[84]

Top Tier Speakers: These are people who have a disproportionate amount of influence over specific communities or audience groups, but do not interact with these groups regularly. Examples include celebrities, politicians, well known activists, media personalities, high level religious leaders, etc. Because they have a high level of influence but don't interact regularly with the audience group you aim to influence, they can publicly support the intervention, lending legitimacy and reaching large audiences, but infrequently.

Community Speakers: These are people who have a strong influence in local communities and interact with the audience regularly. They can incorporate DSI content into their regular communications and offer legitimacy and strength to the messages. Examples include local religious leaders, community opinion leaders, community activists, etc.

Audience Speakers: These are members of the audience you want to target as speakers (for example, members of the "information spreaders" audience group who can reach a particular sector of your audience, such as business owners).

Think about how you can recruit advocates from each of these categories, and how you can recognize and reward their efforts so they will stay engaged. What can you do to make it easy for each of these speakers (especially the audience speakers) to spread your message? What content and mediums can they use to pass along the information they get from your intervention? Can you give them content they can easily spread through mediums they are already using? Can your top tier and community speakers model behaviors (e.g., sharing information on social media, passing out pamphlets, wearing armbands or T-shirts) that audience speakers can easily mimic?

3: Message Content

The content of your messages should engage your audience and influence their behavior. Message content should build on an understanding of the audience and context, and each message should have a clear goal.

Part 1:

Using a Goal-Based & Audience-Centered Approach

Your content development process can generate: (1) content that responds to specific events, (2) general and ongoing content, and (3) insights and guidelines for messaging target groups or for messaging around specific events.

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.

Content Development: Once you have set overarching and specific goals, you can use your existing audience and context analysis and conduct further research to develop content guidelines and actual message content. Developing content guidelines based on insights about types of message content and about message tone and structure can be useful if there is a broad network of speakers who can use these insights to create their own content, or if you want to be prepared to create strategic content in response to unfolding events. You can record, analyze, and revise content and/or content guidelines throughout and after the research process. Be sure to involve your partners in message development so they can gain insights and strategic skills. Consistent feedback about your messages from local partners can help you detect risks.

Part 2:

Concepts and
Approaches to Consider

Your content should be based on audience research, but you can draw on relevant theories to test ideas that you think might work, and to rule out content that might backfire. General guidelines, risks, and strategies are outlined on the following pages.

The Audience's Perspective Comes First: What you intend to say or convey doesn't matter; your intent isn't what makes content succeed or fail. What matters is how the content is understood, interpreted, or acted upon by the audience, so it is important to keep the audience group's perspective in mind during content development. Remember that people often do not behave rationally, so you need to understand the logic and emotions they are using to make decisions. Think about behavioral drivers (that push or pull people to do something) and barriers (that prevent people from doing something) to understand what motivates and constrains your audience's behavior. For example, empathy might be a behavioral driver, but economic incentives may be even more powerful. Your messages are all being inserted into a context and a broader conversation: try to understand how your message fits into the broader context of your audience's life, and into how your audience group members define their identities and values.

Think about Tone: The tone and type of language you use will make an impression. Think about what kind of language (e.g., slang, a specific group's language/mother tongue, significant colors or symbols) is used and understood by the audience groups you aim to influence and what the use of specific language says about your speaker/brand.[85] Remember that speech is not only language; it can also consist of symbols or multimedia such as film or pictures. If you have chosen to use a non-language medium, you can still think about the tone (e.g., colors, font).

Be Aware That:
  • Humanization may be more effective than building empathy
  • Building trust should be a long term process
  • Be aware that challenging people's existing beliefs or feelings may backfire
  • People's natural instinct to self-justify can affect how they respond to interventions

Understanding Empathy and Humanization: It may be more impactful to focus on humanizing a targeted group rather than creating empathy.

  • **Empathy** can help reduce prejudice and aggression between people,

reduce antipathy towards other conflict groups, and increase people's willingness to take action to help another group.[86] This is specifically true of "empathetic concern," which is when someone feels for another person (i.e., feels concern about someone else's suffering), and is the type of empathy that is most likely to motivate someone to take action.[87] Storytelling and narrative media can be used to stimulate empathy.[88]

At the same time, there are several risks for using empathy. First, whom someone feels empathy for matters. Empathy can motivate someone to act on behalf of a group targeted by dangerous speech or on behalf of his/her own group: a person who feels significantly more empathy for his/her group than for the targeted group may actually be more willing to harm the targeted group, especially if he/she believes that group poses a threat (to his/her group).[89] Empathy can also feed into motivated reasoning if the feeling of empathy contradicts a person's strongly-held belief (e.g., if empathy for a targeted group contradicts a belief that the targeted group is responsible for the economic downfall of a person's own group). This type of internal contradiction can actually lead the person to hold on to his/her initial beliefs even more strongly.[90] In this way, provoking empathy for a targeted group may be risky, especially if someone has already taken action against the targeted group or has strong beliefs about threat and grievance. Empathy may also be reduced during conflict and may not be able to motivate action. This is because conflict creates and promotes negative perceptions of targeted groups and their motives, because social and other pressures to participate in group-targeted harm may be stronger than a feeling of empathy, and because feelings of humiliation, loss of dignity, and trauma can reduce the capacity to empathize.[91]

• **Humanization** of the group targeted by dangerous speech means creating an understanding that the group's members are human, and specifically that they have a mind and mental processes.[92] Dehumanization by comparison requires thinking of others as unthinking, unfeeling, and inherently not human at their essence.[93] Humanization can complement or provide an alternative to empathy. One way to humanize a targeted group may be to focus on "secondary emotions." Primary emotions, like fear, happiness,

and physical pain are emotions that humans share with animals. Secondary emotions, like humiliation, guilt, hope, and remorse are unique to humans. Research has shown that people have trouble understanding secondary emotions for other groups, especially those they are in conflict with.[94] A focus on creating an understanding that targeted groups have secondary emotions may be a promising way to humanize members of the targeted group.[95] Other approaches to humanizing targeted groups include broadening the groups people identify as their own, for example by creating an identity that includes members of the targeted group and the group doing the targeting. You can also think about ways to show that the dehumanized group has multiple identities (the group consists of individuals with many different tastes, opinions, behaviors, etc.).[96]

Similarly to empathy, humanization may encounter challenges around motivated reasoning. It may be effective to combine empathy and humanization, for example, by focusing empathy-related content on secondary emotions.

Trust: A Long-term Phenomenon: Violence prevention strategies often focus on increasing trust between groups. If trust is new and somewhat fragile, claims that a group is trustworthy may be easily discredited, and in turn discredit your initiative. Building trust also creates the opportunity for feelings of betrayal, which, unlike anger or fear, can only exist if there is trust first.[97] Trust may need to be created through a slow, long-term process, built durably by example rather than through message content.

Challenging People's Existing Beliefs or Feelings May Backfire: Recall that people's feelings and beliefs affect how they react to new information, often even more than rational reasoning.[98] People can feel psychologically threatened when forced to confront facts and information that directly challenge things they already feel or believe to be true. This is the case especially if the new information challenges people's sense of identity, meaning that challenging existing beliefs and feelings may be particularly difficult if they relate to group identities.[99] In addition to motivated reasoning, new information can backfire by causing people to hold more strongly to their existing beliefs (they will create

counterarguments to the new facts, recall others who share their view, avoid the information entirely, or selectively interpret the new information); or by making false claims more familiar (and increasing their perceived validity).[100]

This means that messages that seem intuitively pro-peace may trigger motivated reasoning, and backfire. For example, saying that a group is not under threat may trigger a defensive response and make that group feel more under threat— "see, they don't even recognize the extreme threat we live under, no one else will be willing to help us: we need to protect ourselves."

To avoid motivated reasoning, it is important not to present new information in a way that triggers a defensive reaction or makes people feel insecure or threatened. Making people feel secure in their identity may make them more open to facts or arguments that directly contradict their strongly-held beliefs.[101] You can do this by creating opportunities for self-affirmation (such as asking people to affirm their existing beliefs and values),[102] or by boosting people's self-esteem and making them feel good about themselves before providing challenging information.[103] You can also choose not to target the most extreme groups, who hold beliefs most strongly, and instead target moderates, who hold the beliefs but aren't as attached to them.
New information can be packaged to build on peoples' existing beliefs and worldview. For example, you might recognize and affirm a grievance while providing an alternative to violence; challenge the ability of violence to achieve group goals; or frame new information or behaviors in terms of strongly held values, beliefs, and/or feelings.

Narrative media or storytelling that challenges beliefs more subtly may be less threatening to audiences and may not trigger the same level of defensive reaction.[104]

To ensure that your efforts to counter false information do not lend it legitimacy or familiarity, you can also:[105]
- Try to correct false information as quickly as possible, because the more familiar people become with misinformation, the more they are likely to believe that it's true.

- Instead of negating false information, use positive framing (for example, instead of saying "Group X isn't plotting against us," you can say "Group X is open to dialogue"). Negations can cause people to remember the false information more clearly.[106]
- Do not repeat false claims or information when countering; this can make them seem more familiar, which makes people more likely to believe them. If you need to repeat a false claim, be sure to give a warning that the information is false beforehand.[107]
- Give an alternative explanation for the false belief (e.g. for a behavior that is being used as evidence that "Group X is plotting against us").[108]
- Be sure that the correction comes from a trusted source, and does not come from a source that the audience associates with biased information.[109]

Be Aware of Self-Justification: When people participate in negative behavior (e.g., dangerous speech or group-targeted harm), they tend to justify that behavior in order to feel good about themselves and their actions.[110] If you want to reach people who have already taken negative actions, it may be helpful to consider how you can enable them to change their behavior without threatening their idea of themselves as good (for example, by providing a narrative they can use to tie their past behavior and image of themselves with a new commitment to non-violence).

Consider Effective Messaging Approaches:
- Create Incentives Through Appeals to Interests
- Create Unifying Identities
- Change the Conversation
- Reinforce Common Goals
- Provide Accountability
- Use Aspirational Messaging
- Correct Descriptive Norms
- Distract
- Get There First: Psychological Preparation
- Use Stories and Narrative Media
- Use the Medium itself to Convey a Message

Create Incentives Through Appeals to Interests: Dangerous speech often promises rewards and/or threatens punishment. Moderates may be targeted for ostracism or harm while participants may be rewarded with inclusion, recognition, looted goods, etc.[111] Is it possible to appeal to audience interests with key incentives, value, or pressure for positive actions? Offering value in exchange for positive action and appealing to audience interests and incentives are key. There are several types of value that a DSI can provide to its audience.

Tangible rewards for positive behavior may be useful. The tangible rewards themselves are unlikely to increase positive action, but can create a culture in which positive action is recognized and rewarded. Giveaways like T-shirts and wristbands may be used to symbolize acceptance and membership to a group. Access to services and a support network may increase feelings of safety and security in speaking out.

Pride, gratitude, and social recognition that makes someone feel valued for a certain action, attribute, or event, can motivate people to action, and can be given through public praise and acknowledgment.[112] Your content can increase pride, social standing, gratitude, and public recognition for people taking positive actions. This will reinforce positive norms and strengthen a group identity where status (pride and recognition) is attained through positive action.

Positive recognition, pre-action, such as thanking and congratulating individuals for their positive actions before they have actually taken them (e.g., "thank you for your support for peace"; "for your patience as this issue is resolved") could influence people's behavior in a positive way and prevent them from acting negatively.[113] It may be most effective if this is done when people are deciding whether to take a new negative or positive action.

Create Unifying Identities: People are more likely to participate in violence or cause harm if they are acting as part of a group instead of alone.[114] Dangerous speech can create strong group identities around violence and enable people to justify actions that would otherwise be seen as immoral. Your messages

can meet people's needs for group membership and belonging by redefining a group identity or creating an alternate identity narrative in which hostility toward another group is not a defining characteristic of the group. For example, DSIs can tap into identities other than people's dangerous speech identity (such as identities related to professions, sports, gender, etc.) or craft new identities to create a sense of belonging in groups that do not fall along conflict lines.[115] You can use symbols (such as colors, armbands, etc.) to let people signal their membership to the group. Strong positive group identities can create a safe space for people to share opinions and ideas that go against dominant dangerous speech narratives. Acting as part of a group may help make people feel less helpless and believe that their actions matter and serve a greater purpose.[116] Prompting people in a group to think about their own personal values may also help.[117]

Change the Conversation: This may go hand-in-hand with creating a unifying identity. You can attempt to change the conversation to promote desired attitudes and behaviors without directly addressing the hostilities that are fueling dangerous speech. For example, you can change the conversation from being about whether or not to use violence against a group to which are the most important needs and grievances to address, and which of several options can be used to address them. This can help shift conversations without appearing initially or overtly threatening. You can also provide new information through affirming statements instead of as negations or part of an existing conversation. For example, instead of countering stereotypes with negations (e.g. this group is not lazy), you can use affirming language and framing (e.g. many members of this group wake up early to fish every day) for better results.[118]

Framing and counterframing techniques can help you change the conversation. People frame issues to promote an interpretation of a problem and a solution.[119] For example, an economic downturn could be framed as the result of corrupt leadership, meaning corrupt leadership is the problem that must be addressed. It could also be framed as the result of Group A taking over the economy, making the problem Group A. Dangerous speech often frames issues and events to promote the idea that the target group

is the problem that must be dealt with. How issues and events are framed influences what actions people believe are appropriate and necessary.[120]

Instead of trying to win an argument within the dangerous speech frame, you can promote a frame that advances a different understanding of the underlying problem and therefore offers different solutions. For example, you can frame a grievance as a socioeconomic issue rather than an ethnic or religious issue. Frames will work best if they build on people's existing values and worldviews.[121] It may be difficult to reach people who have fully bought into frames that promote a target group as the problem, so when new issues arise, being the first to frame them can give you an important advantage.

Creating an Alternate Identity: Muslims in Rwanda: In Rwanda, Muslim leaders proactively prepared their communities to resist violence. They provided a lens through which individuals could anticipate and understand conflict-related events, psychologically prepared them for resistance, and created clear normative and moral expectations for how Muslims should act when violence came. Muslim leaders recognized in advance of the genocide that political and ethnic polarization might lead to conflict and acted to emphasize community values that contradicted the identity narrative promoted by Hutu extremists. "They sensitized the public to reject the hate propaganda. They cited the teachings of the Quran and instructed people not to participate in violence when it came."[122] Teachers countered divisive propaganda in Muslim schools and reminded

students "that all people are equal, ethnicity should not be divisive, and people should not kill but should try to rescue victims."[123] They drew upon passages in the Quran that warned of a time of temptation and prepared people to resist that temptation. Messages also warned "that hard times were coming" and asked people "to adopt positive values and not implicate themselves in the coming events."[124]

Creating an Alternate Identity: Tuzla: In Bosnia, the people of Tuzla created an identity based on their geographic location (the city of Tuzla) instead of their conflict (ethnic) identities. Public statements by the city's leadership, as well as songs commissioned to celebrate the common identity of the people of Tuzla, and a sense of pride about the Tuzla identity (versus other places where people were divided) were promoted to strengthen this identity and resist participation in the conflict.[125]

Reinforce Common Goals: Appealing to shared interests and needs may help increase cooperation and reduce hostility between groups.[126] Shared goals or interests may include economic prosperity or job creation, and you can frame these goals as dependent on cooperation or as placed at risk by violence.

Provide Accountability: DSIs can provide accountability that is often missing when people act in a group.[127] Your content can make it clear that people are paying attention to each other's actions and can remind audience members

about consequences for negative actions (e.g. economic loss, punishment). By making it clear and public that some people are choosing to act morally, messages can serve as a reminder that people have a choice in their actions.

Use Aspirational Messaging: Messages that show the benefits of a positive behavior (the benefits of peace instead of just the cost of violence) can show individuals what they have to gain by acting peacefully and be incorporated into your content strategy.

Correct Descriptive Norms (Norms-Based Campaigns): People behave based on how others in the groups they associate themselves with behave. Correcting misperceptions of descriptive norms (what people think everyone else is doing or approves/disapproves of) can influence behavior.[128] People often overestimate the prevalence of negative behaviors and opinions (e.g., how many people spread or approve of dangerous speech), and act based on this misperception.[129] Norms-based campaigns usually conduct initial research about perceptions (what people think the norms are), and then conduct surveys of actual behavior and opinions. This research identifies gaps between perceptions and reality, and your message content can correct misperceptions.[130] If you decide to target social norms, there are several key insights on norms-based communication campaigns that you can keep in mind:

- Giving people information about actual norms will be most effective if it relates to the most relevant peer group possible (i.e., the people they see as being most like them).[131] In a dangerous speech situation, this may mean developing data about norms within specific groups, such as religious leaders or members of a specific faith. This information should always be true, so that it does not backfire and reduce your credibility.

- Make desired attitudes and behavior active and visible rather than passive and invisible. Dangerous speech tends to be very visible, which can lead people to believe that everyone else agrees with its message. You can increase the visibility of the attitudes and behavior you want to promote and create the perception that many or most people actively endorse and engage in them. Social pressure (e.g., consciousness of being watched by

peers who are known to approve or disapprove of a behavior) can make people behave in line with social norms, even if they don't agree with the norm.[132] You can try to change peoples' perceptions of what other people will think of them if they do or don't engage in a behavior to create social pressure for the attitudes and behavior you seek to promote.[133]

- Communication can be combined with incentives for people to show their support of your intervention – for example, a successful campaign to reduce college binge drinking used monetary incentives to increase student use of campaign pins and posters.[134]

- Norms generally begin to change with a small group of individuals who are committed to a new norm; when enough people buy in, there is large-scale norm change.[135] Pro-dangerous speech norms that are already in place may make people feel unsafe expressing their opinion. Openly challenging these norms and "the pressure to conform" and creating the impression that there is support for countering dangerous speech can reduce social pressure and help prompt norm change.[136] You can also create spaces (this could include specific social connections and networks, not just physical spaces) where people feel safe engaging in DSI.[137]

- One risk with descriptive norm campaigns is the "boomerang effect" in which individuals who are behaving favorably compared to the norm behave worse when they learn that they are over-performing compared to others. You can mitigate this risk by conveying social approval for people's exemplary actions (e.g., congratulating them).[138]

Note: To complement a norms-based approach, you can attempt to change peoples' attitude towards a behavior, specifically what they think the outcome of the behavior will be, and

whether they feel capable of doing the behavior. For example, you could add positive beliefs and target negative beliefs about what will happen if someone speaks out against dangerous speech, and increase their belief that they can do it (for example, to make someone who thinks, "I can't speak out against dangerous speech because I will get nervous and I won't be able to do it," to feeling like speaking out is something they can do).[139] Messages can build on people's positive beliefs in their ability to do the behavior, attack negative beliefs, or provide support (e.g. providing a message, poster, flyer that they can share) to help them do it.[140]

Example from the Field: Within Rwandan Muslim communities, there was a safe space where people could trust that others would collaborate to resist participation in the genocide. The strength of the resistance norm and identity meant that anyone who acted or suggested acting otherwise faced push-back from others.[141] This also facilitated open communication to plan and prepare resistance, while in many other groups this was not possible (e.g., there was fear that a neighbor might turn you in if you discussed a desire to resist).[142]

Distract: In heated moments when individuals are considering what actions to take, your messages can try to distract them or redirect their attention, giving them time to calm down and be more thoughtful in their decision-making.[143]

Get There First: Psychological Preparation: People form opinions based on what they already know and believe, so you can prepare people to resist and counter dangerous speech by giving them a framework to interpret future dangerous speech and related events. If you foresee increased dangerous speech and incitement, work to create a strong identity with clear norms for what actions are approved and normal in relation to dangerous speech. You can also frame the situation in advance, for example by saying that people will try to cause divisions to serve their own interests, and people should not listen when they hear such propaganda. This type of content may be able to change how people interpret dangerous speech. For example, the leadership of both the Muslim community in Rwanda and of Tuzla in Bosnia prepared the community for incitement and set clear expectations for how they should act.[144] If you use this approach, consider the risk of creating alarm and feeding into existing fears, rumors, or propaganda.

Use Stories & Narrative Media: Stories and narrative media can provide information in a non-threatening way and exemplify positive behavior. They can help people consider relevant situations and learn about the risks of dangerous speech. For example, the Musekeweya Radio drama in Rwanda was able to shift social norms around speaking up while Sawa Shabab's use of radio drama provided opportunities for audience members to discuss potential situations and think about what they would do in those situations.[145]

Use the Medium Itself to Convey a Message: Mediums and actions can be infused with meaning and used to send a message. This can mean getting people to wear a certain colored wristband or T-shirt that shows solidarity with opposition to dangerous speech, or infusing meaning into an action (such as lighting a candle in a window, or posting a picture to social media with a specific pose or symbol -- like the peace sign -- the communicates a specific meaning).

Part 3:

Risk Analysis

It is important to analyze risk for each piece of content you create and disseminate. You should consider three types of risks: (1) "Do No Harm" risks; (2) risks to the speaker or initiative; (3) impact risks. As with your Medium and Speaker strategies, identify the risks; then analyze whether they have a high, medium, or low likelihood of occurrence and potential negative impact; then adapt your strategy accordingly.

First Do No Harm

Identification: The most important type of risk analysis for messages is to ensure that there won't be unintended negative consequences (for example, increased risk of violence, or risks to individuals). Could your message or overall strategy increase the risk of violence by creating or contributing to a perception that many or most people are participating in dangerous speech and/or group-targeted harm? By creating alarm, panic, or starting a rumor? In any other way? Also consider whether the message could lend credibility or increase awareness about existing dangerous speech content, or decrease credibility and influence of messages that are currently having a positive impact. Finally, consider whether the message could put any individual, group, or institution at risk.

Mitigation: If there is a medium or high likelihood that a message will create harm, you should not use that message. If there is a low likelihood that the message will cause harm, but the harm would be medium or high, you should not use that message. For other categories, use your best judgment to ensure that messages will not create harm, and consider if there are ways to change the messages or message delivery to minimize risk.

Risks to the Speaker and/or Initiative

Identification: A message can put its speaker (an individual, organization, or your intervention as a whole) at risk of targeted backlash or of being discredited. If a speaker or the intervention overall is discredited, it will be difficult to maintain influence with the relevant audience groups. Consider whether the message poses a risk to the reputation of the speaker and/or intervention (e.g., by making a promise that you cannot ensure will be kept, providing inaccurate information or information that may become inaccurate, or by making the speaker/intervention seem "immoral" to the audience). You can also assess whether the message makes the speaker or the intervention appear biased. Finally, evaluate whether the message could put the speaker, intervention, or any other actor at risk of legal action or any other type of harm or censorship.

Mitigation: You can include your speaker(s) in this decision. If there is a medium or high likelihood of a medium or high impact risk, you should change the message. If there is a low likelihood but high or medium impact, work with the speaker to develop a strategy. For other categories, use your best judgment, and consider if there are ways to change the messages or message delivery to minimize risk.

Impact Risks

Identification: You should screen each message to see if there is a risk that it will prompt motivated reasoning or a "backfire effect."

Mitigation: The best way to ensure that each message has the intended impact is to develop messages based on a clear understanding of the audience it is intended to reach.

Risk Analysis Chart
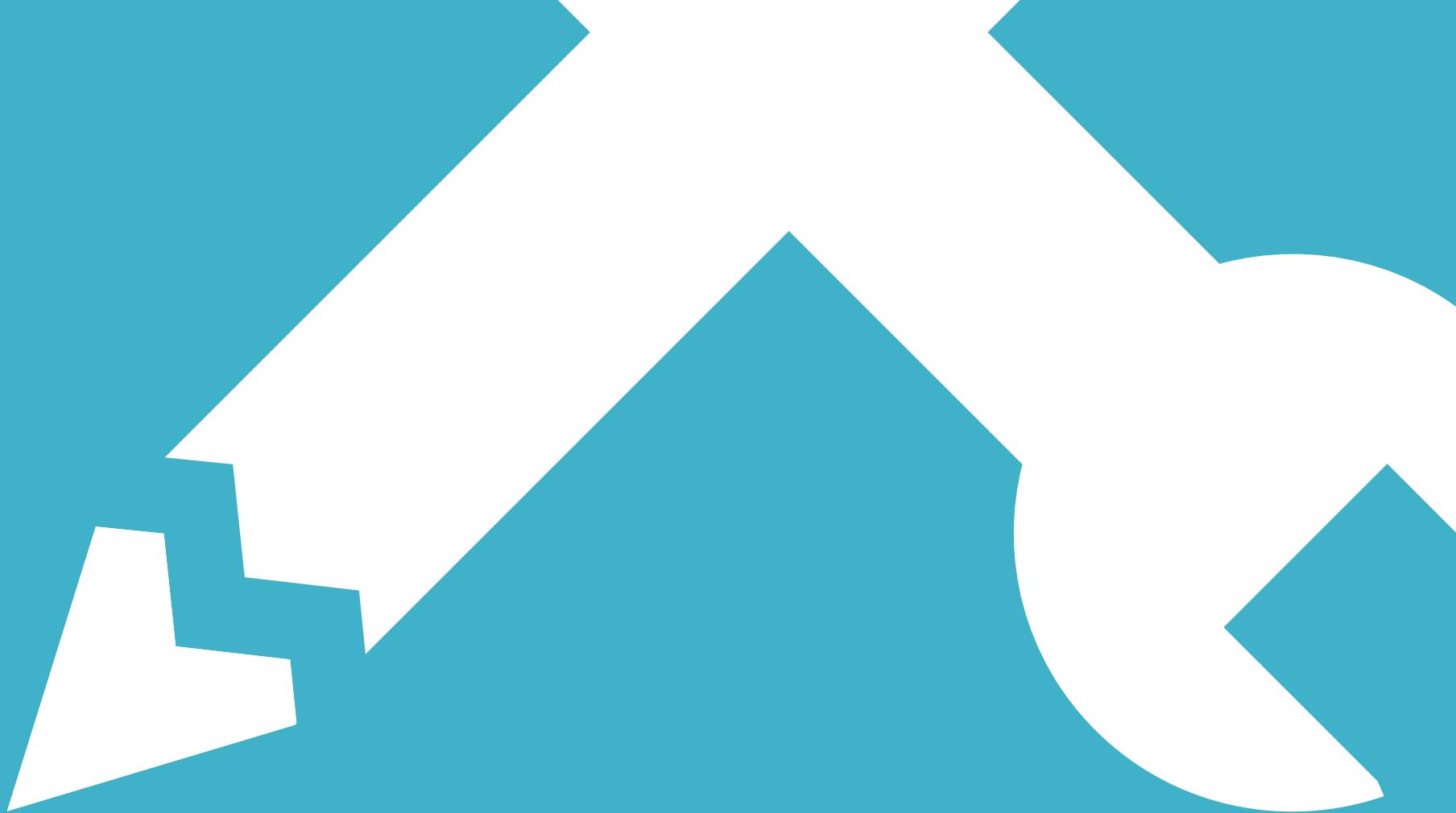
Likelihood:	Impact		
	High	**Medium**	**Low**
High	High Likelihood/ High impact	High Likelihood/ Medium Impact	High Likelihood/ Low Impact
Medium	Medium Likelihood/ High Impact	Medium Likelihood/ Medium Impact	Medium Likelihood/ Low Impact
Low	Low Likelihood/ High Impact	Low Likelihood/ Medium Impact	Low Likelihood/ Low Impact
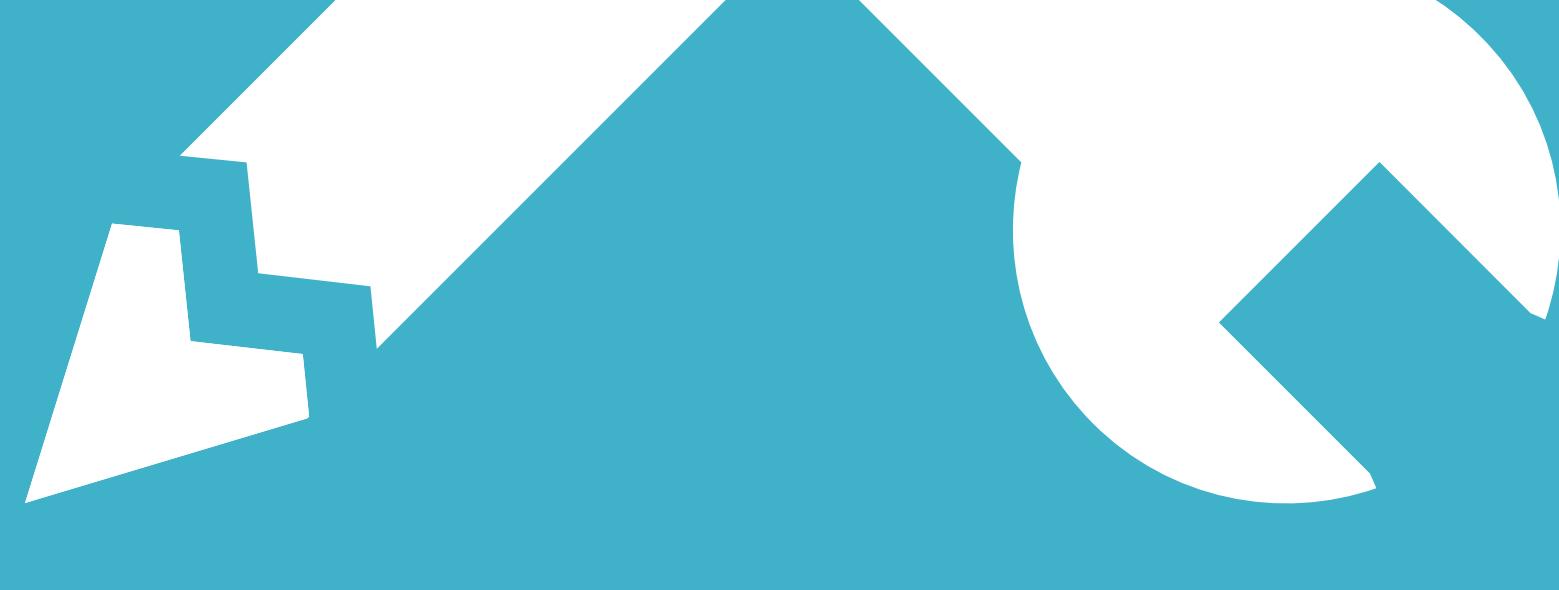
Part 4:

Iteration

As you send out message content, it will be important to find out if it is having its intended impact, and if it is having a negative impact. Are people talking about your content? Making decisions based on it? Using it to fuel rumors or dangerous speech? By setting up feedback loops within the communities your intervention reaches, e.g., through local partners, you can constantly learn more about the impact of your messages and refine and change your messages and assumptions for greater impact.

Toolbox

Focus Groups:[146] Focus groups can be used to:

- Learn about behavioral incentives, drivers, and barriers for different audience groups.
- Learn more about how people discuss dangerous speech, discrimination, and group-targeted harm so you can decide where and how to shape conversations.
- Test message content and get audience feedback. You can see how messages are received, and ask the group to talk about how they think their peers or other communities would react to the messages. You can also test how message content that is developed for a specific audience is interpreted by a different audience group (e.g., to identify risks).
- Focus groups can also be used to co-develop content (see below).

Focus groups bring people together for a group discussion. They generally ask participants to discuss specific topics, questions, or materials, which can range from the group's views of dangerous speech or recent events to their feedback about specific DSI content. A focus group should include 4-8 people so you can get insights from all participants. You can interview participants separately before bringing them together for a joint discussion.

You can identify the types of participants you need (from specific audience groups or multiple audience groups, e.g., "we want five information spreaders from a specific rural area") then identify and recruit people who fit this description. Consider the best way to explain the exercise to participants; if the topic is sensitive, you can design the focus group to cover a different topic or more than one topic. For example, you can design the focus group to consider topics (e.g., the local economy, politics, marriage, religion) that will prompt a conversation that can provide useful information about dangerous speech and conflict dynamics even if it does not discuss them directly.

Friendship Groups:[147] Friendship groups can be used to:

- Learn more about how people discuss dangerous speech, discrimination, and group-targeted harm. Understanding how these conversations play out can help you understand where and how to intervene to try to push them in a different direction.
- Learn how different social groups impact people's behavior and opinions by comparing individual interviews and social conversations.
- Test message content and get feedback, especially about how people will react to messages when they are received in a social setting.

Friendship groups are similar to focus groups, but instead of identifying participants separately, they bring together small groups of friends and family (2-5 people) and you observe them discussing relevant topics. You can facilitate the conversation by asking questions for the group to discuss with each other. These questions should be related to your goals (e.g., get people to talk about stereotypes; grievances and what should be done about them; their group identity; their fears; their hopes for the future). You can create a friendship group for an audience group by identifying a member of that audience group (for example, an information spreader or a reluctant audience member), and asking him/her to invite friends or family members to a discussion. You can interview participants separately then bring them together for a joint discussion. It is important to make all participants feel comfortable, and to ensure that they don't feel judged. Similarly to focus groups, consider how to explain the exercise in order to recruit participants.

Message Co-creation: Message co-creation content can be used in multiple ways:
- The messages that the group creates (and their conversations around these messages) can be analyzed for content (e.g., you can identify common behavioral drivers and barriers for different audience groups), tone, language, and risk, then edited to become final content.
- The content can be edited into templates that you can use and edit as events unfold.
- You can use the content and conversations to identify useful tones, language, incentives, and behavioral drivers and barriers, and incorporate these insights into message guidelines.
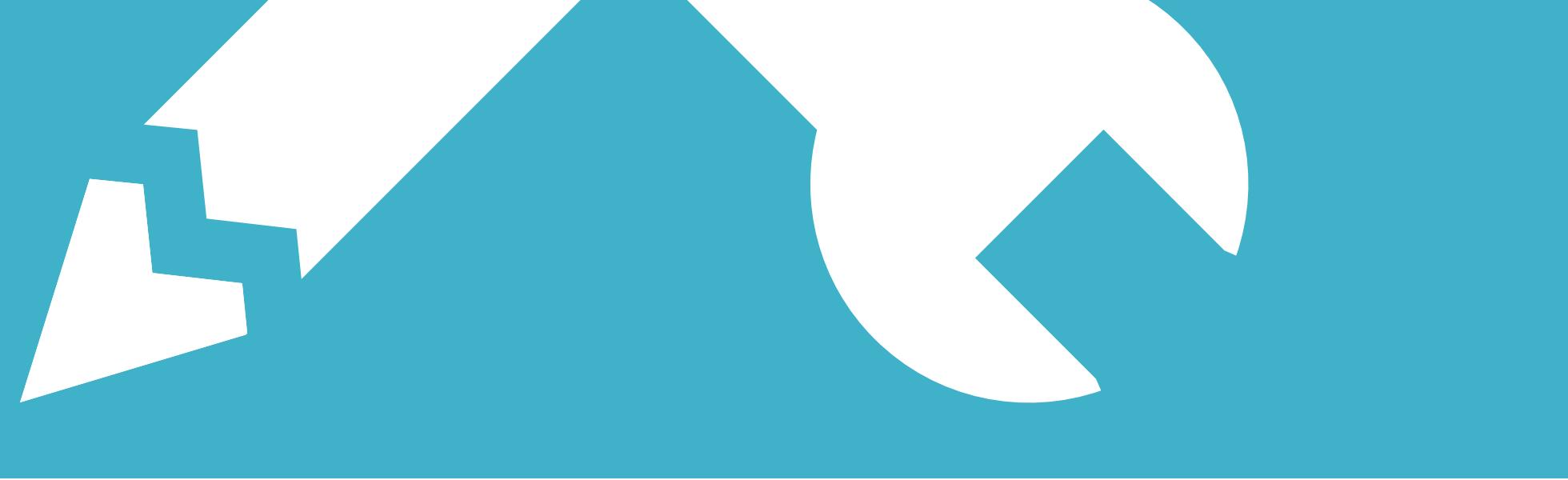
Message co-creation involves developing messages with the audience groups you aim to influence. You present the audience groups with specific scenarios that you want to develop content for, and ask participants what they would say to a peer to persuade him or her to take or not take an action (e.g., not to spread a rumor). There are two main ways you can do this:

1. Developing Messages for Specific Triggers/Decision Moments: You can ask people to develop content that can influence people during specific conflict triggers or decision-making moments. Outline the series of audience behaviors you are trying to change (see Audience Journeys), and create behavior chains for specific situations or scenarios you want to focus on.

Use these behavior chains to pose questions. For example, if you create a behavior chain in which you want an audience group to ask questions about a rumor before spreading it, you can show this behavior chain to participants, and ask: "if you had a friend who was in this situation and was considering whether or not to spread a rumor, what would you say to convince him/her to ask a question about it?" (You can get even more specific, for example by specifying a medium: "What would you want him/her to hear on the radio?" "What would you want his/her pastor to say?"). It's helpful to ask participants what they would tell someone else, because it lets them be honest without taking the risk of talking about themselves directly. You can find out whether each message has its intended behavioral impact based on these behavior chains.

2. Developing Messages for Long-term Change or Education: You can involve participants in creating messages for long-term goal items. For example, if the goal is to debunk a specific stereotype, you could ask people what might make them or a friend reconsider or question that stereotype. If you want to educate people about specific topics, you can ask what people are confused about, and have a group of participants design educational content. You can also have groups review content to make sure that it is accessible for the target group (that the language is easy, understandable, appealing, etc.).

Think about how you want people to develop the content. Do you want them each to come up with content then share it with the group and discuss? Do you want them to work in small groups of 2-3 to come up with the messages, then present back and discuss? When they share, do you want them to present, or do you want everyone to put their content in anonymously then have the facilitator read the content back to the group? As you decide which approach to use, remember that the aim is to make the participants as comfortable, honest, and thoughtful as possible so you can get accurate information.

Ideally you can do this exercise with members of the audience group the messages are targeting. It may be hard to recruit some audience groups: for example, dangerous speech speakers may not be willing to participate. To address this challenge, consider:

- Recruit people who used to be a member of the group but are no longer part of it (e.g., people who spread dangerous speech or participated in violence but don't anymore).
- Recruit people who are at risk (people who are likely to participate

in dangerous speech but haven't yet or people who are participating reluctantly) and have not fully accepted the dangerous speech message.
- Finally, you can recruit people who are close to these individuals (e.g., mothers, friends, etc.) and ask them to come up with messages. For example, "if you had a boyfriend who was unemployed, politically active, and upset about the recent price increases in the market place, and he was approached to attend a rally against Group X, what would you say to persuade him not to do it?"

Observational Research: You can use observational research to figure out the type of language (the language itself, common phrases, slang, and the general tone) that people use, and learn about how people respond to communications and events (e.g., how people are reacting to a radio show or news). It can help you stay up to date on narratives that support or counter dangerous speech and regularly update your content strategy and risk analysis. For example, if a word generally used for peace carries a connotation of "being passive" or "neglecting group grievances," you may want to replace that word in your message.

Interviews: Simple one-on-one interviews with members of audience groups or local partners can help you get feedback on questions and content. People may be more honest outside their social/group setting, especially if there is fear of speaking about these topics in public or in front of peers (for example, reluctant participants in dangerous speech may not want to talk publicly).

Endnotes

[1] Susan Benesch, "Countering Dangerous Speech: New Ideas for Genocide Prevention," (Voices that Poison, 2014). Similar types of speech are sometimes referred to as "extreme hate speech." The United Nations recognizes direct and public incitement to genocide as a crime under the *Convention on the Prevention and Punishment of the Crime of Genocide (Adopted by Resolution 260 (III) A of the United Nations General Assembly on 9 December 1948*; see http://www.hrweb.org/legal/genocide.html) and the International Covenant on Civil and Political Rights notes that "Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law" (http://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx).

[2] Benesch, Countering Dangerous Speech. For additional information about the impact and spread of rumors related to conflict, see Greenhill, Kelly and Ben Oppenheim, "Rumor Has It: The Adoption of Unverified Information in Conflict Zones."

[3] Benesch, Countering Dangerous Speech, pp.7-8.

[4] Stathis Kalyvas, "The Ontology of 'Political Violence': Action and Identity in Civil Wars," Perspectives on Politics 1, no. 3 (September 2003): 475-486

[5] For more information on how to Do No Harm, see the work of Marshall Wallace and the Do No Harm Project. "From Principle to Practice: A User's Guide to Do No Harm" by Marshall Wallace is accessible through the Principle to Practice website at http://www.principletopractice.org/from-principle-to-practice/

[6] Seth Karamage, Program Coordinator, Public Conversations Project, interview by Rachel Brown, by phone (November 6th, 2014); "Interfaith Mediation Center: A Society Free from Violence." Interfaith Mediation Center, ©2015. Web. 20 October 2015. http://www.earlywarningnigeria.org/.

[7] Seema Shah, and Rachel Brown. "Programming for Peace: Sisi ni Amani Kenya and the 2013 Elections: CGCS Occasional Paper Series on ICTs, Statebuilding, and Peacebuilding in Africa, Number 3." Center for Global Communication Studies.

[8] Anderson, Mary B., and Marshall Wallace. *Opting Out of War*: Strategies to Prevent Violent Conflict. Boulder: Lynne Rienner Publishers, Inc., 2013, pp.124-6.

[9] Amra Ali. "The city of colours." <http://www.dawn.com/news/1188776>; Zarina Khan. "Painting the Town Positive - Fighting Extremism with Inspirational Art." <http://mvslim.com/ painting-the-town-positive-fighting-extremism-with-inspirational-art/>.

[10] Mary B. Anderson and Marshall Wallace. *Opting Out of War: Strategies to Prevent Violent Conflict.* (Boulder: Lynne Rienner Publishers, Inc., 2013), pp.157-168.

[11] Theo Dolan. "Countering Hate Speech in South Sudan through Peace Radio." United States Institute of Peace, 7 May 2014. Web. 12 Nov. 2014. <http://www.usip.org/publications/ countering-hate-speech-in-south-sudan-through-peace-radio>.

[12] For example, see Brendan Nyhan and Jason Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions" (Dartmouth College, accessed at http://www.dartmouth.edu/~nyhan/opening-political-mind.pdf, December 16, 2013); and Brendan Nyhan and Jason Reifler, "When Corrections Fail: The persistence of political misperceptions" (Dartmouth, forthcoming in Political Behavior, Accessed at http://www.dartmouth.edu/~nyhan/nyhan-reifler.pdf) pp.2-3,5-6, and pp.28-30. The process of accepting or rejecting information based on whether it is consistent with exiting beliefs is also known as biased assimilation (Pranav Dandekar, Ashish Goel and David T. Lee, "Biased assimilation, homophily, and the dynamics of polarization," PNAS early edition accessed through Stanford.edu, October 2012.). For more information, also see Nichole Argo, Shamil Idriss and Mahnaz Fancy, "Media and Intergroup Relations: Research on Media and Social Change," Research Report (Soliya in partnership with Alliance of Civiliations, 2009), pp.9 - 14

[13] Brendan Nyhan and Jason Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions," pp.2-3,5-6, and 28-30; and Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.9.

[14] Karamage, Seth, Program Coordinator, Public Conversations Project, interview by Rachel Brown.

[15] Deborah J. Terry and Michael A. Hogg, "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification," pp.776; Schultz et al., "The Constructive, Destructive, and Reconstructive Power of Social Norms," pp. 430-33; and Gerber, Green, and Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment," pp.34, 40, 42.

[16] Rui J.P. de Figueiredo, Jr. and Barry R. Weingast, "The rationality of fear: Political opportunism and ethnic conflict." 1999. In *Civil Wars, Insecurity, and Intervention*, eds. Walter and Snyder. New York: Columbia University Press, pp. 261-302. Working Paper accessed online at <http://isites.

harvard.edu/fs/docs/icb.topic1018018.files/CIAO__The_Rationality_of_Fear__Political_Opportunism_and_Ethn...pdf>.

[17] Benesch, Countering Dangerous Speech.

[18] "Norms, Narratives, and Neurons" The Neuroscience and Social Conflict Initiative Meeting Report, 15-16 March 2014 at MIT Media Lab, Beyond Conflict (http://www.beyondconflictint.org/neuroscience-and-social-conflict-initiatives/) with MIT Political Science and Saxelab, citing Kelly Greenhill, p.19.

[19] Information on observational research provided by Alison Demos (interview by Rachel Brown on August 20th, 2014).

[20] Dietram A. Scheufele, "Framing as a Theory of Media Effects," *Journal of Communication* (International Communication Association), Winter 1999: 103-122, p.107, referencing Entman (1993) p.52.

[21] Dennis Chong and James N. Druckman, "Framing Theory," Annual Review of Political Science (Annual Reviews) 10 (2007): pp.103-126, p.112; Jonathan Haidt, *The Righteous Mind: Why Good People are Divided by Politics and Religion* (New York: Vintage Books, 2012), p.104.

[22] Chong and Druckman, "Framing Theory," p.112

[23] Scheufele, "Framing as a Theory of Media Effects," p.107, referencing Entman (1993) p.104.

[24] Alison Demos, interview by Rachel Brown (August 20th, 2014).

[25] For example, Umati, a research project in Kenya that monitored dangerous speech in the lead-up and throughout Kenya's 2013 elections, monitored and categorized social media posts that fell under the category of dangerous speech. (iHub Research. *The Umati Project: Monitoring Dangerous speech Online*. Retrieved December 16, 2014, from http://www.ihub.co.ke/umati).

[26] Research on the role of women in different contexts has shown that women are more likely than men to adopt a broad definition of peace, and thus they might be more likely to accept the role of speech in facilitating violence. Women are likely to be collectively organized, and often undervalue their role in preventing violence. They can play key roles, such as mediating between actors. Women also play negative roles in spreading rumors and misinformation and in supporting men who physically perpetrate violence. Ivan Cardona,

Patricia Justino, Becky Mitchell and Catherine Muller, *From the ground up: Women's roles in local peacebuilding in Afghanistan, Liberia, Nepal, Pakistan and Sierra Leone*, Research report (ActionAid USA (accessed via ActionAid website), Institute of Development Studies, and Womankind Worldwide, September 2012); pp.5-6.

[27] The attitudes and involvement concept is adapted from the marketing concept of "affinity" and "engagement." Information on this concept (and information that helped adapt it to this context) was provided through ongoing conversations with Grant Tudor and interviews with Margot Grover, Kacy Erdelyi, and Alvaro Cabrera.

[28] The conception of Audience Profiles provided in this guide is adapted from the use of "Personas" in advertising. The use and understanding of this concept comes through interviews with a variety of people in the marketing sector, most importantly Alvaro Cabrera and Grant Tudor.

[29] Behavioral drivers and barriers are a concept used in marketing. The author became familiar with these concepts through collaboration with Grant Tudor on Sisi ni Amani Kenya's brand advocacy development.

[30] Elizabeth Levy Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda,", p.576.

[31] Ibid, p.575 (drawing upon the work of D.T. Miller, Monin & Prentice, 2000).

[32] Ibid, p.576.

[33] Ibid.

[34] Ravi Bhavnani. "Ethnic Norms and Interethnic Violence: Accounting for Mass Participation in the Rwandan Genocide."

[35] P. Wesley Schultz, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein and Vladas Griskevicius, "The Constructive, Destructive, and Reconstructive Power of Social Norms," *Psychological Science* 18, no. 5 (May 2007): pp.429-434, p.429.

[36] Ibid, p. 430.

Endnotes

[37] Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," p.575.

[38] Thorson, Emily. "Theory of Planned Behavior Handout."

[39] Alvaro Alvaro interview by Rachel Brown, (August 19th, 2014).

[40] Nyhan and Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions," pp.2-3,5-6, and pp.28-30.

[41] Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," p.576 (surveying additional literature).

[42] See, for example, Omar S. McDoom, "The Psychology of Threat in Intergroup Conflict: Emotions, Rationality, and Opportunity in the Rwandan Genocide." *International Security* 37.2 (2012): 119-55. Web. 3 Nov. 2014. p. 128.

[43] For example, CureViolence, a violence prevention program that uses "interrupters" to interrupt and mediate conflicts, has found that providing a cool down period for emotions is a key component to preventing violent escalation and retaliation. "The Cure Violence Health Model." Cure Violence, 2012. Web. 3 Nov. 2014.

[44] You can include more people if you have specific reasons for needing a larger group, such as wanting to observe broad social interactions.

[45] Margot Grover, interview by Rachel Brown (August 19th, 2014).

[46] Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," p.584; Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.34.

[47] Alvaro Cabrera, interview by Rachel Brown.

[48] For example, during the Rwandan genocide, radio was used to prepare and engage people in killing, but face-to-face engagement played a key role in actual recruitment into killing groups (Lee Ann Fujii, "The Power of Local Ties: Popular Participation in the Rwandan Genocide (2008, September 15). *Security Studies*, 17(3), pp.568-597, p.580.)

[49] Shah and Brown, "Programming for Peace: Sisi ni Amani Kenya and the 2013 Elections: CGCS Occasional Paper Series on ICTs, Statebuilding, and Peacebuilding in Africa, Number 3."

[50] For example, see Sanjana Hattotuwa, "Untying the Gordian Knot: ICT for Conflict Transformation and Peacebuilding," *Dialogue* 2, no. 2 (2004): pp.39-68.

[51] Anderson and Wallace, *Opting Out of War: Strategies to Prevent Violent Conflict*, p. 165.

[52] Ibid.

[53] Amra Ali, "The city of colours." <http://www.dawn.com/news/1188776>; Zarina Khan. "Painting the Town Positive - Fighting Extremism with Inspirational Art." <http://mvslim.com/painting-the-town-positive-fighting-extremism-with-inspirational-art/>."

[54] Seth Karamage, interview by Rachel Brown; "Interfaith Mediation Center: A Society Free from Violence," Interfaith Mediation Center, ©2015. Web. 20 October 2015. http://www.earlywarningnigeria.org/.

[55] Theo Dolan. "Countering Hate Speech in South Sudan through Peace Radio."

[56] Matthew Schissler, "Echo Chambers in Myanmar: Social media and the ideological justifications for mass violence," (2014, September 24).

[57] Ibid.

[58] Ali, Amra. "The city of colours." <http://www.dawn.com/news/1188776>.

[59] For a description of how this plays out in Myanmar, see Schissler, "Echo Chambers in Myanmar: Social media and the ideological justifications for mass violence." For further information about the concept itself, see also Cass R. Sunstein, *Republic.com* 2.0. Princeton University Press (2007).

[60] Ibid.

[61] Friend Me 4 PEACE." *Peace Factory: A bridge of communication between the people in the middle east*. Peace Factory, n.d. Web. 8 Dec. 2014. <http://thepeacefactory.org/friend-me/>.

[62] For example: Francesco Mancini (editor), "New Technology and the Prevention of Violence

and Conflict," Alliance for Peacebuilding, International Coalition for the Responsibility to Protect, UNDP, USAID, and the International Peace Institute (Alliance for Peacebuilding (Accessed at http://www.allianceforpeacebuilding.org/2014/01/new-technology-and-the-prevention-of-violence-and-conflict/); Hattotuwa, "Untying the Gordian Knot: ICT for Conflict Transformation and Peacebuilding," pp.16-17; Jan H. Pierskalla and Florian M. Hollenbach, "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa," *American Political Science Review* (accessed via http://polisci.duke.edu/uploads/media_items/technology-collectiveactioncellphoneviolence.original.pdf) (2013): pp.1-18; Helena Puig Larrauri, "The Use of New Technologies: Expanding Opportunities for Peacebuilding?," Paper, Accessible Online, KOFF: Swiss Peace (KOFF: Swiss Peace, 2014); and Helena Puig Larrauri and Anne Kahl, "Technology for Peacebuilding," *Stability: International Journal of Security & Development* 2, no. 3 (2013): pp. 1-15.

[63] The Do No Harm theories and practical suggestions provided in this Guide (including the concept of "connectors" and "dividers") are based on the work and guidance of Marshall Wallace and the Do No Harm Project. For more information on how to Do No Harm, see the work of Marshall Wallace and the Do No Harm Project. "From Principle to Practice: A User's Guide to Do No Harm" by Marshall Wallace is accessible through the Principle to Practice website at http://www.principletopractice.org/from-principle-to-practice/.

[64] Ibid.

[65] See, for example, Anahi Ayala Iacucci, "The conundrum of digital humanitarianism: when the crowd does harm," *Diary of a Crisis Mapper blog*, November 15, 2013.

[66] For example, see: Mancini (editor), "New Technology and the Prevention of Violence and Conflict"; Hattotuwa, "Untying the Gordian Knot: ICT for Conflict Transformation and Peacebuilding," pp.16-17; and Pierskalla and Hollenbach, "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa."

[67] Puig Larrauri, "The Use of New Technologies: Expanding Opportunities for Peacebuilding?" p.4.

[68] For more information, see Puig Larrauri, Helena, "*The Use of New Technologies: Expanding Opportunities for Peacebuilding?*,", p.4; and Helena Puig Larrauri and Anne Kahl, "Technology for Peacebuilding."

[69] Ibid.

[70] National Science Foundation, "Who's afraid of the HPV vaccine?," (Press Release 10-006), January 13, 2010, accessed at http://www.nsf.gov/news/news_summ.jsp?cntn_id=116186, (quoting Dan Kahan); and Dan Kahan, "Cultural Cognition and Safeguarding the Science Commons" keynote address, accessed via *Yale Climate & Energy Institute* at http:/climate.yale.edu/videos/cultural-cognition-and-safeguarding-science-commons-dan-kahan (Yale, August 4, 2014)

[71] Seth Karamage, interview by Rachel Brown; and "Interfaith Mediation Center: A Society Free from Violence," Interfaith Mediation Center, ©2015. Web. 20 October 2015. http://www.earlywarningnigeria.org/.

[72] Lee Ann Fujii, "The Power of Local Ties: Popular Participation in the Rwandan Genocide (2008, September 15). *Security Studies*, 17(3), pp.568-597.

[73] "Norms, Narratives, and Neurons" The Neuroscience and Social Conflict Initiative Meeting Report, 15-16 March 2014 at MIT Media Lab, Beyond Conflict (http://www.beyondconflictint.org/neuroscience-and-social-conflict-initiatives/) with MIT Political Science and Saxelab, citing Dr. Emily Falk, p.14

[74] Kahan, "Cultural Cognition and Safeguarding the Science Commons," keynote address; Chong and Druckman, "Framing Theory," pp. 111-112.

[75] Nyhan and Reifler, New America Foundation, "Misinformation and Fact-Checking: Research Findings from Social Science," p.14.

[76] David DeSteno, Duane T. Wegener, Richard E. Petty, Derek D. Rucker and Julia Braverman, "Discrete Emotions and Persuasion: The Role of Emotion-Induced Expectancies," *Journal of Personality and Social Psychology* 86, no. 1 (2004): 43-56, pp. 49, 52.

[77] Alison Demos, interview by Rachel Brown (August 20th, 2014).

[78] An example from another subject area is that organizations working in public health have found that influential community leaders engaging in a behavior provides the opportunity for other community members to mimic the behavior, an important step towards creating a social norm. (Peer Pressure can be a Lifesaver, by Helen Coster http://opinionator.blogs.nytimes.com/2014/05/14/peer-pressure-can-be-a-lifesaver/?_php=true&_type=blogs&_r=0).

Endnotes

[79] Benesch, Countering Dangerous Speech.

[80] Schissler, "Echo Chambers in Myanmar: Social media and the ideological justifications for mass violence."

[81] The Do No Harm theories and practical suggestions provided in this Guide (including the concept of "connectors" and "dividers") are based on the work and guidance of Marshall Wallace and the Do No Harm Project. For more information on how to Do No Harm, see the work of Marshall Wallace and the Do No Harm Project. "From Principle to Practice: A User's Guide to Do No Harm" by Marshall Wallace is accessible through the Principle to Practice website at http://www.principletopractice.org/from-principle-to-practice/.

[82] Ibid.

[83] The brand building section in this Guide is based on collaboration with Grant Tudor (Founder & CEO of Populist Group, www.populistgroup.org) and experience building Sisi ni Amani Kenya's brand.

[84] Categories included here are based on work done with Grant Tudor (Founder & CEO of Populist Group, www.populistgroup.org) on Sisi ni Amani Kenya's Brand Advocacy Strategy.

[85] Benesch, Countering Dangerous speech. Benesch notes that dangerous speech was more potent in mother tongue than in the national languages in Kenya

[86] Taya R. Cohen and Chester A. Insko, "War and Peace: Possible Approaches to Reducing Intergroup Conflict," *Perspectives on Psychological Science* 3, no. 2 (2008): 87-93, p.91; Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," pp.575-6; and Thorson, Emily "Some of My Best Friends are Poor? Income Misperceptions and Policy Attitudes," Academic Paper, George Washington University, p.7.

[87] Thorson "Some of My Best Friends are Poor? Income Misperceptions and Policy Attitudes," p.8.

[88] Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," pp.582, 584.

[89] This consideration was outlined in a phone interview with Dr. Emile Bruneau (Bruneau, Emile. Phone interview by Rachel Brown, *PhD, Post-Doctoral Fellow, SaxeLab, MIT,* (August 11, 2014); Bruneau EG, Cikara M, Saxe R, "Minding the Gap: Narrative Descriptions about Mental States Attenuate Parochial Empathy"; Cohen and Insko, "War and Peace: Possible Approaches to Reducing Intergroup Conflict," p.91.

[90] Thorson "Some of My Best Friends are Poor? Income Misperceptions and Policy Attitudes."

[91] This consideration was outlined in a phone interview with Dr. Emile Bruneau (Bruneau, Emile. Phone interview by Rachel Brown, *PhD, Post-Doctoral Fellow, SaxeLab, MIT,* (August 11, 2014); Bruneau EG, Cikara M, Saxe R, "Minding the Gap: Narrative Descriptions about Mental States Attenuate Parochial Empathy"; Cohen and Insko, "War and Peace: Possible Approaches to Reducing Intergroup Conflict," p.91.

[92] This consideration was outlined in a phone interview with Dr. Emile Bruneau (Emile Bruneau, phone interview by Rachel Brown, August 11, 2014). Bruneau EG, Cikara M, Saxe R, "Minding the Gap: Narrative Descriptions about Mental States Attenuate Parochial Empathy."

[93] "Dehumanization in Conflict: Meeting Report 16-17 February, 2013" The Neuroscience and Social Conflict Initiative of Beyond Conflict (http://www.beyondconflictint.org/neuroscience-and-social-conflict-initiatives/) in partnership with the SaxeLab.

[94] Bruneau, Dufour and Saxe, "Social cognition in members of conflict groups: behavioural and neural responses in Arabs, Israelis and South Americans to each other's misfortunes."

[95] This consideration was outlined in a phone interview with Dr. Emile Bruneau (Emile Bruneau, phone interview by Rachel Brown, August 11, 2014). Bruneau EG, Cikara M, Saxe R, "Minding the Gap: Narrative Descriptions about Mental States Attenuate Parochial Empathy."

[96] "Dehumanization in Conflict: Meeting Report 16-17 February, 2013" The Neuroscience and Social Conflict Initiative of Beyond Conflict (http://www.beyondconflictint.org/neuroscience-and-social-conflict-initiatives/) in partnership with the SaxeLab, p.15.

[97] This consideration was outlined in a phone interview with Dr. Emile Bruneau (Emile Bruneau, phone interview by Rachel Brown, August 11, 2014).

[98] For example see Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.14; and Nyhan and Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions," p.4.

99 Nyhan and Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions," pp.2-3,5-6, and 28-30; and Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.9.

100 John Cook and Stephan Lewandowsky. "The Debunking Handbook." University of Queensland, Global Change Institute, and The University of Western Australia, January 23, 2012 (initially published Nov. 2011); Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation; Brendan Nyhan phone interview by Rachel Brown, (June 29th, 2015); and Kelly Greenhill and Ben Oppenheim, "Rumor Has It: The Adoption of Unverified Information in Conflict Zones."

101 Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.9; and Nyhan and Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions."

102 Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.9; Geoffrey L. Cohen, David K. Sherman, Anthony Bastardi, Lillian Hsu and Michelle McGoey, "Bridging the Partisan Divide: Self-Affirmation Reduces Ideological Closed-Mindedness and Inflexibility in Negotiation," *Journal of Personality and Social Psychology* (American Psychological Association) 93, no. 3 (2007): pp.415-430; and Benedict Carey, "Cede Political Turf? Never! Well, Maybe." *New York Times*, November 1, 2010.

103 Nyhan and Reifler, "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions."

104 Argo, Idriss and Fancy, "Media and Intergroup Relations: Research on Media and Social Change," p.35.

105 Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation.

106 Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation.

107 Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation; Cook and Lewandowsky. "The Debunking Handbook," p.5; Greenhill, Kelly and Ben Oppenheim, "Rumor Has It: The Adoption of Unverified Information in Conflict Zones."

108 Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation.

109 Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation.

110 C. McCauley and S. Moskalenko, "Mechanisms of Political Radicalization: Pathways Toward Terrorism," (2008).

111 For example, see Terry and Hogg, "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification," p. 776; Schultz et al., "The Constructive, Destructive, and Reconstructive Power of Social Norms," pp. 430-33; Gerber, Green, and Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment," pp.34, 40, 42; and de Figueiredo and Weingast, "The rationality of fear: Political opportunism and ethnic conflict."

112 Williams and DeSteno, "Pride and Perseverance: The Motivational Role of Pride," pp. 1007-8.

113 Ibid.

114 For example, Mina Cikara, A.C. Jenkins and R. Saxe, "Reduced self-referential neural response during intergroup competition predicts competitor harm," NeuroImage 96 (2014): 36-43.

115 Cohen and Insko, "War and Peace: Possible Approaches to Reducing Intergroup Conflict," p.92.

116 For example, many low-level participants in the Rwandan genocide saw themselves as part of a group, depended on the group to make decisions about their actions, and "tended to see themselves as powerless," while people who openly resisted "did not see themselves as powerless," and made decisions as individuals (Fujii, "The Power of Local Ties: Popular Participation in the Rwandan Genocide," p. 585). DSIs could help provide group identity and guidance for individuals who are likely to look to a group for guidance on how to act and who may feel powerless if acting alone.

117 Cikara, Jenkins, and Saxe, "Reduced self-referential neural response during intergroup competition predicts competitor harm," pp. 37, 42. In addition, in research on micro-level decision-making by the lowest level participants in the Rwandan genocide, Fuji finds that a key difference between people who joined in the killing versus resisted was that "resisters

Endnotes

did not relinquish their agency to that of a group but continued to act as individuals." (Fujii, "The Power of Local Ties: Popular Participation in the Rwandan Genocide," p. 587).

[118] Nyhan and Reifler, "Misinformation and Fact-Checking: Research Findings from Social Science," New America Foundation, p.17.

[119] Scheufele, "Framing as a Theory of Media Effects," p.107, (referencing Entman (1993) p.52).

[120] Chong and Druckman, "Framing Theory," p.112; Haidt, Jonathan, *The Righteous Mind: Why Good People are Divided by Politics and Religion*, p.104.

[121] Ibid.

[122] Anderson, Mary B., and Marshall Wallace. *Opting Out of War: Strategies to Prevent Violent Conflict*. p. 165.

[123] Ibid.

[124] Ibid.

[125] Anderson, Mary B., and Marshall Wallace. *Opting Out of War: Strategies to Prevent Violent Conflict*. pp.124-6.

[126] Cohen and Insko, "War and Peace: Possible Approaches to Reducing Intergroup Conflict," p.92.

[127] Cikara, Jenkins, and Saxe, "Reduced self-referential neural response during intergroup competition predicts competitor harm," pp. 37, 42.

[128] Steve Martin, "98% of HBR Readers Love This Article," *Harvard Business Review*, October 2012; Schultz, et al., "The Constructive, Destructive, and Reconstructive Power of Social Norms,"; Tina Rosenberg, "The Destructive Influence of Imaginary Peers," *New York Times Opinionator*, March 27, 2013; H. Wesley Perkins, David W. Craig and Jessica M. Perkins, "Usings social norms to reduce bullying: A research intervention among adolescents in five middle schools," *Group Processes and Intergroup Relations* 14, no. 5 (2011): 703-722.; "Who's afraid of the HPV vaccine?" National Science Foundation; Carey, "Cede Political Turf? Never! Well, Maybe."

[129] Ibid.

[130] Schultz et al., "The Constructive, Destructive, and Reconstructive Power of Social Norms."

[131] Rosenberg, "The Destructive Influence of Imaginary Peers."; Terry and Hogg, "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification."

[132] Terry and Hogg, "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification," p. 776; Schultz et al., "The Constructive, Destructive, and Reconstructive Power of Social Norms," pp. 430-33; and Gerber, Green, and Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment," pp.34, 40.

[133] Thorson, Emily. "Theory of Planned Behavior Handout."

[134] Rosenberg, "The Destructive Influence of Imaginary Peers"; H. Wesley Perkins, and Alan D. Berkowitz. "Perceiving the community norms of alcohol use among students: some research implications for campus alcohol education programming."

[135] "Norms, Narratives, and Neurons" The Neuroscience and Social Conflict Initiative Meeting Report, 15-16 March 2014 at MIT Media Lab, Beyond Conflict (http://www.beyondconflictint. org/neuroscience-and-social-conflict-initiatives/) with MIT Political Science and Saxelab. p17.

[136] Gerber, Green, and Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment," p.42.

[137] For example, see Bhavnani, "Ethnic Norms and Interethnic Violence: Accounting for Mass Participation in the Rwandan Genocide." p. 659.

[138] Schultz et al., "The Constructive, Destructive, and Reconstructive Power of Social Norms," pp. 430-33.

[139] Terry and Hogg, "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification," p.776; and Thorson Emily, "Theory of Planned Behavior Handout."

[140] Ibid.

[141] Anderson, Mary B., and Marshall Wallace. *Opting Out of War: Strategies to Prevent Violent Conflict*. Boulder: Lynne Rienner Publishers, Inc., 2013, pp.161-2.

142 Ibid, p.167.

143 See, for example, "The Cure Violence Health Model." Cure Violence, 2012. Web. 3 Nov. 2014.

144 Anderson, Mary B., and Marshall Wallace. Opting Out of War: Strategies to Prevent Violent Conflict. Boulder: Lynne Rienner Publishers, Inc., 2013, pp.157-168 and pp. 124-6.

145 Paluck, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," and Dolan, "Countering Hate Speech in South Sudan through Peace Radio."

146 Grover, Margot, interview by Rachel Brown (August 19th, 2014).

147 Ibid.

Works Cited & Consulted

Ali, Amra. "The city of colours." *Dawn* 21 June 2015 [*Pakistan*] . Web. 2 July 2015. <http://www.dawn.com/news/1188776>.

Anderson, Mary B., and Marshall Wallace. *Opting Out of War: Strategies to Prevent Violent Conflict.* (Boulder: Lynne Rienner Publishers, Inc., 2013). Print.

Argo, Nichole, Shamil Idriss and Mahnaz Fancy, "Media and Intergroup Relations: Research on Media and Social Change," Research Report (Soliya in partnership with Alliance of Civilizations, 2009)

Bartlett, Monica Y. and David DeSteno, "Gratitude and Prosocial Behavior: Helping When It Costs You," *Psychological Science* 17, no. 4 (2006): 319-325.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of Processes in Belief: Source Recollection, Statement Familiarity, and the Illusion of Truth. *Journal of Experimental Social Psychology*, 121(4), 446-458.

Benesch, Susan, "Countering Dangerous speech: New Ideas for Genocide Prevention," published by Voices that Poison, 2014. <http://www.VoicesThatPoison.org>.

Bhavnani, Ravi. "Ethnic Norms and Interethnic Violence: Accounting for Mass Participation in the Rwandan Genocide." *Journal of Peace Research* 43.6 (2006): 651-69. Web. 11 Sept. 2014.

Bruneau, Emile G., Nicholas Dufour and Rebecca Saxe, "Social cognition in members of conflict groups: behavioural and neural responses in Arabs, Israelis and South Americans to each other's misfortunes," *Philosophical Transactions of The Royal Society B* (The Royal Society) 367 (2012): 717-730

Bruneau, Emile G. and Rebecca Saxe, "Attitudes toward outgroup are predicted by activity in the precuneus in Arabs and Israelis," *NeuroImage* (Elsevier Inc.) 52, no. 4 (2010): 1704-1711, p. 1705.

Bruneau, Emile, PhD, Post-Doctoral Fellow, SaxeLab, MIT, Phone interview by Rachel Brown (August 11, 2014).

Bruneau, Emile G, Mina Cikara, and Rebecca Saxe, "Minding the Gap: Narrative Descriptions about Mental States Attenuate Parochial Empathy," *PLoS ONE* 10(10) (2015): e0140838. doi:10.1371/journal.pone.0140838.

Cabrera, Alvaro, Senior Partner, Executive Director Creative Strategy, Head of Engagement Planning at OgilvyOne New York, interview by Rachel Brown, New York, New York (August 19th, 2014).

Cardona, Ivan, Patricia Justino, Becky Mitchell and Catherine Muller, "From the ground up: Women's roles in local peacebuilding in Afghanistan, Liberia, Nepal, Pakistan and Sierra Leone," Research report (ActionAid USA (accessed via ActionAid website), Institute of Development Studies, and Womankind Worldwide, September 2012); accessible at http://www.actionaidusa.org/sites/files/actionaid/from_the_ground_up_-_full_report.pdf

Carey, Benedict, "Cede Political Turf? Never! Well, Maybe.," *New York Times*, November 1, 2010

Chong, Dennis and James N. Druckman, "Framing Theory," *Annual Review of Political Science* (Annual Reviews) 10 (2007): 103-126.

Cikara, M., A.C. Jenkins and R. Saxe, "Reduced self-referential neural response during intergroup competition predicts competitor harm," *NeuroImage* 96 (2014): 36-43.

Cohel, Geoffrey L. , David K. Sherman, Anthony Bastardi, Lillian Hsu and Michelle McGoey, "Bridging the Partisan Divide: Self-Affirmation Reduces Ideological Closed-Mindedness and Inflexibility in Negotiation," *Journal of Personality and Social Psychology* (American Psychological Association) 93, no. 3 (2007): 415-430

Cohen, Taya R. and Chester A. Insko, "War and Peace: Possible Approaches to Reducing Intergroup Conflict," *Perspectives on Psychological Science 3*, no. 2 (2008): 87-93

Cook, John, and Stephan Lewandowsky. "The Debunking Handbook." University of Queensland, Global Change Institute, and The University of Western Australia, January 23, 2012 (initially published Nov. 2011).

Coster, Helen, "Peer Pressure Can Be a Lifesaver," *New York Times Opinionator*, May 14, 2014.

Dandekar, Pranav, Ashish Goel and David T. Lee, "Biased assimilation, homophily, and the dynamics of polarization," *PNAS early edition accessed through Stanford.edu*, October 2012.

Dasgupta, Nilanjana and Anthony G. Greenwald, "On the Malleability of Automatic Attitudes: Combating Automatic Prejudice with Images of Admired and Disliked Individuals," *Journal of Personality and Social Psychology* (American Psychological Association, Inc.) 81, no. 5 (2001): 800-814

De Figueiredo, Rui J.P, and Barry R. Weingast. "The rationality of fear: Political opportunism and ethnic conflict." 1999. *In Civil Wars, Insecurity, and Intervention*, eds. Walter and Snyder. New York: Columbia University Press, pp. 261-302. Working Paper accessed online at <http://isites.harvard.edu/fs/docs/icb.topic1018018.files/CIAO__The_Rationality_of_Fear__Political_Opportunism_and_Ethn...pdf>.

"Dehumanization in Conflict: Meeting Report 16-17 February, 2013" The Neuroscience and Social Conflict Initiative of Beyond Conflict (http://www.beyondconflictint.org/neuroscience-and-social-conflict-initiatives/) in partnership with the SaxeLab.

Demos, Alison, Senior Partner and Director of Ethnographic Research at Ogilvy & Mather, interview by Rachel Brown, New York, New York (August 20th, 2014).

DeSteno, David, Duane T. Wegener, Richard E. Petty, Derek D. Rucker and Julia Braverman, "Discrete Emotions and Persuasion: The Role of Emotion-Induced Expectancies," *Journal of Personality and Social Psychology* 86, no. 1 (2004): 43-56

DeSteno, David, Nilanjana Dasgupta, Monica Y. Bartlett and Aida Cajdric, "Prejudice From Thin Air: The Effect of Emotion on Automatic Intergroup Attitudes," Psychological Science 15, no. 5 (2004): 319-324

DeSteno, David, "The power of common ground: Surprisingly tiny factors can warp our judgments of other people. What can we do about it?," *Boston Globe*, September 17, 2011.

Dolan, Theo. "Countering Hate Speech in South Sudan through Peace Radio." *United States Institute of Peace*. United States Institute of Peace, 7 May 2014. Web. 12 Nov. 2014. <http://www.usip.org/publications/countering-hate-speech-in-south-sudan-through-peace-radio>.

Druckman, James N. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23.3 (2001): 225-56. Web. 10 Oct. 2014.

Edry, Ronny, "Israel and Iran: A love story?" Presentation at TEDxJaffa 2012, Filmed September 2012. Video accessed at https://www.ted.com/talks/israel_and_iran_a_love_story.

Entman, Robert, J.B. & M.C. Shapiro Professor of Media and Public Affairs at the George Washington University School of Media and Public Affairs, interview by Rachel Brown, Washington, DC (March 17, 2015).

Erdelyi, Kacy, Senior Partner, WW Group Planning Director, Ogilvy & Mather New York, Interview by Rachel Brown, New York, New York (August 20th, 2014).

"Friend Me 4 PEACE." Peace Factory: *A bridge of communication between the people in the middle east.* Peace Factory, n.d. Web. 8 Dec. 2014. <http://thepeacefactory.org/friend-me/>.

Fujii, Lee Ann, "The Power of Local Ties: Popular Participation in the Rwandan Genocide (2008, September 15). *Security Studies*, 17(3), 568-597.

Gerber, Alan S., Donald P. Green and Christopher W. Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment," *Americal Political Science Review* 102, no. 1 (February 2008): 33-48

Greenhill, Kelly and Ben Oppenheim, "Rumor Has It: The Adoption of Unverified Information in Conflict Zones," *unpublished paper*.

Grover, Margot, Associate Planning Director at Ogilvy & Mather New York, interview by Rachel Brown, New York, New York (August 19th, 2014).

Haidt, Jonathan, *The Righteous Mind: Why Good People are Divided by Politics and Religion* (New York: Vintage Books, 2012).

Hardwig, John. (1985, July). Epistemic Dependence. *The Journal of Philosophy*, 82(7), 335-349. Retrieved from http://web.utk.edu/~jhardwig/EpDep.pdf

Hattotuwa, Sanjana "Untying the Gordian Knot: ICT for Conflict Transformation and Peacebuilding," *Dialogue* 2, no. 2 (2004): 39-68.

Iacucci, Anahi Ayala, "The conundrum of digital humanitarianism: when the crowd does harm," *Diary of a Crisis Mapper blog*, November 15, 2013.

Iacucci, Anahi Ayala. Presentation at United States Institute of Peace Peace Tech Summit: Engineering Durable Peace. Presentation accessed at http://www.usip.org/events/peacetech-summit-engineering-durable-peace, at 23:05 minutes.

iHub Research. "The Umati Project: Monitoring Dangerous speech Online," Retrieved December 16, 2014, from http://www.ihub.co.ke/umati.

IMC Nigeria. Interfaith Mediation Center, n.d. Web. 18 Nov. 2014. <http://www.imcnigeria.org/>.

"Impact of Online Hate" (n.d.). In *Media Smarts: Canada's Centre for Digital and Media Literacy*. Retrieved December 4, 2014, from http://mediasmarts.ca/online-hate/impact-online-hate

"Interfaith Mediation Center: A Society Free from Violence." Interfaith Mediation Center, ©2015. Web. 20 October 2015. http://www.earlywarningnigeria.org/.

"Internews: Local Voices. Global Change." Internews, n.d. Web. 8 Dec. 2014. <https://www.internews.org/>.

"Iranians in TelAviv." *Peace Factory: A bridge of communication between the people in the middle east*. Peace Factory, n.d. Web. 8 Dec. 2014. <http://thepeacefactory.org/iranians/>.

Kahan, Dan "Cultural Cognition and Safeguarding the Science Commons," keynote address, accessed via the *Yale Climate & Energy Institute* at http://climate.yale.edu/videos/cultural-cognition-and-safeguarding-science-commons-dan-kahan (Yale, August 4, 2014)

Kalyvas, Stathis, "The Ontology of 'Political Violence': Action and Identity in Civil Wars," *Perspectives on Politics* 1, no. 3 (September 2003): 475-486

Karamage, Seth, Program Coordinator, Public Conversations Project, interview by Rachel Brown, by phone (November 6th, 2014).

Khan, Zarina. "Painting the Town Positive - Fighting Extremism with Inspirational Art." *MVSLIM*. MVSLIM, May 2015. Web. 2 July 2015. <http://mvslim.com/painting-the-town-positive-fighting-extremism-with-inspirational-art/>.

Kuznar, Lawrence A., Allison Astorino-Courtois, and Sarah Canna, eds. From the Mind to the Feet: *Assessing the Perception-to-Intent-to-Action Dynamic: Strategic Multilayer Assessment*. February, 2011. Air University.

Mancini, Francesco (editor), *New Technology and the Prevention of Violence and Conflict*, Alliance for Peacebuilding, International Coalition for the Responsibility to Protect, UNDP, USAID, and the International Peace Institute (Alliance for Peacebuilding (Accessed at http://www.allianceforpeacebuilding.org/2014/01/new-technology-and-the-prevention-of-violence-and-conflict/).

Martin, Steve, "98% of HBR Readers Love This Article," *Harvard Business Review*, October 2012

McCauley, C., & Moskalenko, S. (2008). Mechanisms of Political Radicalization: Pathways Toward Terrorism. *Terrorism and Political Violence*, 20, 415-433.

McDoom, Omar S. "The Psychology of Threat in Intergroup Conflict: Emotions, Rationality, and Opportunity in the Rwandan Genocide." *International Security* 37.2 (2012): 119-55. Web. 3 Nov. 2014.

National Science Foundation, "Who's afraid of the HPV vaccine?," (Press Release 10-006), January 13, 2010, accessed at http://www.nsf.gov/news/news_summ.jsp?cntn_id=116186,

"Norms, Narratives, and Neurons" The Neuroscience and Social Conflict Initiative Meeting Report, 15-16 March 2014 at MIT Media Lab, Beyond Conflict (http://www.beyondconflictint.org/neuroscience-and-social-conflict-initiatives/) with MIT Political Science and Saxelab.

Nyhan, Brendan, Assistant Professor, Department of Government, Dartmouth College, phone interview by Rachel Brown, (June 29th, 2015).

Nyhan, Brendan and Reifler, Jason "Blank slates or closed minds? The role of information deficits and identity threat in the prevalence of misperceptions" (Dartmouth College, accessed at http://www.dartmouth.edu/~nyhan/opening-political-mind.pdf, December 16, 2013).

Nyhan, Brendan, and Jason Reifler. "Misinformation and Fact-Checking: Research Findings from Social Science." New America Foundation, Media Policy Initiative (*Research Paper*), 2012. Web. 1 July 2015. <www.Newamerica.net>.

Nyhan, Brendan and Reifler, Jason, "When Corrections Fail: The persistence of political misperceptions" (Dartmouth, forthcoming in Political Behavior, Accessed at http://www.dartmouth.edu/~nyhan/nyhan-reifler.pdf)

Paluck, Elizabeth Levy, "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda," *Journal of Personality and Social Psychology* (American Psychological Association) 96, no. 3 (2009): 574-587

Paluck, Elizabeth L. "Peer pressure against prejudice: A high school field experiment examining social network change." *Journal of Experimental Social Psychology* 47: 350-58. Web. 17 Oct. 2014.

Paluck, Elizabeth L., and Hana Shepherd. "The Salience of Social Referents: A Field Experiment on Collective Norms and Harassment Behavior in a School Social Network." *Journal of Personality and Social Psychology* 103.6 (2012): 899-915. Web. 17 Oct. 2014.

Perkins, H. Wesley, David W. Craig and Jessica M. Perkins, "Usings social norms to reduce bullying: A research intervention among adolescents in five middle schools," *Group Processes and Intergroup Relations* 14, no. 5 (2011): 703-722.

Perkins, H. Wesley, and Alan D. Berkowitz. "Perceiving the community norms of alcohol use among students: some research implications for campus alcohol education programming." *Substance Use & Misuse* 21, no. 9-10 (1986): 961-976.

Phillips, Tim, "The Neuroscience of Social Conflict" Presentation at TEDxBoston, November 5 2014. Video accessed from Alliance for Peacebuilding at http://www.allianceforpeacebuilding.org/2014/11/the-neuroscience-of-social-conflict/

Pierskalla , Jan H. and Florian M. Hollenbach, "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa," *American Political Science Review* (accessed via http://polisci.duke.edu/uploads/media_items/technology-collectiveactioncellphoneviolence.original.pdf) (2013): 1-18.

Puig Larrauri, Helena, "The Use of New Technologies: Expanding Opportunities for Peacebuilding?," Paper, Accessible Online, KOFF: Swiss Peace (KOFF: Swiss Peace, 2014).

Puig Larrauri, Helena and Anne Kahl, "Technology for Peacebuilding," *Stability: International Journal of Security & Development* 2, no. 3 (2013): 1-15.

Rosenberg, Tina, "The Destructive Influence of Imaginary Peers," *New York Times Opinionator*, March 27, 2013

Samaratunge, S., & Hattotuwa, S. (2014, September 24). Liking violence: A study of hate speech on Facebook in Sri Lanka. In *Centre for Policy Alternatives (CPA)*. Retrieved October 16, 2014, from http://www.cpalanka.org/liking-violence-a-study-of-hate-speech-on-facebook-in-sri-lanka/

"Sawa Shabab: A ray of hope for young people in South Sudan." *Free Press Unlimited* 22 May 2014. Web. 12 Nov. 2014. <https://www.freepressunlimited.org/en/news/sawa-shabab-a-ray-of-hope-for-young-people-in-south-sudan>.

Scheufele, Dietram A., "Framing as a Theory of Media Effects," *Journal of Communication* (International Communication Association), Winter 1999: 103-122

Schissler, Matthew. (2014, September 24). Echo Chambers in Myanmar: Social media and the ideological justifications for mass violence. *Paper for the Australian National University Department of Political & Social Change Research Colloquium, "Communal Conflict in Myanmar: Characteristics, Causes, Consequences,"* 17-18 March 2014, Yangon, Myanmar.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein and Vladas Griskevicius, "The Constructive, Destructive, and Reconstructive Power of Social Norms," *Psychological Science* 18, no. 5 (May 2007): 429-434

Shah, Seema, and Rachel Brown. "Programming for Peace: Sisi ni Amani Kenya and the 2013 Elections: CGCS Occasional Paper Series on ICTs, Statebuilding, and Peacebuilding in Africa, Number 3." *Center for Global Communication Studies*. Center for Global Communication Studies, Annenberg School for Communication, University of Pennsylvania, Dec. 2014. Web. 21 Sept. 2015. <http://www.global.asc.upenn.edu/app/uploads/2014/12/SisiNiAmaniReport.pdf>.

Terry, Deborah J. and Hogg, Michael A., "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification," *Personality and Social Psychology Bulletin* 22, no. 8 (August 1996): 776-793

"The Cure Violence Health Model." Cure Violence, 2012. Web. 3 Nov. 2014.

"The Prevention Toolbox: Systematising Policy Tools for the Prevention of Mass Atrocities, Policy Brief Series No.4; Countering Ideologies that Justify Mass Atrocities." *Australian Civil-Military Centre*. Australian Civil-Military Centre and Oxford Institute for Ethics, Law and Armed Conflict, n.d. Web. 16 Nov. 2014. <http://www.acmc.gov.au/wp-content/uploads/2014/09/4-The-Prevention-Toolbox-Countering-Ideologies.pdf>.

Thorson, Emily. Assistant Professor of Media and Public Affairs at George Washington University "Theory of Planned Behavior Handout."

Thorson, Emily "Some of My Best Friends are Poor? Income Misperceptions and Policy Attitudes," Academic Paper, George Washington University

Thorson, Emily, Assistant Professor of Media and Public Affairs at the George Washington University School of Media and Public Affairs, interview by Rachel Brown, Washington, DC (July 22, 2014).

Wallace, Marshall, interview by Rachel Brown, Boston, Massachusetts (July 23, 2015).

Wallace, Marshall. From Principle to Practice: *A User's Guide to Do No Harm*. Principle to Practice: The Art and Science of Effective Action. (http://www.principletopractice.org/from-principle-to-practice/). Accessed August 3, 2015.

Williams, Lisa A. and David DeSteno, "Pride and Perseverance: The Motivational Role of Pride," *Journal of Personality and Social Psychology* (Americal Psychological Association) 94, no. 6 (2008): 1007-1017,

Woodrow, Peter and Diana Chigas, "A Distinction with a Difference: Conflict Sensitivity and Peacebuilding," Reflecting on Peace Practice Project, CDA Collaborative Learning Projects (cdacollaborative.org, October 2009).

ACKNOWLEDGEMENTS