
**Countering Dangerous Speech,
Protecting Free Speech: Practical
Strategies to Prevent Genocide**
**Report of the 2014 Sudikoff Annual
Interdisciplinary Seminar**

UNITED STATES
MEMORIAL
**HOLOCAUST
MUSEUM**

SIMON-SKJODT CENTER
FOR THE PREVENTION OF GENOCIDE

About This Report

This report summarizes the discussions and findings of *Countering Dangerous Speech, Protecting Free Speech: Practical Strategies to Prevent Genocide*, the Sudikoff Annual Interdisciplinary Seminar on Genocide Prevention, which was organized by the United States Holocaust Memorial Museum's Center for the Prevention of Genocide in February 2014. The two-day seminar convened three dozen international experts from government, academia, business, law, journalism and civil society to examine strategies for detecting and countering hate speech associated with mass violence without restricting the right to free speech. More information about the seminar is available [here](#). The views expressed in this report do not necessarily reflect those of the United States Holocaust Memorial Museum.

The seminar and this report were made possible by the generous support of the Sudikoff Family Foundation, which funds the Museum's Sudikoff Annual Interdisciplinary Seminar on Genocide Prevention.

Introduction

The Holocaust, Rwanda, and Srebrenica all demonstrated that genocide is often preceded and accompanied by widespread hate speech. The leaders who planned these genocides disseminated ideologies of hatred aimed at their intended targets in order to spur their followers to act, cow bystanders to remain passive, and justify their crimes. In 2008, the Genocide Prevention Task Force co-sponsored by the United States Holocaust Memorial Museum concluded that speech promoting ideologies based on hatred can serve not only as an instrument of genocide but also as a warning sign that a country may be at risk for genocidal violence.

This conclusion raises the possibility that hate speech may present opportunities for preventing genocide, on the one hand by helping to identify places at risk of group-targeted violence and on the other by interventions to prevent hate speech from inciting collective violence. In order to pursue this possibility, it is first necessary to distinguish hate speech that can facilitate and incite collective violence from the ordinary hate speech that occurs daily throughout the world but rarely is accompanied by violence toward its targets.

This groundwork has been performed, principally by Dr. Susan Benesch, the 2013 Edith Everett Fellow of the Museum's Center for the Prevention of Genocide. Drawing from studies of speech that preceded and accompanied episodes of group-targeted mass violence, Dr. Benesch has developed a framework for identifying "Dangerous Speech," i.e., speech that has the capacity to catalyze collective violence.

States' most common approaches to addressing the danger of inflammatory hate speech have been to punish and suppress it. A variety of international conventions and national laws prohibit incitement to genocide, persecution, discrimination or group-targeted hostility. Some states criminalize other types of speech associated with violence that has occurred in their countries, and some censor speech that authorities deem to promote violence or hate. Ending impunity for those who incite collective violence is an important tool for genocide prevention. Too often, however – as shown by an international expert study¹ – laws banning incitement are used to suppress the voices of those most likely to be the targets of discrimination and violence. Protecting the right to free speech and promoting a diversity of voices are also important tools for genocide prevention: hate flourishes best where it lacks competition.

Understanding the factors that make speech dangerous presents opportunities not only for detecting where speech is occurring that signals a risk for collective violence but also for devising interventions that can deprive speech of the power to foment violence without suppressing it. These opportunities were the focus of the Sudikoff Seminar on "*Countering Dangerous Speech, Protecting Free Speech: Practical Strategies to Prevent Genocide*." Held over two days in February 2014 by the Center for the Prevention of Genocide of the United States Holocaust Memorial Museum, the seminar brought together three dozen international

¹Office of the United Nations High Commissioner for Human Rights, "Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence": http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf.

experts from the fields of government, technology, business, civil society, journalism, and academia to address two basic questions:

- How can we detect where hate speech is occurring that threatens to incite mass violence?
- How can we prevent messages of violent hatred from reaching and influencing their intended audience without restricting free speech?

The seminar's first panel opened by discussing the value of using a dangerous speech framework for understanding when hate speech indicates a risk for collective violence that should be addressed. It went on to examine several broad approaches for countering the factors that can give speech the power to catalyze violence, with a particular focus on approaches that do not restrict or punish speech.

The second panel focused on technological developments that can enhance efforts to monitor the "dangerousness" of speech and to counter dangerous speech. It looked at some specific examples of projects that use technology to monitor for incitement and other signs of violence in at-risk communities and to promote peace and counter extremist narratives.

The third panel featured presentations on some of the efforts to prevent mass violence around Kenya's 2013 elections and discussed the lessons to be drawn from Kenya's example for addressing dangerous speech in other situations.

There are many situations in the world today where widespread, inflammatory hate speech is associated with discrimination and violence towards the groups it targets. The fourth panel examined four such situations, using the dangerous speech framework to analyze the factors that are influencing the spread and impact of inflammatory rhetoric.

The seminar's final panel summed up some of the key findings of the seminar and identified some questions and needs that remain to be addressed in order to develop effective approaches for countering dangerous speech.

The following sections summarize the principal points of the seminar's framing papers, presentations and discussions, organized around the topics of the panels. For more information about the seminar and its participants, please see: <http://www.ushmm.org/confront-genocide/speakers-and-events/all-speakers-and-events/2014-sudikoff-seminar>.

I. Dangerous Speech²

Hateful and inflammatory speech is a common occurrence, but mass collective violence associated with such speech is rare. Nevertheless, hateful rhetoric plays an important role not only in inciting group targeted mass violence but also in creating the conditions under which incitement becomes possible, by influencing its audience to view violence against a group as logical, justifiable, and necessary. “Dangerous speech” is the term coined by Dr. Susan Benesch for speech that can facilitate collective violence by conditioning its audience to accept, condone and commit such violence. It includes not only speech that incites but also speech that assembles the tinder to which incitement provides the spark.

The content of speech alone does not make speech “dangerous.” Benesch has identified five factors that in combination enhance the capacity of speech to facilitate collective violence. These factors are: content, speaker, means of dissemination, context and audience.

Content

Like all hate speech, dangerous speech collectivizes fault, that is, it claims that all or most members of a target group think or act in ways that are repugnant to “us” (the audience for the speech). But dangerous speech also portrays the target group as so alien and dangerous to the audience as to be beyond the reach of moral obligations. It commonly does so by dehumanizing the target group, often by comparing it to something disgusting or dangerous, such as rats, lice, cockroaches or snakes. Eliminating such creatures is generally considered a logical and appropriate measure. Dangerous speech also frequently accuses the target group of plotting to harm the audience and portrays the target group as, by its very existence, presenting a threat to the audience. Violence against every member of the target group thus becomes justifiable as necessary self-defense to ensure the audience group’s survival.

Speaker

The ability of dangerous speech to influence its intended audience depends in significant part on the relationship of the audience to the speaker. Audiences are more likely to be receptive to messages from speakers who are popular with the audience or enjoy its respect and trust. When such speakers communicate the views of dangerous speech, audiences are more likely to accept, internalize and repeat those views, which can in turn create a norm that it is logical and acceptable to endorse and express those views.

Means of Dissemination

The impact of dangerous speech upon an audience can be enhanced by its means of dissemination. A message delivered to many people individually or in a forum where opposing views are expressed is likely to influence far fewer people than one delivered to a group that has assembled to support a common interest, such as a religious gathering or a patriotic rally. When outbreaks of collective violence have been preceded and accompanied by dangerous speech, a key means of dissemination has commonly been the principal source of information for the audience, such as state-controlled media.

² For a fuller explanation of the framework, see Susan Benesch, “Dangerous Speech: A Proposal to Prevent Group Violence,” <http://www.dangerousspeech.org/guidelines>.

Context

The context in which hate speech is delivered also influences its capacity to be dangerous. A history of group-targeted discrimination and violence, longstanding competition for resources, and a lack of institutions that people trust to resolve grievances fairly and peacefully are examples of contextual factors that can increase the likelihood that dangerous speech will catalyze collective violence.

Audience

Hate speech cannot be dangerous without an audience that is influenced by its message to accept, condone or commit collective violence against the group that the hate speech targets. Hate speech is most likely to have this effect when it exploits and cultivates the audience's shared grievances and fears, when it resonates with the audience's experiences, and when it appeals to the audience's shared beliefs, especially with respect to its own identity.

II. Non-restrictive Approaches to Countering Dangerous Speech

Under international laws and conventions, restricting the right to free speech is a permissible response to speech that threatens to incite group-targeted hostility and violence. Censorship (i.e., barring the dissemination of prohibited speech) and criminal prosecution for incitement – the most common restrictive approaches to inflammatory hate speech – are not effective for preventing or combatting dangerous speech, however, for a number of reasons. With respect to censorship, it is extremely difficult to erase all trace of speech that has been communicated via today's communications technologies: hateful messages can now be sent to audiences around the world in mere seconds, and those messages can live on even longer than their original authors. Moreover, while narrower than hate speech, dangerous speech is broader than incitement, as it includes not only speech that meets the legal definition of incitement but also speech that, without directly inciting, conditions audiences to accept, condone and commit collective violence. Restrictive approaches also risk enhancing the impact of dangerous speech, such as by lending notoriety to the speaker and drawing broader attention to the message. In practice, moreover, restrictive approaches are too often used to suppress legitimate expressions of grievances or dissent. As was pointed out in the seminar discussions, disagreeable and even hateful speech plays a necessary role in helping societies to air and mediate hostilities and grievances without resort to violence.

Restrictive approaches to combatting dangerous speech primarily target the speaker or the means of dissemination. Non-restrictive approaches are being developed that primarily target the audience, with the aim of preventing audience receptivity to the message of dangerous speech. To draw an analogy from the field of health, if audiences develop resistance to the poison of dangerous speech, then the speech continues to exist but loses the capacity to harm a society's health. Non-restrictive approaches to making audiences resistant to dangerous speech include both longer term efforts to "inoculate" audiences against efforts to condition them to accept and condone violent hate, as well as short-term "injections" of "counterspeech" in dangerous speech situations when violence appears imminent or is already occurring.

Three ways to increase audience resistance to dangerous speech are by fostering habits of critical and skeptical thinking, by instilling empathy for members of other groups, and by enhancing the willingness to speak out and voice dissent. Critical thinking enables audiences to recognize attempts to manipulate them through hate, to foresee the potential consequences of such efforts, and to understand that the consequences may not be in their own interests. Empathy counteracts efforts to dehumanize members of a target group. When audience members express and are exposed to a variety of views within their group, audiences are less likely to perceive the views espoused by promoters of dangerous speech as the norm for their group.

Three of the factors that give power to dangerous speech can also enhance the effectiveness of non-restrictive approaches to countering dangerous speech: the speaker, the means of dissemination, and the audience. Research shows that leaders can diminish the hostility their followers feel for a rival group by making positive or inclusive statements about that group. In particularly tense situations, a strong message from influential leaders rejecting violence can deter and quell violent outbreaks. Consistent speech from influential leaders countering the message of dangerous speech can over time reduce the impact of dangerous speech on its intended audience.

By giving exposure to a diversity of voices and views, media can help to demystify the differences between groups and to counteract the perception that hateful attitudes are the norm for one's own group. In some countries that have experienced collective violence, media organizations have adopted standards and developed training to promote responsible journalism that avoids using inflammatory terminology and giving credence to false or sensational claims.

Media programming can also dilute the effects of dangerous speech on audiences by modeling behaviors that can enhance audience resistance. Some programming has already been specifically designed to inoculate audiences against dangerous speech. This includes the popular Rwandan radio soap opera "Musekweya," which portrays two neighboring villages with a history of hostility and violence. In both villages, some members promote hate to further their own interests, while others work with their counterparts to solve common problems, despite conflicting views and interests. At the heart of the drama is a Romeo and Juliet story. An independent evaluation of this program's effects on its listeners over one year concluded that listeners showed increased empathy for members of other groups and a greater willingness to express their views and voice dissent. To help prevent a repeat of collective violence in Kenya in the period surrounding the 2013 election, one of its longest-running television shows included programming in 2012 that demonstrated the link between dangerous speech and violence and portrayed how some leaders use dangerous speech to enhance their own power. The programming was evaluated as having improved viewers' ability to identify and resist dangerous speech.

A number of voices speaking together can also reduce the impact of dangerous speech, even when the individual speakers do not enjoy notable popularity or influence with the audience. A lone audience member who questions a leader's message of hate is likely to experience ostracism or other negative consequences. As the number of audience members who question the message increases, however, the likelihood of negative consequences decreases. By demonstrating that the message of dangerous speech is not universally accepted by the audience, this kind of

“counterspeech in unison” can prevent the message from being internalized as a norm by the audience.

Non-restrictive approaches to countering dangerous speech can help national and international government policy makers fulfill their responsibilities for civilian protection under international law and norms. The International Convention on the Prevention and Punishment of Genocide, for example, includes incitement to genocide among the activities that its signatories are obliged to prevent and punish. Similarly, the World Summit Outcome Document of 2005, which sets out the Responsibility to Protect, specifies that governments have a responsibility not only to protect populations from genocide, war crimes, crimes against humanity and ethnic cleansing but also to prevent incitement to those crimes. While dangerous speech includes speech that does not meet the elements of the crime of incitement, effectively countering dangerous speech can prevent incitement by depriving inflammatory speech of the power to incite.

III. Potential Uses of Technology to Monitor and Counter Dangerous Speech

Technology and especially Information and Communication Technologies (ICTs) are playing increasingly important roles in the human rights field generally and specifically in efforts to identify and monitor signs of mass human rights violations. Presentations at the seminar demonstrated how emerging technology, including data mining of online and social media and crowd sourcing through use of ICTs, can be used to collect and analyze the “dangerousness” of speech, using the dangerous speech framework. While these technologies enhance the speed with which threats can be detected and addressed, they continue to be heavily reliant on human analysis to interpret and confirm the data collected. This is especially true when it comes to monitoring speech. Whether words like “cockroach” or “snake” are being used in a hateful and dehumanizing manner depends upon the context in which they are used, and it has not yet been possible to automate that kind of textual analysis.

Other technology fields show promise for monitoring dangerous speech. In particular, social network analysis, predictive analytics and sentiment analysis present possibilities for: tracking the dissemination of dangerous speech; identifying the most prolific propagators and analyzing their influence; predicting trends in dangerous speech and associated violence; and gauging changes in the dangerousness of discourse over time.

The role of ICTs in disseminating hateful and dangerous speech has gained considerable attention. Some argue that the internet even favors extremist narratives. The same technologies, however, can be used to counter dangerous speech in a variety of ways. For example, the seminar’s participants heard how communities whose youth are subject to recruiting efforts by religious extremists are using social media to communicate and involve youth in a narrative of community identity that counters the narrative of extremism. There have also been a number of popular online campaigns against hate speech, though the degree to which they reach and influence the audiences most receptive to messages of hate is unclear. In some places at risk of collective violence, ICTs have been used to disseminate information that debunks rumors, a form of communication frequently used to disseminate dangerous and inciting speech. Again, the effect of such efforts on the audience most likely to give credence to the rumors is unclear.

Similarly, in particularly tense situations, messages have been disseminated via ICTs that target a specific audience with the goal of persuading its members to abstain at least temporarily from actions that could lead to violence. There is anecdotal evidence supporting the effectiveness of this approach.

As these examples indicate, ICTs can be used both for long-term efforts to “inoculate” audiences against dangerous speech and also to “inject” speech into situations where collective violence is occurring or appears imminent in order to counter dangerous speech and incitement. It was noted at the seminar that in several situations of rising tensions around elections in post-conflict societies, social media have actually been dominated by voices calling for calm and denouncing dangerous rhetoric. On Twitter, there has been an increasing tendency for users to confront those who post hateful messages, sometimes causing the posters to delete their messages and even apologize. This demonstrates that ICTs can empower users to create and enforce discourse norms that reject hateful and dangerous speech.

Using ICTs to counter dangerous speech raises a number of potential risks, such as harassment and threats targeting those who disseminate countering messages and blocking or hijacking of websites by hackers. In all uses of ICTs to monitor or counter dangerous speech, moreover, it is essential to protect the right to privacy. Careful risk assessment and informed consent of participants should be incorporated into projects.

IV. Monitoring and Countering Dangerous Speech in Practice: The Example of Kenya in 2013

In 2007 and 2008, during and following presidential elections marred by allegations of widespread fraud, Kenya experienced outbreaks of mass ethnic violence that left more than 1,000 dead, displaced more than a half million people, and devastated the economy. Dangerous speech had been pervasive preceding and during the violence, with inciting messages widely disseminated via both the traditional media and ICTs, especially text messaging. Of the four Kenyans charged by the International Criminal Court for crimes against humanity in connection with the violence, one was a radio broadcaster. Fears that the violence might recur around the 2013 elections sparked extensive efforts by the international community and, more importantly, by Kenyans themselves to ensure that the elections proceeded peacefully.

As the seminar’s participants heard, some of the efforts to prevent a recurrence of mass violence in Kenya around the 2013 elections focused on monitoring for and countering dangerous speech, especially as it was being communicated via ICTs. For example, in advance of the elections, Umati, a project of iHub Research, monitored online media and social media to collect hateful and inflammatory speech and to analyze it using the dangerous speech framework. One quarter of the 5,683 examples analyzed were classed as “very dangerous,” and the vast majority of these were on Facebook. Of the posters of hateful or dangerous messages online, 94% used either their own name or an identifiable pseudonym. The Umati Project tracked messages in seven languages, but the hateful or dangerous messages it captured were primarily in English, Swahili or Sheng, the slang popular in Nairobi. While the volume of online dangerous speech rose

around the election and provided real cause for concern, there was a much greater volume of messages promoting peace and calling upon users to refrain from hateful and dangerous rhetoric.

The rising level of online dangerous speech that Umati was tracking prompted it to initiate a project to counter the spread of false rumors, a form of dangerous speech that played an important role in inciting violence in Kenya in 2007-2008. Called Nipe Ukweli (“Gimme Truth”), it aimed to educate people in areas that had experienced violence to recognize and reject dangerous speech and to question and refuse to pass on rumors that could incite violence. Nipe Ukweli’s material and messages were posted online on Facebook and Twitter, broadcast on community radio stations, and disseminated in person via community forums and outreach to local youth, activists, and religious, political and social leaders.

The role that text messaging played in inciting and organizing violence in 2007-2008 caused some peace activists to consider how mobile technology might be used to “incite peace” instead. The organization Sisi ni Amani (“We are Peace”) Kenya created an SMS subscription service to help local peace activists prevent incitement and violence. It conducted extensive outreach in communities that had experienced ethnic violence, working with local organizations to identify and recruit subscribers, targeting not only trusted leaders but also key communicators. It worked with a marketing firm to build a brand identity and used focus groups to craft effective messages. Much of its messaging consisted of useful civic education information around the election, but it was also able to monitor for signs of incitement or impending violence and send targeted messages to the affected communities in the language of the subscribers there. These targeted messages did not simply call for peace but aimed at specific goals, such as to get the recipients to calm down, to question rumors, or to consider the potential consequences of violence. In a post-election survey of 7,350 subscribers, 92% of respondents stated they believed the messages had helped to prevent violence.

Following the mass violence of 2007-2008, Kenyan media engaged in soul-searching over the role they had played in escalating tensions and inciting violence. The result was efforts to raise journalistic standards through establishing a revised code of conduct for journalists and providing professional training for hundreds of journalists. Several seminar participants described the massive peace campaign that was waged in the Kenyan media around the 2013 elections. In addition to political, religious and popular figures, journalists and broadcasters openly called upon their audiences to remain calm and refrain from violence. For a time, one popular television news program dedicated a nightly segment to “naming and shaming” those who engaged in hate speech, prompting a number of those named to make public apologies. On the day polling started, a major Kenyan daily ran a banner headline reading “Never Again” over an article warning that violence could lead to the country’s destruction.

The 2013 Kenyan elections were highly controversial and the outcome was not decided for nearly a month.³ Although tensions were high and there were isolated violent incidents, mass violence did not break out. Dangerous speech was evident throughout the country, especially on social media, but the widespread perception was that the chorus of voices calling for peace and

³ Five days after the March 4 polling, Uhuru Kenyatta was declared the victor, but his main rival, Raila Odinga, challenged the results in court. On March 30, Kenya’s Supreme Court certified the election results, and Odinga conceded.

rejecting hateful rhetoric was far louder. This instance of “counterspeech in unison” is only one of numerous factors that may have influenced Kenyans to refrain from violence, chief among the others being the political alliance between the leaders of two of Kenya’s largest ethnic groups, a new constitution, and the still fresh memory of what the country had suffered. While the peace campaign has been hailed by many as a success, some have criticized it, noting that by creating a norm that rejected hateful discourse, the campaign also suppressed legitimate dissent and caused the media to abandon its watchdog role. They point out that the winners of the 2013 elections have done little to heal the deep divisions in Kenyan society and have instituted restrictions on press freedom and on the activities of civil society organizations.

V. Dangerous Speech in the World Today: Four Case Studies

The signs of dangerous speech are evident in many areas of the world. The seminar’s participants heard about four countries where widespread hate speech is occurring in association with harm and violence toward the groups that the hate speech targets.

In Bahrain, a hate speech campaign against the majority Shia population is being waged by the Sunni royal family and certain influential Sunnis close to the family. It uses dehumanizing language against the Shia; accuses the Shia of being foreigners who are loyal to Bahrain’s enemy, Iran, and of plotting to undermine the state; and calls for the departure or removal of the Shia from Bahrain. The hate speech is being disseminated via traditional media -- which are heavily censored by the state and largely controlled by the royal family and those close to them - - and also via social media.

No violence between Bahrain’s native Sunni and Shia civilians has occurred in association with this hate speech. Key audiences for the speech appear to be Bahrain’s security forces and Sunni immigrants whom the state is bringing in from Yemen, Syria and Pakistan and often incorporating into the security forces. The speech is being used to justify both the state’s crackdown on those seeking democratic reforms – who are equated with Shia terrorists – and its violent and persecutory actions against the Shia, particularly those of Persian descent.

In Burma, widespread anti-Muslim hate speech has already facilitated violent attacks against Muslim communities in which hundreds have been killed and thousands of Muslim homes, mosques and businesses have been destroyed. The extremely virulent hate speech campaign is being propagated by Buddhist monks, who constitute the most trusted and respected group in the country. It uses vile and dehumanizing language to describe Muslims; accuses Burma’s Muslim population of plotting to wipe out Buddhism and impose sharia law in Burma; and maintains that, because of their supposedly high birthrate, Muslims’ very presence in the country threatens eventually to overwhelm the Buddhist population (90% of Burma’s population is Buddhist; Muslims comprise an estimated 5%). In Rakhine State, where the Muslim Rohingya minority resides, political leaders also engage in dangerous and inciting anti-Muslim rhetoric, even openly calling for ethnic cleansing of the state’s Muslim population.

The contextual factors that enhance the impact of anti-Muslim dangerous speech on audiences in Burma include: widespread racial and religious prejudice against Muslims, who are largely of

South Asian descent; an educational system that actively discourages critical thinking; lack of knowledge about Burma's history and its many ethnic groups; media that were completely state-controlled until 2012 and that are still subject to state interference; an extremely low rate of media literacy; lack of training for journalists; active mistrust of the government and legal system. Audiences are therefore credulous and receptive toward messages of hate that appear to come from monks and that are being disseminated via public rallies, mass produced CDs, DVDs and stickers, and television and radio. Social media, which are new but gaining popularity, are being used to disseminate particularly inflammatory anti-Muslim hate speech, replete with sensational false claims and photo-shopped images.

Burma's central government has defended the anti-Muslim hate campaign and supported its political proposals while failing to prosecute those who have openly incited violence. It actively pursues policies that discriminate against and persecute Muslims, particularly the Rohingya. The police and security forces have looked on without acting during a number of violent attacks against Muslims. Although the government still frequently censors and jails Burmese who speak out or demonstrate in favor of rights and reforms, the leaders of the anti-Muslim hate speech campaign are able to speak freely and receive permission for their rallies. The campaign has increasingly targeted the leader of Burma's democracy movement, Aung San Suu Kyi, claiming that she is sympathetic toward Muslims, that Muslims dominate her party, and that the further reforms she is seeking could lead to a Muslim takeover of the country. The anti-Muslim hate speech campaign's freedom to operate and political message, combined with its ability to provide mass transportation to its rallies and distribute tens of thousands of CDs, DVDs and stickers for free, lead some to surmise that it is being sponsored by certain elements of the regime in order to prevent or even reverse democratic reforms.

Hungary has in recent years seen a meteoric rise in the popularity of far-right parties and movements that espouse anti-Semitism and especially anti-Roma hate. The anti-Roma hate speech has been accompanied by outright violence against Roma and their communities. The speech refers to the Roma as animals, claims that they are genetically disposed to criminality, and portrays their presence as threatening the security and economic well-being of ethnic Hungarians.

The economic crisis of 2008 and two subsequent recessions have helped to fuel the popularity of inflammatory hate speech and those who preach it. The far-right Jobbik party, which engages in anti-Semitic and anti-Roma hate rhetoric, has become the third most popular in Hungary. Their message appeals to widespread prejudice toward Roma among ethnic Hungarians, to growing anti-European Union sentiment, and to the perception that Hungary has been victimized by Western banks and by liberal socialism, both often portrayed in anti-Semitic rhetoric as elements of a Jewish conspiracy. The anti-Semitic and anti-Roma hate speech is being disseminated via social media, traditional media, public demonstrations, and recruitment activities of paramilitary and other far-right organizations. Particular audiences are young Hungarian men and university students.

The Hungarian government has not only failed to speak out in any effective way against the increasingly dangerous rhetoric of hate, but has facilitated it. The ruling Fidesz Party actively courts the support of Jobbik and the far right in its quest to make permanent its hold on power

through changes in laws and the constitution. The government's increasing control over the media and retaliation against those who oppose it have favored hateful rhetoric, which is becoming the norm. It has allowed the formation and operations of paramilitary organizations that have intimidated, laid siege to and violently attacked Roma communities with impunity. Although the government has responded in a few cases when these organizations' violations of the law have become particularly visible, the steps it has taken have had little deterrent effect. Although Hungary has laws criminalizing incitement, they have only been enforced against Roma. Similarly, while hate crime laws are rarely used to prosecute crimes against Roma, charges of anti-Hungarian hate crime have been brought against Roma who responded to paramilitaries' provocations.

Although India has been a secular democracy for over sixty years, it has seen numerous instances of incitement to violence, especially against religious groups. Political leaders in particular have used violent hate speech to win support from voters who share their faith. They disseminate this speech via the media outlets that are most popular with their intended audience as well as through public speeches and on social media. Their messages appeal to the prejudices, fears and grievances of their audiences. Both Hindu and Muslim politicians have engaged in hateful and inciting rhetoric. Currently, Hindu nationalism is becoming increasingly popular (about 80% of India's population is Hindu), and some of its leaders promote anti-Muslim hate and have been associated with violence against Muslims. Their speech dehumanizes and demonizes Muslims, accusing them of being loyal to Pakistan, India's enemy, and of plotting to take over India through both violence and their supposedly high birthrate.

India has laws that ban not only incitement but also speech that promotes enmity between groups and offends religious feelings. Nevertheless, political leaders who promote and incite hate and violence toward religious groups have never been held to account. Instead, these laws are most often used to silence writers, artists and scholars whose work is deemed by some to be offensive.

As these examples demonstrate, dangerous or potentially dangerous speech occurs under a variety of conditions and types of regimes. While there are differences in the content, context, audience and target of the speech in these examples, there are similarities as well. Thematically, the speech in all these examples ascribes negative traits and intentions to every member of the target group. It also presents the target group as being less than human and as posing an existential threat to the audience. The role of the state in these examples is particularly striking. In none of these situations is there an effort by the state or influential political leaders to counter inflammatory hate speech. Existing laws against incitement are not being used to sanction actual incitement, but speech restrictions are being imposed on those who do not engage in hate speech and even those who are most likely to be the victims of incitement. In all of these examples, moreover, the propagators or facilitators of the speech include influential political leaders who are exploiting group-targeted hate to advance or preserve their power. Fortunately, in all of the countries examined, there is active civil society that is willing or already attempting to counter hateful and dangerous speech.

One goal of the seminar panel that examined current dangerous speech situations could not be fulfilled: discussing possible strategies and tools that could be effective in countering dangerous speech in each situation, building on the strategies and tools presented during previous panels.

Seminar participants recognized that interventions to counter dangerous speech in a given situation must be based on a more detailed analysis of the speech than was possible within the parameters of the seminar.

VI. Lessons Learned, Outstanding Questions, and Next Steps

The seminar discussions led to consensus on certain points with respect to the value of the dangerous speech concept and framework for combating incitement to mass atrocities and to the need and possibilities for countering dangerous speech without infringing upon the right to free expression. Seminar participants also identified needs and questions that must be addressed in order to develop effective approaches for countering dangerous speech and ensuring that policy makers understand and use them.

The Dangerous Speech Framework

The genocide prevention field recognizes that genocide and mass atrocities result from processes that develop over time. The later in this process that intervention is undertaken, the fewer the available options and the greater the likelihood that the intervention will be costly, coercive and controversial. Two keys to prevention are therefore early warning – the ability to identify countries at risk of mass atrocities long before violence occurs – and early intervention, when the available tools and strategies are more numerous and generally less costly and controversial.

Incitement is commonly cited as one of the contributing factors to genocide and mass atrocities. Like the actions it aims to incite, incitement occurs at the end of a developing process – a speech can only create an imminent likelihood of violence if the audience that is to commit the violence has already been conditioned to view violence as logical, justifiable and necessary. To prevent incitement, then, it is necessary to identify where speech is occurring that could have this conditioning effect and to intervene before the conditioning process is complete. The seminar participants agreed that the dangerous speech framework can help fill both needs. It provides an analytical tool for gauging the capacity of speech to catalyze collective violence and for identifying the specific factors that enhance this capacity. This analysis can in turn inform efforts that prevent speech from catalyzing violence by targeting the specific factors that make the speech dangerous.

Efforts to monitor for speech that could lead to incitement using the dangerous speech framework show the promise of this tool. Many questions remain about how to apply this tool in the many differing conditions under which dangerous speech occurs. Seminar participants saw a need for further study and monitoring of inflammatory hate speech in a variety of places at risk for mass atrocities, which should improve understanding of the ways speech can impact the risk for violence and of the relative importance of the dangerous speech factors under different conditions. For example, is there a difference in the factors that make speech dangerous online as opposed to offline? While recent technological developments greatly expand the capacity to monitor for dangerous speech, it is important to insure that monitoring efforts do not infringe upon the right to privacy.

Effective Approaches for Countering Dangerous Speech

Dangerous speech is occurring in many places throughout the world and under a variety of conditions. Seminar participants noted that efforts to counter dangerous speech should be tailored to the specific conditions under which it is occurring and should therefore be based on rigorous analysis of those conditions. Monitoring dangerous speech can contribute essential data for such an analysis and should be part of the process. More work is needed to develop an analytical framework for determining what countering efforts would be possible and effective under the specific conditions in which dangerous speech is occurring. Case studies of past instances of dangerous speech situations and the effect of efforts to prevent such speech from catalyzing collective violence can advance the development of an analytical framework for countering dangerous speech.

Seminar participants agreed that while it is necessary to prevent and punish direct incitement to violence, there is a need to develop effective, non-restrictive approaches that can help prevent the conditions under which incitement becomes possible. Approaches that aim to prevent audiences from being receptive to the message of dangerous speech show particular promise for preventing incitement while protecting the right to free speech. Special note was taken of the role that discourse norms play in spreading or countering dangerous speech. Exposing audiences to a diversity of voices can lessen the likelihood that hateful rhetoric will be internalized as the norm for discourse. Fostering the development of social norms that reject inflammatory hate speech can be more effective than legal sanctions for deterring dangerous speech and incitement.

To be effective, efforts to counter the impact of dangerous speech need to reach and influence the audiences that are susceptible to the speech. Countering messages will be most effective when they come from speakers whom the audience trusts or admires and when they resonate with the audience's beliefs, values and experiences. Countering efforts therefore need to involve actors within the communities where dangerous speech is occurring, as they will best understand the intended audience and what messages will appeal to it. In many of the places where dangerous speech is occurring, there is active civil society that is willing to address the problem but needs training and tools to do so.

Countering Dangerous Speech as a Focus of Atrocity Prevention Policymaking

By identifying where dangerous speech is occurring and by countering its impact on its intended audience, policy makers can protect populations from the mass atrocities that dangerous speech can facilitate, and they can do so without infringing on freedom of speech. There is a need to provide policy makers clear examples and actionable recommendations demonstrating how dangerous speech raises the risk for mass atrocities and how that risk can be mitigated. For example, through laws and regulations, governments can encourage media access and representation for a diversity of voices. Governments can ensure that educational curricula do not reflect prejudice and stereotypes, and they can develop curricula that promote critical thinking, including media literacy, and present the culture and accomplishments of minorities. State and political leaders can also counter the impact of dangerous speech by disseminating positive or inclusive messages regarding groups targeted by dangerous speech and by openly refuting and rejecting hateful rhetoric.

Too often, however, state and political leaders are the propagators or facilitators of dangerous speech. Government leaders have special obligations under international law and norms to defend universal human rights, to condemn and eliminate racial discrimination, and to protect their populations from genocide, war crimes, crimes against humanity and ethnic cleansing. The approaches for countering dangerous speech without infringing upon the right to free expression provide a means to fulfill these obligations. There is a need to develop international standards and norms rejecting state propagation and facilitation of dangerous speech, and non-restrictive approaches to countering dangerous speech should be incorporated into international efforts to enhance states' capacity to prevent atrocity crimes.

The Simon-Skjodt Center for the Prevention of Genocide

of the United States Holocaust Memorial Museum works to prevent genocide and related crimes against humanity. The Simon-Skjodt Center is dedicated to stimulating timely global action to prevent genocide and to catalyze an international response when it occurs. Our goal is to make the prevention of genocide a core foreign policy priority for leaders around the world through a multi-pronged program of research, education, and public outreach. We work to equip decision makers, starting with officials in the United States but also extending to other governments, with the knowledge, tools, and institutional support required to prevent—or, if necessary, halt—genocide and related crimes against humanity.

UNITED STATES
HOLOCAUST
MEMORIAL
MUSEUM

100 Raoul Wallenberg Place, SW
Washington, DC 20024-2126

