

**Written Testimony of NADINE STROSSEN,<sup>1</sup> John Marshall Harlan II Professor Law,  
New York Law School  
Hearing before Committee on Homeland Security, U.S. House of Representatives,  
June 26, 2019  
“Examining Social Media Companies’ Efforts to Counter Online Terror Content and  
Misinformation”**

**INTRODUCTION**

Chairman Thompson, Ranking Member Rogers, and distinguished Members of this Committee:

I am honored to join my esteemed co-panelists in addressing this hearing’s important topics. I appreciate the Committee members’ and panelists’ commitments to counter the potential<sup>i</sup> serious adverse impacts of the pertinent online expression: expression that could promote terrorism; and misinformation, which could defraud individuals and distort elections and other democratic processes and institutions.

I thank the Committee for exercising its important oversight functions to examine the content moderation policies of the powerful social media companies that are represented today.<sup>ii</sup> Even though any direct regulation of these policies would raise serious concerns about abridging the companies’ First Amendment rights, it is essential to consider how the companies should exercise those rights in ways that promote the free speech and other rights of the rest of us – and in ways that promote democracy, security, and other important concerns.

I understand that I was invited to complement this panel of social media leaders, despite my relative lack of specific experience with social media in particular, because of my longstanding scholarship and advocacy about freedom of speech for potentially harmful speech in general (including speech with terror content and misinformation) in many contexts, including online media.<sup>2</sup> For example, I was deeply involved in the developments leading to the historic 1997 Supreme Court case that first considered – and upheld -- First Amendment free speech rights online: *Reno v. ACLU*.<sup>3</sup> I was the national ACLU President throughout all the pertinent developments, including lobbying and litigating against Congress’s first online censorship law (enacted in 1996), which the high Court struck down, essentially unanimously. The Court celebrated the Internet as “a unique...medium of worldwide human communication,” whose “content...is as diverse as human thought.”<sup>4</sup>

Today’s discussion can gain much from the teachings of this landmark case, and from other past efforts to restrict various media expression feared to potentially cause harm. Not only did the Court strike down Congress’s first Internet censorship law in *Reno v. ACLU*, but it also struck down Congress’s revised version of that law in subsequent rulings.<sup>5</sup> Likewise, in its most recent decision about online expression, two years ago, the Court again unanimously struck down a law restricting such expression (in that case, a state law).<sup>6</sup> Moreover, the Court again hailed the unique importance of online communications, declaring:

While in the past there may have been difficulty in identifying the most important places...for the exchange of views, today the answer is clear. It is cyberspace--the "vast democratic forums of the Internet" in general, ... and social media in particular.<sup>7</sup>

---

<sup>i</sup> I deliberately refer to the “potential” adverse impacts of expression with terrorist content and misinformation because many experts have concluded that such expression will not necessarily contribute to the feared potential harms, and that non-censorial strategies such as the ones I discuss can significantly reduce that potential danger.

<sup>ii</sup> I am confining my focus to the dominant large companies, including the three that are represented at this hearing (Facebook, Google, and Twitter). They exercise outsize influence, thus as a practical matter requiring many people to use their services. Concerning smaller social media companies, potential users retain real choices about whether or not to participate. Accordingly, such smaller companies should (as a normative matter) have more latitude to choose content and define communities (again, as a legal matter, all of these companies have such latitude).

As the preceding outline of relevant Supreme Court rulings indicates, my support for online free expression is largely paralleled by the Court’s speech-protective rulings, and those in turn reflect the views of Justices across the ideological spectrum. Despite all the polarization in our political system and society, these particular issues about online expression should garner broad consensus in the other branches of government, as they have on the Court. Notwithstanding how divided our views might be on contested public policy issues, we all have the same stake in preserving the most robust freedom of speech for all such views -- no matter how extreme, controversial, or generally feared such views might be. In fact, those of us who are engaged in public policy debates have the greatest stake in strong freedom of speech. As the Court consistently has held, speech on public policy issues is the most important expression in our political system, essential not only for individual freedom and equality, but also for our very democracy itself. In its words: “Speech concerning public affairs is more than self-expression; it is the essence of self-government.”<sup>8</sup>

The speech at issue in today’s hearings – speech with terrorist<sup>iii</sup> content and misinformation – certainly concerns public affairs; indeed, that is precisely why it is potentially so harmful, as well as undeniably important. As the Court likewise has explained, such speech deserves the strongest protection not *despite* its potential serious harmful impact, but rather precisely *because of* such powerful potential. Let me quote a 2011 decision upholding freedom for extremely controversial, provocative speech (this decision was nearly unanimous, with only one dissenting vote):

Speech is powerful. It can stir people to action, move them to tears of both joy and sorrow, and -- as it did here -- inflict great pain. [W]e cannot react...by punishing the speaker. As a Nation we have chosen a different course -- to protect... speech on public issues to ensure that we do not stifle public debate.<sup>9</sup>

## **OVERVIEW**

I will first set out my three major conclusions about the important, challenging issues raised by today’s hearing. I will then lay out some more specific points that reinforce these conclusions.

## **THREE MAJOR CONCLUSIONS**

**FIRST. Any effort to restrict online terror content and misinformation will be at best ineffective and at worst counterproductive in achieving the important goal of such efforts: to counter the expression’s potential adverse impacts.**

As the Electronic Frontier Foundation [“EFF”] recently concluded:

[C]ontent moderation was never meant to operate at the scale of billions of users.... [A]s pressure from lawmakers and the public to restrict various types of speech – *from terrorism to fake news* -- grows, companies are desperately looking for ways to moderate content at scale. They won’t succeed – at least if they care about protecting online expression.<sup>10</sup>

EFF and others who have monitored content moderation efforts for years have consistently reached the same conclusion. For example, in a 2017 report, EFF stated:

Over the years, we’ve found that companies’ efforts to moderate online content almost always result in overbroad content takedowns or account deactivations. We therefore are justifiably skeptical [about] the latest efforts...to combat pro-terrorism content.<sup>11</sup>

Concepts such as “terror content” and “misinformation” are inherently, inescapably vague and broad. Therefore, anyone who decides whether particular social media posts should be so classified, and hence restricted, inevitably exercises enormous discretion. Enforcers of any such concepts will necessarily exercise this discretion in accordance with subjective values – their own, or those of their

---

<sup>iii</sup> The Committee’s designated topic for this hearing uses the phrase “*terror* content”; in this testimony, I also use the phrase “*terrorist* content” interchangeably.

social media employers, or those of powerful political and other established interests. As the old saying observes, “One person’s terrorist is another’s freedom fighter.” Likewise, one person’s “misinformation” or “fake news” is someone else’s cherished truth.

The definitions of prohibited terrorist or extremist content that Twitter and Facebook have enforced were cited as examples of these inevitable definitional problems of vagueness and overbreadth, in an important 2018 Report by the United Nations Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, David Kaye [“UN Special Rapporteur’s Report”].<sup>12</sup>

The unavoidable indeterminacy of these elastic concepts means that their enforcement will be arbitrary at best, discriminatory at worst. Predictably, these concepts will be disproportionately enforced against marginalized, unpopular, dissident individuals and groups, those that lack political power.

I will now cite a few illustrations of the foregoing general, inevitable problems specifically concerning the particular expression at issue: social media speech with terrorist content and misinformation.

**When social media target speech with terrorist (or “extremist”) content, they inevitably suppress much valuable speech, including human rights advocacy**

These problems were detailed, for example, in a May 30, 2019 joint report by the Electronic Frontier Foundation, Syrian Archive, and Witness.<sup>13</sup> Syrian Archive engages in “documentation related to human rights violations committed by all sides involved in the conflict in Syria,” and Witness promotes effective video advocacy for human rights. Noting that social media “companies have come under increasing pressure” to restrict extremist or terrorist expression, the report explained that both algorithmic and human content moderation techniques have “caught in the net” “not only content deemed extremist, but also...useful content like human rights documentation,” with “mistakes at scale that are decimating human rights content.” As the report elaborated:

[I]t is difficult for human reviewers – and impossible for machines -- to consistently differentiate activism, counter-speech, and satire about extremism from extremism itself. Blunt content moderation systems at scale inevitably make mistakes, and marginalized users are the ones who pay for those mistakes.

The report documented multiple examples of such counterproductively suppressed marginalized speakers, including groups advocating for the independence of the Chechen Republic of Iskeria, groups advocating for an independent Kurdistan, satirical commentary, and conflict documentation by journalists and human rights defenders in Syria, Yemen, and Ukraine.

In the same vein, a 2017 *New York Times* story described how YouTube’s “effort to purge extremist propaganda from its platform” had led it to “inadvertently remove[] thousands of videos that could be used to document atrocities in Syria, potentially jeopardizing future war crimes prosecutions.”<sup>14</sup> Given the breakdown of independent media since the start of the Syrian conflict, individuals and civil society organizations have subsequently used YouTube to document the war, including atrocities and human rights violations. Since some of the “disappeared” videos cannot be restored, we are losing “the history of this terrible war,” and “the richest source of information about human rights violations in closed societies,” according to experts whom the *Times* quoted.

This persistent problem, inherent in the irreducibly vague, overbroad concepts of terrorist or extremist content (as well as misinformation), had previously been documented in a 2017 EFF report, which cited further illustrations, including that “Facebook...deactivated the personal accounts of Palestinian journalists” on the ground that “they were involved in `terrorist activity,’” and temporarily banned a journalist from the United Arab Emirates “for posting a photograph of Hezbollah leader Hassan Nasrallah with an LGBTQ pride flag overlaid on it – a clear case of parody counter-speech that Facebook’s content moderators failed to grasp.”<sup>15</sup>

**Suppressing speech with terrorist content may well not promote counter-terrorism efforts and could even undermine them**

The Electronic Frontier Foundation has assembled expert testimony about the strategic downsides of suppressing this expression, even beyond the adverse impact that such suppression has on free speech:

[T]he question is *not* whether terrorists are using the Internet to recruit new operatives – the question is whether taking down pro-terrorism content and accounts will *meaningfully contribute* to the fight against global terrorism. Governments have not sufficiently demonstrated this to be the case. And some experts believe this absolutely *not* to be the case.<sup>16</sup>

Let me quote just a few of the many experts who have reached this negative conclusion, for the multiple reasons indicated.

**Censorship of terrorist content doesn't promote national security**

Michael German, a former FBI agent with counter-terrorism experience, who is now a fellow at the Brennan Center for Justice, stated: “Censorship has never been an effective method of achieving security, and...suppressing online content will be as unhelpful as smashing printing presses.”<sup>17</sup>

**Keeping terrorist content online may provide opportunities for constructive engagement that could avert terrorist acts**

For example, a Kenyan government official opposed shutting down an Al Shabaab Twitter account, because “Al Shabaab needs to be engaged positively and [T]witter is the only avenue.”<sup>18</sup>

More generally, this conclusion was reached by a United Nations report on “The Use of the Internet for Terrorist Purposes”:

Online discussions provide an opportunity to present opposing viewpoints or to engage in constructive debate, which may have the effect of discouraging potential supporters. Counter-narratives with a strong factual foundation may be conveyed through online discussion forums, images and videos. Successful messages may also demonstrate empathy with the underlying issues that contribute to radicalization, such as political and social conditions, and highlight alternatives to violent means of achieving the desired outcomes.<sup>19</sup>

A powerful specific example of the effective use of social media platforms to counter online terrorist propaganda comes from the U.S. Center for Strategic Counterterrorism Communications. Noting that the Center uses Facebook and YouTube for such purposes, the UN report cited one illustration of the touted strategy of “reducing radicalization and extremist violence by identifying in a timely manner extremist propaganda...on the Internet and responding swiftly with targeted counter-narratives”:

For instance, in May 2012, the Center...responded, within 48 hours, to banner advertisements promoting extremist violence posted on various websites by Al-Qaida in the Arabian Peninsula, with counter-advertisements on the same websites featuring an altered version of that same message that was intended to convey that the victims of the terrorist organization's activities were Yemeni nationals.<sup>20</sup>

**Keeping terrorist content online facilitates intelligence gathering and counter-terrorism efforts**

Let me again quote the above-cited UN report:

While terrorists have developed many ways to use the Internet in furtherance of illicit purposes, their use of the Internet also provides opportunities for the gathering of intelligence and other activities to prevent and counter acts of terrorism, as well as for the gathering of evidence for the prosecution of such acts. A significant amount of knowledge about the functioning, activities and sometimes the targets of terrorist organizations is derived from ...Internet

communications. Further, increased Internet use for terrorist purposes provides a corresponding increase in the availability of electronic data which may be compiled and analysed for counter-terrorism purposes. Law enforcement, intelligence and other authorities are developing increasingly sophisticated tools to proactively prevent, detect and deter terrorist activity involving use of the Internet.<sup>21</sup>

**Social media companies’ restrictions on misinformation likewise have suppressed much valuable information, and also have reinforced misinformation**

Efforts to clearly, consistently define and enforce prohibited “misinformation” are at least as futile as those to define prohibited “terror content.” The UN Special Rapporteur’s Report stressed the inevitable vagueness and overbreadth of restrictions on “disinformation,” warning that some such “measures, particularly those that...restrict[...]news content, may threaten independent and alternative news sources or satirical content.”<sup>22</sup> Likewise, EFF’s May 1, 2019 report concluded that “when tech companies ban an entire category of content” such as “disinformation,” “they have a history of overcorrecting and censoring accurate, useful speech – or, even worse, reinforcing misinformation.”<sup>23</sup>

One especially ironic illustration of the latter problem is a 2018 incident in which Facebook’s training materials used an egregious example of disinformation that was incendiary to boot. It was a photograph that depicted dozens of Buddhist monks surrounded by piles of dead, barely clothed bodies, which was captioned as “The Bod[ies] of Muslims slaught[er]ed by Buddhist[s].” Facebook’s training materials described this image as “a newsworthy exception” to Facebook’s general ban on nudity (another inherently vague, overbroad concept of restricted speech) because it depicted “the victims of violence in Burma [Myanmar].” In fact, though, this image actually depicted the aftermath of an earthquake in another country years earlier.<sup>24</sup>

**SECOND. Social media companies’ most effective strategies for countering the potential adverse impact of terrorist content and misinformation are non-censorial, including: altering the algorithmic curation that amplifies some potentially dangerous content; and empowering users with more individualized tools to understand and control the content they see, and to assess its credibility.**

As stated by Vera Eidelman, a staff attorney with the ACLU’s Speech, Privacy, and Technology Project: “Rather than focus[ing] their resources and innovation on how best to censor, [social media] companies should invest in user controls and enabling third party innovation re[garding] user controls and content moderation.”<sup>25</sup>

Likewise, in a May 1, 2019 report critiquing social media restrictions on “disinformation,” the EFF endorsed two interrelated technological approaches that social media should pursue to empower all of us to make our own voluntary, informed choices about what online material to view, and what not to view, consistent with our own interests and values: “addressing the algorithmic ‘megaphone’ at the heart of the problem and giving users control over their own feeds.”<sup>26</sup> Although this particular report focused on disinformation, its conclusions apply fully to other potentially problematic online content, including terrorist material:

Algorithms like Facebook’s Newsfeed or Twitter’s timeline make decisions about which...content to promote and which to hide. That kind of curation can play an amplifying role for some types of incendiary content, despite the efforts of platforms like Facebook to tweak their algorithms to “disincentivize” or “downrank” it. Features designed to help people find content they’ll like can too easily funnel them into a rabbit hole of disinformation. That’s why platforms should examine the parts of their infrastructure that are acting as a megaphone for dangerous content and address the root cause of the problem rather than censoring users.

Transparency about how a platform’s algorithms work, and tools to allow users to...create their own feeds, are critical...[Facebook’s] [r]ecent transparency improvements in this area are encouraging, but don’t go far enough...

Users shouldn’t be held hostage to a platform’s proprietary algorithm. Instead of...giving users just a few opportunities to tweak it, platforms should open up their APIs<sup>iv</sup> to allow users to create their own filtering rules for their own algorithms. News outlets, educational institutions, community groups, and individuals should all be able to create their own feeds, allowing users to choose who they trust to curate their information and share their preferences with their communities.

### **Additional non-censorial approaches**

In addition to the foregoing essential user empowerment strategies, other non-censorial approaches can also curb the potential adverse impact of terrorist content and misinformation more effectively than restricting such expression. These include:

- enforcing the many existing laws against actual terrorist and fraudulent conduct; and
- increasing media literacy, so consumers of online expression learn how to avoid terrorist and fraudulent communications, and how to find and generate effective “counterspeech,” refuting and responding to such problematic communications, dissuading other people from accepting their messages, and perhaps even dissuading those who issued the communications (as has happened in significant instances).

PEN America, which advocates for writers and free speech, has issued two recent reports about fraudulent news and disinformation (in March 2019 and October 2017),<sup>27</sup> which strongly endorse media literacy skills as the ultimate antidote to the potential serious adverse impact of such expression. As its 2019 report concluded: “[T]he most effective proactive tactic against fraudulent news is a citizenry that is well-equipped to detect, and reject, fraudulent claims.”<sup>28</sup> Correspondingly, that report concluded that “the spread of fraudulent news must not become a mandate for government or corporate censorship.”<sup>29</sup> Non-censorial steps that social media companies *should* take, according to PEN America, include “empower[ing] consumers with easy-to-use tools...to gauge the credibility of information disseminated through the platform.”<sup>30</sup>

### **THIRD. While social media companies have the legal right to engage in content moderation -- including efforts to restrict terrorist content and misinformation -- they should do so in ways that are consistent with universal human rights norms, including those governing freedom of expression. At a minimum, they should follow procedural standards that promote accountability, fundamental fairness/due process, and transparency.**

The Guiding Principles on Business and Human Rights, adopted by the United Nations Human Rights Council in 2011, urge companies to adhere to international human rights standards throughout their operations and wherever they operate.<sup>31</sup> Although these Principles are non-binding, the “overwhelming role” that the giant social media companies play “in public life globally argues strongly for their...implementation” of these Principles, according to the UN Special Rapporteur’s Report.<sup>32</sup>

In terms of free speech norms, the UN Special Rapporteur’s Report maintained that these companies should permit “users to develop opinions, express themselves freely and access information of all kinds in a manner consistent with human rights law.”<sup>33</sup> The applicable human rights law substantially overlaps with core U.S. free speech principles; it requires that any speech restriction should be clearly and

---

<sup>iv</sup> “API” is an abbreviation for “application program interface,” which is a set of routines, protocols, and tools for building software applications.

narrowly defined, and demonstrated to be both necessary and proportionate to avert specific, serious harm that the speech would directly cause. For speech that is feared to have a more indirect, speculative harmful potential, we should respond with non-censorial measures, as outlined above.

Concerning minimal procedural standards, a starting point is the “Santa Clara Principles On Transparency and Accountability in Content Moderation,” which were adopted in 2018 by a group of civil liberties organizations and individual experts.<sup>34</sup> These minimum procedural principles have also been endorsed by the UN Special Rapporteur’s Report, and at least their general “spirit” has been endorsed by many major social media companies, including all three companies represented at this hearing.<sup>35</sup>

The Santa Clara Principles spell out detailed steps that social media companies should take to pursue the following broader initiatives:

- 1) Publishing the numbers of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines;
- 2) Providing notice to each user whose content is taken down or whose account is suspended about the reason for such action; and
- 3) Providing a meaningful opportunity for timely appeal of any content removal or account suspension.

### **MORE SPECIFIC POINTS THAT REINFORCE THESE MAJOR CONCLUSIONS**

#### **1<sup>st</sup>) Throughout history, we have seen a constant impulse to disproportionately blame expression for societal problems, which is understandable but misguided.**

Correspondingly, it seems to be intuitively appealing to seek to suppress expression of ideas that one disfavors or fears to be potentially dangerous. As former Supreme Court Justice Oliver Wendell Holmes memorably put it:

Persecution for the expression of opinions seems to me perfectly logical. If you have no doubt of your premises or your power, and want a certain result with all your heart, you naturally express your wishes in law, and sweep away all opposition.<sup>36</sup>

Nonetheless, Holmes even more memorably explained why that tempting speech-blaming/speech-suppressive instinct is inconsistent with individual liberty and democracy, setting out the “emergency principle” that all modern Justices have embraced:

[W]e should be eternally vigilant against attempts to check the expression of opinions that we loathe and believe to be fraught with death, unless they so imminently threaten immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country.<sup>37</sup>

The general pattern of scapegoating expression for allegedly fostering societal problems is especially pronounced concerning expression that is conveyed by any new media. Each new, powerful communications medium raises concerns about its ability to transmit controversial expression to vulnerable individuals and groups, provoking fear – even panic – about potential ensuing harm. Accordingly, throughout history, censorial efforts have greeted each new communications medium: from the printing press to the Internet.

Let me list some examples of media-blaming for a range of serious social problems, just within my adult lifetime:

- Sexual expression in all media, including online, has been blamed for undermining everything from women’s equality and safety to the “traditional American family.”
- So-called “hate radio” has been blamed for fomenting domestic extremism and terrorism.
- Violent videos have been blamed for instigating school shootings.

- Rap and rock lyrics have been blamed for instigating sexual assaults against women and also shootings of police officers.

**2<sup>nd</sup>) With 20-20 hindsight, we have consistently come to realize that this scapegoating of expression as a purported major cause of social ills – and the associated calls for censorship as a purported solution -- have multiple interrelated flaws.**

- This approach wrongly regards individuals as passive automatons who lack autonomy to make our own choices about what expression to view and what not to, and to avoid being passively “brainwashed” by what we do view.
  - To be sure, as discussed above, social media companies and others should take affirmative steps to maximize individual freedom of choice in this sphere. We should ensure that all media consumers have the educational and technological resources to make truly independent, informed, voluntary decisions about our communications. In the social media context, this means *not* directing users to increasingly extreme content, unbeknownst to them. It *does* mean developing and deploying technology that will empower each user to make maximally individualized choices about what content to view and to communicate, and what to avoid or block.
- Scapegoating expression also diverts attention and resources from the real problems: underlying attitudes and actual conduct. Suppressing expression is a superficial, cheap “quick fix” for complex, deep-rooted problems, which actually fixes nothing. To the contrary, pushing the feared ideas underground may well make it harder to counter those ideas, and also harder to prevent those who hold them from engaging in harmful conduct.
- This censorial approach may not only make it harder to recruit advocates of the feared ideas/actions away from their ideologies, but it may well also increase attention, sympathy, and support for such ideologies among members of the broader public. This pattern is so common that several terms have been coined to describe it, including “the forbidden fruits effect,” “the boomerang effect,” and “the Streisand effect” (the latter term was coined when Barbra Streisand sought to block online photographs of her Malibu home, thus increasing exponentially the viewing of such photographs). We should focus instead on persuading people to reject dangerous ideas, and preventing people from engaging in harmful conduct.

**3rd) Social media companies are not constrained by the First Amendment’s Free Speech Clause, which limits only government actors, not private sector actors. To the contrary, social media companies have their own First Amendment rights, including the right to decide which speakers and expression to permit – or not to permit – on their platforms. However, these platforms should provide the same free speech opportunities that government is required to provide, consistent with both compelling policy concerns and global human rights norms applicable to business.**

As I noted at the outset of this testimony, social media companies have their own First Amendment rights to adopt and enforce content moderation policies they choose, and any government regulation of such policies – whether *prescribing* or *proscribing* – any such policies – would raise serious concerns about abridging the companies’ freedom of speech. However, it is eminently appropriate for Congress and other government actors, as well as civil society groups and users, to encourage these companies to implement content moderation policies, and to take other actions, that promote their users’ free speech, as well as promoting other essential concerns, including national security and democracy.

As a practical matter, social media platforms now constitute the most important forums for exchanging information and ideas, including between “We the People” (to quote the Constitution’s

opening words) and the political candidates and officials who are accountable to us. In a 2017 Supreme Court decision that unanimously struck down a state law that restricted access to social media by convicted sex offenders who had served their prison terms, Justice Anthony Kennedy’s majority opinion stressed the social media’s stature as the preeminent platform for expression. If we do not have equal, open access to these forums, to convey and receive communications, then for all practical purposes, our freedom of speech – and, accordingly, our equal stature as sovereign citizens -- is curtailed. As Justice Kennedy declared:

A fundamental principle of the First Amendment is that all persons have access to places where they can speak and listen, and then, after reflection, speak and listen once more. The Court has sought to protect the right to speak in this spatial context. A basic rule, for example, is that a street or a park is a quintessential forum for the exercise of First Amendment rights.... While in the past there may have been difficulty in identifying the most important places (in a spatial sense) for the exchange of views, today the answer is clear. It is cyberspace...and social media in particular.<sup>38</sup>

Moreover, as discussed above, social media companies should adhere to the UN Human Rights Council’s Guiding Principles on Business and Human Rights, which include respect for free speech. As the UN Special Rapporteur urged, social media companies should engage in content moderation that permits “users to express themselves freely and access information of all kinds in a manner consistent with human rights law.”<sup>39</sup>

**4<sup>th</sup>) A core free speech principle that social media companies should honor, consistent with both U.S. and human rights law, is “content neutrality” or “viewpoint neutrality”: that speech should not be restricted solely due its disfavored content – i.e., its viewpoint, message, or ideas.**

No matter how loathed or feared such content may be, by no matter how many of us, we must respond to it with non-censorial counter measures, including education and persuasion. Measures that discriminate against speech based solely on its disfavored content or viewpoint are almost automatically unconstitutional. The Supreme Court has hailed content neutrality as “the bedrock principle” undergirding constitutional freedom of speech.<sup>40</sup>

This fundamental free speech principle is reflected in international human rights norms. Accordingly, the UN Special Rapporteur’s Report expressly urges social media companies to enforce their content moderation policies consistent with a “non-discrimination” standard, rather than through “heavy-handed viewpoint-based regulation.”<sup>41</sup>

**5<sup>th</sup>) Social media companies should additionally honor the complementary “emergency” principle, which is also integral to both U.S. and human rights law.**

When we move beyond the *content* of speech and consider its *context*, speech may be restricted if it satisfies the emergency test: when, under all the facts and circumstances, the speech directly causes certain specific, imminent, serious harm, which cannot effectively be countered through non-censorial measures.

This key principle is also reflected in the global human rights requirements of “necessity” and “proportionality.” As the UN Special Rapporteur’s Report explained, proponents of any speech restriction “must demonstrate that the restriction imposes the least burden on the exercise of” free speech, “and actually protects, or is likely to protect, the legitimate...interest at issue.” Proponents of the restriction “may not merely assert necessity but must demonstrate it, in the restriction of specific expression.” Moreover, social media “[c]ompanies should ...demonstrate the necessity and proportionality of any content actions (such as removals or account suspensions).”<sup>42</sup>

Applying these standards to terror content and misinformation, social media companies should not restrict such expression unless they could demonstrate that the restriction was “necessary” and “proportional” for averting the potential harms of such expression. This showing would be hard to make concerning either terror content or misinformation, in light of the evidence discussed above, which demonstrated the inherent overbreadth of such speech restrictions, and called into question whether they are even effective in averting the potential harms, let alone necessary.

**6<sup>th</sup>) The Supreme Court has designated several narrowly defined categories of speech that may be restricted consistent with the content neutrality and emergency principles, including two that are pertinent to the two types of speech at issue in these hearings: speech with terrorist content and misinformation. It would be appropriate for social media companies to restrict these narrowly defined subcategories of speech with terrorist content and misinformation: speech that satisfies the standards for punishable incitement, fraud, or defamation.**

- The Court has barred government from restricting speech that contains terrorist content or speech that is feared to potentially contribute to terrorism unless, in context, it satisfies the following, appropriately strict, standards: it intentionally incites imminent violent or criminal conduct, and it is actually likely to do so imminently. Accordingly, the Court has struck down even restrictions on explicit advocacy of violent or criminal conduct, including terrorism, when it falls short of the foregoing strict intentional incitement standard.<sup>43</sup>
- The Court has barred government from punishing many kinds of “misinformation” and even outright lies, except in certain situations when intentional falsehoods directly cause certain specific imminent serious harms, including by defrauding an individual who has reasonably relied on the falsehood in a way that causes demonstrable tangible injury; or by defaming an individual about a matter of private concern in a way that injures her reputation and causes demonstrable tangible injury. When the defamatory falsehood pertains to a public official or public figure, it may not be punished unless the complainant can establish, by “clear and convincing evidence,” that the speaker knowingly or recklessly lied.<sup>44</sup>

**7<sup>th</sup>) Speech that does not satisfy the emergency test may still cause serious harm; that is true for speech with terrorist content and misinformation. However, the modern Court has consistently enforced the content neutrality and emergency principles because it is even more harmful to grant enforcing authorities latitude to punish speech that does not satisfy the emergency test. This is true regardless of who the enforcing authorities are, including social media companies.**

As Supreme Court Justice Oliver Wendell Holmes famously recognized, “Every idea is an incitement.”<sup>45</sup> He did not mean that government may therefore suppress every idea, but rather the opposite: if every idea that could potentially incite harmful conduct or consequences could be suppressed, all ideas could be suppressed. Accordingly, to shield freedom to express ideas – potentially inciting and potentially dangerous as they are – we should confine censorial power only to ideas and expression that satisfy the emergency test: directly causing specific, imminent, serious harm.

If censorial power could be exercised under a looser, broader standard, the resulting discretion would inevitably lead to suppressing valuable speech, and would disproportionately target speech by relatively disempowered, marginalized individuals and groups, including those who challenge the status quo.

- For example, before the Supreme Court adopted the strict “intentional incitement” test, government regularly enforced legal restrictions on “terrorist” and other feared expression against politically unpopular, relatively powerless speakers, including abolitionists, socialists,

women’s suffragists, pacifists, anti-war and anti-draft demonstrators, and civil rights advocates.

- Likewise, before the Supreme Court adopted strict standards limiting punishable defamation, national media outlets and civil rights leaders and organizations were regularly targeted with defamation lawsuits that (absent the Court’s invalidation) would have led to speech-suppressive damage awards, preventing information and advocacy about the civil rights movement from reaching the critically important nationwide audience.
- When expression may be restricted short of the emergency standard, the restrictions are often counterproductive: suppressing expression that would actually promote the countervailing goals at issue.
- As detailed above, these general, inevitable problems have – predictably -- specifically afflicted social media companies’ enforcement of their standards that restrict terrorist content and misinformation.

### CONCLUSION

In closing, I would like to invoke an apt observation by the famed twentieth century journalist H.L. Mencken: “For every complex problem, there is a solution that is clear, simple – and wrong.”

How to effectively counter the serious potential adverse impact of terror content and misinformation is certainly a complex problem. While restricting such expression might appear to be a clear, simple solution, it is in fact neither – and, moreover, it is wrong. We must focus on the non-censorial strategies I have discussed, including user empowering education and technology. Although these approaches are also not simple, they are far more promising than censorship.

- 
1. Nadine Strossen is the John Marshall Harlan II Professor of Law at New York Law School and the immediate past national President of the American Civil Liberties Union (1991-2008). She gratefully acknowledges the following NYLS students for providing valuable assistance with this testimony, including the preparation of endnotes: Managing Research Assistant Marc D. Walkow, and Research Assistants Aaron Hansen and Serene Qandil.
  2. See, e.g., Nadine Strossen, *HATE: Why We Should Resist It with Free Speech, Not Censorship* (New York: Oxford University Press, 2018).
  3. *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997). Justice O’Connor authored a partial dissent, in which Justice Rehnquist joined, but this concerned only a narrow particular application of the statute (as applied to an online communication involving only one adult and one or more minors, such as when an adult knowingly sends an email to a minor); both of these Justices agreed with the majority’s broad holdings about the law’s general unconstitutionality, making the decision essentially unanimous. *Reno*, 521 U.S. at 886 (O’Connor, J., concurring in part and dissenting in part).
  4. *Reno*, 521 U.S. at 850, 852.
  5. *Ashcroft v. American Civil Liberties Union*, 535 U.S. 564 (2002); *Ashcroft v. American Civil Liberties Union*, 542 U.S. 656 (2004), *cert. denied* *Mukasey v. American Civil Liberties Union*, 2009 U.S. LEXIS 598 (2009).
  6. *Packingham v. North Carolina*, 137 S. Ct. 1730 (2017).
  7. *Packingham*, 137 S. Ct. at 1735.
  8. *Garrison v. Louisiana*, 379 U.S. 64, 74-75 (1964).
  9. *Snyder v. Phelps*, 562 U.S. 443, 460-61 (2011).
  10. Jillian C. York and Corynne McSherry, “Content Moderation is Broken. Let Us Count the Ways,” *Electronic Frontier Foundation*, Apr. 29, 2019, <https://www.eff.org/deeplinks/2019/04/content-moderation-broken-let-us-count-ways>.
  11. Sophia Cope, Jillian C. York and Jeremy Gillula, “Industry Efforts to Censor Pro-Terrorism Online Content Pose Risks to Free Speech,” *Electronic Frontier Foundation*, July 12, 2017, <https://www.eff.org/deeplinks/2017/07/industry-efforts-censor-pro-terrorism-online-content-pose-risks-free-speech>.

- 
12. David Kaye, *Report of the Special Rapporteur to the Human Rights Council on online content regulation*, ¶ 26, U.N. Doc. A/HRC/38/35 (April 6, 2018), <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf>.
  13. Jillian C. York, "Caught in the Net: The Impact of 'Extremist' Speech Regulations on Human Rights Content," *Electronic Frontier Foundation*, May 30, 2019, <https://www.eff.org/wp/caught-net-impact-extremist-speech-regulations-human-rights-content>.
  14. Malachy Browne, "YouTube Removes Videos Showing Atrocities in Syria," *The New York Times*, Aug. 22, 2017, <https://www.nytimes.com/2017/08/22/world/middleeast/syria-youtube-videos-isis.html>.
  15. Cope, York, and Gillula, "Industry Efforts."
  16. Cope, York, and Gillula, "Industry Efforts."
  17. Jenna McLaughlin, "The White House Asked Social Media Companies to Look for Terrorists. Here's Why They'd #Fail.," *The Intercept*, Jan. 20, 2016, <https://theintercept.com/2016/01/20/the-white-house-asked-social-media-companies-to-look-for-terrorists-heres-why-theyd-fail>.
  18. Jillian C. York and Trevor Timm, "U.S. Government Threatens Free Speech With Calls for Twitter Censorship," *Electronic Frontier Foundation*, Jan. 6, 2012, <https://www.eff.org/deeplinks/2012/01/us-government-calls-censor-twitter-threaten-free-speech>.
  19. United Nations Office on Drugs and Crime, *The use of the Internet for terrorist purposes* (Vienna: United Nations, 2012), 12.
  20. UNODC, *Use of the internet*, 13.
  21. UNODC, *Use of the internet*, 12.
  22. Kaye, *Report of the Special Rapporteur*, ¶ 31.
  23. Jillian C. York, David Greene, and Gennie Gebhart, "Censorship Can't be the Only Answer to Disinformation Online," *Electronic Frontier Foundation* (May 1, 2019), <https://www.eff.org/deeplinks/2019/05/censorship-cant-be-only-answer-disinformation-online>.
  24. Joseph Cox, "Facebook's Own Training Materials Fell for Fake News," *Motherboard/Tech by Vice* (Sep. 5, 2018), [https://www.vice.com/en\\_us/article/j5ny5d/facebook-training-manuals-documents-fell-fake-news](https://www.vice.com/en_us/article/j5ny5d/facebook-training-manuals-documents-fell-fake-news).
  25. Vera Eidelman, email message to Nadine Strossen, May 28, 2019.
  26. York, Greene, and Gebhart, "Censorship Can't be the Only Answer."
  27. "Truth on the Ballot: Fraudulent News, the Midterm Elections, and Prospects for 2020," *PEN America* (Mar. 13, 2019), <https://pen.org/wp-content/uploads/2019/03/Truth-on-the-Ballot-report.pdf>; "Faking News: Fraudulent News and the Fight for Truth," *PEN America* (Oct. 12, 2017), <https://pen.org/wp-content/uploads/2017/11/2017-Faking-News-11.2.pdf>.
  28. PEN America, "Truth on the Ballot," 7.
  29. PEN America, "Truth on the Ballot," 48.
  30. PEN America, "Faking News," 27.
  31. Kaye, *Report of the Special Rapporteur*, ¶ 6.
  32. Kaye, *Report of the Special Rapporteur*, ¶ 5.
  33. Kaye, *Report of the Special Rapporteur*, ¶ 15.
  34. "Santa Clara Principles on Transparency and Accountability in Content Moderation," *The Santa Clara Principles* (May 2018), <https://santaclaraprinciples.org>. The authors of the Principles were the ACLU of Northern California, The Center for Democracy & Technology, Electronic Frontier Foundation, New America's Open Technology Institute, Irina Raicu, Nicolas Suzor, Sarah T. Roberts, and Sarah Myers West.
  35. Gennie Gebhart, "Who Has Your Back? Censorship Edition 2019," *Electronic Frontier Foundation* (June 12, 2019), <https://www.eff.org/wp/who-has-your-back-2019>.
  36. *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).
  37. *Ibid.*
  38. *Packingham*, 137 S. Ct. at 1735.
  39. Kaye, *Report of the Special Rapporteur*, ¶ 39.
  40. *Texas v. Johnson*, 491 U.S. 397, 414 (1988).
  41. Kaye, *Report of the Special Rapporteur*, ¶¶ 48, 66.
  42. Kaye, *Report of the Special Rapporteur*, ¶¶ 7, 28, 66.
  43. *Brandenburg v. Ohio*, 395 U.S. 444 (1969).
  44. *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964).
  45. *Gitlow v. New York*, 268 U.S. 652, 673 (1925) (Holmes, J., dissenting).

---