UNIVERSITY *of* PENNSYLVANIA
JOURNAL *of* LAW & PUBLIC AFFAIRS

Vol. 5 January 2020 No. 2

INCITEMENT IN AN ERA OF POPULISM: UPDATING *BRANDENBURG* AFTER CHARLOTTESVILLE

Richard Ashby Wilson[*] *& Jordan Kiper*[**]

We live in an era of populism, characterized by political polarization, inciting speech on social media, and an escalation in hate crimes. The regulatory framework for direct incitement to imminent lawless action established fifty years ago in Brandenburg *is showing signs of severe strain. One of the central frailties of* Brandenburg*'s three-part test is the lack of guidance on how courts should evaluate the probability that an inciting speech act will cause an imminent offense. In the absence of clear direction on analyzing risk, judges often rely on outdated heuristics and misleading metaphors. This article is the first to draw on behavioral research to construct a systematic evidence-based framework for assessing the likelihood that inciting speech will result in imminent lawless action. This matrix is then applied to the fact pattern in* Sines v. Kessler*, a civil suit arising from the events in Charlottesville, Virginia in 2017.*

[*] Gladstein Distinguished Professor of Human Rights and Professor of Law and Anthropology at the University of Connecticut School of Law.
[**] Assistant Professor of Anthropology at the University of Alabama, Birmingham.
This research was supported in part through a visiting scholarship from the Russell Sage Foundation. Any opinions expressed are those of the author and should not be construed as representing the opinions of the Foundation. Sincere thanks to colleagues Leslie Levin, David Richards, Alex Tsesis, and the participants at seminars at Columbia University, Max Planck Institute-Halle, the New School for Social Research, New York University, Pace School of Law, and Yale Law School for their constructive comments. Thanks also to Sarah Hamilton, Allaina Murphy, Brendan McPherson, Anne Rajotte, and Katie Winograd for their assistance. Any errors of fact or interpretation are our own.

> *"[W]hen a speaker incites a crowd to violence, his incitement does not receive constitutional protection."*
> –Bible Believers v. Wayne County, Michigan

INTRODUCTION

We live in an era of populist politics that is characterized by intense emotional attributions of blame to elites,[1] and an increase in political communication that emphasizes threat, anger, and fear.[2] While populism is

[1] *See* Nadia Urbinati, *Political Theory of Populism,* 22 ANN. REV. POL. SCI. 111, 119 (2019) (discussing the role of anti-elite populism in recent political elections).
[2] *See* Michael Hameleers, Linda Bos & Claes H. de Vreese, *"They Did It": The Effects of Emotionalized Blame Attribution in Populist Communication,* 44 COMM. RES. 870, 871-72 (2017) (analyzing how anger, fear, and emotions as used in populist communication affect different groups of people).

time-honored, what is new about the current wave of populism is the degree to which polarizing and caustic messages are amplified daily on social media platforms that moderate the content of online speech under a First Amendment framework.[3] Vigorous political communication, even that which is offensive and reprehensible to some, is protected by the First Amendment and much political speech can be met with a variety of non-censorial measures such as counterspeech.[4]

However, a segment of political speech may constitute incitement to imminent physical attacks on political opponents or particular social groups.[5] Since "[t]he First Amendment does not protect violence,"[6] incitement is not constitutionally protected,[7] and implies a framework of regulation. The thesis advanced here is that our current legal system is ill-equipped to deal with inciting speech because it does not possess a systematic framework to evaluate which speech causes the greatest risk of imminent violence.

One corollary of current populist politics in the United States since 2016 has been an increase in bias-motivated crimes against persons or property. The 2016 presidential election occurred against the backdrop of one

---

[3] *See generally* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1599 (2018) (explaining how social media companies' content moderation policies are shaped by the First Amendment).

[4] *See* NADINE STROSSEN, HATE: WHY WE SHOULD RESIST IT WITH FREE SPEECH, NOT CENSORSHIP 158 (2018) (discussing the importance of counterspeech as a response to hate speech).

[5] *See* Marc Rohr, *Grand Illusion? The* Brandenburg *Test and Speech that Encourages or Facilitates Criminal Acts*, 38 WILLAMETTE L. REV. 1, 14 (2002) (explaining the imprecision of the *Brandenburg* test, which is intended to articulate the limits of freedom of speech); Alexander Tsesis, *Inflammatory Speech: Offense Versus Incitement,* 97 MINN. L. REV. 1145, 1145-49 (2010) (distinguishing offensive speech from incitement); Leslie Kendrick, Note, *A Test for Criminally Instructional Speech*, 91 VA. L. REV. 1973, 1987-88 (2005) (discussing the role of imminence in differentiating between criminal instructive speech, incitement, and political advocacy). *But see* Eugene Volokh, *Crime-Facilitating Speech*, 57 STAN. L. REV. 1095, 1103-06 (2005) (analyzing speech that facilitates the actual commission of crimes). *See generally* Steven Gey, *The* Brandenburg *Paradigm and Other First Amendments*, 12 U. PA. J. CONST. L. 971 (2010) (examining the applicability of the *Brandenburg* paradigm to nonpolitical speech).

[6] NAACP v. Claiborne Hardware Co., 458 U.S. 886, 916 (1982).

[7] "Speech that falls within th[e] category of incitement is not entitled to First Amendment protection." James v. Meow Media, Inc., 300 F.3d 683, 698 (6th Cir. 2002) (considering negligence claims against producers of violent video games and movies). An inciter is a person who "counsels, commands or advises the commission of a crime," according to GLANVILLE WILLIAMS, CRIMINAL LAW 612 (2d ed. 1962).

of the largest recorded increases in hate crimes[8] in U.S. history, second only to the upsurge after the terrorist attacks of September 11, 2001.[9] In 2016, the nation's law enforcement agencies[10] reported 6121 single-bias incidents of hate crimes involving 7509 victims, an increase of 5% in the number of victims, even as overall crime decreased.[11] In 2017, the trajectory of hate crimes continued to climb steeply, and agencies reported 7106 single-bias incidents, an increase of 17% on the year before.[12] Of note in 2017 was the 31% rise in hate crimes committed on the basis of religious animus, primarily against Jews and Muslims.[13]

Social media is often identified as a contributing factor for much that ails modern society, including political polarization, but the mere fact that so much political discourse occurs online has allowed researchers to analyze more precisely than before the relationship between online speech and offline hate crimes. Their findings have been sobering.[14] For instance, Müller and

---

[8] The Hate Crime Statistics Act, 34 U.S.C. § 41305 (2018) defines hate crimes as "crimes that manifest evidence of prejudice based on race, gender or gender identity, religion, disability, sexual orientation, or ethnicity." The Bureau of Statistics' National Crime Victimization Survey (NCVS) classifies a crime as a hate crime when the victim reports "at least one of three types of evidence that the act was motivated by hate: (1) the offender used hate language, (2) the offender left behind hate symbols, or (3) police investigators confirmed that the incident was hate crime." *Hate Crime*, BUREAU OF JUSTICE STATISTICS, DEP'T OF JUSTICE, https://www.bjs.gov/index.cfm?ty=tp&tid=37 [https://perma.cc/P833-J35U] (last visited Jan. 28, 2019).

[9] *See* Griffin Edwards & Stephen Rushin, The Effect of President Trump's Election on Hate Crimes 6-7 (Jan. 18, 2018) (unpublished manuscript), https://ssrn.com/abstract=3102652 [https://perma.cc/WW6G-YL6K] (analyzing the effects of political speech during the U.S. presidential elections).

[10] Hate crimes are underreported by law enforcement agencies, and economists find evidence for as many as 50,000 hate crimes a year in the United States. *See* Dhammika Dharmapala & Richard H. McAdams, *Words That Kill? An Economic Model of the Influence of Speech on Behavior (with Particular Reference to Hate Speech),* 34 J. LEGAL STUDS. 93, 94 n.3 (2005) (noting the difference between statistics published by the FBI and the Southern Poverty Law Center, which estimates there are 50,000 hate crimes annually).

[11] CRIMINAL JUSTICE INFO. SERV. DIV., FED. BUREAU OF INVESTIGATION, UNIFORM CRIME REPORT, HATE CRIMES STATISTICS, INCIDENTS & OFFENSES 1-2 (2016) [hereinafter 2016 UNIFORM CRIME REPORT].

[12] *Id.*; CRIMINAL JUSTICE INFO. SERV. DIV., FED. BUREAU OF INVESTIGATION, UNIFORM CRIME REPORT, HATE CRIMES STATISTICS, INCIDENTS & OFFENSES 1-2 (2017) [hereinafter 2017 UNIFORM CRIME REPORT].

[13] In 2016, the FBI reported 1,538 religious bias incidents and in 2017, it reported 1,679. 2016 UNIFORM CRIME REPORT, *supra* note 11; 2017 UNIFORM CRIME REPORT, *supra* note 12.

[14] *See generally* Karsten Müller & Carlo Schwarz, Fanning the Flames of Hate: Social Media and Hate Crime (Feb. 19, 2018) (unpublished manuscript), https://ssrn.com/abstract= 3082972 [https://perma.cc/RY4Z-ZLBS] (comparing the context of online hate speech in Germany and

Schwarz identify a statistically significant correlation between anti-immigrant as well as anti-Muslim tweets and actual attacks on Muslims and immigrants in 2016.[15] The unprecedented conjuncture of populism and social media has created new challenges for the regulation of incitement and problematic internet content more broadly,[16] given that the present legal regime was designed in the 1960s, during the epoch of television, mass-circulation newspaper, and leaflets passed out by hand on the urban street-corner.

We can identify these new challenges in incitement cases currently winding their way through the U.S. courts.[17] One illuminating case is *Sines v. Kessler*, a civil suit filed in the aftermath of the events in Charlottesville, Virginia, on August 11 and 12, 2017.[18] The case relies on the defendants' speech on Discord, an "invitation only" social media platform, in the days before the violence in which a protestor was killed.[19] Through posts on the web platform, the defendants planned and coordinated the march of white

---

the USA in 2016 & 2017); Jonathan Rothwell & Pablo Diego-Rosell, Explaining Nationalist Political Views (Nov. 2, 2016) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2822059 [https://perma.cc/UZ57-HKAC] (examining support for nationalist policies and racial resentment); Edwards & Rushin, *supra* note 9.

[15] Karsten Müller & Carlo Schwarz, From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment 25-28 (Nov. 2, 2019) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3149103 [https://perma.cc/TDY2-86F4].

[16] *See, e.g.*, Nina I. Brown & Jonathan Peters, *Say This, Not That: Government Regulation and Control of Social Media*, 68 SYRACUSE L. REV. 521, 527 (2018) (arguing that there is little the federal government can do regarding content posted on social media); Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 454 (2017) (noting that "efforts to hold social media companies responsible under the civil provisions of the federal material-support statute have consistently failed"); Gill Grassie, *The Campaign Against Hate Crime Online—Can Lessons Be Learned from the IP Takedown Experience to Date?*, 21 J. INTERNET L. 3, 3 (2017) (discussing government efforts to regulate social media companies with respect to racist or defamatory content); James Grimmelmann, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 55-76 (2015) (providing a taxonomy of moderation in online communities); Klonick, *supra* note 3, at 1635 (examining how content is moderated ex ante vs. ex post); Kyle Langvardt, *Regulating Online Content Moderation,* 106 GEO. L.J. 1353, 1357 (2018) (considering different policies on content moderation); Alexander Tsesis, *Social Media Accountability for Terrorist Propaganda*, 86 FORDHAM L. REV. 605, 628 (2017) (discussing how Congress should look to international legislative models when drafting material-support statutes).

[17] *See, e.g.*, Sines v. Kessler, 324 F. Supp. 3d 765 (W.D. Va. 2018) (alleging white supremacists conspired to commit racial violence and asking whether the defendants' incitement laden social media posts violated Virginia's hate crimes statute); Gersh v. Anglin, 353 F. Supp. 3d 958 (D. Mont. 2018) (seeking punitive damages for the defendant's inciting and threatening posts against a Jewish woman on the Daily Stormer website).

[18] *Sines*, 324 F. Supp. 3d at 765.

[19] *Id*. at 776, 779.

nationalists and white supremacists in Charlottesville, provided videos on fighting techniques, and urged marchers to bring weapons such as semiautomatic rifles, handguns and knives.[20] Most significantly, the defendants encouraged unlawful acts of violence, posting explicit calls for violence against protestors and making approving remarks about running over protestors with a vehicle, which in fact occurred.[21]

On the basis of the march organizers' social media posts, the U.S. District Court for the Western District of Virginia has confirmed the injuries suffered by the protestors were "reasonably foreseeable"[22] in advance. This raises the question, could they have been prevented if there had been an adequate risk assessment prior to the march? Answering requires that we delve more deeply into *Brandenburg*, the constitutional precedent on incitement.[23] In *Brandenburg*, a 1969 case concerning a Ku Klux Klan leader from Ohio, the U.S. Supreme Court established the precedent that speech could not be suppressed unless it was "directed to inciting or producing imminent lawless action and is likely to incite or produce such action."[24] *Brandenburg*'s three-part test abjures the suppression of mere advocacy of an offense, and requires that the speaker (1) intended for a crime to be committed and advocated an offense that is both (2) imminent and (3) likely to occur.[25]

While subsequent jurisprudence has gone some way to spelling out the meaning of the "advocacy" and "imminence" elements, courts have provided very little direction regarding how likely a crime must be, and what criteria should be utilized in risk analysis. Incitement law therefore demands that courts assess the risks accompanying hazardous public utterances but has, thus far, not furnished the necessary tools to assess those risks. *Brandenburg*'s lack of precision on the imminence and likelihood prongs hinders its responsiveness to a resurgence of discriminatory animus in political discourse. We are in a similar moment to the mid-2000s, when some legal scholars claimed that the "War on Terror" had stretched *Brandenburg* to its breaking point in cases involving post-9/11 terrorist propagandists.[26]

This Article is the first to draw on recent behavioral research on persuasion, political communication, and dehumanizing speech to create a

---

[20] *Id.* at 776, 804-05.
[21] *Id.* at 796.
[22] *Id.* at 797.
[23] Brandenburg v. Ohio, 395 U.S. 444 (1969) (per curiam).
[24] *Id.* at 447.
[25] *Id.*
[26] *See generally* Laura K. Donohue, *Terrorist Speech and the Future of Free Expression*, 27 CARDOZO L. REV. 233, 240 (2005) (discussing the implications of terrorist speech on free speech norms, and the strength of *Brandenburg* as a matter of precedent); Thomas Healy, *Brandenburg in a Time of Terror*, 84 NOTRE DAME L. REV. 655 (2009) (examining the continuing strength of Brandenburg in an age of political jihadism).

rigorous and empirically-tested framework to guide legal actors as they assess the probability that speech will culminate in imminent lawless action. Given the uncertainty surrounding the imminence and likelihood elements of incitement doctrine and the widespread occurrence of bias-motivated crimes, it is a propitious time to turn to behavioral research for guidance on the types of speech and contexts that are most likely to end in violent confrontations.

This Article proceeds in four parts. Part I provides background on incitement in the twentieth century and how the "bad tendency" and "clear and present danger" tests were excessively restrictive of political speech, and neither could operationalize elements of imminence and the context of utterances. Part II considers *Brandenburg*'s test for incitement and parses the three elements, observing that direct advocacy has received the lion's share of the attention of courts, while imminence and probability have received short shrift. Part III develops an original critique of current incitement jurisprudence, observing how judges are prone to rely on personal hunches, heuristics, and common metaphors to describe the likely causal effects of speech. Part IV offers a solution to the current challenge of uninformed risk assessment of speech by distilling the settled social science research on inciting speech to identify ten primary factors, which, if present, provide for joint sufficiency and indicate an elevated risk for imminent lawless action.

There are compelling reasons to revisit incitement law at our current political juncture. Along with conspiracy and attempt, incitement is one of the few offenses in criminal law that is an inchoate crime. Where the probability is high that crimes will ensue, the speech advocating a crime is already a punishable act, and law enforcement agencies need not wait for a deleterious chain of events to unfold before they act.[27] Incitement is a crime of prevention and deterrence, but incitement law requires reform to better fulfill its purpose of protecting citizens, and especially some of our most vulnerable citizens, from imminent harm and injury.

---

[27] As described in *United States v. White,*

> In the case of a criminal solicitation, the speech—asking another to commit a crime—is the punishable act. Solicitation is an inchoate crime; the crime is complete once the words are spoken with the requisite intent, and no further actions from either the solicitor or the solicitee are necessary.

610 F.3d 956, 960 (7th Cir. 2010).

I. "EVERY IDEA IS AN INCITEMENT": REPRESSIVE JURISPRUDENCE PRE-*BRANDENBURG*

A. The Context of Speech

Since the history of incitement law in the early-mid twentieth century is well-rehearsed, this Section focuses particularly on one strand of that narrative; the unsuccessful effort to shift the constitutional emphasis from the content of speech to the context of inciting speech, and in particular to the elements of imminence and probability.

In the midst of Civil Rights marches, anti-Vietnam war protests and widespread social upheaval, the U.S. Supreme Court ruled that political advocacy could no longer be banned on the basis of "mere advocacy" alone.[28] Courts are required to assess the probability that a crime is imminent, an activity that implies an assessment of context.[29] Risk analysis is therefore at the heart of contemporary First Amendment jurisprudence, but it took fifty years for this idea, originally advanced by Justice Holmes in *Schenck*, to be accepted.[30] Even after it was formally endorsed, courts have proved neither adept nor eager to conduct a risk analysis of inciting speech, an inherently difficult and contentious exercise that implies provisional predictions about future behavior. Importantly, courts have not been willing to set out the risk analysis criteria they use to determine whether the crime being advocated is imminent or likely.

For most of our country's history, the U.S. federal government and state governments[31] prohibited speech considered a threat to government authority or public order, including speech that simply criticized individual

---

[28] *Brandenburg*, 395 U.S. at 449.

[29] *Id.* at 447.

[30] In 1969, *Brandenburg* upheld the principle of risk analysis, understood as the evaluation of the likelihood that advocacy of a crime will result in imminent lawless action. *Id.* In his opinion fifty years earlier in *Schenck,* Holmes opined that,

> The question in every case is whether the words used are used in such circumstances and are of such a nature as to create a clear and present danger that they will bring about the substantive evils that Congress has a right to prevent. It is a question of proximity and degree.

Schenck v. United States, 249 U.S. 47, 52 (1919). The call to review the circumstances as well as the proximity and degree of speech represents an instruction to courts to engage in risk analysis, and People v. Rubin, 96 Cal. App. 3d 968, 979 (Cal. Ct. App. 1979), understood "degree" as the risk-assessing "likelihood of producing such action."

[31] *See, e.g.*, PHILLIP BLUMBERG, REPRESSIVE JURISPRUDENCE IN THE EARLY AMERICAN REPUBLIC: THE FIRST AMENDMENT AND THE LEGACY OF ENGLISH LAW 337 (2010) (describing the statutes in Southern states in the early-mid 1800s suppressing abolitionist speech, as well as a House of Representatives rule with the same effect).

office holders.[32] The legal doctrines of criminal libel, blasphemy and out-of-chamber contempt of the courts and legislatures were all inherited from the monarchical system of Great Britain, and epitomized what Phillip Blumberg accurately terms the "repressive jurisprudence" of the early American Republic.[33]

Severe restrictions on political speech were unwound between the mid-1930s and the early 1950s as the U.S. Supreme Court swept away blasphemy,[34] contempt of court for out-of-court speech,[35] and criminal and civil libel involving public persons.[36] Incitement was one of the last elements of the regime of repressive First Amendment jurisprudence inherited from English common law to be liberalized. Each loosening of the restrictions on political speech incrementally elevated the status of context, including the circumstances of speech and the risk that it would imminently result in a crime.

B. Bad Tendencies

Until the end of the First World War, the mere tendency of speech to encourage unlawful acts was sufficient basis to warrant its suppression, no matter how remote the risk that the offense advocated would actually occur.[37] For example, the indictment in *Pierce v. United States* asked not whether the low-circulation pamphlet could undercut a national program of war conscription, but simply whether "the statements contained in the pamphlet were not such as would naturally produce the forbidden consequences."[38]

---

[32] *Id.* at 3-6 (describing the early years of the American republic and the consequences of the 1798 Sedition Act). *See generally* John F. Wirenius, *The Road to Brandenburg: A Look at the Evolving Understanding of the First Amendment*, 43 DRAKE L. REV. 1 (1994) (reciting the history of First Amendment jurisprudence up into the twentieth century).
[33] BLUMBERG, *supra* note 31, at 1.
[34] Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495, 506 (1952).
[35] Nye v. United States, 313 U.S. 33, 42 (1941); Bridges v. California, 314 U.S. 252, 271 (1941).
[36] *See generally* New York Times Co. v. Sullivan, 376 U.S. 254 (1964) (applying a new actual malice standard to a civil libel case involving Martin Luther King, Jr.); Garrison v. Louisiana, 379 U.S. 64 (1964) (analyzing the applicability of a criminal libel statute in the context of allegedly disparaging remarks made by a prosecutor about local judges).
[37] *See generally* KENT GREENAWALT, SPEECH, CRIME AND THE USES OF LANGUAGE 221-22 (1989) (laying out the bad tendency doctrine); Rodney A. Smolla, *Should the* Brandenburg v. Ohio *Incitement Test Apply in Media Violence Cases?* 27 N. KY. L. REV. 1 (2000) (applying the *Brandenberg* defense to a series of media violence hypotheticals); Wirenius, *supra* note 32, at 17 n.103 (summarizing early twentieth-century developments in the caselaw); Zechariah Chafee, Jr., *Book Review*, 62 HARV. L. REV. 891, 901 (1949) (reviewing ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT (1948)).
[38] Pierce v. United States, 252 U.S. 239, 243-44 (1920).

In 1919, antiwar protestors were jailed under the "bad tendency" standard in *Debs v. United States* and *Frohwerk v. United States*, with *Debs* stating that the jury "could not find the defendant guilty for advocacy of any of his opinions unless the words used had as their natural tendency and reasonably probable effect to obstruct the recruiting service . . . ."[39] Even a tendency to produce an outcome that was not in itself an offense could be sufficient to warrant criminal censure.[40]

The "bad tendency" standard was formally repudiated in First Amendment jurisprudence in 1919,[41] but its ghost still haunts modern municipal and state ordinances.[42] Here, it is worth observing that First Amendment jurisprudence remains perpetually untidy, even as standards are gradually liberalized.[43] Standards emphatically renounced in one constitutional moment can resurface decades later at the municipal and state level.

With the entry of the United States into the First World War and the passage of the Espionage Act of 1917,[44] a more flexible standard was needed to protect political dissent, while acknowledging that the government might legitimately restrict speech that undermined the war effort. Any dissenting speech could conceivably stir up a *tendency* to law-breaking, and what was additionally required was an assessment of risk that weighed the gravity and proximity of the crime being urged.

C. Repressing Political Dissent: Clear and Present Danger

Schenck v. United States and *Abrams v. United States*, both decided in 1919, announced a new test for criminal advocacy and are widely considered the *fons et origo* of modern jurisprudence on incitement and

---

[39] Debs v. United States, 249 U.S. 211, 216 (1919); *see also* Frohwerk v. United States, 249 U.S. 204, 209 (1919) (observing that "it is impossible to say that it might not have been found that the circulation of the paper was in quarters where a little breath would be enough to kindle a flame").

[40] Fox v. Washington, 236 U.S. 273, 275 (1915). *Fox* upheld a Washington statute preventing the printing and circulation of written matter "which shall tend to encourage or advocate disrespect for law or for any court or courts of justice," even though "disrespect for the law" was not a crime at the time. *Id.*

[41] *See, e.g.*, Ashcroft v. Free Speech Coal., 535 U.S. 234, 253 (2002) ("[T]he mere tendency of speech to encourage unlawful acts is not a sufficient reason for banning it.").

[42] *See* Bible Believers v. Wayne Cty., Mich., 805 F.3d 228, 237 (6th Cir. 2015) (describing criminal accountability under state law and local ordinances "for conduct which has the tendency to incite riotous behavior or otherwise disturb the peace").

[43] Gey, *supra* note 5, at 971 (discussing the "fracturing" of First Amendment law).

[44] *See* Dennis v. United States, 341 U.S. 494, 533-34 (1951) (providing a review of Espionage Act cases).

the First Amendment.[45] In *Schenck*, the defendants were convicted of violations of the Espionage Act for printing and circulating an antiwar leaflet to men who had been called up for military service.[46] Justice Oliver Wendell Holmes replaced the overly capacious "bad tendency" standard with the ostensibly more stringent "clear and present danger test,"[47] which he articulated as:

> The question in every case is whether the words used are used in such circumstances and are of such a nature as to create a clear and present danger that they will bring about the substantive evils that Congress has a right to prevent. It is a question of proximity and degree.[48]

Despite the fact that Holmes took great pains in *Schenck* to add the qualifiers of "proximity" and "degree," which have been parsed by courts as tests of imminence and likelihood,[49] subsequent First Amendment cases emphasized advocacy and content, rather than context and immediate circumstances. There was no reference to gravity, proximity, or likelihood in *Frohwerk*, a fourth Espionage Act case from 1919 which instead harked back to the "bad tendency" standard.[50]

Any consideration of "proximity" (imminence) or "degree" (likelihood) was also spurned in *Gitlow* (1925).[51] In the place of a careful risk analysis, *Gitlow* conjured up metaphors of fire and conflagration:

> And the immediate danger is none the less real and substantial, because the effect of a given utterance cannot be accurately foreseen. The State cannot reasonably be required to measure

---

[45] Abrams v. United States, 250 U.S. 616, 627 (1919); Schenck v. United States, 249 U.S. 47 (1919); *see, e.g.*, DANIEL FARBER, THE FIRST AMENDMENT 67-68 (4th ed. 2014) (asserting that the *Schenck* and *Abrams* decisions in 1919 represented "one of the great turning points in the development of First Amendment Doctrine"); Healy, *supra* note 26, at 711 (observing how Brandenburg's application of strict scrutiny of advocacy has its origins in *Schenck*); Tsesis, *Inflammatory Speech*, *supra* note 5, at 1156 (describing how *Schenck* "established the groundwork for contemporary doctrine").

[46] *Schenck*, 249 U.S. at 48-49.

[47] *See id.* at 52; *see also Abrams*, 250 U.S. at 627 (using the expression "clear and imminent danger").

[48] *Schenck*, 249 U.S. at 52.

[49] "Proximity" could be straightforwardly read as "imminence" if we consider Holmes's dissenting comment in *Abrams* indicating that the public expression of opinions remains protected speech unless the opinions "so imminently threaten immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country." *Abrams*, 250 U.S. at 630. Subsequent courts found that "degree" denoted the "likelihood of producing such action." *See* People v. Rubin, 96 Cal. App. 3d 968, 979 (Cal. Ct. App. 1979).

[50] Frohwerk v. United States, 249 U.S. 204, 209 (1919).

[51] Gitlow v. New York*,* 268 U.S. 652, 669 (1925).

> the danger from every such utterance in the nice balance of a jeweler's scale. A single revolutionary spark may kindle a fire that, smouldering for a time, may burst into a sweeping and destructive conflagration.[52]

The Supreme Court Justices in *Gitlow* stated their disdain for predictive forecasting: "the question whether any specific utterance coming within the prohibited class is likely, in and of itself, to bring about the substantive evil, is not open to consideration."[53] The modern reticence to engage in risk analysis, then, has identifiable historical antecedents.

In rejecting predictive forecasting, incitement law slipped backwards into "bad tendency" territory once again in which all seditious advocacy was treated as an existential threat to stability and order. Justice Holmes, joined by Justice Brandeis, dissented from the majority's (mis)interpretation of the "clear and present danger" standard that Holmes had created six years previously, remarking that there was no real danger that the "Left Wing Manifesto" published by the defendant would culminate in the overthrow of the government.[54] Holmes and Brandeis memorably protested the overly inclusive application of incitement law to a low-circulation and insignificant socialist pamphlet, on the grounds that under this standard, "[e]very idea is an incitement."[55]

After the Second World War, Red Scare legislation such as the Smith Act strictly curtailed political speech.[56] In 1951, *Dennis v. United States*, a watershed case in which the Supreme Court reviewed the Smith Act, held that the overall threat of communist revolution was a "sufficient evil"[57] that unshackled the courts from restraining criteria of probability or immediacy: "The damage which such attempts create both physically and politically to a nation makes it impossible to measure the validity in terms of the probability of success, or the immediacy of a successful attempt."[58] By *Dennis*, the "clear and present danger" standard had become so enfeebled that it provided little protection from Red Scare mania,[59] and permitted convictions for merely reading and discussing books by Stalin, Marx and Engels, and Lenin.[60]

---

[52] *Id.*
[53] *Id.* at 670.
[54] *Id.* at 655.
[55] *Id*. at 673 (Holmes, J., dissenting).
[56] Alien Registration Act of 1940, Pub. L. No. 76-670, 54 Stat. 670; *see, e.g.*, 18 U.S.C. § 2385 (2018) (criminalizing activity which could be construed as supporting "overthrowing or destroying the government of the United States" or any state governments).
[57] Dennis v. United States, 341 U.S. 494, 509 (1951).
[58] *Id*.
[59] *See generally* ELLEN SCHRECKER, MANY ARE THE CRIMES: MCCARTHYISM IN AMERICA (1999) (discussing the Red Scare and the impact of McCarthyism).
[60] *Dennis*, 341 U.S. at 582 (Douglas, J., dissenting).

Dennis was riven with sharp disagreement, and Justice Douglas wrote an angrily worded dissent[61] which held that there is a "wide difference"[62] between advocacy and incitement, and observed that communists in America are "miserable merchants of unwanted ideas; their wares remain unsold."[63] Imminence[64] and probability—in short, risk analysis—constituted Douglas's answer to the majority's unwarranted suppression of a marginal political doctrine that patently lacked the resources to mount an existential challenge. Probability is at the heart of the question of whether advocacy of a crime constitutes incitement; namely the probability that an incitement to lawbreaking sufficiently enhances the chances that a violation will occur.[65] But what is sufficient?

D. Free Speech in the Early Civil Rights Era

By 1957, McCarthyism was a spent political force and Justice Douglas' plea for a more contextual approach to incitement began to gain traction. In that year, *Yates v. United States* reversed the convictions of communist party officials under the Smith Act and affirmed that speakers could not be punished merely for advocating seditious political beliefs.[66] *Yates* drew a sharp line between abstract advocacy that was protected speech and incitement to illegal action that was unprotected.[67]

Noto v. United States, decided during the tumult of the Civil Rights era, demanded narrower criteria for incitement, reversing the conviction of a communist activist convicted under the Smith Act for advocating the overthrow of the government.[68] *Noto* raised the actus reus threshold of incitement to include imminence. Advocacy of crimes, such as sabotage to achieve a revolutionary moment, must represent "present advocacy,"[69] not just urge revolution in a remote future. Advocacy of rebellion was insufficient, and there had to be some chance that the incitement would achieve the undesirable result.[70] *Noto* was the stepping stone to *Brandenburg* and signified the turning point when requirements of proximity and likelihood became part of the consensus opinion.

---

[61] *Id.* at 579.
[62] *Id.* at 586.
[63] *Id.* at 589.
[64] *Id.* at 585.
[65] *Id.*
[66] Yates v. United States, 354 U.S. 298, 318-27 (1957), *overruled by* Burks v. United States, 437 U.S. 1 (1978).
[67] *Id.* at 318.
[68] Noto v. United States, 367 U.S. 290 (1961).
[69] *Id.* at 298.
[70] *Id* at 298-99.

To summarize: before *Brandenburg*, government regulation of speech was frequently oppressive and served to preserve state interests. During two World Wars and one Cold War, legislators and judges regulated dissent that opposed conscription for war or advocated revolutionary socialism. That the Supreme Court lifted historic restrictions on speech in the cauldron of the civil rights and anti-Vietnam war era was not the result of a road-to-Damascus conversion of justices to the cause of dissent, but because the state was secure enough to tolerate seditious speech, given the limited ideological appeal and feeble organizational capacity of American socialism.[71]

II. *BRANDENBURG V. OHIO*: CONTEXTUALIZING SPEECH AND ASSESSING RISK

A. The Contours of Brandenburg

Brandenburg v. Ohio, decided in 1969, announced a new standard that required courts to evaluate the risks of political speech and embraced the contextual elements presaged in *Schenck*, combining direct advocacy of a crime with an assessment that the offense will be committed imminently. The *Brandenburg* test has been settled law for five decades.[72]

That said, *Brandenburg* is an odd decision. Barely four pages long, it is terse and unsigned, having been issued per curiam. It defines incitement with a novel three-part test that adds imminence and likelihood to a prior advocacy standard but provides no guidance on the three elements. It is contradictory in places, question-begging in others and gives the overall impression of being undercooked. This resulted from the fact that it was drafted by two authors, each with distinct objectives. Justice Abe Fortas commenced the initial draft but then resigned from the Court amidst an ethics scandal, and the judgement was revised and completed by Justice Brennan.[73]

---

[71] "[I]t is impossible for me to say that the Communists in this country are so potent or so strategically deployed that they must be suppressed for their speech." Dennis v. United States, 341 U.S. 494, 589 (1951). We might also recall Holmes's dissent in *Abrams* when he refers to the defendants as "poor and puny anonymities." Abrams v. United States, 250 U.S. 616, 629 (1919).

[72] *See* Gey, *supra* note 5, at 977 (declaring that the *Brandenburg* standard "is now one of the most well-established aspects of modern constitutional doctrine"); Tsesis, *Inflammatory Speech*, *supra* note 6, at 1159 (stating that *Brandenburg* represents "The Modern Test"). The *Brandenburg* test has been augmented by the principle that the First Amendment does not protect political speech or expressive conduct that materially supports foreign terrorist organizations. Holder v. Humanitarian Law Project, 561 U.S. 1, 4 (2010).

[73] Gey, *supra* note 5, at 977; *see also* Bernard Schwartz, *Justice Brennan and the Brandenburg Decision—A Lawgiver in Action*, 79 JUDICATURE 24, 27-28 (1995) (providing a history and analysis of the Court's decision in *Brandenberg*).

Scholars are frequently critical of *Brandenburg*, and Gey calls it "murky and inelegant."[74] Their criticisms are comprehensible but perhaps miss the point. *Brandenburg* only makes sense as a stealth precedent. Like Justice Harlan in *Noto*, Justice Brennan rolled out a new test covertly, all the while professing that the new schema was simply an extension of *Dennis* and *Yates*.[75] *Brandenburg* represents, in fact, a clear break with First Amendment law up until that point.

B. The Facts and the Test

Defendant Clarence Brandenburg was a Ku Klux Klan leader who invited a television crew to a farm outside Cincinnati where he addressed a small group of Klan members in 1964,[76] the year in which the Civil Rights Act was passed.[77] After the footage aired on Cincinnati and national television stations, he was prosecuted under Ohio's antilabor syndicalism statute which made it illegal to advocate an offense or violence.[78]

Ohio prosecutors relied on two films taken by the television crew as evidence.[79] In the first, twelve hooded figures carrying firearms and ammunition stood by a burning cross.[80] The second film featured six armed figures wearing Klan regalia.[81] According to the lower court, the defendant's most inciting utterance occurred in the first film, "We're not a revengent [sic] organization, but if our President, our Congress, our Supreme Court, continues to suppress the white, Caucasian race, it's possible that there might have to be some revengeance [sic] taken."[82] Clarence Brandenburg added that the Klan was planning a march six days later to Congress, and, after that, to Florida and Mississippi.[83] In the second film, he made a similar speech that included extreme animus and threats against Jews and African Americans.[84]

Brandenburg's conviction in a jury trial resulted in a fine of $1000 and a sentence of one to ten years in prison,[85] and the Ohio Supreme Court

---

[74] Gey, *supra* note 5, at 977.
[75] Brandenburg v. Ohio, 395 U.S. 444, 448, 453, 457 (1969) (per curiam) (citing Yates v. United States, 354 U.S. 298 (1957)); *see* Schwartz, *supra* note 73 (documenting Brennan's authoring of *Brandenburg*).
[76] *Brandenburg*, 395 U.S. at 444-45.
[77] The Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241.
[78] *Brandenburg*, 395 U.S. at 444-45.
[79] *Id.* at 445.
[80] *Id.* at 445-46.
[81] *Id.* at 447.
[82] *Id.* at 446.
[83] *Id.*
[84] *Id.* at 447.
[85] *Id.* at 445.

affirmed the trial verdict.[86] The U.S. Supreme Court then reversed the conviction on the grounds that the defendant's statements represented "mere abstract teaching of the moral propriety" of racism and possible future action against the government and others.[87] *Brandenburg* thus proclaimed its new test:

> the principle that the constitutional guarantees of free speech and free press do not permit a State to forbid or proscribe advocacy of the use of force or of law violation except where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.[88]

On this basis, and citing *Noto*, *Brandenburg* held that the Ohio syndicalism statute violated the First Amendment because it failed to distinguish between "mere advocacy" and "incitement to imminent lawless action."[89]

Brandenburg's innovation was to add two contextual conditions—imminence and likelihood—to the long-established element of criminal advocacy. *Brandenburg*, however, provides no guidance on any of the three elements of the test. Advocacy was already a fundamental (if not the sole) element of incitement law before *Brandenburg*, and therefore it had already been the subject of extensive judicial interpretation as we just saw in the cases just reviewed, including *Schenck*, *Abrams*, *Gitlow*, *Dennis*, and *Noto*. Imminence was indeterminate at the time and has received only a modicum of elucidation since *Brandenburg*. Likelihood has been barely defined at all. Healy observes that such conceptual gaps "have been largely glossed over by courts and scholars."[90]

Furthermore, *Brandenburg* did not apply its own criteria of imminence and likelihood to the fact pattern in the case. It never asked whether the Ku Klux Klan posed a danger at the time of the cross-burning event, or whether the risk of violence had been elevated sufficiently by televising a Klan leader's speech to a wide audience. Clarence Brandenburg's televised speech included a call to march on Washington, D.C. and the statements, "[t]his is what we are going to do to the n-----rs," and, "[b]ury the n----rs," which were uttered next to a burning cross surrounded by hooded and armed Klansmen, at a time (1964) when violent clashes were a frequent

---

[86] *Id.*
[87] *Id.* at 448-49 (citing Noto v. United States, 367 U.S. 290, 297-98 (1961)).
[88] *Id.* at 447.
[89] *Id*. at 448-49. Later cases have interpreted the standard in *Brandenburg* as requiring "violence or physical disorder in the nature of a riot." White v. Lee, 227 F.3d 1214, 1228 (9th Cir. 2000) (citing *Brandenburg*, 395 U.S. at 448, to define "imminent lawless action").
[90] Healy, *supra* note 26, at 660; *see also* Rohr, *supra* note 5, at 7, 14 (discussing the peculiarities and ambiguities of the *Brandenburg* framework).

occurrence on the streets of the United States.[91] There is reason to believe that if the Court had applied the contextual approach it was advocating exactingly, then Brandenburg's utterances may not have qualified as protected speech.

Our thesis here is that the lacunae in *Brandenburg*, particularly with respect to imminence and likelihood, hinder its application to political speech that incites violence,[92] and effectively stymie any preventative function incitement law might have. The next Sections explore what subsequent courts have understood by advocacy, imminence, and likelihood, with special attention to the least developed of the three; likelihood, or probability. Since 1969, *Brandenburg* has only been cited thrice by the U.S. Supreme Court: in *Hess v. Indiana* (1973), *NAACP v. Claiborne Hardware* (1982), and *Texas v. Johnson* (1989). The Supreme Court has not applied the decision in nearly thirty years, and therefore any comprehensive discussion of incitement doctrine must incorporate rulings by lower courts.

C. The Elements of Incitement: Direct Advocacy

In the United States, it is axiomatic that robust political speech, including that which is offensive, caustic, and even coercive is protected by the First Amendment.[93] As Justice Stevens wrote in *Claiborne*,

> [s]trong and effective extemporaneous rhetoric cannot be nicely channeled in purely dulcet phrases. An advocate must be free to stimulate his audience with spontaneous and emotional appeals for unity and action in a common cause. When such appeals do not incite lawless action, they must be regarded as protected speech.[94]

The key phrase here is "incite lawless action," and according to the language in *Brandenburg*, political speech is protected unless it constitutes "advocacy" that is "directed to inciting or producing imminent lawless action."[95] There is a great deal of guidance about what kind of speech constitutes "advocacy," and it is

---

[91] *Brandenburg*, 395 U.S. at 446 n.1.

[92] The Fourth Circuit has limited the application of *Brandenburg* to political speech. Rice v. Paladin Enterps., Inc., 128 F.3d 233, 262 (4th Cir. 1997). Not all courts, however, have followed this interpretation.

[93] *See, e.g.*, NAACP v. Claiborne Hardware Co., 458 U.S. 886, 910 (1982) (protecting coercive speech); Watts v. United States, 394 U.S. 705, 707-08 (1969) (protecting robust criticism of, and non-true threats towards, the President); New York Times Co. v. Sullivan, 376 U.S. 254, 270 (1964) (protecting caustic speech).

[94] *Claiborne*, 458 U.S. at 928.

[95] *Brandenburg*, 395 U.S. at 447.

widely recognized that the advocacy prong is the most salient of the three prongs of the *Brandenburg* test.[96]

As with most crimes, incitement contains a mens rea requirement in which the speaker must intend to advocate a criminal offense.[97] A mere tendency for the speech to prompt lawless action is not sufficient to remove constitutional protection.[98] Constitutional scholars generally regard the reference to advocacy that is "directed to inciting or producing" as enacting an intent requirement,[99] and Chemerinsky maintains that "Brandenburg contains an intent requirement: the speech must be directed to causing the harm."[100] Subsequent decisions have made this explicit. For instance, *State v. Beasley* opines, "[w]hen considering the offense of inciting to riot, . . . the language used must clearly intend to incite a breach of the peace."[101]

What is repeated again and again in the post-*Brandenburg* era is that criminal advocacy in the abstract is not enough in itself to trigger a sanction.[102] Contextual evaluation has become more relevant in recent years, as courts are forced to adapt First Amendment law to the fast-moving advertising, internet, and social media environment.[103]

A recent case illustrates the principle that advocacy alone is insufficient. In *American Freedom Defense Initiative v. Metropolitan Transportation Authority*, the court found that bus advertisements that portrayed a man whose head and face were mostly covered by a head scarf and a recited quote from a Hamas television channel, "Killing Jews is Worship that draws us close to Allah" and stated underneath the quote, "That's His Jihad. What's yours?" were

---

[96] *See* Smolla, *supra* note 37, at 15 (stating that advocacy/intent is the most significant of the three prongs of *Brandenburg*).

[97] MODEL PENAL CODE §2.02, §5.02 (AM. LAW INST. 1962); *see* Kent Greenawalt, *Speech and Crime*, 5 AM. BAR FOUND. RES. J. 645, 652 (1980) (elaborating on the requirement that the speaker have the purpose of inciting a crime).

[98] Ashcroft v. Free Speech Coalition, 535 U.S. 234, 253 (2002).

[99] Gey, *supra* note 5, at 1021; Healy, *supra* note 26, at 698; Smolla, *supra* note 37, at 10.

[100] Erwin Chemerinsky, *The Rhetoric of Constitutional Law,* 100 MICH. L. REV. 2008, 2017 (2002).

[101] State v. Beasley, 317 So.2d 750, 753 (Fla. 1975); *see also* Bible Believers v. Wayne Cty., Mich.*,* 805 F.3d 228, 245-46 (6th Cir. 2015) (discussing intention and advocacy).

[102] *See* Connick v. Myers, 461 U.S. 138, 147-48 (1983) (stating that the determination of whether an employee's speech is a matter of public concern depends on the "content, form, and context" of the statement); NAACP v. Claiborne Hardware Co., 458 U.S. 886, 927 (1982) (reiterating that "mere advocacy of force or violence does not remove speech from the protection of the First Amendment").

[103] *See generally* Tsesis, *Social Media Accountability for Terrorist Propaganda*, *supra* note 16 (discussing the role of context in determining the liability of intermediaries for social media posts that incite terrorism).

protected by the First Amendment.[104] The plaintiffs argued that the advertisements advocated violence, but the court countered that the target of the purported advocacy was not clear, and that the ads were not advocating any action.[105] The constitutional protection afforded to abstract advocacy means that speech that fails to advocate any action at all cannot constitute incitement, no matter how repugnant or reprehensible the views advocated.[106]

Next, it is settled First Amendment law that the suppression of speech cannot be content—or viewpoint—based, and any such restrictions are subject to strict scrutiny in which the government bears the burden of proof that the proposed regulation is necessary. [107] The only allowable constraints on speech are "time, manner[,] or place" restrictions that are content-neutral.[108] The paradigmatic case instantiating this principle is *R.A.V. v. City of St. Paul, Minnesota*. In *R.A.V.*, Justice Scalia, writing for the majority, reversed and remanded the conviction of a defendant who had been convicted for burning a cross on the lawn of an African American family under a city ordinance that prohibited bias-motivated disorderly conduct.[109] St. Paul's statute was substantially overbroad[110] and violated the First Amendment by censoring speech on the basis of content (i.e., bias-motivated content) because it only placed prohibitions on speech that discriminated on the basis of "race, color, creed, religion[,] or gender."[111] While obscenity, defamation and fighting words can be regulated because of their content, the St. Paul ordinance went "beyond mere content, to actual viewpoint discrimination."[112]

A 2003 cross-burning case that represents an exception to the anticontent discrimination rule is *Virginia v Black*. This case invoked "true threat," rather than

---

[104] Am. Freedom Def. Initiative v. Metro. Transp. Auth., 70 F. Supp. 3d 572, 582-83 (S.D.N.Y. 2015), *vacated*, 109 F.Supp.3d 626 (S.D.N.Y. 2015), *aff'd*, 815 F.3d 105 (2d Cir. 2016) (vacating the decision below because the transit authority changed its polices to prohibit all political advertising).

[105] *Id.* at 582.

[106] *See* Hess v. Indiana, 414 U.S. 105, 109 (1973) (per curiam) (noting the requirement that the inciter advocate lawless action that is imminent); *see also* Bible Believers v. Wayne Cty., Mich., 805 F.3d 228, 245 (6th Cir. 2015) (noting that the advocacy of the Bible Believers group did not call for any action at all).

[107] *See Bible Believers*, 805 F.3d at 247 (reiterating the importance of content-neutral speech restrictions); *see also* Harper v. Poway Unified Sch. Dist., 445 F.3d 1166, 1184-85 (9th Cir. 2006) (discussing how public schools are excepted from the usual injunction on viewpoint discrimination and may prohibit student speech that disrupts school activities or violates the rights of other students).

[108] *Bible Believers*, 805 F.3d at 247.

[109] R.A.V. v. City of St. Paul, Minn., 505 U.S. 377, 379, 396 (1992).

[110] *Id*. at 391.

[111] *Id*.

[112] *Id*.

incitement.[113] In *Black*, three individuals (including a leader of the Ku Klux Klan) were convicted separately under a Virginia statute banning cross-burning and their convictions were reversed by the Virginian Supreme Court, citing *R.A.V.* The Supreme Court, however, noted that the Virginia statute was analytically distinguishable from the statute at issue in *R.A.V.* since it was content and viewpoint neutral, simply making it a felony "for any person . . . with the intent of intimidating any person or group . . . to burn . . . a cross on the property of another, a highway or other public place,"[114] regardless of viewpoint on race, religion, gender, political affiliation, or any other axis of identity.[115]

Justice O'Connor, writing for the majority, observed that cross-burning has been a symbol of the Ku Klux Klan's reign of terror in the South since 1866,[116] and concluded that a state may ban cross-burning carried out with intent to intimidate, consistent with the First Amendment.[117] Importantly, the Court confirmed that the First Amendment does not extend to true threats.[118] Intriguingly for our discussion, *Black* permits the suppression of true threats against collective groups, be they families or racial or religious groups. It is possible that under *Black*'s definition of group threat, Clarence Brandenburg's declarations of his violent intentions towards African Americans, uttered using an ethnic slur and alongside armed Klansmen and in front of a burning cross, may have constituted a true threat.[119] The ruling in *Black* could constitute a charter for how state legislatures may write statutes banning incitement and true threats that call for violent action.

Of the three elements of the *Brandenburg* test, advocacy has received the most interpretation and guidance in the subsequent constitutional jurisprudence. The speaker must directly advocate lawless action and not simply express a provocative or unpopular view that may result in public disorder.[120] This is not an objective test because the speaker must intend that others commit an offense, and their intentionality need not be expressed via an instruction for the specific offense committed but may be implicitly

---

[113] Virginia v. Black, 538 U.S. 343, 359-66 (2003). *Black* relies primarily on the jurisprudence related to threats but cites *Brandenburg* twice: for guidance on incitement and whether cross burning is protected expression. *Id*. at 359, 366.

[114] *Id*. at 348.

[115] The Court held as unconstitutional that section of the Virginia statute that claimed that cross burning constitutes, in itself, prima facie evidence of intent to intimidate. *Id*. at 345. Virginia's statute read, "[a]ny such burning of a cross shall be prima facie evidence of an intent to intimidate a person or group of persons," *id*. at 348, on the grounds that this unnecessarily suppresses ideas and abjures consideration of any contextual factors. *Id.* at 345.

[116] *Id.* at 352-57.

[117] *Id.* at 362.

[118] *See id*. at 360 (explaining that a prohibition on true threats protects individuals from the fear of violence and the disturbances that fear creates).

[119] Brandenburg v. Ohio, 395 U.S. 444, 446 n.1 (1969) (per curiam).

[120] *Id.* at 447-48.

understood.[121] As established in *R.A.V.*, restrictions on speech cannot be content- or viewpoint-based.[122]

D. The Elements of Incitement: Imminence

Imminence is the second element of the *Brandenburg* test, and, combined with the likelihood prong, it distinguishes modern incitement from the repressive jurisprudence of the "bad tendency" and "clear and present danger" tests. Along with conspiracy and attempt, incitement is an inchoate crime.[123] The standard justification for inchoate crimes is that they are crimes of prevention,[124] designed for circumstances in which the offense is impending, and there is not sufficient time to expose "falsehoods and fallacies"[125] in the marketplace of ideas,[126] or to allow deliberation on the part of the listener, or give police enough time to take appropriate measures.[127]

By requiring imminence, *Brandenburg* built on *Schenck*'s language of "proximity and degree," but *Brandenburg* did not explain precisely how imminent the lawless action must be to warrant suppression of the speech. Proximity remained undefined until another per curiam Court decision, *Hess v. Indiana*.[128] Gregory Hess was convicted for disorderly conduct during an anti-Vietnam war demonstration on the campus of Indiana University. After police forced a group of about 100 student demonstrators blocking traffic to move to the sidewalk, Hess said in a normal voice and to no one in particular, "We'll take the fucking street later [or again]."[129] The Supreme Court reversed the conviction, stating that Hess's words could not be considered

---

[121] Bible Believers v. Wayne Cty., Mich., 805 F.3d 228, 245, 246 n.11 (6th Cir. 2015).
[122] R.A.V. v. City of St. Paul, Minn., 505 U.S. 377, 391 (1992).
[123] JOSHUA DRESSLER & STEPHEN P. GARVEY, CASES AND MATERIALS IN CRIMINAL LAW 772 (6th ed. 2012); Larry Alexander & Kimberly D. Kessler, *Mens Rea and Inchoate Crimes*, 87 J. CRIM. L. & CRIMINOLOGY 1138, 1139 (1997).
[124] For more context regarding inchoate crimes, see generally ANDREW ASHWORTH & LUCIA ZEDNER, PREVENTATIVE JUSTICE (2014) and Michael T. Cahill, *Defining Inchoate Crime: An Incomplete Attempt*, 9 OHIO ST. J. CRIM. L. 751 (2012).
[125] As Justice Brandeis eloquently stated, "If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence." Whitney v. California, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring).
[126] *See* Abrams v. United States, 250 U.S. 616, 630 (Holmes, J. dissenting) ("[T]he best test of truth is the power of the thought to get itself accepted in the competition of the market . . . .").
[127] *See generally* Healy, *supra* note 26, at 680 (examining imminence and police intervention in the *Al-Timimi* case).
[128] Hess v. Indiana, 414 U.S. 105 (1973) (per curiam); *see* Smolla, *supra* note 37, at 22-30 (analyzing imminence in *Hess* and the development of imminence after *Brandenburg*).
[129] *Hess*, 414 U.S. at 107.

obscenity,[130] fighting words, or insult since they were not directed at any particular person.[131]

According to the Court, Hess's utterance was ambivalent and did not call for any immediate action.[132] Lawless action was only encouraged "at some indefinite future time,"[133] which was insufficient to justify a conviction for incitement, because, citing *Brandenburg*, the lawless action must be imminent. *Hess* did not demarcate any clear temporal boundaries of imminence, but analysts have inferred that "imminent" meant immediately or within a few hours, not a day or several days or weeks later.[134] Greenawalt sees the "imminent" in *Hess* as "momentarily" or in the "very near future."[135]

Imminence was further defined by the U.S. Supreme Court in *NAACP. v. Claiborne Hardware Co*, a civil case brought by white store owners against African American leaders who organized a boycott of white-owned stores in Claiborne County, Mississippi in 1966 and 1967.[136] One of the defendants, NAACP field secretary Charles Evers, had threatened African Americans in the community with physical harm if they patronized white-owned stores, saying "If we catch any of you going in any of them racist stores, we're gonna break your damn neck."[137]

Reversing the convictions, the Court made clear "that mere *advocacy* of the use of force or violence does not remove speech from the protection of the First Amendment" unless that violence is imminent.[138] Acts of violence and intimidation were prevalent after Evers's 1966 speech, including shots fired into a home, a brick thrown through a windshield and beatings of boycott-breakers, but these happened some "weeks or months" later, and therefore could not be considered proximate to the speech act.[139] As in *Hess*, the Supreme Court refrained from specifying exactly how soon after Evers' speech any violence would have had to have occurred for the defendant to be held liable.

The temporal duration of imminence was extended significantly in 1979 by the California Court of Appeal in *People v. Rubin*,[140] in which the

---

[130] By 1973, the word "fuck" was no longer considered obscene. *Id.* at 107 (citing Cohen v. California*,* 403 U.S. 15 (1971), in which the defendant wore a jacket bearing the antiwar message, "Fuck the Draft").
[131] *Id.*at 107-08.
[132] *Id.* at 108-09.
[133] *Id.*
[134] *See* Healy, *supra* note 26, at 715-17 (explaining that as time passes, predicting the likelihood of a crime becomes more difficult as conditions and variables may change).
[135] GREENAWALT, *supra* note 37, at 267.
[136] NAACP v. Claiborne Hardware Co., 458 U.S. 886, 889-91 (1982).
[137] *Id.* at 902.
[138] *Id.* at 927 (emphasis in original).
[139] *Id.* at 928.
[140] People v. Rubin, 96 Cal. App. 3d 968, 979 (Cal. Ct. App. 1979).

defendant, the national director of the Jewish Defense League, offered a reward for the murder or serious injury of any Nazi participant in a march planned in Skokie, Illinois, some five weeks later. At a press conference, Irving Rubin held up five hundred-dollar bills and said he would give them to any person who "kills, maims, or seriously injures a member of the American Nazi Party. . . . And if they bring us the ears, we'll make it a thousand dollars. The fact of the matter is, that we're deadly serious. This is not said in jest, we are deadly serious."[141]

The California appeals court determined that five weeks was sufficiently imminent to trigger the application of *Brandenburg*: "We think solicitation of murder . . . even though five weeks away, can qualify as incitement to imminent lawless action."[142] The court refrained from any categorical statement that imminence hinges on gravity, however. Instead, using metaphorical and elliptical language, the court tentatively suggested that imminence is a relative concept and suggested that it varied according to the "nature" (i.e., gravity) of the crime being incited: "But time is a relative dimension and imminence a relative term, and the imminence of an event is related to its nature. A total eclipse of the sun next year is said to be imminent. An April shower thirty minutes away is not."[143] Such lyrical and indirect speech is common in First Amendment cases and is utterly unconstructive.

In the current digital era, imminence is coming under new scrutiny. In *United States v. Fullmer*, the Third Circuit found that the act of scheduling a specific time for an unlawful act committed online, even if that time is not immediate, was relevant in determining imminence. In *Fullmer*, an animal rights group's website advocated electronic sit-ins, where many people access a website at the same time, causing it to crash. The advocacy group's website provided a schedule for sit-ins and updates on ongoing sit-ins. Because the group provided information about sit-ins at specified times, the court found that the information about sit-ins was intended to incite an imminent unlawful act.[144]

E. The Elements of Incitement: Likelihood

In 1919, Justice Holmes's formulation of the "clear and present danger" test was accompanied by an insistence that courts should assess the probability that a crime could ensue from dissenting speech. *Debs* demanded an assessment of the "reasonably probable effect" of speech,[145] and Justice Holmes dissented

---

[141] *Id.* at 982.
[142] *Id.* at 979.
[143] *Id.* at 978.
[144] United States v. Fullmer, 584 F.3d 132, 155 (3d Cir. 2009).
[145] Debs v. United States., 249 U.S. 211, 216 (1919).

in *Abrams* because he felt that the antiwar speech in question had little chance of having a tangible effect on the war effort since it represented little more than "the surreptitious publishing of a silly leaflet by an unknown man."[146]

Courts applying the "clear and present danger" test soon jettisoned the element of probability, and the majorities in *Gitlow* and *Dennis* took great pains to discount its relevance.[147] Risk assessment reentered incitement doctrine via the constitutional law of fighting words, and specifically *Cantwell*. This case interpreted the "clear and present danger" test stringently by refusing to suppress insults, profanities or verbal abuse, and banning only those words "likely to provoke violence and disturbance of good order, even though no such eventuality be intended."[148] What is intriguing about *Cantwell* is that it elevates probability over intentionality, an inversion of their conventional framing in First Amendment law.

As noted, *Brandenburg* does not tell us how much a speech act must elevate the risk of an offense to justify its banning, and neither do *Hess* and *Claiborne*, subsequent cases that addressed the imminence prong. A combing of lower court cases turns up rather meager findings. In *United States v. White*, a 2010 case of a white-supremacist website that solicited the murder of the foreperson of the jury that had convicted a white nationalist leader, the Seventh Circuit considered the defendant's web posting in the context of the extremist community of Neo-Nazis frequenting the website (Overthrow.com), where enemies were frequently identified for assassination.[149] *White* quotes the defendant's own admission that his online solicitation to murder had "so great a potential for action."[150] Because the call reached a multitude of white supremacist readers, the Seventh Circuit held that "*someone* could kill or harm Juror A."[151]

Rubin promised to define the likelihood prong of *Brandenburg* by titling a whole section of the decision, "Degree: Likelihood of Producing Action," but the section simply restates the obvious in an anecdotal manner, observing that emotional appeals to political violence are given respectability when transmitted by reputable news media and that words uttered with a serious intention are more likely to produce violence than those delivered in jest.[152] *Rubin* offers no generalizable statement on probability, including the criteria on which to base any risk analysis or how much the speech must elevate the chances of violence to lose constitutional protection.

---

[146] Abrams v. United States, 250 U.S. 616, 628 (1919) (Holmes, J., dissenting).
[147] Dennis v. United States, 341 U.S. 494, 527 (1951); Gitlow v. New York, 268 U.S. 652, 669 (1925).
[148] Cantwell v. Connecticut, 310 U.S. 296, 309 (1940).
[149] United States v. White, 610 F.3d 956, 957 (7th Cir. 2010).
[150] *Id.*
[151] *Id.* at 1016 (emphasis in original).
[152] People v. Rubin, 96 Cal. App. 3d 968, 979-80 (Cal. Ct. App. 1979).

Thus far, there has been no systematic discussion of the class of speech acts or the contextual factors most likely to incite imminent lawless action. The caselaw is anecdotal and has thus far abjured any comprehensive or rigorous statement of a generalizable principle of risk analysis. Courts are left to reach back to cases with wildly different fact patterns, with the result that principles gleaned from earlier trials are often misapplied.[153]

First Amendment scholars have offered guidance on the requisite threshold of probability, and Smolla suggests a "more probable than not" standard that violence will ensue from inciting speech,[154] but he gives no basis for his formulation, other than it seems like a promising place to start. Healy discusses likelihood under an article subheading on imminence, and suggests that the probability standard should be a "reasonable chance" that the harm will result,[155] drawing on Greenawalt who recommends "reasonable likelihood," an imprecise formulation that simply adds "reasonable" to "likelihood" to render the standard appear more, well, reasonable.[156] Healy mulls over a hypothetical scenario that could elevate or depress the likelihood of a crime occurring if, for instance, a speaker encourages a crowd to storm the city hall in five hours, but draws few general principles.[157]

---

[153] For instance, the appellants in *Nwanguma v. Trump* misapply *Bible Believers* in their brief to the Sixth Circuit when they fail to recognize that in *Bible Believers*, the irate crowd attacked the speakers rather than attacking other members of the audience who were targeted by the speaker. *See* Opening Brief of Appellants Donald J. Trump and Donald J. Trump for President, Inc. at 34-35, Nwanguma v. Trump, 903 F.3d 604 (6th Cir. 2018) (No. 17-6290). *See generally* Bible Believers v. Wayne Cty., Mich., 805 F.3d 228 (6th Cir. 2015).

[154] *See* Smolla, *supra* note 38, at 10 ("The 'likelihood' prong of *Brandenburg* appears to mean simply that it is more probable than not that violence will ensue as a result of the defendants' action.").

[155] *See* Healy, *supra* note 26, at 715 n.358 ("Greenawalt has proposed a similar standard, arguing that there must be a 'reasonable likelihood' that criminal advocacy will lead to crime."); *see also id.* at 721 ("As a middle position, therefore, we might conclude that the government can prohibit advocacy of extraordinary harm if there is a 'reasonable chance' that the harm will result.").

[156] *See* GREENAWALT, *supra* note 37, at 267-68 (stating that the likelihood should be "substantial" and vary according to the gravity of the crime).

[157] In Healy's own words,

> As the time frame expands outward, however, the prediction becomes increasingly difficult because of the many unknowable variables involved. If a speaker urges a rowdy audience to storm city hall five hours later, it may initially appear likely that they will do so. But many things could happen between now and then to dissuade them.

Healy, *supra* note 26, at 716.

And that is all we have to go on with respect to the likely meaning of *Brandenburg*'s "likely to incite or produce."[158] It is not an overstatement to conclude that the likelihood prong of the *Brandenburg* three-part test is woefully underdeveloped and largely anecdotal, to the extent that it hampers the meaningful application of the element, and possibly interferes with the application of incitement law *tout court*. According to Gey, the end result of the *Brandenburg* test is "a system guaranteeing virtually absolute protection of free speech within the realm of political advocacy."[159]

The manifest lack of protection for the persons targeted by inciting speech is related in part to the dearth of guidance on key elements of the *Brandenburg* test, and incitement law could greatly benefit from more precise and accurate guidance on the elements of imminence and probability. Thus far, to our knowledge, there has been no comprehensive and contemporary statement on the conditions that predict uptake of inciting speech by listeners.

III. CAUSATION, IMMINENCE, AND LIKELIHOOD

Why does it matter that the *Brandenburg* test only provides meager guidance on its contextual elements? Stated plainly, until incitement doctrine includes an unambiguous definition of probability and a systematic framework for risk analysis, courts are still in the terrain of the "clear and present danger" test, or worse, "bad tendency" doctrine. For the past one hundred years, the U.S. Supreme Court and the lower courts have aspired to go beyond the suppression of mere advocacy and to evaluate the potential harms of speech in its context; hence the "proximity and degree" language of the "clear and present danger" standard, and the imminence and likelihood elements of the *Brandenburg* test.

In a democracy, a systematic framework of risk assessment is required to evaluate political speech, but this project has progressed in fits and starts and is still incomplete. As long as the nonadvocacy prongs remain undefined, the legal regulation of political speech will remain unpredictable and inconsistent, and this will undermine its legitimacy. Without a reliable and defensible guide to predicting what categories of public speech acts in what kinds of contexts are more likely to prompt violence, the courts cannot fulfill the main rationale of incitement law, namely the prevention of substantial evils that Congress and state legislatures have a right to prevent.

---

[158] Furthermore, to our knowledge there is no discussion at all of the phrase "or produce," either in the caselaw or in legal scholarship.

[159] Gey, *supra* note 5, at 975.

If the state cannot adequately protect its citizenry from one another and prevent foreseeable harms, then it has failed in its most elementary duty.[160]

At this point, it is worth reviewing first principles. Like conspiracy and attempt, incitement is an inchoate crime,[161] where the verbal communication completes the offense and there need be no uptake on the part of the listener to commit the crime nor indeed any consequences at all. Inchoate crimes are crimes of prevention, designed to interdict a harmful chain of causation once a substantial step has been taken towards commission. Charging a crime ab initio is the only way to prevent the offense,[162] for instance when there is no time for police action, persuasion, and counterspeech to do their work.

What is apparent is that if it is to suppress inciting speech, law enforcement must necessarily engage in risk analysis based upon the context of the incitement. Prevention is simply impossible without predictive forecasting of likely consequences. In every incitement trial, the question is

---

[160] In Karl Popper's liberal democratic theory of the state, the state exists to protect citizens' freedoms from the aggression of others. KARL POPPER, THE OPEN SOCIETY AND ITS ENEMIES 110-11 (5th ed. 1966). Consistent with Thomas Hobbes' version of the social contract, Popper conceives the origins of the state in its function as a crime prevention society based upon a social contract with its citizens to provide for their security. *See id.* ("I am perfectly ready to see my own freedom of action somewhat curtailed by the state, provided I can obtain protection of that freedom which remains . . . . [T]he state should be considered as a society for the prevention of crime i.e. aggression.").

[161] BLACK'S LAW DICTIONARY defines an inciter as:

> [O]ne who counsels, commands or advises the commission of a crime. It will be observed that this definition is much the same as that of an accessory before the fact. What, then, is the difference between the two? It is that *in incitement the crime has not (or has not necessarily) been committed,* whereas a party cannot be an accessory in crime unless the crime has been committed. An accessory before the fact is party to consummated mischief; an inciter is guilty only of an inchoate crime.

Incitement, BLACK'S LAW DICTIONARY (10th ed. 2014) (emphasis added) (quoting GLANVILLE WILLIAMS, CRIMINAL LAW: THE GENERAL PART 612 (Stevens & Sons eds., 2d ed. 1961)).

[162] Incitement overlaps conceptually with attempt, and Justice Douglas's concurring opinion in *Brandenburg* recalls that in *Schenck*, the defendant was charged with attempts to cause insubordination in the military and to obstruct enlistment. Brandenburg v. Ohio, 395 U.S. 444, 450 (1969). In incitement, the substantial step that warrants intervention is the direct advocacy of a crime where the commission is imminent and likely. Gideon Yaffe elaborates on this observation, noting,

> In cases of this kind, whether a solicitation is enough for the act element of the attempt depends on two things: on the nature of the crime and the circumstances, and on the test that the court in question employs for determining whether the defendant's act was enough for the act element of the attempt.

GIDEON YAFFE, ATTEMPTS 197 (2010).

the same: does the speech act directly advocate an offense and if so, then does the advocacy sufficiently elevate the risk that an offense will ensue, to the extent that the speech loses its constitutional protection? Once the advocacy prong is fulfilled and it is established that the speaker possessed the requisite intention to advocate the commission of the crime, then liability hinges on whether it is likely that imminent lawless action will occur.

The likelihood that a crime will be committed depends on the circumstances of the speech act: context is everything. In First Amendment jurisprudence, it is now received wisdom that, "the character of every act depends upon the circumstances in which it is done."[163] The same words that are uncontroversial in a peaceful context may be inciting in a violent scenario; "[An] utterance in a context of violence can lose its significance as an appeal to reason and become part of an instrument of force."[164] After *Brandenburg*, imminence and likelihood represent the required contextual elements that define inciting speech. The actus reus elements of incitement are located therefore in the actual speech act itself, and in a situation where the commission of the crime is both imminent and likely. The mens rea of incitement resides in the direct advocacy of a crime, but advocacy is benign unless it occurs in a hazardous setting where there is a heightened probability that the crime will be committed imminently.

Together, imminence and likelihood characterize the potential nexus between the speech act and the future crime, and therefore could be considered tests of causation, but because incitement is an inchoate crime, they are tests of potential, future, or *probable* causation. The incitement caselaw says nothing about probable causation, so what exactly is it? Legal philosophers such as Michael Moore have argued that all causation involves the elevation of probability, or is, in his words, "chance-raising."[165] For Moore, chance-raising is a reasonable test of causation, and he observes that in law, an increase in the probability of effectuating a result is part of the very definition of cause: "a cause is the raising of the probability of its effect."[166]

---

[163] Schenck v. United States, 249 U.S. 47, 52 (1919); *see also* Bose Corp. v. Consumers Union of U.S., Inc., 466 U.S. 485, 504-05 (1984) ("In each of these areas, the limits of the unprotected category, as well as the unprotected character of particular communications, have been determined by the judicial evaluation of special facts that have been deemed to have constitutional significance.").

[164] Milk Wagon Drivers Union of Chi., Local 753 v. Meadowmoor Dairies, Inc., 312 U.S. 287, 293 (1941).

[165] *See* MICHAEL S. MOORE, CAUSATION AND RESPONSIBILITY: AN ESSAY IN LAW, MORALS AND METAPHYSICS 307-09 (2009) (noting that "risk-based" liability for raising the probability of a harm or crime is widely recognized in both criminal and civil responsibility doctrines).

[166] *Id.* at 307.

For their part, courts have studiously avoided all discussion of probable causation and the relationship between imminence and likelihood. Nor have they provided any general guidance on how to determine probable causation as opposed to the conventional garden-variety (completed) causation usually required in criminal law. As criteria that indicate an enhanced risk of causation, there is a certain element of redundancy in the inclusion of both likelihood and imminence, since a bad result which is imminent is also likely and that which is likely is, probabilistically, more imminent (or at least more imminent than it was when it was unlikely). Imminence and likelihood are conceptually interrelated and partially define one another, according to the expression: "more likely, more imminent," or the formula, >L ⇄ >I.

Even though the *Brandenburg* precedent is fifty years old, contemporary courts still invoke opaque criteria when evaluating probable causation and when distinguishing this task from conventional, backward-looking appraisals of causation. Recalling the adage attributed to the physicist Niels Bohr, "It is difficult to make predictions, especially about the future," it is worth acknowledging that assessing the risk of a future harm is inherently tentative, and is an exercise fraught with, well, risk. And yet in incitement cases, the police, prosecutors, and the courts currently possess few analytical tools to engage in risk assessment.

In practice, judges put their faith in conventional heuristics, or mental shortcuts allowing decisions about complex material, that are part of the habitual parlance of judging and derived from "common sense."[167] Stated less generously, judges are deciding incitement cases on the basis of cognitive illusions, hunches, intuitions, and implicit biases.

As a result of incitement doctrine's conceptual lacunae, courts have fallen into three unhealthy habits with respect to the causation dimensions of incitement. First, they have reverse-engineered probable causation from the concrete consequences of a speech act, also known as the logical fallacy of *post hoc ergo propter hoc*.[168] Second, they have advanced confusing and misleading metaphors of speech to connect public speech to ensuing harms. Finally, they have enunciated

---

[167] Amos Taversky and Daniel Kahneman explain this phenomenon further, stating

> This judgmental heuristic is called availability. Availability is a useful clue for assessing frequency or probability, because instances of large classes are usually reached better and faster than instances of less frequent classes. However, availability is affected by factors other than frequency and probability. Consequently, the reliance on availability leads to predictable biases . . . .

Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 3, 11 (Daniel Kahneman et al. eds., 1982).

[168] Literally, "after this, therefore, because of this."

general theories of the relationship between speech and social behavior that lack any empirical substantiation. We address these three forms of flawed reasoning in turn.

A. Post Hoc Ergo Propter Hoc

In key incitement cases, there has been a tendency to analyze risk on the basis of the putative consequences of speech, and to construe these effects as an intrinsic and inexorable result of speech. In philosophy as in law, *post hoc ergo propter hoc* is a logical fallacy[169] that holds (incorrectly) that that if *Y* occurred after *X*, then *X* must have caused *Y*. Since the Scottish Enlightenment philosopher David Hume, however, it is accepted that chronology does not prove causation. In their landmark treatise *Causation in the Law*, Hart and Honoré affirmed that "not all events which follow each other in invariable sequence are causally related."[170]

Unquestionably, harms may occur for reasons that are completely unrelated to a speech act, for instance when the intended injury is performed by a person who did not hear the original incitement. Furthermore, a speech act that directly advocates a crime and substantially raises the risk of imminent lawless action to a threshold that justifies its banning under law may not actually trigger the harm. Causation is, after all, probabilistic, and a communicative input may not generate a criminal outcome in every instance.

Illustrating this phenomenon, the Supreme Court reverse-engineered imminence and likelihood in *Claiborne Hardware* when it found that Charles Evers' public threat to break the "damn neck" of any individual violating the boycott of white stores[171] was protected speech, but with the qualification that "[i]f that language had been followed by acts of violence, a substantial question would be presented whether [the speaker] could be held liable for the consequences of that unlawful conduct."[172]

As it happened, there were multiple acts of violence against African Americans accused of violating the NAACP's boycott in Claiborne County, Mississippi, but the Court discounted this because the violence occurred some "weeks or months" after Evers' speech.[173] Here, the inchoate character of the crime of incitement is extinguished when the elements of probable causation are determined ex post facto by reference to the ensuing consequences of speech, rather than by reference to the degree to which the

---

[169] Along with other logical fallacies such as generalization from one occurrence, claim-begging, and tautological argumentation.
[170] H.L.A. HART & TONY HONORÉ, CAUSATION IN THE LAW 15 (Oxford Univ. Press 2d ed. 1985).
[171] NAACP v. Claiborne Hardware Co., 458 U.S. 886, 902 (1982).
[172] *Id*. at 928.
[173] *Id.* at 904, 928.

criminal advocacy elevated the risk of violence in the moment and circumstances in which the advocacy occurred.

Similarly, in *McCoy v. Stewart*, the Ninth Circuit held that the defendant's argument that his advice on how to commit criminal offenses to members of a street gang called the "Bratz" was mere idle talk would be nullified,

> if the state could prove that the speech actually caused imminent lawless action. Here, however, McCoy correctly observes that his words did not actually incite anyone to commit a crime. There is no evidence in the record that the Bratz engaged in any crime as a result of his advice. Indeed, there is no evidence that McCoy's speech played any part at all in any crime committed by the Bratz.[174]

Finally, in *American Freedom Defense Initiative v. Metropolitan Transportation Authority*, the court held that there was no element of imminence present in the ads that referred to the murder of Jews because no acts of violence could be attributed to the ads.[175] As in *Claiborne*, the courts seem to be saying that incitement can only be proven if there is evidence of actual harm, rather than judging the facts by reference to the ex ante imminence and likelihood elements of the *Brandenburg* test.

More examples could be found. What is apparent in each instance is that the court is mistakenly deciding the elements of imminence and probability on the basis of what happened next, that is, actual causation specific to the facts of the case, rather than probable causation at the moment of the utterance in a particular context. As argued, this constitutes specious reasoning about the probable causation pertaining to an inchoate crime, which must be performed ex ante rather than ex post facto.

Determining likelihood and imminence on the basis of outcomes underlines the deep-seated tension between the probable causation elements of incitement and the inchoate nature of crime, an unresolved tension that explains the inconsistencies in reasoning and doctrine. When the putative consequences of a speech act are invoked in determining the two actus reus elements of the crime of incitement—imminence and likelihood—then actual causation becomes an element of the crime and incitement can no longer be considered an inchoate crime.

---

[174] McCoy v. Stewart, 282 F.3d 626, 631 n.6 (9th Cir. 2002).

[175] *See* Am. Freedom Def. Initiative v. Metro. Transp. Auth., 70 F. Supp. 3d 572, 582-83 (S.D.N.Y. 2015) ("The defendants also cannot point to any objective evidence to support their concerns that the advertisement is an imminent incitement of violence . . . . Therefore, these ads—offensive as they may be—are still entitled to First Amendment protection."), *vacated*, 109 F. Supp. 3d 626 (S.D.N.Y. 2015), *aff'd*, 815 F.3d 105 (2d Cir. 2016).

Logically, this defeats the preventative ends of making incitement a crime in the first place, and it compels the potential victims of speech that incites hate crime to absorb all the risks of the calls to violence against them. If there is an unspoken consensus that courts will only find incitement when the crime being incited occurs, and if indictments must demonstrate a causal nexus between inciting speech and criminal acts, then criminal prosecutors are likely to wait until a crime has been completed before charging an individual for incitement. The approach of the courts has thwarted the ends of crime prevention.

B. Misleading Metaphors of Causation

First Amendment jurisprudence is a metaphor-rich environment in which judges unleash their lyrical passions. The doyen of First Amendment poetic license is Justice Oliver Wendell Holmes, who penned such memorable lines as "the word is . . . not transparent . . ., it is the skin of a living thought . . . ."[176] Even more renowned is Holmes's marketplace of ideas metaphor which construes public, political speech as a consumer item that can be freely bought and sold.[177] Even though it was coined nearly a hundred years ago, the marketplace metaphor still has currency in recent First Amendment cases.[178]

Some metaphors in incitement rulings are conspicuously causal and describe the nexus between speech and subsequent action. Causal metaphors are more than just an attractive adornment to a decision. They do actual work by verifying the connection between words and acts that may not be easily discernible in the evidence presented to the court. For instance, in *Dennis*, a Red Scare case in which the defendants were convicted simply of reading and discussing Marxist works, the judgment approvingly cited Judge Learned Hand in *Masses Publishing Co. v. Patten*: "One may not counsel or advise

---

[176] Towne v. Eisner, 245 U.S. 418, 425 (1918).

[177] Holmes called upon this analogy in *Abrams v. United States*, writing in dissent,

> But when men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out.

250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

[178] *See, e.g.*, Citizens United v. Fed. Election Comm'n, 558 U.S. 310, 354 (2010) (asserting that a previous opinion, overruled by this decision, "interferes with the 'open marketplace' of ideas protected by the First Amendment"); Bible Believers v. Wayne Cty., 805 F.3d 228, 243 (6th Cir. 2015), ("The First Amendment offers sweeping protection that allows all manner of speech to enter the marketplace of ideas.").

others to violate the law as it stands. Words are not only the keys of persuasion, but the triggers of action . . . ."[179]

Metaphors of fire and conflagration abound in incitement decisions, and were particularly a feature of the "bad tendency" or "advocacy alone is sufficient" standard of incitement.[180] In *Frohwerk*, the U.S. Supreme Court acknowledged that the defendant's publication of antiwar and pro-Germany sentiments in the weekly newspaper *Missouri Staats-Zeitung* seemingly posed a minimal risk because it was a local Kansas City newspaper published in German, but the Court convicted him nonetheless on the grounds that "it is impossible to say that it might not have been found that the circulation of the paper was in quarters where a little breath would be enough to kindle a flame and that the fact was known and relied upon by those who sent the paper out."[181]

The conviction of Socialist Party of America member Benjamin Gitlow relied on a similar combustion metaphor regarding the potential harms of political speech when it held that "[a] single revolutionary spark may kindle a fire that, smouldering for a time, may burst into a sweeping and destructive conflagration."[182] Fire and conflagration metaphors abound in recent cases such as *United States v. White*, in which the Seventh Circuit held that the white supremacist website posting the name of a jury foreman and inviting readers to harm him was "playing with fire,"[183] adding yet more redolent imagery to the district court's metaphor of speech as "powerful medicine" in that case.[184]

---

[179] Dennis v. United States, 341 U.S. 494, 571 (1951) (quoting Masses Publ'g Co. v. Patten, 244 F. 535, 540 (S.D.N.Y. 1917)).

[180] They did not, therefore, originate in another Holmes's aphorism that "[t]he most stringent protection of free speech would not protect a man in falsely shouting fire in a theater and causing panic." Schenck v. United States, 249 U.S. 47, 52 (1919).

[181] Frohwerk v. United States, 249 U.S. 204, 209 (1919).

[182] Gitlow v. New York, 268 U.S. 652, 669 (1925).

[183] United States v. White, 698 F.3d 1005, 1014 (7th Cir. 2012) ("The fact that White made an effort to discourage assassination attempts against Juror A when law enforcement moved against his website shows at a minimum that he knew he was playing with fire.").

[184] As Judge Adelman held in another *United States v. White* opinion,

> The constitutional right of free expression is powerful medicine in a society as diverse and populous as ours. It is designed and intended to remove governmental restraints from the arena of public discussion, putting the decision as to what views shall be voiced largely into the hands of each of us . . . .

638 F. Supp. 2d 935, 956 (N.D. Ill. 2009) (quoting Cohen v. California, 403 U.S. 15, 24 (1971)).

Some legal scholars have objected to the prevalence of metaphorical and figurative language in judgments.[185] Felix Cohen famously excoriated the U.S. Supreme Court's figurative language in deciding where a company resided (and therefore could be sued) in *Tauza v. Susquehanna Coal Company*.[186] Cohen condemned the Court's use of metaphor and analogy in determining the factual question, "where is a corporation?" on the grounds that symbolic discourse interfered with a rational and scientific deliberation of the facts of the case.[187] Since Cohen, philosophers have developed a critique of metaphorical language and argued that metaphors are not propositional in character,[188] are confusing, emotive, and unsuited to serious scientific inquiry,[189] and are "empty of guidance for the court."[190]

Metaphors are not simply adornments in First Amendment law, and as the linguistic philosopher J.L. Austin observed, they do things.[191] Metaphors convert the complex relationship between a speech act and its

---

[185] *See* Felix Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809, 812 (1935) (describing the use of metaphor in jurisprudence a distraction from the actual "social forces which mold the law and the social ideas by which the law is to be judged"). *See generally* RICHARD ASHBY WILSON, INCITEMENT ON TRIAL: PROSECUTING INTERNATIONAL SPEECH CRIMES 152-65 (2017) (detailing historical ambivalence about the use of metaphor in jurisprudence and concluding that such the rhetorical device is often used "in a way that is misleading and rests on a defective psychology").

[186] *See generally* Tauza v. Susquehanna Coal Co., 115 N.E. 915 (1917) (describing the relevant corporation's objection that it could not be sued in New York, even though it regularly conducted business there, because it was headquartered and established somewhere else). Responding to the decision in *Tauza*, Cohen wrote,

> Yet it is exactly in these terms of transcendental nonsense that the Court of Appeals approached the question of whether the Susquehanna Coal Company could be sued in New York State. 'The essential thing,' said Judge Cardozo, writing for a unanimous court, 'is that the corporation shall have come into the State.' Why this journey is essential, or how it is possible, we are not informed. The opinion notes that the corporation has an office in the State, with eight salesmen and eleven desks, and concludes that the corporation is really 'in' New York State. From this inference it easily follows that since a person who is in New York can be sued here, and since a corporation is a person, the Susquehanna Coal Company is subject to suit in a New York court

Cohen, *supra* note 185, at 811-12.

[187] Cohen, *supra* note 185, at 811 (critiquing Judge Cardozo's choice of transcendental language to describe the legal location of a corporation).

[188] *See* Donald Davidson, *What Metaphors Mean*, *in* INQUIRIES INTO TRUTH AND INTERPRETATION 262 (2d ed. 2001) (denying that metaphors stand for specific facts or have "specific cognitive content").

[189] *Id.* at 247.

[190] HART & HONORÉ, *supra* note 170, at 97.

[191] J.L. AUSTIN, *How to Do Things With Words, in* THE WILLIAM JAMES LECTURES DELIVERED AT HARVARD UNIVERSITY IN 1955 6 (1962) ("[T]o utter the sentence . . . is not to *describe* my doing of what I should be said in so uttering to be doing or to state that I am doing it: it is to do it.").

context into a set of physical objects that make up the material-object world, transforming what are in fact social relations (how a speaker might goad others into acts of violence) into material-object-world relations. Metaphors of fire or medicine allow judges to posit a mechanistic causal relationship between advocacy and behavior. When courts perceive speech as analogous to a smoldering fire or "triggers of action" or "powerful medicine," they characterize the possible mental causation between persons in the paradigm of physical laws of cause and effect, laws that more easily conform to criminal law's standard model of causation.

Yet this analogy is built on a false premise. Unlike fire or medicine, humans possess intentionality and agency,[192] and the intersubjective dynamics of persuasion between persons is more complicated than a fire spreading through a forest or a medicine being administered to a patient. In addition to raising valid due process concerns (how does one appeal against a conviction written in flights of figurative fancy?), metaphors hinder our comprehension of how exactly and realistically an advocacy might have increased the likelihood of an imminent offense. As an antirealist language, metaphors should not serve as the basis of legal reasoning, especially when they contemplate the contextual elements of speech.[193] Metaphorical language of the type that is often present in calls to violence ought to be met with sober deliberation in the courts, not yet more metaphorical speech.

C. Baseless Theories of Speech

> *"[W]hen we don't have enough knowledge to decide a case in an informed way, we necessarily fall back on how we 'feel' about the case."*
>
> –Judge Richard A. Posner[194]

---

[192] By "agency" we refer to intentionality and a minimum capacity to take preliminary steps towards the fulfillment of the intention. *See* JOHN R. SEARLE, RATIONALITY IN ACTION 95-96 (2001) (identifying the following feature of a rational agent: she is conscious, persists through time, operates with reasons, and is capable of deciding, initiating and carrying out actions and is responsible for at least some of her behavior).

[193] *See* Berkey v. Third Ave. Ry. Co. 155 N.E. 58, 61 (N.Y. Ct. App. 1926) ("Metaphors in law are to be narrowly watched, for starting as devices to liberate thought, they end often by enslaving it."). *See generally* Steven L. Winter, *Transcendental Nonsense, Metaphoric Reasoning, and the Cognitive Stakes for Law*, 137 U. PA. L. REV. 1105, 1162–63 (1989) (summarizing critiques of metaphors in legal reasoning).

[194] Joel Cohen, *An Interview with Judge Richard A. Posner,* A.B.A. J. (July 1, 2014, 10:20 AM), http://www.abajournal.com/magazine/article/an_interview_with_judge_richard_a._posner [https://perma.cc/7868-JQT5].

This Section demonstrates that in First Amendment cases, judges are often prone to making generalized assertions about social behavior, and in particular, about the likely consequences of speech that provokes a violent response, or which denigrates or threatens a social group. These claims are seldom, if ever, accompanied by empirical substantiation of any kind, and appear to be based on the judges' intuitions, unconscious biases, and rules of thumb.

Some judicial statements about the effects of speech are plausible, others are utterly implausible, but whatever the case, judicial declarations are often made without any supporting evidence, and therefore are unverified. And yet these claims can play an essential role in the judges' reasoning in hard cases. For example, Justice Brandeis claimed in *Whitney* that "It is the function of speech to free men from the bondage of irrational fears,"[195] although that would obviously depend on the kind of speech. Self-evidently, political speech can inculcate irrational fears and instigate public disorder and violence, otherwise there would be no basis for laws regulating incitement, true threats, and fighting words.

Further in this vein, in deciding that the burning of the American flag was constitutionally protected speech, the Supreme Court stated in *Texas v. Johnson* that taking "serious offense" at an expression is not "necessarily likely" to provoke a disturbance of the peace.[196] The conjoining of "necessarily" and "likely" seems like clever hedging on the part of the Justices, and leaves imprecise the Court's opinion on the nature of the probabilistic relationship between taking serious offence and subsequent disorder. Of course, the taking of offense does not necessarily cause violence in all instances. However, behavioral research on the defilement of national symbols indicates that taking serious offense does increase the chances that there will be disorder. For instance, moral foundations psychologist Jonathan Haidt writes about how the flag is conceived as "one of the sacred pillars supporting the community" and an object of "infinite value," the desecration of which is sure to lead to a response that is "swift, emotional, collective and punitive."[197]

Judicial decisions frequently pronounce on the nature of the harms experienced by historically disadvantaged communities subjected to true threats and incitement in an evidentiary vacuum. Justice Scalia, writing for the majority in the cross-burning case of *R.A.V. v. City of St. Paul*, struck

---

[195] Whitney v. California, 274 U.S. 357, 376 (1927) (Brandeis, J., concurring).

[196] *See* Texas v. Johnson, 491 U.S. 397, 408 (1989) ("The State's position, therefore, amounts to a claim that an audience that takes serious offense at particular expression is necessarily likely to disturb the peace and that the expression may be prohibited on this basis. Our precedents do not countenance such a presumption.").

[197] *See* JONATHAN HAIDT, THE RIGHTEOUS MIND: WHY GOOD PEOPLE ARE DIVIDED BY POLITICS AND RELIGION 174 (2012) (describing emotional and punitive responses to the desecration of the flag).

down the city ordinance banning cross-burning and daubing swastikas by reasoning that injury caused by an expression of group hatred is not "qualitatively different" from injury which does not invoke group hatred.[198] No evidence is provided for this assertion, which set off a lively debate between the majority and Justice Stevens, who wrote a concurring opinion, about whether the denigration of a person's race is of a qualitatively different order than speech that disparages their personal characteristics.[199]

This question has been tested empirically, and the social science literature has identified the uniquely harmful effects of racist speech. In two experiments, psychologists Boeckman and Liew tested Asian American students' responses to racist speech relative to personal forms of insult, and their reactions to offenses that are motivated by group hatred (e.g., insulting speech) to those that are not (e.g., petty theft).[200] They found that hate speech depressed collective self-esteem[201] and caused more extreme emotional reactions than insults not based on group hatred. They concluded that hate speech is "distinctive"[202] and has a "broader harmful impact . . . than other forms of insult or property crime."[203] In the light of this study and others,[204] Justice Scalia's claims run contrary to the scientific evidence and therefore represent mere personal opinion.

Then there are the customary yardsticks used by the Court to measure the likely responses to provocation and therefore the likelihood of imminent lawless action. One objective standard found in fighting words doctrine is how the "average person" might respond, a thesis that originates in the "reasonable man" test for standards of care and foreseeability in tort law, as

---

[198] R.A.V. v. City of St. Paul, Minn., 505 U.S. 377, 392-93 (1992) (discounting Justice Stevens's claim that there was a "qualitatively different" type of injury targeted by the ordinance relevant for first amendment analysis).

[199] *See id.* at 392 (providing Justice Scalia's opinion on the meaning of "qualitatively different" in context, which conflicts with Justice Stevens' view); *see also id.* at 424-25 (Stevens, J., concurring) (providing an alternative approach to Justice Scalia's analysis).

[200] Robert J. Boeckmann & Jeffrey Liew, *Hate Speech: Asian American Students' Justice Judgments and Psychological Responses,* 58 J. SOC. ISSUES 363, 363 (2002).

[201] *Id.* at 377.

[202] *Id.* at 379.

[203] *Id.* at 365.

[204] *See, e.g.*, Laura Beth Nielsen, *Subtle, Pervasive, Harmful: Racist and Sexist Remarks in Public as Hate Speech,* 58 J. SOC. ISSUES 265, 279 (studying the prevalence and severe effects of racist and sexist speech in public places); *see also* Brian Mullen & Joshua M. Smyth, *Immigrant Suicide Rates as a Function of Ethnophaulisms: Hate Speech Predicts Death,* 66 PSYCHOSOMATIC MED. 343, 343 (2004) (studying the correlation between ethnic slurs and suicide in immigrant populations).

introduced by Justice Holmes.[205] In First Amendment cases, the average person standard can serve as a legal fiction that cloaks the opinions and prejudices of the majority as to the kind of speech acts that give offense and prompt public disorder.

In *Bible Believers*, the Sixth Circuit held that carrying a severed pig's head on a stick and placards stating that the Prophet Muhammed is a fraud and a pedophile was not speech "likely to provoke the average person."[206] Who exactly is the "average person" here? They do not seem to be a follower of the prophet Muhammed and the religion of Islam. The Court would have been well advised to avoid adopting a majoritarian (i.e., Christian) religious vantagepoint from which to evaluate how the Muslim community of Dearborn, Michigan might (and in fact, did, repeatedly) react to the evangelical group's provocations. Instead, the court's standard for inciting public disorder might begin with the attitudes and values of the members of the group who are the target of abusive and inciting speech. That is, how are most Muslims in Dearborn, Michigan likely to react to religiously based insults from the Bible Believers evangelical group? Indeed, an empirically verifiable standard of the risk of provocation was available to the Court, insofar as the Bible Believers' speech had already prompted public disorder at the Arab International Festival in 2011, the year preceding the events in question.[207] In response, the Bible Believers group ratcheted up their religious baiting at the 2012 cultural festival by parading with placards that expressed even more offensive insults about the prophet Muhammed and by carrying a severed pig's head on a stick.

It has been established in other cases that when evaluating whether a speech act constitutes incitement, it is imperative to understand how the speaker's intended audience, not the "average person" or the general public, would have understood her words.[208] There is precedent in First Amendment law for evaluating the likely consequences of a speech act on the basis of its ordinary meaning for its intended audience, rather than the legal fiction of the "average person." For instance, *Gitlow* affirmed that,

---

[205] In *The Common Law*, Holmes sought to resolve the problem of negligence in torts law by creating an objective standard to determine what outcomes are foreseeable from individual actions, so as to attribute liability in a credible manner. OLIVER WENDELL HOLMES, THE COMMON LAW 93 (1881) (describing the reasonable man test for foreseeable "consequences" that ought result in liability in tort).

[206] Bible Believers v. Wayne Cty., Mich., 805 F.3d 228, 246 (6th Cir. 2015) (citing Street v. New York, 394 U.S. 576, 592 (1969), which held that no advocacy can constitute fighting words unless it is "likely to provoke the average person to retaliation").

[207] *Id.* at 236.

[208] The same words may have entirely different meanings depending on the specific audience. *Compare* Saylor v. Bd. of Educ. of Harlan Cty, Ky., 118 F.3d 507, 509 (6th Cir. 1997) (describing the statement "I'll take care of him" being taken to mean corporal punishment), *with* United States v. Franklin, 415 F.3d 537, 550 (6th Cir. 2005) (describing defendant's promise to "take care of [him]," which witness understood as sharing proceeds of robbery).

> The court, among other things, charged the jury, in substance, that they must determine what was the intent, purpose and fair meaning of the Manifesto; that its words must be taken in their ordinary meaning, as they would be understood by people whom it might reach.[209]

As we have demonstrated, judges often rely on metaphors that are misleading and non-falsifiable in incitement cases. They make generalized statements unsupported by any empirical research and employ standards to gauge likely responses that are not objective and as we just saw in *Bible Believers*, can convey majoritarian religious biases. Personal experience and the evidence presented in the trial are not be a sound basis for calculating risk. Determinations of fact that rely on a general model of speech and behavior and that lack empirical substance are therefore specious. In their desire to prioritize normative principles, decisions in incitement cases such as *Bible Believers*, *R.A.V.*, and *Texas v. Johnson* have made proclamations about social behavior that are contradicted by behavioral research.

A sounder basis for determining imminence and likelihood in incitement cases would involve disaggregating the brute facts from the guiding principles and accepting the facts of the case on the basis of the observed pattern of behavior, informed by relevant social science research on the topic. When the facts are congruent with the extant policy, then all is well. If the facts contradict the policy, then it is inadvisable, as courts have occasionally done, to reconfigure the facts to conform to the preferred policy.

Of course, even if the behavioral research does not support current policy, there may still exist compelling normative reasons to maintain the policy anyway. Where the facts of the matter are inconsistent with the desired policy of the court, for instance, arguendo, that flag burning creates widespread offense and offense is more likely to lead to violence, courts and legislatures may still decide to permit flag burning for other reasons, for instance, because they value vigorous political protest of a symbolic kind. We term this elevation of policy considerations over countervailing material facts in incitement law a "normative override."[210] Rather than concocting an implausible interpretation of the facts to suit the desired policy, it is preferable to acknowledge explicitly that the courts are engaging in a normative override.

There is much work to be done with respect to the probable causation elements of incitement doctrine and procedure. The consistency and predictability of the law of incitement would be enhanced if courts were transparent in advance about the criteria they use to evaluate the content and

---

[209] Gitlow v. New York, 268 U.S. 652, 661 (1925).

[210] To our knowledge, "normative override" has not been previously used in this way in legal scholarship.

context of speech to determine whether it could foreseeably result in imminent lawless action.

IV. A MATRIX FOR ASSESSING RISK

> *"How are we going to substitute a realistic, rational, scientific account of legal happenings for the classical theological jurisprudence of concepts?"*
> –Felix Cohen[211]

For much of U.S. history, the crime of incitement was an instrument to suppress seditious and antiwar speech.[212] In the mid-twentieth century, courts began to orient doctrine to protecting individuals from harm motivated by group bias.[213] This transition is still incomplete. In order to complete it, First Amendment analysis in incitement cases must engage more fully in a contextual risk analysis.

The idea of risk assessment was introduced in the "clear and present danger" test, but its promise was never fully realized, and the test became a repressive tool.[214] The fetters were lifted in *Brandenburg* which announced the elements of imminence and likelihood but did not define them.[215] Consequently, judges and other legal actors have relied on models of speech that are metaphorical and inaccurate.[216] As a result, the elements remain obscure, leaving the targets of incitement unprotected. The incitement matrix below offers an evidence-based approach to risk assessment informed by the latest research on the relationship between speech and behavior.[217] The factors are organized under three general headings: the attributes of the speaker, the content of the message, and the

---

[211] *See* Cohen, *supra* note 185, at 821.
[212] *See supra* Subsection I.C.
[213] *See supra* Subsection I.D.
[214] *See supra* Section II.D.
[215] *See supra* Section II.E.
[216] *See* Paul Horwitz, *Free Speech as Risk Analysis: Heuristics, Biases, and Institutions in the First Amendment,* 76 TEMP. L. REV. 1, 4-5 (2003) (proposing an application of behavioral analysis to risk in First Amendment law).
[217] For a summary chart, see supra Appendix. Susan Benesch presents her own matrix for direct and public incitement to genocide. *See* Susan Benesch, *Vile Crime or Inalienable Right: Defining Incitement to Genocide*, 48 VA. J. INT'L L. 485, 493-94 (2008) (proposing a new spectrum and definition for evaluating indictment to genocide). Her matrix is a dramatic improvement on what came before in the international law of direct and public incitement to commit genocide. Our matrix differs from hers in two respects: it is informed by social research and it focuses on *Brandenburg* and U.S. law rather than international criminal law. *See also* Richard Ashby Wilson, *Inciting Genocide with Words*, 36 MICH. J. INT'L L. 277, 300-04 (2015) (evaluating Benesch's framework).

context in which the message is delivered. Each is known to elevate the probability that listeners will act on criminal advocacy.

The proposed incitement scale delivers a more transparent and rational basis for evaluating the risks associated with specific speech acts than the hodge-podge of folk theories currently circulating in First Amendment law. It does, however, raise the question of the appropriate relationship between law and social science. After *Brown* cited psychological research to reject the notion of "separate but equal," there emerged a voluminous literature on the place of social science in law, and space constraints do not allow us to review it comprehensively.[218] Our sympathies are with legal scholars such as Cass Sunstein who embrace behavioral research and aspire to make government logical.[219]

It is worth noting that risk assessment based on behavioral research is already an existing component of the criminal justice system and shapes judicial decisions on a daily basis.[220] Although not an incitement case, *Harper v. Poway Unified School District* cited numerous social science studies to justify the decision that a high school could suspend a student for wearing a T-shirt with a message expressing religious condemnation of homosexuality.[221] The statistical evidence *Harper* cited was particularly compelling: "[A]mong teenage victims of anti-gay discrimination, 75% experienced a decline in academic performance, 39% had truancy problems and 28% dropped out of school."[222]

The Ninth Circuit concluded that "Harper's wearing of his T-shirt 'colli[des] with the rights of other students' in the most fundamental way,"[223] citing eight studies that demonstrate that verbal attacks on students because of their sexual orientation lead to a significant and measurable decline in future

---

[218] *See generally* Rachel F. Moran, *What Counts as Knowledge? A Reflection on Race, Social Science, and the Law*, 44 L. & SOC'Y REV. 515 (2010) (reviewing social science in the law since *Brown*).

[219] Cass R. Sunstein, *Making Government Logical*, N.Y. TIMES (Sept. 19, 2015), https://www.nytimes.com/2015/09/20/opinion/sunday/cass-sunstein-making-government-logicalhtml.html [https://perma.cc/9C8R-26FK]; *see also* NAT'L RESEARCH COUNCIL, USING SCIENCE AS EVIDENCE IN PUBLIC POLICY (Kenneth Prewitt et al. eds., 2012) (discussing the use of behavioral research in federal agency programs).

[220] *See* JOHN MONAHAN & LAURENS WALKER, SOCIAL SCIENCE IN LAW: CASES AND MATERIALS 398-409 (7th ed. 2010) (discussing the uses of social science in the determination of future facts). For a critical view of prediction in criminal law, see also Sandra G. Mayson, *Collateral Consequences and the Preventive State*, 91 NOTRE DAME L. REV. 301, 303-04 (2015) (providing a critical view of the uses of social science in the determination of future facts, especially in a criminal law context).

[221] Harper v. Poway Unified School Dist., 445 F.3d 1166, 1178-79, 1184 (9th Cir. 2006).

[222] *Id.* at 1179.

[223] *Id.* at 1178 (citing Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 508 (1969)).

academic performance and their potential success in later life.[224] Subsequently, the passage of *Harper* citing relevant social science studies has been reproduced verbatim in *Gillman*, another First Amendment case.[225] While our advocacy of the relevance of social science to legal problems is robust, our approach is cautious and prudent, and we advise a limited application of our ten-point incitement scale. The matrix only applies to the imminence and probability prongs of the three-part *Brandenburg* test, after it has been established that there is advocacy of imminent lawless action.

Furthermore, the objection might be raised that the matrix contemplates a review of the status of the speaker and the content of their message as relevant contextual factors, and in *Citizens United* and *Reed v. Town of Gilbert*, the U.S. Supreme Court countenances neither content-viewpoint discrimination nor speaker-identifying factors in First Amendment analysis.[226]

With respect to possible content or viewpoint discrimination, our scale does not replace existing tests for criminal advocacy. The risk assessment matrix is only pertinent once it has been established that the speech in question advocates a crime. Furthermore, the three content-related factors in the matrix all contain a call to violence, speech that is already prohibited under existing law of true threat and incitement. The factors are therefore consistent with prevailing prohibitions on criminal advocacy. Additionally, we advise that the presence of one or more factors is not in itself predictive of imminent lawless action. Instead, it is the ensemble of factors that may be jointly sufficient. The presence of one to three factors indicates low risk, four to six factors indicates medium risk and seven to ten factors indicates high risk. Thus, the matrix can be exculpatory as well as inculpatory.

With respect to speaker-identifying factors, consideration of the speaker's corporate identity was permissible for many years on the basis of *Austin v. Michigan Chamber of Commerce*, which allowed the regulation of corporate funding of elections from general funds to prevent the appearance of corruption in politics.[227] While space does not allow us to develop the argument here, we favor the approach adopted in *Austin*, which was overruled by *Citizens United*, a decision that has generated more negative comment than most. That being said, the reasons given by the Court for excluding consideration of the speaker's identity in *Citizens United* are not present in our matrix. That decision states that courts cannot restrict political speech on

---

[224] *Id.* at 1179.
[225] Gillman v. Sch. Bd. for Holmes Cty., Fla., 567 F. Supp. 2d 1359, 1370 (N.D. Fla. 2008).
[226] *See* Citizens United v. Fed. Election Comm'n, 558 U.S. 310, 315 (2010) (opining that political speech may not be banned on the basis of the speaker's identity); *see also* Reed v. Town of Gilbert, Ariz., 135 S. Ct. 2218, 2222 (2015) (holding that laws restricting speech based on its content are presumptively unconstitutional); R.A.V. v. City of St. Paul, Minn., 505 U.S. 377, 395 (1992) (stating that laws regulating speech must be content neutral).
[227] Austin v. Mich. Chamber of Commerce, 494 U.S. 652, 668 (1990).

the basis of the speaker's corporate identity,[228] but our scale does not specify any identity, corporate or otherwise. Nor may the government distinguish between speakers as a way of controlling content according to *Citizens United*,[229] but our scale does not seek to suppress content via a speaker's identity.

Finally, *Citizens United* holds that government restrictions may not disfavor certain types of speakers,[230] but our scale also does not do that; it simply notes that the more authoritative, credible and charismatic speakers are for their audience (i.e., regardless of ideological viewpoint or their specific political party membership), the more likely it is that their words will influence audience behavior. In our view, it is impossible to evaluate a speech act in its context and the probability that it will provoke imminent lawless action without considering the attributes of a speaker and his or her relationship with an audience. Such analysis is compatible with, and indeed, demanded by, existing First Amendment jurisprudence.

Research on the effects of speech on listeners has made enormous strides since *Brandenburg* was decided in 1969, and we know much more about the ways that directly inciting speech can impact behavior than before. The next Section of this Article offers an incitement scale of ten factors to assist courts and policymakers as they assess the risk that an incitement could lead to imminent lawless action. It distills the most up-to-date behavioral and social research on persuasion, political communication and denigrating speech into a checklist that can be operationalized by legal actors.

A. The Attributes of the Speaker

1. The speaker occupies an official position of authority within government or a political party or political movement.

If an authority figure exhorts violence or other persons appear to accept violent exhortations, then a majority of individual listeners will likely conform to the authority's message, even if they disagree with it. The power of an authority figure's words was clearly demonstrated in Milgram's famous experiments in which otherwise normal and empathetic individuals ("teachers") would obey an authority's instructions to administer seemingly deadly electrical shocks to innocent victims

---

[228] *Citizens United*, 558 U.S. at 315.
[229] *Id.* at 312.
[230] *Id.* at 313-14.

("learners").[231] For decades, psychologists asked "Would people still obey today?" Burger replicated Milgram's results in 2009, finding that nearly three-quarters of participants delivered "shocks" as instructed.[232] In recent variation on Milgram's study, Mermillod and colleagues found that most participants comply with an authority's destructive requests, such as insulting another person, even when there is little or no pressure to do so.[233] Focusing on the underlying cognition of obedience, Grzyb and colleagues discovered that obedience to an authority can be elicited with as little as the would-be follower sharing an aspect of the authority's social identity, such as gender, or possessing a high need for cognitive closure (i.e., desire to remove ambiguity from the world).[234]

These findings extend to numerous studies of social groups. In Asch's experiments on intragroup behavior, many participants submitted to the collective's patently false beliefs about the world if they were outnumbered or under the observation of an authority figure.[235] Zimbardo's famous Stanford Prison Experiment elicited similar behavior when participants ("guards") physically abused others ("prisoners") or conformed to an authority figure's instructions to control prisoners.[236] In a real-world scenario, Straus recently analyzed twenty-four cases of mass violence in Africa, discovering that the most critical factor in each case was government authorities who encouraged violence and also coordinated it through their speech acts.[237] Other recent studies on mass violence corroborate Straus's findings and show that the greater the

---

[231] *See* Stanley Milgram, *Liberating Effects of Group Pressure*, 1 J. PERSONALITY AND SOC. PSYCHOL. 127, 134 (1965) (discussing obedience to authority figures and the moderating role of group conformity).

[232] Jerry M. Burger, *Replicating Milgram: Would People Still Obey Today?*, 64 AM. PSYCHOLOGIST 1, 8 (2009).

[233] Martial Mermillod et al., *Destructive Obedience Without Pressure: Beyond the Limits of the Agentic State*, 46 SOC. PSYCHOL. 345, 350 (2015).

[234] *See generally* Tomasz Grzyb et al., *Cognitive Structuring and Obedience Toward Authority*, 135 PERSONALITY & INDIVIDUAL DIFFERENCE 115 (2018) (discussing personal traits that make individuals susceptible to obedience to authority).

[235] *See* Solomon E. Asch, *Opinions and Social Pressure,* 193 SCI. AM. 31, 32-33 (1955) (discussing social conformity to the opinions of others).

[236] *See generally* Craig Haney, Curtis Banks & Philip Zimbardo, *Interpersonal Dynamics in A Simulated Prison*, 1 INT'L J. CRIMINOLOGY & PENOLOGY 69 (1973); Phil Banyard, *Tyranny and the Tyrant: From Stanford to Abu Ghraib*, 20 THE PSYCHOLOGIST 494, 494 (2007) (discussing film footage of the experiment that shows Zimbardo behaving as an authority figure, demanding that the guards restrict the prisoners' freedom, a demand to which they conformed).

[237] SCOTT STRAUS, MAKING AND UNMAKING NATIONS: WAR, LEADERSHIP, AND GENOCIDE IN MODERN AFRICA 92, 99 (2015).

authority of the speaker, the more likely a community will conform to the individuals verbal instructions to commit violent acts.[238]

2. The speaker is perceived by supporters as credible or charismatic.

Whereas authority is associated with an individual's formal position in a social hierarchy, credibility is a communicative construct that derives from a leader's personal qualities such as expertise, relatability, and trustworthiness. To be effective, a credible group leader must also balance multiple commitments while signaling dedication to his or her base,[239] usually through making politically appropriate promises and delivering on them, while also not benefiting other groups more.[240] This raises the question of how leaders maintain credibility while leading their group into conflict. Leaders in conflict situations or societies with a history of intractable conflicts often frame precarious circumstances in culturally salient ways that resonate with their followers.[241] As an illustration, Benford and Snow's review of conflicts and social movements revealed that leaders who sustain their credibility are often those who speak with a "narrative fidelity," such that their words reinforce their group's ideologies and myths.[242]

While credibility promotes trust, a leader's charisma provides him or her with power, and even the perception of having spiritual, if not magical-like

---

[238] SARAH SORIAL, SEDITION AND THE ADVOCACY OF VIOLENCE: FREE SPEECH AND COUNTER-TERRORISM 2, 84-85, 115-16 (2012).

[239] *See generally* Richard J.B. Bosworth & Joseph A. Maiolo, *Introduction to Volume II*, and *Introduction to Part I*, *in* 2 THE CAMBRIDGE HISTORY OF THE SECOND WORLD WAR 1-20, (Michael Geyer & Adam Tooze eds., 2015) (addressing the role of leaders and their signaled devotion to constituents when discussing the role of ideological commitments in the World War II); Roseanne W. McManus, *Making it Personal: The Role of Leader-Specific Signals in Extended Deterrence,* 80 J. POL. 982, 982-95 (2018) (discussing the balance leaders must strike between signaling commitments to their own country and weaker states).

[240] *See, e.g.*, James H. Read & Ian Shapiro, *Transforming Power Relationships: Leadership, Risk, and Hope*, 108 AM. POL. SCI. REV. 40, 45-47 (2014) (using the case study of Mandela and de Klerk to illustrate the importance of their credibility and risk-taking in leading their constituencies to a peaceful solution in South Africa).

[241] *See* Stuart J. Kaufman, *Symbols, Frames, and Violence: Studying Ethnic War in the Philippines*, 55 INT'L. STUD. Q. 937, 943-46 (2011); *see also* Daniel Bar-Tal, Neta Oren, & Rafi Nets-Zehngut, *Sociopsychological Analysis of Conflict-Supporting Narratives: A General Framework*, 51 J. PEACE RES. 662-75 (2014) (analyzing the recurrence of conflict-supporting narratives that draw upon culturally accepted stories or frames to justify intergroup conflicts and thereby satisfy the basic psychological needs of the collective).

[242] *See* Robert D. Benford & David A. Snow, *Framing Processes and Social Movements: An Overview and Assessment*, ANN. REV. SOC. 611, 622 (2000).

abilities, by his or her followers.[243] Even when lacking expertise, communicators who have charisma—perceived authenticity, honesty, and physical presence—can attain positions of authority and become quite persuasive, especially during periods of crisis or rapid social change.[244] As Weber observed, charismatic leaders often attract followers full of "enthusiasm, or of despair and hope."[245] Since charismatic leaders often contest traditional or bureaucratic authorities, they can typically break with established norms.[246] Despite breaking traditional norms, charismatic leaders are often highly successful and persuasive because they are masters of nonverbal and verbal communication. For instance, they often modulate radically the volume and tempo of their speeches.[247] Additionally, charismatic leaders frame their goals in terms of the success of the ingroup,[248] using "us" and "them" rhetoric and stressing the vital importance of giving oneself to the group. [249] The latter is known as "ego surrender" and is a common appeal made by charismatic leaders of sociopolitical movements, as well as religious communities and cults, which encourage followers to find purpose and meaning in surrendering to the group and, by extension, the leader.

3. The speaker has regular access to means of mass communication, or the ability to control information, or to suppress alternative sources of information.

---

[243] *See generally* S. Abbruzzese, *Charisma: Social Aspects of*, *in* 8 INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL & BEHAVIORAL SCIENCES 1653 (Paul B. Baltes & Neil J. Smelser eds., 2001); Caitlin Andrews-Lee, *The Revival of Charisma: Experimental Evidence from Argentina and Venezuela*, 52 COMP. POL. STUD. 688-95 (2019); Yasmeen Yousif Pardesi & Yousif Pardesi, *Charismatic Leadership: A Critical Analysis*, 2 GOV. ANN. RES. J. POL. SCI. 71-76 (2013).

[244] MAX WEBER, ON CHARISMA AND INSTITUTION BUILDING 49, xiv-lvi, 18-27, 254-57 (1968).

[245] *Id* at 49.

[246] *Id.* at 24.

[247] Oliver Niebuhr, Jana Voße & Alexander Brem, *What Makes a Charismatic Speaker? A Computer-Based Acoustic-Prosodic Analysis of Steve Jobs' Tone of Voice,* 64 COMPUTERS HUM. BEHAV. 366, 376-78 (2016).

[248] *See generally* Saul L. Miller, Jon K. Maner, & D. Vaughn Becker, *Self-Protective Biases in Group Categorization: Threat Cues Shape the Psychological Boundary Between "Us" and "Them,"* 99 J. PERSONALITY SOC. PSYCHOL. 62-77 (2010) (addressing the vulnerability of groups to engage in bias toward others from outgroup categorization); Henri Tajfel, *Social Identity and Intergroup Behavior,* 13 SOC. SCI. INFO. 65 (1974) (providing a general definition of "in-group" for the social sciences as a group of people who see themselves as sharing an identity and interest, and generally a sense of otherness or exclusivity as regards all nonmembers of the group. An out-group is any group of people who are seen as not belonging to the in-group of the speaker and his or her primary audience).

[249] *See* Loretta S. Wilson & Susan Kwileck, *Are These People Crazy, or What? A Rational Choice Interpretation of Cults and Charisma*, 19 HUMANOMICS 29, 30 (2003).

Since Arendt's *The Origins of Totalitarianism*,[250] scholars have advanced the claim that authoritarian leaders seek to control information by hindering alternative political voices, while totalitarians strive to exercise personal and absolute control over information.[251] A leader may suppress alternative sources of information indirectly by placing party loyalists in official positions overseeing mass communication or directly by attacking the news media, including encouraging attacks on dissenting journalists.[252] For Maynard and Benesch, potentially dangerous leaders are those who attempt to control or erode systems of alternative information, since this practice correlates with violent regimes and is a necessary (but not sufficient) condition for genocide.[253] For instance, in a survey of over one hundred states from 1960 to 1999, Smith found that authoritarian repression, including stifling alternative sources of information, correlates with political violence unless a regime has a durable natural resource such as oil.[254]

Most countries have moral and legal injunctions against interpersonal violence, [255] but if citizens hear repeated exhortations from their leaders to commit violence against an outgroup, they may conclude that the usual strictures have been lifted. Moreover, persons may weigh the consequences of their involvement and come to believe that such violence is justified, necessary, and likely to go unpunished. Repetitive antisemitic exhortations in multiple domains

---

[250] *See generally* HANNAH ARENDT, THE ORIGINS OF TOTALITARIANISM (Harcourt 1973) (explaining how the roots of totalitarianism lie in a strong leader's charisma, ideology, and ability to sweep a people under the total arm of a regime by promising a utopia, while authoritarianism centers on total power).

[251] *See, e.g.*, JUAN J. LINZ, TOTALITARIAN AND AUTHORITARIAN REGIMES 67-70 (2000) (contrasting the methods by which authoritarian and totalitarian regimes seek to control their population).

[252] NATASHA EZROW & ERICA FRANTZ, DICTATORS AND DICTATORSHIPS: UNDERSTANDING AUTHORITARIAN REGIMES AND THEIR LEADERS 9-10, 192-93, 223 (2011) (addressing how the structure and effects of political regimes leaning towards or embracing dictatorship share the characteristic of elites mutually profiting from abuse and corruption such as controling or monopolizing the media and military).

[253] Jonathan Leader Maynard & Susan Benesch, *Dangerous Speech and Dangerous Ideology: An Integrated Model for Monitoring and Prevention,* 9 GENOCIDE STUD. & PREVENTION 70, 83 (2016); *see also* Susan Benesch, *The New Law of Incitement to Genocide: A Critique and a Proposal,* DANGEROUS SPEECH PROJECT 9-10 (Apr. 1, 2009), https://dangerouspeech.org/new-law-of-incitement-to-genocide/ [https://perma.cc/GZ9B-TX9S] (laying out Benesch's preconditions for incitement to genocide).

[254] *See* Benjamin Smith, *Oil Wealth and Regime Survival in the Developing World, 1960–1999*, 48 AM. J. POL. SCI. 232, 234, 241 (2004) ("Highly authoritarian regimes actually experienced considerably higher levels of protest than did others… [but] mechanisms other than repression drive the relative respite from protest that oil-rich states enjoy.").

[255] By *moral injunction* we mean a societal imperative informed by a sacred text or accepted standard of conduct that resembles an ethical or social normative imperative, while by *legal injunction* we mean criminal or civil rules that compel or prohibit behavior.

of media contributed to the Nazification of Europe in the 1930s.[256] Joseph Goebbels, the Nazi Minister of Propaganda, and Walther Schulze-Wechsungen, another prominent Nazi propagandist, effectively promoted the repetition of Nazi ideology as the key to propagandizing the masses.[257] Indeed, mass violence against a recognizable civilian population is often preceded by a sustained media campaign that repeats messages of hatred and/or violence.[258] Exposure to repeated messages often creates the illusion that the message is, in fact, true.[259] Exposure to repeated messages also decreases the likelihood in which listeners are able to discern whether the speaker or media source is disseminating false information.[260]

A free media is vital not only for democratic processes but also for sustained peace. The less control a leader has over the means of communication, the more likely the public will discover and respond to the leader's injustices and thereby remove him or her from office.[261] More critically, the more a leader controls the means of communication, the more it strengthens his or her regime [262] and, during times of state weakness, the greater the likelihood of armed conflict.[263] In these moments, leaders who control information and have a ready outlet for their own messages often

---

[256] Nico Voigtländer & Hans-Joachim Voth, *Nazi Indoctrination and Anti-Semitic Beliefs in Germany*, 112 PROC. NAT'L ACAD. SCI. 7931, 7931 (2015).

[257] *See* NICHOLAS O'SHAUGHNESSY, SELLING HITLER: PROPAGANDA AND THE NAZI BRAND 102-08 (2016) (discussing how part of the Nazi theory of mass persuasion involved repetition, as valued by its highest propagandists, including Goebbels and Schulze-Wechsungen).

[258] *See, e.g.*, GARTH S. JOWETT & VICTORIA O'DONNELL, PROPAGANDA AND PERSUASION 46 (5th ed. 2012) (describing how political leaders incited inter-ethnic hatred in the 1991–1995 Balkans conflict).

[259] *See generally* Lisa K. Fazio et al., *Knowledge Does Not Protect Against Illusory Truth*, 144 J. EXPERIMENTAL PSYCHOL. 993 (2015) (discussing the effects that repeated statements have on the illusion of truth).

[260] *See generally* Jason D. Ozubko & Jonathan Fugelsang, *Remembering Makes Evidence Compelling: Retrieval from Memory Can Give Rise to the Illusion of Truth*, 37 J. EXPERIMENTAL PSYCHOL. LEARNING, MEMORY & COGNITION 270 (2011) (discussing how repeated messages create an illusion of truth due to the fluency in retrieved information from memory).

[261] *See generally* Joan E. Cho, Jae Seung Lee & B.K. Song, *Media Exposure and Regime Support Under Competitive Authoritarianism: Evidence from South Korea*, 17 J.E. ASIAN STUD. 145 (2017) (discussing media control and authoritarian rule in South Korea under Park Chung Hee).

[262] *Id.* at 148-49.

[263] *See* Edward D. Mansfield & Jack Snyder, *Democratic Transitions, Institutional Strength, and War,* 56 INT'L ORG. 297, 298-99 (2002) (discussing the unique opportunities for armed conflict during the democratization process).

distract the public from their misdeeds, such as appropriating resources for themselves or dividing the public against any political opposition.[264]

B. The Content of the Message

1. The speaker's message contains explicit or implicit calls for violent acts against members of an outgroup.

Yanagizawa-Drott provides compelling evidence that explicit calls for violence on RTLM radio had a significant impact on the Rwandan genocide, concluding that: "the main radio station broadcasting anti-Tutsi propaganda during the Rwandan genocide significantly increased participation … [by] approximately 10% overall."[265] However, advocates of violence seldom use direct language, and coded or euphemistic speech is more the norm. Jean Paul Akayesu, for instance, incited mass violence against Tutsis, who he called *inyenzi* ("cockroaches"),[266] using the expression "go to work," which meant "go kill the Tutsis and Hutu political opponents."[267] Coded or implicit speech relies upon the linguistic and cultural competency of members of the ingroup, including symbolic and nonlinguistic background information, allowing only cultural insiders to infer what the speaker intends.[268]

For example, an international tribunal found in 2018 that Serb nationalist politician Vojislav Šešelj instigated persecution (including forcible displacement and deportation) of Croatians in the Yugoslav Wars using coded speech particular to Serb nationalism that contained ethnic and religious slurs and asserted an imaginary boundary of Serbian territory (viz. "The Karlobag-Karlovac-Ogulin-Virovitica Line").[269] When

---

[264] *See* JACK SNYDER, FROM VOTING TO VIOLENCE: DEMOCRATIZATION AND NATIONALIST CONFLICT 56 (2000) (describing the conditions within democratizing societies that foster the acceptance of nationalist ideas).

[265] David Yanagizawa-Drott, *Propaganda and Conflict: Evidence from the Rwandan Genocide*, 129 Q.J. ECON. 1947, 1989 (2014).

[266] Prosecutor v. Ruggiu, ICTR-97-32-I, Judgment and Sentence, ¶ 44(iii) (June 1, 2000).

[267] *Id.* at ¶ 44(iv).

[268] John Searle, *Indirect Speech Acts*, *in* 3 SYNTAX & SEMANTICS 60-61 (Peter Cole & Jerry L. Morgan eds., 1975).

[269] Prosecutor v. Vojislav Šešelj, MICT -16-99-A, Appeals Proceeding, ¶129-30, 166 (April 11, 2018); *see also* Anthony Oberschall, *Propaganda, Hate Speech and Mass Killings*, *in* PROPAGANDA, WAR CRIMES TRIALS AND INTERNATIONAL LAW: FROM SPEAKERS' CORNERS TO WAR CRIMES 171-200 (Predrag Dojčinović ed., 2012) (analyzing propaganda techniques for inciting mass violence and documenting how Šešelj's propaganda likely incited mass violence

compared to direct speech, indirect speech may have an equally powerful, but more subtle, effect on audiences. In a series of experiments, Eerland and colleagues observed that participants vividly remembered direct speech after exposure, while indirect speech was not easily recalled but participants still showed signs of having encoded the speech's message into memory.[270] Similarly, Gubler and Kalmoe found that citizens of Israel and India exposed to indirect calls for violence showed a significant increase in support for policies harming the outgroup.[271] Remarkably, these effects even influenced persons with low outgroup prejudice and low aggressive personality traits, and persisted even when persons were presented with speeches in favor of helping the outgroup.[272] Accordingly, indirect language may be just as effective if not more effective than direct language in promoting prejudice or harm against an outgroup.

2. The message dehumanizes an outgroup, or expresses disgust for an outgroup, or calls for acts of revenge against an outgroup.

In law, dehumanizing language is often seen as the most egregious form of hate speech, a view with origins in the Nuremberg trials, and especially the verdict against Nazi propagandist Julius Streicher who "termed the Jew a germ and a pest, not a human being, but 'a parasite, an enemy, an evil-doer, a disseminator of diseases who must be destroyed in the interest of mankind.'"[273] Dower, writing about the conduct of war in the Pacific during the Second World War, famously opined: "The dehumanization of the Other contributed immeasurably to the psychological distancing that facilitates killing."[274] Waldron places a strong emphasis on language that characterizes members of other racial groups as "bestial or subhuman."[275] Further, psychological studies have

---

in the Yugoslav Wars). *See generally* ANTHONY OBERSCHALL, VOJISLAV ŠEŠELJ'S NATIONALIST PROPAGANDA: CONTENTS, TECHNIQUES, AIMS AND IMPACTS, 1990–1994. (2006), http://www.baginst.org/uploads/1/0/4/8/10486668/vojislav_seseljs_nationalist_propaganda-_contents_techniques_aims_and_impacts.pdf [https://perma.cc/W2WS-PW5U] (providing an overview of techniques of mass persuasion used by history's most notorious propagandists to induce the acceptance of or participation in mass violence, and discussing how Šešelj's propaganda parallels mass persuasion intended to incite mass violence).

[270] Anita Eerland, Jan A. Engelen & Rolf A. Zwaan, *The Influence of Direct and Indirect Speech on Mental Representations*, 8 PLOS ONE, June 12, 2013, at 1, 6, 8.

[271] Joshua R. Gubler & Nathan P. Kalmoe, *Violent Rhetoric in Protracted Group Conflicts: Experimental Evidence from Israel and India*, 68 POL. RES. Q. 651, 651-52 (2015).

[272] *Id.* at 661.

[273] BETH VAN SCHAACK & RONALD C. SLYE, INTERNATIONAL CRIMINAL LAW AND ITS ENFORCEMENT: CASES AND MATERIALS 883 (Robert C. Clark et al. eds., 2d ed. 2010).

[274] JOHN W. DOWER, WAR WITHOUT MERCY: RACE AND POWER IN THE PACIFIC WAR 11 (1986).

[275] JEREMY WALDRON, THE HARM IN HATE SPEECH 66 (2012).

shown that dehumanizing language predicts intergroup violence.[276] This view of the deleterious effects of dehumanizing language is the accepted and arguably standard model of hate speech. U.S. courts convicting individuals of hate crimes have emphasized the dehumanizing language of defendants, and in the 2018 case of the three Kansas men convicted of a plot to bomb Somali immigrants, the court heard how one defendant referred to Muslims as "cockroaches" and to himself as the "Orkin man," referencing a pest extermination company.[277]

Disgust is a central element of moral reasoning and regulating social behavior generally,[278] and can be mobilized by certain speakers to dehumanize the outgroup. Higher order cognition, such as the executive functions in normal human interactions with other persons, exhibit low levels of disgust, as opposed to human interactions with polluting substances, such as human waste and corpses, which are inanimate and typically elicit high levels of disgust.[279] When feelings of disgust are elicited toward other people, ordinary projection of an imagined other who possesses a mind and human feeling is reduced. The result is a dismissal of the agency, intentionality, and subjectivity of the target of disgust.[280] According to Harris and Fiske, when persons project low warmth and incompetence onto a stigmatized group, such as the disabled, the poor, drug addicts and immigrants, they also experience high levels of disgust, thus indicating that perceptions of low warmth and incompetency go hand-in-hand with disgust and dehumanization.[281] Therefore, speech acts that portray an outgroup with disgust-inducing stimuli, such as pests or diseases, can stimulate feelings of social disengagement toward the outgroup.

---

[276] *See generally* Lasana T. Harris & Susan T. Fiske, *Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups,* 17 PSYCHOL. SCI. 847 (2006) (discussing the propensity of dehumanization to lower empathy for an outgroup); Tage S. Rai, Piercarlo Valdesolo & Jesse Graham, *Dehumanization Increases Instrumental Violence, but Not Moral Violence*, 114 PROC. NAT'L ACAD. SCI. 8511 (2017) (analyzing experimental effects of dehumanization and finding that dehumanizing language enables violence that perpetrators see as wrong but necessary).

[277] *See* Mitch Smith, *Kansas Trio Convicted in Plot to Bomb Somali Immigrants,* N.Y. TIMES (Apr. 18, 2018), https://www.nytimes.com/2018/04/18/us/kansas-militia-somali-trial-verdict.html [https://perma.cc/FPU7-FB2C] (reporting on a jury trial in which the government argued unsuccessfully that the First Amendment protected defendants' dehumanizing speech).

[278] *See* HAIDT, *supra* note 197, at 21-26 (describing a cross-cultural study in which participants found "harmless-taboo violations [are] universally wrong" even if no one is harmed by the violations).

[279] *See* Harris & Fiske, *supra* note 275, at 852 (finding lower medial prefrontal cortex activity and higher levels of disgust when participants were shown pictures of extreme outgroups that were consistent with the way they viewed objects).

[280] *Id.* at 848.

[281] *Id.* at 852.

Revenge is a powerful psychological motivation for intra- and intergroup conflict because it has been naturally selected to deter or discourage future conspecific threats, to preclude transgressions, and to reinforce social cooperation.[282] These functions of revenge are evident in studies of punishment and cooperation in game theoretical simulations.[283] Studies also show that ingroup members experience a vicarious sense of moral justification when retribution is carried out on members of an outgroup for assault or provocation.[284] In an experiment using Serb nationalist Vojislav Šešelj's speeches, Lillie et al. discovered that revenge speech lowered the propensity of participants to empathize with the outgroup and increased their willingness to justify violence morally.[285] Moreover, revenge speech consolidated ingroup identity to the same extent as highly nationalistic rhetoric.[286] Recently, Böhm, Rusch, and Gürerk found that the most direct and causal motivation for revenge was the felt need to protect the group, so much so that groups often opt for preemptive force when they believe that they are victims.[287] Such findings are supported by postconflict ethnographies, such as Mamdani's, which found that deep-seated notions of victimhood and revenge motivated many Hutu genocidaires.[288]

3. The message identifies a direct threat to the ingroup and identifies a clear and foreseeably violent course of action that can be imminently taken by listeners to remove the source of the threat.

For political entrepreneurs, a time-honored strategy for mobilizing supporters is manufacturing indignation about a problem and mobilizing collective action to address it. Yet, unlike ordinary political entrepreneurs, authority figures who incite hatred engage in what George calls "contentious collective action," in which a speaker convinces followers that they are

---

[282] *See* MICHAEL E. MCCULLOUGH, BEYOND REVENGE: THE EVOLUTION OF THE FORGIVENESS INSTINCT 49 (2008) (discussing revenge as an evolutionary adaptation).
[283] *Id.* at 99-103.
[284] Brian Lickel et al., *Vicarious Retribution: The Role of Collective Blame in Intergroup Aggression*, 10 PERSONALITY SOC. PSYCHOL. REV. 372, 386 (2006).
[285] Jordan Kiper et al., Propaganda, Empathy and Support for Intergroup Violence: the Moral Psychology of International Speech Crimes 18-19 (Mar. 20, 2019) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2580521 [https://perma.cc/S2A9-Z2Y8].
[286] *Id.*
[287] Robert Böhm, Hannes Rusch & Özgür Gürerk, *What Makes People Go to War? Defensive Intentions Motivate Retaliatory and Preemptive Intergroup Aggression*, 37 EVOLUTION & HUM. BEHAV. 29, 32-33 (2016).
[288] *See* MAHMOOD MAMDANI, WHEN VICTIMS BECOME KILLERS: COLONIALISM, NATIVISM, AND THE GENOCIDE IN RWANDA 32-33, 59, 259 (2001) (describing the history and formation of Hutu and Tutsi as political identities that laid the foundation for the Rwandan genocide).

wronged by a group and justified in taking action against them.[289] Furthermore, the speaker cultivates feelings of righteous indignation around a master narrative, which often draws from the ingroup's religion or worldview, making contentious collective action appear morally right.[290]

While mob violence and spurious attacks could ensue from such contentious speech, it is not incitement unless the speaker openly advocates for the unlawful oppression of, or an explicit attack on, the vilified group.[291] Rarely, though, is violence "sparked" by a single inciting speech, but rather stems from a combination of factors, including a speech event that draws from the narratives already cultivated by the speaker within his or her contentious movement.[292] A relevant example here is terrorist speech: the messages used by a hate group or terrorist organization when recruiting, training, planning, financing, and coordinating its activities.[293] At the core of terrorist speech is violent action, framed within a central narrative—an ideology that justifies violence on account of group threats or "fighting fire with fire" to achieve a political goal[294]—as advocated by the cell's political entrepreneurs. Thus, when a leader calls for violence, would-be perpetrators are ready to act.[295]

A study by Halperin, Canetti-Nisim and Hirsch-Hoefler found that the combination of group-based hatred with a sense of direct threat is one of the most significant antecedents for political intolerance and violence in Israel.[296] More broadly, a meta-analysis of psychological studies on ingroup threats

---

[289] *See* Cherian George, *Hate Spin: The Twin Political Strategies of Religious Incitement and Offense-Taking*, 27 COMM. THEORY 156, 157-58 (2017) (developing the concept of "hate spin," defined as "a twin political strategy of incitement and manufactured indignation, exploiting group identities to mobilize supporters and coerce opponents.").

[290] *Id.*

[291] *Id.* at 159.

[292] *See id.* at 163 (arguing that the "spark" frame of hate speech is counterproductive to understanding and combatting religious intolerance); *see also* SIDNEY TARROW, POWER IN MOVEMENT: SOCIAL MOVEMENTS AND CONTENTIOUS POLITICS 21 (2d ed. 1998) ("By drawing on inherited collective identities and shaping new ones, challengers delimit the boundaries of their respective constituencies and define their enemies. . . .").

[293] *See* Steven Beale, *Online Terrorist Speech, Direct Government Regulation, and the Communications Decency Act*, 16 DUKE L. & TECH. REV., 333-34 (2018) (proposing modifications to the Communications Decency Act to better regulate terrorist hate speech on social media and online).

[294] *See* Kurt Braddock & John Horgan, *Towards a Guide for Constructing and Disseminating Counternarratives to Reduce Support for Terrorism*, 39 STUD. CONFLICT & TERRORISM 381, 383 (2015) (describing the narrative form of terrorist messages).

[295] *See* Stephen Reicher & Alexander Haslam, *Fueling Extremes*, 27 SCI. AM. MIND 34, 39 (2016) (discussing psychological behaviors in terrorist cells).

[296] Eran Halperin, Daphna Canetti-Nisim & Sivan Hirsch-Hoefler, *The Central Role of Group-Based Hatred as an Emotional Antecedent of Political Intolerance: Evidence from Israel*, 30 POL. PSYCHOL. 93, 115-16 (2009).

and outgroup attitudes discovered that negative outgroup attitudes, such as hatred and distrust, were significantly influenced by realistic or symbolic threats posed by an outgroup.[297] Such threats increase fears about an outgroup and support for aggressive policies toward them. A study by Weisel and Zultan, for example, found that a group under threat is more likely to contribute materially to a conflict when called upon by others to do so.[298] While threatening messages elicit fear, it is worth noting that fear is itself a complex and unpredictable response. Even if a person is genuinely frightened by a message, a shift in attitude alone is often not enough to create a behavioral change, since an affective and cognitive component is necessary for behavior.[299] Consequentially, fear messages are most likely to enact behavioral changes when there is an identifiable path of action for the listener to take in order to remove the threat and, thus, the source of fear.[300]

C. The Context of the Speech

1. There is a history of intergroup conflict between the ingroup and outgroup, and the number of instances of intergroup violence has increased overall in the previous twelve months.

A history of conflict between the ingroup and outgroup often precedes genocide and crimes against humanity and is likely to contribute to mass atrocities because of unresolved grievances, dysfunctional justice systems, land disputes, and competition over natural resources.[301] Mamdani's study of the role of victimhood in the Rwandan genocide[302] has convinced scholars to focus on the combined effect of a history of conflict and a collective memory of victimhood as factors that indicate a susceptibility to incitement to

---

[297] Blake M. Riek, Eric W. Mania & Samuel L. Gaertner, *Intergroup Threat and Outgroup Attitudes: A Meta-Analytic Review,* 10 PERSONALITY & SOC. PSYCHOL. REV. 336, 349 (2006).
[298] Ori Weisel & Ro'i Zultan, *Social Motives in Intergroup Conflict: Group Identity and Perceived Target of Threat*, 90 EUR. ECON. REV. 122, 130 (2016).
[299] *See* Susan T. Fiske & Cydney Dupree, *Gaining Trust as Well as Respect in Communicating to Motivated Audiences about Science Topics*, 111 PROC. NAT'L ACAD. SCI. 13593, 13594 (2014) (evaluating perceptions of climate scientists as a potential barrier to public trust in their message).
[300] RICHARD M. PERLOFF, THE DYNAMICS OF PERSUASION: COMMUNICATION AND ATTITUDES IN THE 21ST CENTURY 200-01 (4th ed. 2010) (discussing the behavioral changes caused by fear in circumstances where "people perceive that they are capable of averting the threat").
[301] Leader Maynard & Benesch, *supra* note 252*,* at 78.
[302] MAMDANI, *supra* note 287.

genocide.[303] A sense of collective victimhood is associated with increased feelings of vulnerability and mistrust, an expectation of hostility in the future and a fear of physical or symbolic annihilation; and it is also associated with reduced guilt and willingness to forgive the outgroup.[304] Accordingly, groups with a history of conflict and an ingroup having a collective memory of victimization are more likely than others to support aggression towards the outgroup.[305]

When small-scale intergroup conflict increases, violence becomes more likely because, in general, there is a decrease in the quality of intergroup contact, and the ingroup is less likely to oppose leaders who wish to place restrictions on the outgroup in the name of self-defense.[306] For example, Pettigrew and Tropp's meta-analysis of over 500 studies on group interactions found that increased prejudice and violence toward an outgroup strongly correlated with diminished intergroup contact.[307] Given that frequent intergroup contact reduces the likelihood of mass violence and increased conflicts or perceived threats diminish intergroup contact, an increase in small-scale incidents of violence can render mass violence more imminent by reducing the mediating effects of intergroup conflict.[308]

A widely accepted finding on retaliatory violence is that persons who are identity-fused with their group—i.e., see their personal and group identity as equivalent—are more likely than others to support or volunteer for violence against a seemingly perpetrating outgroup.[309] For instance, studies

---

[303] *See generally* Noa Schori-Eyal et al., *The Shadows of the Past: Effects of Historical Group Trauma on Current Intergroup Conflicts*, 43 PERSONALITY & SOC. PSYCHOL. BULL. 538 (2017) (discussing associations between collective memories of trauma, victimhood, and enduring group violence).

[304] *Id.* at 538.

[305] *Id.* at 540 (arguing that perpetual ingroup victim orientation "entails a commitment to the defense of the ingroup, and consequently greater support for aggressive measures against enemy outgroups").

[306] Justin T. Pickett et al., *Contact and Compromise: Explaining Support for Conciliatory Measures in the Context of Violent Intergroup Conflict*, 51 J. RES. CRIME & DELINQ. 585, 588, 591, 604-05 (2014) (describing how threat-oriented beliefs in a study of Israeli Arabs and Jews were associated with decreased support for compromise whereas high quality contact between those groups was associated with increased support for compromise).

[307] Thomas F. Pettigrew & Linda R. Tropp, *A Meta-Analytic Test of Intergroup Contact Theory,* 90 J. PERSONALITY & SOC. PSYCHOL. 751, 766 (2006).

[308] Ananthi Al Ramiah & Miles Hewstone, *Intergroup Contact as a Tool for Reducing, Resolving, and Preventing Intergroup Conflict*, 68 AM. PSYCHOLOGIST 527, 531-32 (2013).

[309] Leah A. Fredman, Brock Bastian & William B. Swann, *God or Country? Fusion with Judaism Predicts Desire for Retaliation Following Palestinian Stabbing Intifada*, 8 SOC. PSYCHOL. & PERSONALITY SCI. 882, 882-83, 885 (2017) ("[I]dentity fusion has been implicated in conflicts in the Middle East. For example, strongly fused persons were especially inclined to serve as front-line combatants during the 2011 Libyan revolution.").

of Israelis during the 2015 Palestinian "Stabbing Intifada"[310] and Libyans during the 2011 revolution[311] found that persons who felt fused with their group predicted endorsement of, or participation in violence. Persons who experience an exclusive sense of victimhood after an attack on their group—i.e., that their group alone comprises true victims—are more likely to have strong negative intergroup attitudes.[312] Given human cognitive capacity for revenge,[313] retaliation in a nation-state or region can become especially dangerous, leading to a cycle of deepening group commitments and escalating intergroup conflict.[314] Such violence can lead to collective blame and vicarious retribution,[315] and in the context of a civil war, weaken institutions and ultimately challenge the control of the State.[316]

2. There is a major national political election in the next twelve months or there was a major national political election in the last twelve months.

For any country, an election year is an important period, but for countries where elites vie for political dominance amid a system of corrupt and fraudulent politics, election years can witness dramatic spikes in violence. The spike is linked to the activity of militias, who often align themselves with armed wings

---

[310] *Id.* at 885.

[311] Harvey Whitehouse, Brian McQuinn, Michael Buhrmester & William B. Swann, *Brothers in Arms: Libyan Revolutionaries Bond like Family*, 111 PROC. NAT'L ACAD. SCI. 17783, 17784 (2014).

[312] *See, e.g.*, Johanna Ray Vollhardt & Rezarta Bilali, *The Role of Inclusive and Exclusive Victim Consciousness in Predicting Intergroup Attitudes: Findings from Rwanda, Burundi, and DRC,* 36 POL. PSYCHOL. 489, 489-506 (2015) (examining the theories and evidence regarding victim consciousness and empirical support for the hypothesis that exclusive victimhood consciousness predicts negative intergroup attitudes, while inclusive victimhood predicts positive intergroup attitudes).

[313] *See generally* Michael E. McCullough, Robert Kurzban & Benjamin A. Tabak, *Cognitive Systems for Revenge and Forgiveness*, 36 BEHAV. & BRAIN SCI. 1 (2013) (drawing from findings in cognitive science to posit that humans have a cognitive system for retaliation, which was selected to deter harm, and a cognitive system for computing relational costs, which can function to deter costly conflict).

[314] *See, e.g.,* Robert S. Walker & Drew H. Bailey, *Body Counts in Lowland South American Violence*, 34 EVOLUTION & HUM. BEHAV. 29, 29-34 (2013) (analyzing violence in South American societies and finding that attacks become deadlier when revenge is involved, and that revenge raids become deadlier than the previous grievance, leading to cycles of violence).

[315] *See* Lickel et al., *supra* note 283 (providing a theoretical framework for the phenomenon of collective blame in which someone is attacked by virtue of their apparent shared identity with the original perpetrator).

[316] *See* Kristian Skrede Gleditsch, *Transnational Dimensions of Civil War*, 44 J. PEACE RES. 293, 304 (2007) (finding "that countries recently involved in conflict or new states are substantially more likely to experience conflict").

of political parties in the run-up to elections.[317] Under the guise of removing threatening regimes or protecting threats to the state, militias are often sought by politicians and loyalists to intimidate or attack political opponents, if not control entire regional populations.[318] Numerous studies reveal, however, that coordinated attempts to intimidate political opponents do not cause the targeted communities to retreat from political participation but to engage in it more vigorously.[319] Consequentially, intergroup tensions mount quickly, and violence can escalate as an instrument of political change.

When election violence is "bottom-up," it typically involves intermittent and minor clashes between groups of political oppositions, as witnessed in the U.S. 2016 presidential election race.[320] Election violence is most violent, however, when it is "top-down" or sponsored by ruling governments who attempt to suppress oppositional groups, reduce political competition, and diminish voter turnout, thereby maintaining or even extending their political power.[321] Although such top-down election violence often entails a victory for incumbents, their victory is usually temporary, as cycles of violence often follow with, for example, oppositional forces resorting to violence to discredit the government or incumbents further restricting peoples' rights.

To illustrate, a study by Hafner-Burton, Hyde, and Jablonski examined 458 elections from 1981 to 2004, finding that election violence increases the likelihood that incumbents remain in power, but also finding that their society is at a greater risk for longterm cycles of violence.[322] Post-election violence is often greater in magnitude than the initial pre-election violence;[323] leading to enduring struggles over land resources, ethicized party

---

[317] Sabine C. Carey & Neil J. Mitchell, *Progovernment Militias*, 20 ANN. REV. POL. SCI. 127, 139 (2017).
[318] *Id.* at 132.
[319] *See* Sandra Ley, *To Vote or Not to Vote: How Criminal Violence Shapes Electoral Participation*, 62 J. CONFLICT RESOL. 1963, 1963-64 (2018) (discussing studies that show victims of political violence are more likely than nonvictims to participate in politics).
[320] *See* Cynthia Akwei, *Mitigating Election Violence and Intimidation: A Political Stakeholder Engagement Approach*, 46 POL. & POL'Y. 472, 491–92 (2018) (noting that these incidents were not just the conference attendees creating negative feelings on their own, but that the political candidates helping to incite bad feelings were not ruling governments and still part of the "top-down" process).
[321] Emilie M. Hafner-Burton, Susan D. Hyde & Ryan S. Jablonski, *When Do Governments Resort to Election Violence?*, 44 BRIT. J. POL. SCI. 149, 150 (2013).
[322] Emilie M. Hafner-Burton, Susan D. Hyde, & Ryan S. Jablonski, *Surviving Elections: Election Violence, Incumbent Victory and Post-Election Repercussions*, 48 BRIT. J. POL. SCI. 459, 467, 482 (2018).
[323] *See, e.g.*, Henrik Angerbrandt, *Deadly Elections: Post-Election Violence in Nigeria*, 56 J. MOD. AFR. STUD. 143, 144 (2018) (describing how violence in reaction to election results, such as those in Nigeria in 2011, can be of a much greater magnitude).

formation, reduced voter turnout, declines in democracy and perceptions of insecurity;[324] and even to journalistic self-censorship.[325] Given the social volatility of elections, if a speaker incites violence during or soon after an election period, his or her words are likely to be effective.

3. There is significant polarization of political organizations along religious, ethnic, or racial lines.

Political polarization within a community or between communities is known to increase intolerance and violence.[326] When schisms occur along religious,[327] ethnic,[328] or racial lines,[329] communities can experience even greater escalations in violence, largely due to growing levels of mutual mistrust, fear, and hostility. Political polarization along these lines, in turn, can increase narratives of exclusive or competitive victimhood—i.e., "my group has suffered more than anyone else's"—and is linked to aggressive policies toward an outgroup.[330] As a result, polarization in the social and historical context of a community makes it more likely that an audience will be receptive to speakers who incite hatred or violence.[331]

Polarized communities are thus vulnerable to calls for seemingly justified collective violence. Inciting authority figures and politicians, for instance, often seek to amplify collective grievances and biases towards an outgroup, whose members are generally portrayed as being responsible to some degree for the ingroup's travails.[332] Communities are more sensitive to scapegoating an outgroup during periods of acute loss, displacement, or injustice in addition to heightened intergroup tensions over collective memories of social trauma or unresolved political injustices. For example, a

---

[324] *Id*. at 144-45.

[325] *See* Lisa Weighton & Patrick McCurdy, *The Ghost in the News Room: The Legacy of Kenya's 2007 Post-Election Violence and the Constraints on Journalists Covering Kenya's 2013 General Election*, 11 J.E. AFR. STUD. 649, 662 (2017).

[326] Debra H. Swanson, *"All We Are Saying is Give Peace a Chance": A Sociological Response to Violence and Political Polarization*, 50 SOC. FOCUS 291, 298 (2017).

[327] *See generally* Dov Waxman, *A Dangerous Divide: The Deterioration of Jewish-Palestinian Relations in Israel*, 66 MIDDLE E.J. 11 (2012).

[328] *See generally* DONALD L. HOROWITZ, ETHNIC GROUPS IN CONFLICT (1985).

[329] *See generally* Reeve D. Vanneman & Thomas F. Pettigrew, *Race and Relative Deprivation in the Urban United States,* 13 RACE 461 (1972).

[330] Luca Andrighetto, Silvia Mari, Chiara Volpato & Burim Behluli, *Reducing Competitive Victimhood in Kosovo: The Role of Extended Contact and Common Ingroup Identity*, 33 POL. PSYCHOL. 513, 513-14 (2012).

[331] Leader Maynard & Benesch, *supra* note 252, at 78.

[332] *See* Herbert Blumer, *Race Prejudice as a Sense of Group Position*, 1 PAC. SOCIOL. REV. 3, 6 (1958) (discussing the major influence authorities, social elites, and public figures have in shaping public discussions about group position and prejudice).

famous study by Hovland showed that antiblack violence (including lynching) in the southern United States between 1882 and 1930 significantly and robustly correlated with poor economic conditions of perpetrator communities.[333] Exploiting grievances, then, in a politically polarized environment along religious, ethnic or racial lines can easily stir up animosities toward an outgroup.

4. The emotional state of the audience at the time of the speech appears heightened and predisposed towards violent activity.

While outgroup aggression is often prompted by cognitive beliefs that one's ingroup is threatened by a targeted outgroup, it is usually emotionally expressed as righteous indignation, such that the ingroup feels justified in "self-defense" against the outgroup.[334] Remarkably, political ideologies that cultivate feelings of anger render groups more likely than others to experience righteous indignation, thus suggesting the important role of a speaker in cultivating anger prior to the onset of a threat.[335] Testing this hypothesis, Kuppens and colleagues undertook two studies in which groups were induced with anger, finding that anger indeed caused a spike in intergroup biases and hostilities—but only for outgroups that were portrayed as threatening.[336] Hence, anger seems to indicate that some persons are more susceptible than others to channeling their frustrations into righteous indignation and possibly violent actions toward a threatening outgroup.

D. Applying the Matrix: Sines v. Kessler

This Section illustrates how the matrix for conducting a risk analysis of inciting speech can be applied to a concrete and recent case involving incitement to violence presently being adjudicated in the courts, in order to evaluate whether the framework could have predicted the eventual outcome. The case is *Elizabeth Sines v. Jason Kessler*, a civil suit filed against the

---

[333] *See generally* Carl Iver Hovland & Robert R. Sears, *Minor Studies of Aggression: VI. Correlation of Lynchings with Economic Indices*, 9 J. PSYCHOL 301 (1940) (discussing the correlation between experienced economic deprivation in communities of white perpetrators and mass violence against African Americans).

[334] George, *supra* note 288, at 158-59.

[335] *See* Roni Porat, Eran Halperin & Maya Tamir, *What We Want Is What We Get: Group-Based Emotional Preferences and Conflict Resolution*, 110 J. PERSONALITY & SOC. PSYCHOL. 167, 168-69 (2016).

[336] Toon Kuppens, Thomas V. Pollet, Cátia Teixeira, Stéphanie Demoulin, S. Craig Roberts & Anthony C. Little, *Emotions in Context: Anger Causes Ethnic Bias but Not Gender Bias in Men but Not Women*, 42 EUR. J. SOC. PSYCHOL. 432, 439-40 (2012).

organizers of a white supremacist rally in Charlottesville, Virginia on August 11–12, 2017. The facts of the case are provided in the memorandum opinion issued on July 9, 2018 by District Judge Norman K. Moon allowing the suit to proceed, in which the court took all factual allegations in the complaint as true.[337]

The central questions considered here are, could law enforcement agencies have anticipated that the organizers' direct advocacy of violence on social media would lead to imminent lawless action, specifically to the injury of dozens of protestors and the death of one protestor, Heather Heyer? Could a systematic risk analysis conducted a priori have justified earlier intervention by the authorities, and therefore prevented the harms and fatality from occurring?

Each of the ten factors in the matrix are reviewed in the light of the available evidence, to determine whether they were present before the white nationalists' torchlit march on the evening of August 11, 2017. The evidence comes primarily from exchanges on social media and includes speech (for instance, dehumanizing language) that may in itself be protected under the First Amendment.

1. The speaker occupies an official position of authority within government or a political party or political movement.

Present: The organizers of the march are leaders in the white nationalist or white supremacist or neo-Nazi movement, holding official positions of authority within their respective organizations. Defendants include the leader of the Traditionalist Worker Party,[338] the two leaders of the League of the South,[339]leaders of the Nationalist Front, and the leader and "Commander" of the National Socialist Movement, a white supremacist organization with a paramilitary structure.[340]

2. The speaker is perceived by supporters as credible or charismatic.

Present: Defendant Richard Spencer has arguably the highest public profile of any white supremacist in America.[341] Spencer is a new breed of media-friendly white supremacist, holding degrees from the Universities of Virginia and Chicago, and attired in expensive three-piece suits, gold

---

[337] Sines v. Kessler, 324 F.Supp.3d 765, 774 (W.D. Va. 2018).
[338] *Id.* at 775.
[339] *Id*. at 775-76.
[340] *Id*. at 776, 793.
[341] John Woodrow Cox, '*Let's Party Like It's 1933': Inside the Alt-Right World of Richard Spencer,* WASH. POST (Nov. 22, 2016), https://www.washingtonpost.com/local/lets-party-like-its-1933-inside-the-disturbing-alt-right-world-of-richard-spencer/2016/11/22/cf81dc74-aff-11e6-840f-e3ebab6bcdd3_story.html [https://perma.cc/8JDU-2XQV].

cufflinks, and Swiss watches.[342] He has single-handedly redefined white supremacist politics in the current era, and coined the euphemistic expression "alt-right."[343] He holds unparalleled sway in the white nationalist movement, illustrated by his ability to bring together at Charlottesville a number of disparate and sometimes antagonistic white nationalist groups. Spencer has a devoted following who see him as a fearless and charismatic leader who they venerate with chants of "Hail Spencer! Hail victory!"[344] Defendant National Socialist Movement leader Jeff Shoep also commands the devotion of his followers, as expressed in online statements such as, "So much respect for my Commander Jeff Schoep. I will go into battle with you anytime Sir."[345]

3. The speaker has regular access to means of mass communication, or the ability to control information, or to suppress alternative sources of information.

Present: Several defendants manage prominent websites or podcasts that reach a wide audience of white nationalists. Defendants Andrew Anglin and Robert Ray run the Daily Stormer website, and defendant Michael Peinovich hosts a podcast called the Daily Shoah.[346] Defendants Kessler and Mosely coordinated the white nationalist march via a moderated social media platform called "Discord" that is widely used by white nationalists.[347] Much of what the Court knows about the defendants' conversations preceding the white nationalist rally comes from a record of the exchanges on Discord.[348] The Daily Stormer maintained a live feed of the events in Charlottesville, where viewers made comments such as "We have an army! This is the beginning of a war!"[349] The fact pattern indicates that the organizers had more than just a few social media accounts; they were publishers of content on America's leading white nationalist sites and had the ability to control information conveyed to their mass audience.[350]

4. The message contains explicit or implicit calls for violent acts against members of an outgroup.

---

[342] *Id.*
[343] *Id.*
[344] Sines v. Kessler, 324 F. Supp. 3d 765, 786 (W.D. Va. 2018)
[345] *Id.* at 793-94.
[346] *Id.* at 775.
[347] *Id.* at 776.
[348] *Id.*
[349] *Id.* at 789.
[350] *Id.* at 784-95.

Present: The conversation on Discord included planning details, racist "jokes," and both explicit and implicit calls for violence against counter-protestors. With respect to coded speech, we identified fake advertisements for a pepper spray that would kill African Americans "on contact"[351] and questions about whether it was legal to run over protestors blocking a roadway.[352] In response, one Discord user posted a picture of a bus running over protestors.[353]

Other statements on social media explicitly called for or countenanced violence, such as "I'm ready to crack skulls,"[354] references to marchers as "warriors,"[355] and exhortations to bring weapons and to wear "a good fighting uniform."[356] Instructional videos posted on Discord showed how to fight in military formation and use shields.[357] Members of Vanguard America discussed on social media which weapons they should bring, including firearms, batons and knives.[358] *Sines* concludes that such statements went beyond mere advocacy of violence and included specific instructions to carry out violent acts.[359]

Additionally, and more seriously, images of extreme violence and animus circulated on Discord in the run-up to the march, including a drawing of one defendant, Matthew Heimbach, with "kill tallies" of communists, the words "n- killer," and an image of decapitated black men.[360] The phrase "run them over" was popularized by websites such as the *Daily Caller*, on the mass-circulation service Fox Nation, and by the defendants on Discord.[361]

5. The message dehumanizes an outgroup, expresses disgust for an outgroup, or calls for acts of revenge against an outgroup.

Present: Posts on Discord used dehumanizing language to express extreme racial animus against African American and Jewish individuals,[362] with one Defendant telling a reporter that white nationalists "outnumbered the anti-white, anti-American filth."[363] Another Defendant called the counter-

---

[351] *Id.* at 777.
[352] *Id.* at 796.
[353] *Id*.
[354] *Id.* at 784.
[355] *Id.* at 776.
[356] *Id.* at 786, 803.
[357] *Id.* at 790.
[358] *Id.* at 788.
[359] *Id.* at 803.
[360] *Id.* at 791.
[361] *Id.* at 796.
[362] *Id.* at 780.
[363] *Id.* at 789.

protestors "savages."[364] Revenge speech was also present, although it came after the events and therefore does not contribute to the risk analysis. Defendant Kessler remarked on the intentional killing of protestor Heather Heyer, agreeing with the Loyal White Knights that she was a communist and adding that, "Communists have killed 94 million. Looks like it was payback time."[365] Another defendant stated after Heyer was killed, "[O]ur rivals are just a bunch of stupid animals."[366]

6. The message identifies a direct threat to the ingroup and identifies a clear and foreseeably violent course of action that can be taken by listeners imminently to remove the source of the threat.

Present: White nationalist and supremacist groups have consistently portrayed white Christians as under threat of replacement by Jews.[367] Defendant Eli Mosley of the Identity Evropa white supremacist organization led the marchers in the chant "Jews will not replace us."[368] Mosley declared that the defendants would not be replaced "without a fight."[369] Therefore, organizers of the march identified Jews as a threat to their livelihoods, and encouraged their followers to engage in a torchlit march in which they would fight, literally and figuratively, to remove the threat.

7. There is a history of intergroup conflict between the ingroup and outgroup, and the number of instances of intergroup violence has increased overall in the previous twelve months.

Present: As seen in the introduction to this Article, after a decade of decline, documented hate crimes increased sharply in 2016 and 2017, indicating a general climate of heightened conflict between white nationalists (ingroup) and nonwhites (outgroup) in the United States.[370] If we focus more specifically on the actions of white nationalist organizations, earlier in 2017, white nationalist groups clashed with protestors in Berkeley, California over the course of several months.[371] A number defendants at Charlottesville

---

[364] *Id.* at 794.
[365] *Id.* at 796.
[366] *Id.* at 787.
[367] *Id.* at 785.
[368] *Id.*
[369] *Id.* at 790.
[370] *See supra* notes 8-13 and accompanying text.
[371] Natasha Lennard, *The Violent Clashes in Berkeley Weren't 'Pro-Trump' Versus 'Anti-Trump'*, ESQUIRE (April 16, 2017), https://www.esquire.com/news-politics/news/a54564/the-violent-clashes-in-berkeley-werent-pro-trump-versus-anti-trump/ [https://perma.cc/EC7H MMAA].

participated in the Berkeley skirmishes and one of them, Identity Evropa founder Nathan Damigo, referred to Berkeley as a "test run" for Charlottesville.[372]

8. There is a major national political election in the next twelve months or there was a major national political election in the last twelve months.

Present: One of the most contentious national elections in the country's history took place only nine months prior to the August 2017 conflict. Defendant Matthew Heimbach, secretary of the white supremacist Traditionalist Workers' Party and an organizer of the Charlottesville march, maintained an active and violent presence at Trump presidential election rallies and pled guilty to disorderly conduct after assaulting protestors at a Trump rally in Louisville, Kentucky in 2016.[373]

9. There is significant polarization of political organizations along religious, ethnic, or racial lines.

Present: White nationalist and white supremacist organizations are by definition polarized along ethnic, racial or religious lines, according to the ideology of these groups, which either seek white dominance or a white ethno-nationalist state.[374] Defendant National Socialist Movement openly espouses a neo-Nazi ideology,[375] and defendant Richard Spencer advocates a "peaceful ethnic cleansing" that would remove all non-whites from American soil.[376]

10. The emotional state of the audience at the time of the message appears heightened and predisposed towards violent activity.

---

[372] *Sines*, 324 F. Supp. 3d at 790.

[373] Thomas Novelly & Matthew Glowicki, *White Nationalist Pleads Guilty to Disorderly Conduct at Trump Rally*, COURIER J. (Louisville) (July 19, 2017), https://www.courier-journal.com/story/news/crime/2017/07/19/white-nationalist-pleads-guilty-harassment-charge-trump-rally/492219001/ [https://perma.cc/TB4B-X6UP].

[374] *See* Amanda Taub, '*White Nationalism,' Explained,* N.Y. TIMES (Nov. 21, 2016), https://www.nytimes.com/2016/11/22/world/americas/white-nationalism-explained.html [https://perma.cc/RS7Z-PUGA] (describing how white nationalists seek the dominance of white people); *White Nationalist*, S. POVERTY L. CTR., https://www.splcenter.org/fighting-hate/extremist-files/ideology/white-nationalist [https://perma.cc/BL7V-3BMK] (last visited Jan. 28, 2019) (reviewing the ideology of white nationalist groups in America).

[375] *See Neo-Nazi*, S. POVERTY L. CTR., https://www.splcenter.org/fighting-hate/extremist-files/ideology/neo-nazi [https://perma.cc/HX7L-BS89] (reviewing the ideology of the National Socialist Movement).

[376] Taub, *supra* note 373.

Present: The chorus of approval of violence on social media before the march had identifiable effects on the emotional state of the white nationalist marchers. Defendant Cantwell said to a reporter before the march that he was "trying to make [himself] more capable of violence."[377] When asked by a reporter if he was armed, Cantwell produced two semi-automatic rifles, three handguns, and a knife.[378] The emotional state of white nationalist marchers appears to have been heightened in the pre-march planning phase, as they were instructed to enter Emancipation Park in military formations, some carrying pre-made banners with statements such as, "Gas the k---s, race war now!"[379] They shouted antisemitic slogans as they passed a synagogue.[380] At the torchlit march on August 11, white nationalists gave Nazi salutes, chanted "Blood and Soil," and made monkey sounds at the protestors.[381]

In summary, ten of the ten risk factors were present before the events in Charlottesville on August 11 and 12, 2017, indicating that imminent violence was highly likely. As we now know, the defendants' speech and other acts did culminate in actual violence. Dozens of assaults occurred on August 11, and one protestor, Heather Heyer, was killed on August 12, when defendant James Fields drove his car into a crowd of peaceful protestors.[382] *Sines* maintains that the statements on Discord indicate the defendants were involved in a conspiracy to commit racial violence[383] and that the violence was "reasonably foreseeable," citing as evidence the inciting speech on Discord and other white nationalist sites, the instructions to bring lethal weapons, and the post hoc approval of James Fields's killing of a protestor.[384]

CONCLUSION

Drawing on behavioral and humanistic research, this Article provides an urgently needed framework to determine whether lawless action is imminent and likely to result from inciting speech. This ten-factor matrix is of course open to revision as our knowledge improves. Skeptics may ask, "why should we rely on social science when it's always developing and changing and there never seems to be unqualified certainty?" The obvious reply is that legal procedure and doctrine is always developing and changing

---

[377] *Sines,* 324 F. Supp. 3d at 786.
[378] *Id.* at 796.
[379] *Id.* at 779.
[380] *Id*.
[381] *Id.* at 785.
[382] *Id.* at 778-79.
[383] *Id.* at 799.
[384] *Id.* at 795-96.

too, and may at times lack certainty, as our review of First Amendment law demonstrates. Evidently, both law and social research are moving targets. They transform as societal norms shift and as our knowledge improves. It behooves courts to ensure that they are adequately informed, which necessarily brings law and science into dialogue with each other.

Given that federal and state prosecutors exercise wide discretion in charging defendants for incitement, a more transparent risk assessment framework is not in itself sufficient to transform incitement law, although it is a necessary component of any reform. Here, it should be noted that legal change often occurs through strategic litigation brought by pressure groups. First Amendment law applies equally to criminal prosecution and to civil litigation, and in each setting, the question is the same: "is this speech constitutionally protected?" Now we have a framework for approaching that question in a way that relies less on hunches and heuristics than empirically based findings in historical, social science, and behavioral research.

APPENDIX: THE INCITEMENT MATRIX

1. The speaker occupies an official position of authority within government or a political party or political movement.
2. The speaker is perceived by supporters as credible or charismatic.
3. The speaker has regular access to means of mass communication, or the ability to control information, or to suppress alternative sources of information.
4. The message contains explicit or implicit calls for violent acts against members of an outgroup.
5. The message dehumanizes an outgroup, or expresses disgust for an outgroup, or calls for acts of revenge against an outgroup.
6. The message identifies a direct threat to the ingroup and identifies a clear and foreseeably violent course of action that can be taken by listeners imminently to remove the source of the threat.
7. There is a history of intergroup conflict between the ingroup and outgroup, and the number of instances of intergroup violence has increased overall in the previous twelve months.
8. There is a major national political election in the next twelve months or there was a major national political election in the last twelve months.
9. There is significant polarization of political parties along religious, ethnic, or racial lines.
10. The emotional state of the audience at the time of the message appears heightened and predisposed towards violent activity.