

2021

## Hate Speech on Social Media: Content Moderation in Context

Richard A. Wilson

*University of Connecticut Human Rights Institute, richard.wilson@uconn.edu*

Molly Land

*University of Connecticut School of Law*

Follow this and additional works at: [https://opencommons.uconn.edu/law\\_papers](https://opencommons.uconn.edu/law_papers)



Part of the [First Amendment Commons](#)

---

### Recommended Citation

Wilson, Richard A. and Land, Molly, "Hate Speech on Social Media: Content Moderation in Context" (2021).  
*Faculty Articles and Papers*. 535.

[https://opencommons.uconn.edu/law\\_papers/535](https://opencommons.uconn.edu/law_papers/535)

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### Hate Speech on Social Media: Content Moderation in Context

RICHARD ASHBY WILSON & MOLLY K. LAND

*For all practical purposes, the policy of social media companies to suppress hate speech on their platforms means that the longstanding debate in the United States about whether to limit hate speech in the public square has been resolved in favor of vigorous regulation. Nonetheless, revisiting these debates provides insights essential for developing more empirically-based and narrowly tailored policies regarding online hate.*

*First, a central issue in the hate speech debate is the extent to which hate speech contributes to violence. Those in favor of more robust regulation claim a connection to violence, while others dismiss these arguments as tenuous. The data generated by social media, however, now allow researchers to empirically test whether there are measurable harms resulting from hate speech. These data can assist in formulating evidence-based policies to address the most significant harms of hate speech, while avoiding overbroad regulation.*

*Second, reexamining the U.S. debate about hate speech also reveals the serious missteps of social media policies that prohibit hate speech without regard to context. The policies that social media companies have developed define hate speech solely with respect to the content of the message. As the early advocates of limits on hate speech made clear, the meaning, force, and consequences of speech acts are deeply contextual, and it is impossible to understand the harms of hate speech without reference to political realities and power asymmetries. Regulation that is abstracted from context will inevitably be overbroad.*

*This Article revisits these debates and considers how they map onto the platform law of content moderation, where emerging evidence indicates a correlation between hate speech online, virulent nationalism, and violence against minorities and activists. It concludes by advocating specific recommendations to bring greater consideration of context into the speech-regulation policies and procedures of social media companies.*

## ARTICLE CONTENTS

INTRODUCTION.....	1031
I. DEBATING HATE SPEECH IN THE UNITED STATES.....	1034
A.    THE CASE FOR RESTRICTING HATE SPEECH.....	1035
B.    THE FREEDOM OF EXPRESSION RIPOSTE.....	1037
II. THE HARMS OF ONLINE HATE SPEECH.....	1039
III. THE PLATFORM LAW OF HATE SPEECH.....	1045
A.    DEFINING HATE SPEECH ONLINE.....	1046
B.    THE LIMITS OF PLATFORM LAW.....	1053
IV. TOWARDS CONTEXT-SPECIFIC CONTENT MODERATION.....	1061
A.    THE ROLE OF CONTEXT.....	1061
B.    LAW WITHOUT CONTEXT.....	1063
C.    RECOMMENDATIONS.....	1069
CONCLUSION.....	1075



# Hate Speech on Social Media: Content Moderation in Context

RICHARD ASHBY WILSON & MOLLY K. LAND \*

## INTRODUCTION

Hate speech and hate crimes are trending. In the past five years, there has been an upsurge in extreme nationalist and nativist political ideology in mainstream politics globally. In the United States, the President regularly mobilizes a political constituency by vilifying Mexican immigrants as “criminals” and “rapists” who “infest” America,<sup>1</sup> and by promoting a “zero tolerance” policy at the border that punitively separates children from their parents, including persons exercising their right to apply for asylum.<sup>2</sup> Data suggest a connection between this rise in rhetoric to increases in hate crimes in the United States.<sup>3</sup> Similar trends are evident abroad as well. In the United Kingdom, the 2016 Brexit referendum elicited conspicuous expressions of anti-Muslim and anti-immigrant sentiment and coincided with the sharpest increase in religiously and racially motivated hate crimes ever recorded in British history.<sup>4</sup>

In the United States, there has been vigorous debate on the regulation of hate speech for decades, but the dominant legal frameworks for addressing the harms in hate speech were created in a world without the internet and urgently need updating. Historically, those advocating for the suppression of hate speech concentrated their efforts on measures that

---

\* Richard Ashby Wilson is the Gladstein Chair of Human Rights and Professor of Law and Anthropology at the University of Connecticut School of Law. Molly Land is the Catherine Roraback Professor of Law and Human Rights at the University of Connecticut School of Law and Associate Director of the Human Rights Institute. We extend our thanks to Nadine Strossen for her astute comments on our arguments, and to the editors of the *Connecticut Law Review* for their careful feedback and editing. We are grateful to Allaina Murphy and Danielle Nadeau for research assistance.

<sup>1</sup> Andrés Oppenheimer, *Immigrant Families Don't “Infest” America—But Trump's Racist Rhetoric Does*, MIAMI HERALD (June 20, 2018), <https://www.miamiherald.com/news/local/news-columns-blogs/andres-oppenheimer/article213543104.html>.

<sup>2</sup> Catherine E. Shoichet, *‘Zero Tolerance’ a Year Later: How the US Family Separations Crisis Erupted*, CNN (Apr. 5, 2019), <https://www.cnn.com/interactive/2019/04/us/immigrant-family-separations-timeline>.

<sup>3</sup> Griffin Edwards & Stephen Rushin, *The Effect of President Trump's Election on Hate Crimes 3* (Jan. 14, 2018) (unpublished manuscript), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3102652](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3102652).

<sup>4</sup> May Bulman, *Brexit Vote Sees Highest Spike in Religious and Racial Hate Crimes Ever Recorded*, INDEPENDENT (July 7, 2017), <https://www.independent.co.uk/news/uk/home-news/racist-hate-crimes-surge-to-record-high-after-brexit-vote-new-figures-reveal-a7829551.html>.

compelled gatekeepers, such as district attorneys, newspaper editors, publishers, or university provosts, to restrict certain classes of speech. Now, in the cacophony on social media, gatekeeping itself has been transformed.<sup>5</sup> Governments still attempt to regulate speech through gatekeepers such as social media platforms, but the sheer volume of speech involved requires operationalizing the definition of hate speech through algorithmic processes and tens of thousands of human content moderators.

Governments are no longer the primary regulators of speech.<sup>6</sup> Their regulatory capacity has been far outstripped by some of the largest companies in the world by public stock valuation,<sup>7</sup> which together regulate the speech of 3.7 billion active social media users.<sup>8</sup> Facebook, which moderates the online speech of over 2.4 billion active monthly users, is the largest publisher of content in human history.<sup>9</sup> In a reversal of the historic roles, private corporations have even become the de facto regulators of government speech, as when Facebook banned the Commander-in-Chief of Myanmar's military from the platform and removed over 400 other news, entertainment, and lifestyle pages linked to the military.<sup>10</sup>

For all practical purposes, the decision of social media companies to prohibit hate speech on their platforms means that the longstanding debate in the United States about whether to limit hate speech in the public square has been resolved in favor of regulation. The new realities of the internet do not mean that the prior debates on hate speech are irrelevant, however. Instead, we contend that a reexamination of the debates over hate speech that occurred in the United States in the 1980s and 1990s can help chart a course toward a more empirically-based and narrowly tailored policy regarding online hate speech.

Social media policy might be informed by the terms of the U.S. hate speech debate in two vital ways. First, social media creates data that allow us to determine whether hate speech that nonetheless falls short of direct incitement can still contribute to violence. Those in favor of more robust

---

<sup>5</sup> EMILY B. LAIDLAW, *REGULATING SPEECH IN CYBERSPACE: GATEKEEPERS, HUMAN RIGHTS AND CORPORATE RESPONSIBILITY*, at xi–xii (2015) (describing gatekeeping by private companies on the internet).

<sup>6</sup> See DAVID KAYE, *SPEECH POLICE: THE GLOBAL STRUGGLE TO GOVERN THE INTERNET* 112 (2019) (noting that “a few private companies” control social media).

<sup>7</sup> See *FB Facebook, Inc. Class A Common Stock*, NASDAQ, <https://www.nasdaq.com/market-activity/stocks/fb> (last visited Jan. 14, 2020) (listing Facebook’s common stock value); *GOOG Alphabet Inc. Class C Capital Stock*, NASDAQ, <https://www.nasdaq.com/market-activity/stocks/goog> (last visited Jan. 14, 2020) (listing Google’s capital stock value).

<sup>8</sup> SIMON KEMP, *DIGITAL 2019: Q4 GLOBAL DIGITAL STATSHOT* 6, [https://www.slideshare.net/slideshow/embed\\_code/key/NSBCWzcfbsbGXTg](https://www.slideshare.net/slideshow/embed_code/key/NSBCWzcfbsbGXTg) (last visited Jan. 30, 2020).

<sup>9</sup> *Facebook Q3 2019 Results*, FACEBOOK, [https://s21.q4cdn.com/399680738/files/doc\\_financials/2019/q3/Q3-2019-Earnings-Presentation.pdf](https://s21.q4cdn.com/399680738/files/doc_financials/2019/q3/Q3-2019-Earnings-Presentation.pdf) (last visited May 1, 2020).

<sup>10</sup> *Removing Myanmar Military Officials from Facebook*, FACEBOOK (Aug. 28, 2018), <https://about.fb.com/news/2018/08/removing-myanmar-officials/>.

regulation of hate speech have long claimed a connection to violence, while free speech advocates have dismissed such claims as too tenuous to support regulation.<sup>11</sup> The data now available on social media, however, allow researchers to empirically test whether there are visible, measurable harms resulting from hate speech. Studies thus far indicate that speech that denigrates and generates discriminatory animus against social groups such as immigrants or religious minorities does increase the risk of real-world violence against them.<sup>12</sup> Thus, contrary to justifications underlying the highly deferential standard of the U.S. First Amendment, it is no longer possible to discount the real-world risks of hate speech that falls short of incitement or true threat.

Second, reexamining the U.S. debate about hate speech also reveals the serious shortcomings of social media policies that seek to regulate content without reference to context. The policies that social media companies have developed define hate speech based on content alone. Currently, they have no mechanism for taking into account the context of the speech. However, if there is one thing that we can learn from Critical Race Theorists who sought to ban “assaultive speech,”<sup>13</sup> it is that words and meaning are deeply contextual. Defining hate speech in ways that are abstracted from local realities, relationships, and power will inevitably be overbroad.

Third, our approach draws from theories of legal pluralism to highlight how prevailing social norms about the regulation of speech have been as influential as legal doctrine in shaping the regulatory frameworks that have emerged in the past five years.<sup>14</sup> Legal pluralism embraces the mutually constitutive nature of law and social norms and shares with Critical Race Theory an account of how law is embedded in and often reinforces inequality based on status. The greater part of political discourse currently takes place online on privately-owned fora where standard international and domestic legal protections in general do not apply.<sup>15</sup> In that space,

---

<sup>11</sup> See John T. Bennett, *The Harm in Hate Speech: A Critique of the Empirical and Legal Bases of Hate Speech Regulation*, 43 HASTINGS CONST. L.Q. 445, 498 (2016) (“The harm caused by words is nebulous, easily exaggerated, and readily contrived.”).

<sup>12</sup> See *infra* Part II.

<sup>13</sup> See Charles R. Lawrence III et al., *Introduction to WORDS THAT WOUND: CRITICAL RACE THEORY, ASSAULTIVE SPEECH, AND THE FIRST AMENDMENT I*, 1 (Mari J. Matsuda et al. eds., 1993).

<sup>14</sup> See Paul Schiff Berman, *Global Legal Pluralism*, 80 S. CAL. L. REV. 1155, 1157 (2007) (discussing the applicability of global legal pluralism); Sally Engle Merry, *Legal Pluralism*, 22 LAW & SOC’Y REV. 869, 875–79 (1988) (reviewing the history and relevance of the theory of legal pluralism); Franz von Benda-Beckmann, *Who’s Afraid of Legal Pluralism?*, 47 J. LEGAL PLURALISM & UNOFFICIAL L. 37, 39 (2002) (defending the idea of legal pluralism from its critics).

<sup>15</sup> Companies have limited direct obligations under the human rights standards that govern freedom of expression. See Molly K. Land, *Toward an International Law of the Internet*, 54 HARV. INT’L L.J. 393, 444–48 (2013); see also Molly K. Land, *Against Privatized Censorship: Proposals for Responsible Delegation*, 60 VA. J. INT’L L. 363, 392–93 (2020) (examining application of international

changing social norms around hate speech have become a driver of increased regulation. Hate speech is being censored at a scale that was unthinkable twenty-five years ago. On the one hand, this may well be justified given the emerging consensus in social science research that there is a correlation between hate speech on social media and offline violence. However, the hate speech policies of social media platforms go beyond what was urged by even the most ardent proponents of banning hate speech in the 1980s and 1990s.

This Article examines this debate in Part I, recalling both the arguments of those who advocated in favor of greater regulation of hate speech, as well as the responses of those who worried that such restrictions would stifle dissent. Part II then examines how these debates play out against the background of social media, where emerging evidence indicates a correlation between hate speech online, virulent nationalism, and violence against minorities and activists. In Part III, we evaluate the platform law of hate speech, arguing that social media regulation is abstracted from context and power and, thus, is substantially overbroad. Finally, in Part IV, we suggest recommendations for how social media platforms might alter their policies and procedures to better account for the context of online speech.

## I. DEBATING HATE SPEECH IN THE UNITED STATES

Although legal interpretations of the U.S. First Amendment have changed significantly over the last century, current jurisprudence is highly deferential to speech. A high water mark for the First Amendment was *New York Times Co. v. Sullivan* in 1964, which applied a heightened standard for public officials in defamation lawsuits.<sup>16</sup> This was followed by *Brandenburg v. Ohio* in 1968, in which the U.S. Supreme Court held unconstitutional a statute that had been the basis for the conviction of a KKK member for inciting speech that included racist and anti-Semitic elements.<sup>17</sup> In the decades that followed, a debate raged about whether the pendulum had swung too far and whether the robust protections of the First Amendment for racist speech carried too many risks for vulnerable minority groups.<sup>18</sup>

---

law to the speech-regulating activities of social media companies).

<sup>16</sup> 376 U.S. 254, 283 (1964).

<sup>17</sup> *Brandenburg v. Ohio*, 395 U.S. 444, 448–49 (1969) (per curiam).

<sup>18</sup> See L. BOLLINGER, *THE TOLERANT SOCIETY: FREE SPEECH AND EXTREMIST SPEECH IN AMERICA* (1986) (defending *Brandenburg's* approach to racist speech); Richard Delgado, *Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling*, 17 HARV. C.R.-C.L. L. REV. 133 (1982) (advocating a tort remedy for racist speech); Mari J. Matsuda, *Public Response to Racist Speech: Considering the Victim's Story*, 87 MICH. L. REV. 2320, 2321 (1989) (arguing that public prosecution is an appropriate response to racist speech of the Ku Klux Klan).

### A. *The Case for Restricting Hate Speech*

In the hate speech debate in the United States in the 1980s and 1990s, the first challenge was to define the type of speech that required restriction. Advocates argued that speech should be considered hate speech when it is used as a weapon to “ambush, terrorize, wound, humiliate, and degrade.”<sup>19</sup> Mari Matsuda, a leading advocate of restrictions on “words that wound,” argued that racist speech presents ideas “so historically untenable, so dangerous, and so tied to perpetuation of violence and degradation of the very classes of human beings who are least equipped to respond that it is properly treated as outside the realm of protected discourse.”<sup>20</sup> Another influential definition conceived hate speech capaciously, as “any form of expression through which speakers primarily intend to vilify, humiliate, or incite hatred against their targets,” and proposed defining “hate speech” as attacks “so virulent that an observer would have great difficulty separating the message delivered from the attack against the victim.”<sup>21</sup>

Proponents of restricting hate speech focused their attention on changing the culture of political dialogue at universities in the United States. This was a response to a genuine problem, namely the perception that universities were an unreceptive environment for members of racial and ethnic minorities, female students, and other historically disadvantaged groups who had only gained significant access to elite public and private universities in the previous decade.<sup>22</sup> They encouraged university administrators to promulgate campus speech codes, citing studies that racist or sexist language and conduct can create a hostile learning environment.<sup>23</sup> Campus speech codes were often broadly worded and restricted speech that was already prohibited (e.g., true threats, incitement) along with speech protected by the First Amendment and prevalent in public and private discourse (e.g., racist speech, or speech that intends to humiliate).<sup>24</sup> In the five years between 1986 and 1991, approximately 137

---

<sup>19</sup> Lawrence et al., *supra* note 13, at 1.

<sup>20</sup> Matsuda, *supra* note 18, at 2357.

<sup>21</sup> Kenneth D. Ward, *Free Speech and the Development of Liberal Virtues: An Examination of the Controversies Involving Flag-Burning and Hate Speech*, 52 U. MIAMI L. REV. 733, 765–66 (1998).

<sup>22</sup> Mary Ellen Gale, *Reimagining the First Amendment: Racist Speech and Equal Liberty*, 65 ST. JOHN’S L. REV. 119, 164–65 (1991) (“If we can ever construct rules that successfully balance the rights and liberties of dominant and subordinate groups in hopes of creating a more just and equal society, the university, increasingly a site of racist and sexist incidents, seems like an appropriate and necessary place to begin.”).

<sup>23</sup> *Id.* at 177 n.223 (“[The National ACLU] urges each university . . . ‘to develop comprehensive plans aimed at reducing prejudice, responding promptly to incidents of bigotry and discriminatory harassment, and protecting students from any such further incidents.’”).

<sup>24</sup> See Jon Gould, *The Triumph of Hate Speech Regulation: Why Gender Wins but Race Loses in America*, 6 MICH. J. GENDER & L. 153, 158–59 (1999) (providing examples of prohibited “discriminatory harassment” at the University of Michigan and University of Wisconsin).



colleges and universities adopted new speech codes.<sup>25</sup> Their half-life was rather short, and all the campus speech codes challenged in the courts by the American Civil Liberties Union (ACLU) were struck down on First Amendment grounds.<sup>26</sup> The message from the courts was clear: State universities must permit all speech that the government cannot itself censor.

Current approaches adopted by proponents of hate speech regulation bear little resemblance to the early campus codes. Jeremy Waldron's influential book, for example, emphasizes the broader dignitary harms in hate speech, illustrated not through social science studies or other empirical evidence, but through hypothetical scenarios—such as one in which a Muslim father has to explain to his son and daughter why there is an ethnic slur daubed on a wall in their neighborhood blaming Muslims for 9/11.<sup>27</sup> Waldron reviews various formulations of hate speech without endorsing any of them, and instead builds his case around the concept of group libel or defamation, as contemplated by the *Beauharnais* case of 1952.<sup>28</sup> Chris Demaske similarly makes an expansive argument for regulating hate speech based upon a positive case to achieve greater social equality.<sup>29</sup>

International law, in turn, is more permissive of regulation than the U.S. First Amendment but does not go as far as Waldron or Demaske, and requires that hate speech express an explicit intent to incite harm or possess an identifiable causal nexus to actual harm. For instance, a 2019 report by the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression emphasized the empirical connection between hate speech and incitement to violence in contemplating companies' hate speech policies.<sup>30</sup> Others have proposed definitions of "hate speech" closer to incitement or threat, forms of speech that are already banned by law in the United States and most other countries.

---

<sup>25</sup> *Id.* at 158; see also Steven R. Glaser, *Sticks and Stones May Break My Bones, but Words Can Never Hurt Me: Regulating Speech on University Campuses*, 76 MARQ. L. REV. 265, 267 (1992) (developing a critical view of campus hate speech codes); Azhar Majeed, *Defying the Constitution: The Rise, Persistence, and Prevalence of Campus Speech Codes*, 7 GEO. J.L. & PUB. POL'Y 481, 486 (2009) (analyzing the legality of speech codes on college campuses).

<sup>26</sup> NADINE STROSSEN, HATE: WHY WE SHOULD RESIST IT WITH FREE SPEECH, NOT CENSORSHIP 31 (2018). Relevant cases include: *DeJohn v. Temple Univ.*, 537 F.3d 301, 320 (3d Cir. 2008); *Dambrot v. Cent. Mich. Univ.*, 839 F. Supp. 477, 482 (E.D. Mich. 1993); *UWM Post, Inc. v. Bd. of Regents of the Univ. of Wis.*, 774 F. Supp. 1163, 1181 (E.D. Wis. 1991).

<sup>27</sup> JEREMY WALDRON, THE HARM IN HATE SPEECH 1 (2012).

<sup>28</sup> *Id.* at 51–52 (citing *Beauharnais v. Illinois*, 343 U.S. 250, 283–84 (1952) (Reed, J., dissenting); 343 U.S. at 292 (Jackson, J., dissenting); and 343 U.S. at 271–72 (Black, J., dissenting)).

<sup>29</sup> Chris Demaske, *Social Justice, Recognition Theory and the First Amendment: A New Approach to Hate Speech Restriction*, 24 COMM. L. & POL'Y, 347, 393 (2019).

<sup>30</sup> David Kaye, *Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 8, U.N. Doc. A/74/486 (Oct. 9, 2019).

Synthesizing free speech theory and Critical Race Theory, Andrew Sellars identifies eight traits indicative of hate speech: (1) the speech targets a group, or individual as a member of a group; (2) the content in the message expresses hatred; (3) the speech causes harm; (4) the speaker intends harm; (5) the speech incites harmful acts beyond the speech itself; (6) the speech is either public or directed at a member of the group; (7) the context makes a violent response possible; and (8) the speech has no redeeming purpose.<sup>31</sup> In sum, modern legal definitions of hate speech have moved away from prohibiting merely offensive speech and concentrated more on incitement or threat, that is, forms of speech that are already banned by most national legal systems.

### B. *The Freedom of Expression Riposte*

The U.S. Supreme Court has eschewed a constitutional definition of hate speech and repeatedly rejected efforts to establish a hate speech exception to the First Amendment.<sup>32</sup> Thus, much hate speech in the United States, even that which is highly offensive, is protected speech. Current constitutional law only suppresses hate speech that also represents a “true threat” or incitement to imminent lawless acts.<sup>33</sup> Even then, any regulation of inciting or threatening speech must be consistent with the long-established principles of viewpoint and content neutrality that prevent the state from disfavoring certain messages, ideas, subject matter, or opinions, for instance, simply because they are offensive or disagreeable.<sup>34</sup>

Nadine Strossen’s book, *Hate: Why We Should Resist It with Counterspeech, Not Censorship*, distills the modern “free speech advocacy” standpoint. Strossen starts with the observation that there is no clear and consistent definition of “hate speech.”<sup>35</sup> Hate speech is not a term of legal art and, in her view, it is factually wrong to assert that “hate speech

---

<sup>31</sup> ANDREW F. SELLARS, BERKMAN KLEIN CTR. FOR INTERNET & SOC’Y RESEARCH, DEFINING HATE SPEECH 25–31 (2016), <https://ssrn.com/abstract=2882244>.

<sup>32</sup> See, e.g., *Matal v. Tam*, 137 S. Ct. 1744, 1764 (2017) (“[T]he proudest boast of our free speech jurisprudence is that we protect the freedom to express ‘the thought that we hate.’”); *Snyder v. Phelps*, 562 U.S. 443, 461 (2011) (“As a Nation we have chosen a different course—to protect even hurtful speech on public issues to ensure that we do not stifle public debate.”).

<sup>33</sup> *Virginia v. Black*, 538 U.S. 343, 359 (2003); *Brandenburg v. Ohio*, 395 U.S. 444, 448–49 (1969) (per curiam). See Richard Ashby Wilson & Jordan Kiper, *Incitement in an Era of Populism: Updating Brandenburg After Charlottesville*, 5 U. PA. J.L. & PUB. AFF. 56, 62–72 (2020) (reviewing the legal history of incitement and true threat).

<sup>34</sup> See *R.A.V. v. City of St. Paul*, 505 U.S. 377, 395 (1992) (stating that laws restricting speech based on its content are presumptively unconstitutional); *Reed v. Town of Gilbert*, 135 S. Ct. 2218, 2222 (2015) (holding that laws may not restrict speech based on its content). See David L. Hudson, Jr., *The Content-Discrimination Principle and the Impact of Reed v. Town of Gilbert*, 70 CASE W. RES. L. REV. 259, 260 (2019) (reviewing the central analytical role of the content-discrimination principle in First Amendment jurisprudence).

<sup>35</sup> STROSSEN, *supra* note 26, at 3.

is not free speech.”<sup>36</sup> According to Strossen, the term is deployed to demonize views people find offensive and to punish a broad swathe of expression, including political discourse that is integral to democratic deliberation.<sup>37</sup> Reviewing hate speech laws in the United States and globally, Strossen concludes that it is not possible to draft hate speech laws that are not “unduly vague,” overbroad, and counter-productive.<sup>38</sup> She critiques the legal systems of Germany, France, and other European countries, which convict hundreds of defendants a year for offenses as imprecise as “incitement to hatred,” including a number of cases that seem disproportionately chilling of political speech.<sup>39</sup>

Strossen correctly observes that speech regulations tend to disfavor minority viewpoints. Indeed, there is a long, repressive history of government censorship in the United States going back to the Early Republic.<sup>40</sup> In the 1830s, Southern states banned abolitionist speech on the grounds that it had the potential to incite violence and rebellion.<sup>41</sup> For much of the twentieth century, incitement law was used to stifle socialist thought and speech, including the mere reading and discussion of books by Marx and Engels.<sup>42</sup> Strossen observes that the Republican National Committee and some state legislatures have included the Black Lives Matter movement in resolutions condemning hate speech.<sup>43</sup> She concludes that government censorship of political speech is a greater threat than individuals expressing discriminatory bias, and she endorses the aphorism, “[t]he cure is worse than the disease” when it comes to hate speech regulation.<sup>44</sup>

There is ample recent evidence to support the observation that censorious regimes frequently abuse incitement laws to stifle political and religious dissent. For instance, in 2019, the Chinese government convicted a Christian pastor for incitement of subversion of state power and sentenced him to nine years in prison for saying that the Chinese Communist Party’s ideology was “morally incompatible with the Christian

---

<sup>36</sup> *Id.* at 1, 3.

<sup>37</sup> *Id.* at 1.

<sup>38</sup> *Id.* at 13.

<sup>39</sup> *See id.* at 83–86 (discussing hate speech laws around the world).

<sup>40</sup> *See, e.g.,* PHILLIP I. BLUMBERG, REPRESSIVE JURISPRUDENCE IN THE EARLY AMERICAN REPUBLIC: THE FIRST AMENDMENT AND THE LEGACY OF ENGLISH LAW 2 (2010) (describing the repressive speech laws of the first fifty years of the republic).

<sup>41</sup> *See id.* at 340 (discussing prohibitions on abolitionist speech).

<sup>42</sup> *See* Wilson & Kiper, *supra* note 33, at 65–69 (reviewing the legal suppression of socialists and communists in the United States during the twentieth century). *See generally* John F. Wirenius, *The Road to Brandenburg: A Look at the Evolving Understanding of the First Amendment*, 43 *DRAKE L. REV.* 1, 2–3 (1994) (reciting the history of First Amendment jurisprudence in the twentieth century).

<sup>43</sup> STROSSEN, *supra* note 26, at 17.

<sup>44</sup> *Id.* at 14.

faith.”<sup>45</sup> In the same year, the Governor of South Dakota proposed new legislation banning “incitement to riot” to stamp out protest by indigenous and environmental activists at the Keystone XL pipeline.<sup>46</sup> Around the world, defamation laws are used to target and silence human rights defenders.<sup>47</sup>

Strossen’s thesis, which is representative of free speech advocacy more generally, can be summarized in three points: (1) there is insufficient evidence that hate speech falling short of incitement or true threat causes the harms attributed to it; (2) hate speech laws do not effectively reduce such putative harms; and (3) even if hate speech laws did reduce the feared harms, the damage to democratic legitimacy and freedom of speech would be too great to justify them.<sup>48</sup> By the early 2000s, the free speech advocates had largely won the argument, at least in U.S. courts. Although a number of scholars continued to write about the need to mitigate the harms of hate speech, their efforts at changing policy were largely unsuccessful, and municipalities, states, and state universities generally refrained from proclaiming hate speech codes.<sup>49</sup> The apparent success of free speech advocacy was, however, short-lived.

## II. THE HARMS OF ONLINE HATE SPEECH

Debates about the scope of acceptable political speech are adamant and date back at least to Aristotle’s *Rhetoric*,<sup>50</sup> but what is new and different about the present moment is the outsized role played in popular political discourse by a small number of global technology corporations. In the past ten years, the rise of social media companies and their content moderation policies have transformed the hate speech debate. We are now in a brave new world in which the First Amendment applies to only a fraction of public discourse about the issues of the day. Democratic

---

<sup>45</sup> Eva Dou, *Activist Chinese Pastor Gets Nine-Year Prison Sentence*, WALL ST. J. (Dec. 30, 2019), <https://www.wsj.com/articles/activist-chinese-pastor-gets-nine-year-prison-sentence-11577698307>.

<sup>46</sup> Stephen Groves, *South Dakota Governor Plans Revision of Riot-boosting Laws*, ASSOCIATED PRESS (Dec. 16, 2019), <https://apnews.com/42f4533aac4913f2bb686d8d06c14b7b>; see also *South Dakota Governor Drops Anti-Protest Laws in Settlement Agreement with ACLU*, ACLU (Oct. 24, 2019), <https://www.aclu.org/press-releases/south-dakota-governor-drops-anti-protest-laws-settlement-agreement-aclu> (announcing settlement agreement with ACLU).

<sup>47</sup> OSCE OFFICE FOR DEMOCRATIC INSTS. & HUMAN RIGHTS, HUMAN RIGHTS GUIDELINES ON THE PROTECTION OF HUMAN RIGHTS DEFENDERS 70 (2014).

<sup>48</sup> STROSSEN, *supra* note 26, at 36.

<sup>49</sup> See, e.g., Richard Delgado, *Toward a Legal Realist View of the First Amendment*, 113 HARV. L. REV. 778, 798–99 (2000) (reviewing STEVEN H. SHIFFRIN, *DISSENT, INJUSTICE, AND THE MEANINGS OF AMERICA* (1999)) (advocating for solutions such as hate speech codes and the application of tort principles in order to limit hate speech on college campuses).

<sup>50</sup> See ARISTOTLE, *RHETORIC* 3–6 (Paul Negri & Jenny Bak eds., W. Rhys Roberts trans., Dover 2004) (1910) (outlining Aristotle’s views on political speech).

deliberation now happens online, and in particular, on Facebook, Twitter, YouTube, and Reddit, which are owned by private corporations that set their own terms and conditions.<sup>51</sup> Across the world, social media is the town square in which current affairs are discussed and political opinions hashed out.<sup>52</sup>

There is much that is positive about the relationship between democracy and social media. Information flows more freely and quickly in the body politic. Government and citizens can communicate directly, without the filter of traditional media, and in this way social media shrinks the space between citizens and their official representatives, permitting average people to pose questions to officials and state their views on policy issues. Social media can also enhance accountability of state institutions by providing video evidence of violence and violations of rights committed by police or the security forces.<sup>53</sup>

At the same time, it is indisputable that social media is being exploited to promote populist and nativist politics around the globe. The number of populist governments has doubled worldwide since the early 2000s, and populist leaders have been elected in Brazil, Guatemala, Hungary, India, the Philippines, Poland, Turkey, the United Kingdom, the United States, and Venezuela, among other places.<sup>54</sup> As part of a “cultural backlash” against elites, populists win approval by harshly criticizing and even ridiculing established political elites and nostalgically harking back to bygone days when (it is imagined) things were better for working people.<sup>55</sup>

---

<sup>51</sup> See Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 22 (2004) (arguing that private corporations that control online platforms for deliberation will try to limit the participation and expression of users in the pursuit of their own economic interests).

<sup>52</sup> See NANJALA NYABOLA, *DIGITAL DEMOCRACY, ANALOGUE POLITICS: HOW THE INTERNET ERA IS TRANSFORMING POLITICS IN KENYA 1* (2018) (showing the connectedness of Kenyan citizens via social media platforms such as WhatsApp and Twitter); Danielle Keats Citron & Neil M. Richards, *Four Principles for Digital Expression (You Won't Believe #3!)*, 95 WASH. U. L. REV. 1353, 1355–56 (2018) (“Rather than just the virtual town square, the Internet is bound up in everything we do and everywhere we do it . . .”).

<sup>53</sup> Sam Gregory, *Cameras Everywhere Revisited: How Digital Technologies and Social Media Aid and Inhibit Human Rights Documentation and Advocacy*, 11 J. HUM. RTS. PRAC. 373, 374 (2019).

<sup>54</sup> Paul Lewis et al., *Revealed: The Rise and Rise of Populist Rhetoric*, GUARDIAN (Mar. 6 2019, 9:00 AM), <https://www.theguardian.com/world/ng-interactive/2019/mar/06/revealed-the-rise-and-rise-of-populist-rhetoric>; Richard Javad Heydarian, *Understanding Duterte's Mind-Boggling Rise to Power*, WASH. POST (Mar. 20, 2019, 11:55 AM), <https://www.washingtonpost.com/news/worldpost/wp/2018/03/20/duterte/>; Clarence Page, *Comedians as Populist Leaders? Ukrainian Comic—And Our Own President—Proves It's No Joke*, CHI. TRIB. (Apr. 23, 2019, 2:05 PM), <https://www.chicagotribune.com/columns/clarence-page/ct-perspec-page-ukraine-president-donald-trump-comedian-slovenia-guatemala-20190423-story.html>.

<sup>55</sup> Ronald F. Inglehart & Pippa Norris, *Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash 3* (Harvard Kennedy Sch. Faculty, Working Paper No. RWP16-026, 2016), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2818659](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2818659).

Populist speech is often graphic and brimming with crude insults, racial resentment, chauvinistic intolerance, and xenophobia.<sup>56</sup>

For some observers, the simultaneous rise of social media and populist politics is no coincidence, as social media “privileges discourse that is simple, impulsive, and uncivil” and debases serious-minded democratic deliberation.<sup>57</sup> Observers blame social media for the creation of “echo chambers” of like-minded individuals and the deepening of political polarization that allows populist politics to take root and thrive.<sup>58</sup> Others go a step further and claim that online disinformation campaigns have undermined democratic elections and influenced the outcome of the Brexit vote and the 2016 presidential election in the United States. Susan Morgan, for example, notes that, “days before the election, messages circulated on social media that Hillary Clinton had died. And in some key battlegrounds, messages were targeted at Democrat voters claiming that the date of the election had changed.”<sup>59</sup>

Once populist leaders achieve political power, they frequently take advantage of the affordances of social media—including the ability to transmit and amplify their messages directly to a wide audience—to consolidate their hold on government.<sup>60</sup> In some cases, they mobilize troll armies to flood social media with disinformation and angry rhetoric targeting their political opponents, a phenomenon that has been termed “digital populism.”<sup>61</sup> “Digital authoritarianism” is another newly coined term that describes the more sinister uses of social media by states or state-aligned actors that includes surveilling, harassing, threatening, and inciting violence against their political opponents.<sup>62</sup>

<sup>56</sup> *Id.* at 7; see also Jonathan T. Rothwell & Pablo Diego-Rosell, *Explaining Nationalist Political Views: The Case of Donald Trump* 1 (Working Paper, 2016), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2822059](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2822059) (examining support for Donald Trump and racial resentment).

<sup>57</sup> Brian L. Ott, *The Age of Twitter: Donald J. Trump and the Politics of Debasement*, 34 *CRITICAL STUD. MEDIA COMM.* 59, 59 (2017).

<sup>58</sup> JOSHUA A. TUCKER ET AL., *SOCIAL MEDIA, POLITICAL POLARIZATION, AND POLITICAL DISINFORMATION: A REVIEW OF THE SCIENTIFIC LITERATURE* 16 (2018), <https://www.hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf>.

<sup>59</sup> Susan Morgan, *Fake News, Disinformation, Manipulation and Online Tactics to Undermine Democracy*, 3 *J. CYBER POL'Y* 39, 40 (2018); see also Martin Innes et al., *Disinformation and Digital Influencing After Terrorism: Spoofing, Truthing and Social Proofing*, *CONTEMP. SOC. SCI.*, Jan. 2019, at 1, 1–15 (discussing the effects of disinformation campaigns in the aftermath of terrorist attacks); Nicholas Thompson & Issie Lapowsky, *How Russian Trolls Used Meme Warfare to Divide America*, *WIRED* (Dec. 17, 2018), <https://www.wired.com/story/russia-ira-propaganda-senate-report/> (discussing a Senate report that shows how Russia’s online disinformation campaign might have affected the 2016 election).

<sup>60</sup> Ergin Bulut & Erdem Yörük, *Digital Populism: Trolls and Political Polarization of Twitter in Turkey*, 11 *INT’L J. COMM.* 4093, 4094 (2017).

<sup>61</sup> *Id.* at 4093.

<sup>62</sup> CARLY NYST & NICK MONACO, *INST. FOR THE FUTURE, STATE-SPONSORED TROLLING: HOW GOVERNMENTS ARE DEPLOYING DISINFORMATION AS PART OF BROADER DIGITAL HARASSMENT*

The fact that hate speech has moved online means that its character and effects are much easier to study than when it was confined to traditional media.<sup>63</sup> Indeed, social scientists have begun to identify a correlation between online hate speech and offline violence.<sup>64</sup> The preponderance of studies conducted during the recent upsurge of right-wing populism pinpoint both direct and systemic harmful effects of hate speech. The Irish human rights organization Front Line Defenders writes that, “[t]here is a well-evidenced link between defamatory attacks online and in pro-government media and an escalation to physical attacks on individuals and their families.”<sup>65</sup>

Establishing that hate speech on social media directly causes or even correlates with hate crimes is always difficult, but there is a growing body of evidence that widespread attacks on immigrants and other minorities in recent years have been instigated online. Social scientists such as Edwards and Rushin have found evidence for what they call the “Trump effect”—a correlation between Trump’s election and an increase in hate crimes in the United States during 2016.<sup>66</sup> Using panel regressions, they find that “counties that voted for President Trump by the widest margins in the presidential election also experienced the largest increases in reported hate crimes.”<sup>67</sup> They conclude that Trump’s election may have resulted in approximately 410 additional hate crimes nationally per quarter, or 2048 additional hate crimes since his electoral victory.<sup>68</sup> Additionally, Müller and Schwarz identify a statistically significant correlation between anti-immigrant and anti-Muslim tweets and actual physical attacks on

---

CAMPAIGNS 1 (2018), [https://www.iftf.org/fileadmin/user\\_upload/images/DigIntel/IFTF\\_State\\_sponsored\\_trolling\\_report.pdf](https://www.iftf.org/fileadmin/user_upload/images/DigIntel/IFTF_State_sponsored_trolling_report.pdf) (describing state sponsored threats and incitements); Gregory, *supra* note 53, at 380; *see generally* MUNA ABBAS ET AL., AM. BAR ASS’N, INVISIBLE THREATS: MITIGATING THE RISK OF VIOLENCE FROM ONLINE HATE SPEECH AGAINST HUMAN RIGHTS DEFENDERS IN GUATEMALA (2019), [https://www.americanbar.org/groups/human\\_rights/reports/invisiblethreats-online-hate-speech/](https://www.americanbar.org/groups/human_rights/reports/invisiblethreats-online-hate-speech/) (describing incitement against human rights defenders in Guatemala).

<sup>63</sup> Neurolinguistic analysis of social media has involved the design of classifiers for hate speech detection with a high level of accuracy. *See* Mai ElSherief et al., *Hate Lingo: A Target-based Linguistic Analysis of Hate Speech in Social Media*, ASS’N FOR ADVANCEMENT ARTIFICIAL INTELLIGENCE (Apr. 11, 2018), <https://arxiv.org/abs/1804.04257>.

<sup>64</sup> Peter Neumann, *Foreword* to JOHANNES BALDAUF ET AL., HATE SPEECH AND RADICALISATION ONLINE: THE OCCI RESEARCH REPORT 5 (2019); Josh Jacobs, *Does Online Hate Drive Anti-Migrant Violence?*, FIN. TIMES (Sept. 24, 2018), <https://www.ft.com/content/a3d4b800-9bf3-11e8-88de-49c908b1f264>.

<sup>65</sup> FRONT LINE DEFS., GLOBAL ANALYSIS 2018, at 6 (2019), [https://www.frontlinedefenders.org/sites/default/files/global\\_analysis\\_2018.pdf](https://www.frontlinedefenders.org/sites/default/files/global_analysis_2018.pdf).

<sup>66</sup> Edwards & Rushin, *supra* note 3, at 3; *see also* Karsten Müller & Carlo Schwarz, *From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment 2* (Nov. 2, 2019) (unpublished manuscript), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3149103](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3149103) (analyzing the relationship between Twitter posts of Donald J. Trump and offline violence against minorities).

<sup>67</sup> Edwards & Rushin, *supra* note 3, at 3.

<sup>68</sup> *Id.* at 13.

Muslims and immigrants in Germany and the United States in 2016.<sup>69</sup> In a study of 100 U.S. cities between 2011-2016, Relia et al. find that hate crimes correlate with tweets containing targeted discrimination on the basis of race, ethnicity, and national-origin.<sup>70</sup>

The genocide of the Rohingya in Myanmar stands out as the most egregious and harmful use of social media by a government thus far. According to a report on the situation in Myanmar issued by the Office of the Prosecutor of the International Criminal Court, anti-Muslim rhetoric and threats of violence against Rohingya were posted on Facebook by high-ranking members of Myanmar's military and Buddhist nationalists beginning in 2012.<sup>71</sup> In two waves of violence in 2016 and 2017, military and non-Rohingya civilians violently attacked Muslim Rohingya villages in Rakhine State.<sup>72</sup> The most conservative estimate is that the Myanmar armed forces ("*Tatmadaw*") killed 10,000 Rohingya civilians,<sup>73</sup> but some observers put the number higher.<sup>74</sup> As of September 2018, there had been 6097 documented incidences of gender-based violence, including sexual violence,<sup>75</sup> and approximately 780,000 Rohingya had been forcibly displaced to neighboring Bangladesh.<sup>76</sup>

Posts on Facebook are believed to have both incited specific acts of violence against the Rohingya and contributed to a general climate that made the genocide possible. A UN-sponsored fact-finding mission documented over 150 public social media accounts, pages, and groups on Facebook that "regularly spread messages amounting to hate speech against Muslims in general or Rohingya in particular."<sup>77</sup> The report noted a number of these accounts were "particularly influential" in light of "the number of followers (all over 10,000, but some over 1 million), the high levels of engagement of the followers with the posts (commenting and sharing), and the frequency of new posts (often daily, if not hourly)."<sup>78</sup> The Myanmar government deployed online narratives evoking several different themes about the supposed threat posed by the Rohingya Muslims,

---

<sup>69</sup> Karsten Müller & Carlo Schwarz, *Fanning the Flames of Hate: Social Media and Hate Crime* 27, 42 (Nov. 3, 2019) (unpublished manuscript), <https://ssrn.com/abstract=3082972>.

<sup>70</sup> Kunnal Relia et al., *Race, Ethnicity and National Origin-based Discrimination in Social Media and Hate Crimes Across 100 U.S. Cities* (Jan. 21, 2019) (unpublished manuscript), <https://arxiv.org/abs/1902.00119> (testing the relationship between social media and hate crimes).

<sup>71</sup> Office of the Prosecutor, ICC-01/19, *Situation in the People's Republic of Bangladesh/Republic of the Union of Myanmar*, ¶ 59 (July 4, 2019).

<sup>72</sup> *Id.* ¶¶ 5, 67–68.

<sup>73</sup> *Id.* ¶ 90.

<sup>74</sup> *Id.* ¶ 89.

<sup>75</sup> *Id.* ¶ 94.

<sup>76</sup> *Id.* ¶ 69.

<sup>77</sup> Report of the Detailed Findings of the Independent International Fact-Finding Mission on Myanmar, U.N. Doc. A/HRC/39/CRP.2, ¶ 1310 (Sept. 17, 2018) [hereinafter Myanmar Report].

<sup>78</sup> *Id.*



including descriptions of them as “an existential threat to the country,” “a threat to Burmese racial purity,” and “a threat to Buddhist religious sanctity.”<sup>79</sup> According to the UN mission, “[s]uch narratives latch onto long-standing anti-Muslim prejudices and stereotypes; they are designed to stoke fear.”<sup>80</sup>

Anecdotal evidence indicates that these campaigns were influential in inciting specific attacks. For example, while a widespread and systematic attack on the Rohingya civilian population was underway, young soldiers in the Myanmar military posted derogatory and inciting messages on Facebook containing the same narratives deployed by propaganda accounts. On August 26, 2017, a soldier wrote that he could not wait to be deployed to Rakhine State because the “Muslim dogs” represented a threat to the citizens, and the next day he posted, “on the battlefield, whoever is quick will get to eat you, Muslim dogs.”<sup>81</sup> On August 27, 2017, a police officer involved in forcibly removing and killing Rohingya civilians posted that he “[h]a[s] been wanting to kill these ‘Kalar’ for so long. Only got to kill them just now.”<sup>82</sup> In its assessment of the situation, the Office of the Prosecutor of the International Criminal Court noted: “A lieutenant of the 33rd LID deployed to Rakhine in early August 2017 posted a ‘smirking emoji’ on Facebook, along with the commentary, ‘[i]f they’re Bengali [. . .] they’ll be killed.’”<sup>83</sup> In a report released in 2018, Reuters documented more than 1000 posts, comments, images, and videos on Facebook attacking the Rohingya or Muslims generally in extremely derogatory terms and threatening them with violence.<sup>84</sup> All the posts were in the Burmese language. Further, the UN mission affirmed that in addition to direct incitement, anti-Rohingya speech on social media “contributed to and exacerbated a climate in which hate speech thrives and in which individuals and groups may be more receptive to calls of incitement to violence.”<sup>85</sup>

The impact of these derogatory messages was amplified by the lack of alternative channels for information and low levels of digital literacy in Myanmar. The dependence of the population on Facebook for internet access is so great that, “[t]he word ‘Facebook’ has become synonymous with the internet itself in Myanmar.”<sup>86</sup> The Report of the Independent

---

<sup>79</sup> *Id.* ¶¶ 1315–17.

<sup>80</sup> *Id.* ¶ 1314.

<sup>81</sup> Office of the Prosecutor, *supra* note 71, ¶ 176.

<sup>82</sup> “Kalar” is an anti-Muslim racial epithet. *Id.*

<sup>83</sup> *Id.*

<sup>84</sup> *Id.* ¶ 59.

<sup>85</sup> Myanmar Report, *supra* note 77, ¶ 1327.

<sup>86</sup> Megha Rajagopalan et al., *How Facebook Failed the Rohingya in Myanmar*, BUZZFEED NEWS (Aug. 27, 2018, 6:46 PM), <https://www.buzzfeednews.com/article/meghara/facebook-myanmar-rohingya-genocide>.

International Fact-Finding Mission on Myanmar concluded, “[t]he role of social media is significant. Facebook has been a useful instrument for those seeking to spread hate, in a context where, for most users, Facebook is the Internet.”<sup>87</sup> The use of Facebook by Myanmar’s military to incite and pursue genocide is not in dispute, and Facebook has since publicly accepted responsibility for its central role in the genocide.<sup>88</sup>

Faced with the evidence from the above scenarios, it is no longer credible to maintain that hate speech that would be protected under the First Amendment (such as ethnic slurs or speech that denigrates a group but does not directly advocate violence) does not have the potential to cause substantial harm and injury. There is ample research showing that in certain circumstances hate speech correlates with, and possibly has a causal nexus to, measurable harms. Of course, an empirical consensus on the deleterious effects of denigrating and inciting speech does not in itself determine the normative question of hate speech regulation, since one might reasonably support freedom of expression for other reasons, including its value in the process of democratic decision-making. However, given the mounting evidence of the systemic and corrosive effects of hate speech, including speech that does not rise to the level of incitement or threat, we need to be asking what kinds of speech have which effects, and what would be the cost and consequences of taking action to address those effects.

### III. THE PLATFORM LAW OF HATE SPEECH

David Kaye, the former UN Special Rapporteur for the Promotion and Protection of Freedom of Opinion and Expression, has used the term “platform law” to refer to the rules that social media companies have developed to determine what user-generated online content may be posted to their platforms.<sup>89</sup> Platform law accurately captures the degree to which private corporations have established stable definitions of key terms like hate speech, operational rules about how to handle harmful or abusive content, and enduring mechanisms to enforce those rules. This Part discusses first how these platforms have defined “hate speech” and then introduces two significant limitations of this approach: the lack of

---

<sup>87</sup> Report of the Independent International Fact-Finding Mission on Myanmar, U.N. Doc. A/HRC/39/64, ¶ 74 (Sept. 12, 2018).

<sup>88</sup> Paul Mozur, *A Genocide Incited on Facebook, with Posts from Myanmar’s Military*, N.Y. TIMES (Oct. 15, 2018), <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>; Alexandra Stevenson, *Facebook Admits It Was Used to Incite Violence in Myanmar*, N.Y. TIMES (Nov. 6, 2018), <https://www.nytimes.com/2018/11/06/technology/myanmar-facebook.html>.

<sup>89</sup> David Kaye, *Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 1, U.N. Doc. A/HRC/38/35 (Apr. 6, 2018).

transparency and accountability in how platforms apply their rules, and their failure to consider the context of speech.

### A. *Defining Hate Speech Online*

Although it is true that some platforms such as Gab, 4Chan, and 8Chan engage in little or no moderation of the content posted on their platforms,<sup>90</sup> most mainstream platforms actively moderate content. Platforms have responded to the criticism that they are too permissive of online hate speech by creating infrastructures of regulation, including algorithms that remove content and mechanisms by which users can flag potentially abusive content that is then sent to tens of thousands of moderators for review.<sup>91</sup> The companies have wide discretion in designing their content moderation policies since they are not government institutions and therefore not directly subject to the First Amendment.<sup>92</sup> Although they have been regulated in many places around the world and are increasingly subject to much more stringent legal requirements in Europe, they otherwise operate like private clubs that set their own terms of service, to which all users must agree in order to participate on the platform.

In the early years, platform policies on the regulation of hate speech were strongly influenced by the First Amendment, but current hate speech policies of these private entities more closely resemble a European approach to speech regulation.<sup>93</sup> The amount of hate speech that internet

---

<sup>90</sup> See Jane Coasten, *Gab, the Social Media Platform Favored by the Alleged Pittsburgh Shooter, Explained*, VOX (Oct. 29, 2018, 11:00 AM), <https://www.vox.com/policy-and-politics/2018/10/29/18033006/gab-social-media-anti-semitism-neo-nazis-twitter-facebook> (“[U]sing little to no moderation, Gab has become the go-to social networking site for the alt-right and, moreover, the furthest fringes of the far right.”); Emily Stewart, *8chan, a Nexus of Radicalization, Explained*, VOX (Aug. 5, 2019, 12:56 PM), <https://www.vox.com/recode/2019/5/3/18527214/8chan-walmart-el-paso-shooting-cloudflare-white-nationalism> (“4chan’s content moderation policies are, to put it lightly, pretty lax, but 8chan’s are nearly nonexistent.”).

<sup>91</sup> TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA 4 (2018); James Grimmelmann, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 55–76 (2015) (providing a taxonomy of moderation in online communities); Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1635 (2018) (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).

<sup>92</sup> The U.S. Supreme Court held in *Manhattan Community Access Corp. v. Halleck* that merely hosting speech is not a traditionally exclusive public function and on its own does not transform private entities into state actors subject to First Amendment constraints. 139 S. Ct. 1921, 1930 (2019). Citing this precedent, the Ninth Circuit established that YouTube is a private forum, not a public forum subject to judicial scrutiny under the First Amendment. *Prager Univ. v. Google LLC*, 951 F.3d 991, 995 (9th Cir. 2020).

<sup>93</sup> See David Kaye, *Speech Police: The Global Struggle to Govern the Internet*, YOUTUBE (June 6, 2019), <https://www.youtube.com/watch?v=W6PDZ-o5Khg> (quotation at 39:34) (“[O]ne of the interesting things about the platforms, and our reaction to the platforms, is how much there is a clamoring, in our public, for the taking down of hate speech. . . . There may be a difference between

platforms remove is staggering, given that Facebook has over 2.4 billion active monthly users and Twitter has over 330 million monthly users<sup>94</sup> sending over 500 million tweets per day.<sup>95</sup> In 2019, social media companies including Facebook, Google, and Twitter, removed seventy-two percent of hate speech on their platforms that violated the European Commission's code of conduct.<sup>96</sup> Facebook reported in 2017 that it was deleting nearly 288,000 hate speech posts per month,<sup>97</sup> and this tripled to nearly 1 million hate speech posts per month by the third quarter of 2018.<sup>98</sup> In the first quarter of 2020, Facebook reported taking action on 9.6 million posts, or 3.2 million per month.<sup>99</sup> The rate of removal can spike even higher during critical incidents, such as when the New Zealand mosque shooter streamed on Facebook Live in 2019.<sup>100</sup> In the twenty-four hours after that tragedy, Facebook reported that it prevented the uploading of more than 1.2 million videos of the attack and removed approximately 300,000 additional copies after they were posted.<sup>101</sup>

Platform law employs capacious and far-reaching definitions of hate speech. For example, YouTube's policy is to "remove content promoting violence or hatred against individuals or groups based on any of the following attributes: age, caste, disability, ethnicity, gender identity and expression, nationality, race, immigration status, religion, sex/gender, sexual orientation, victims of a major violent event and their kin, veteran

---

our legal norms and our social norms, and the social norms in the United States, at this particular moment, may not be all that different from European norms. So, as the companies move to restrict more and more what they call 'hate speech' or 'hateful conduct,' I don't know if there'll be much of a real difference between European and American attitudes.").

<sup>94</sup> KEMP, *supra* note 8, at 24.

<sup>95</sup> *Twitter Usage Statistics*, INTERNET LIVE STATS, <https://www.internetlivestats.com/twitter-statistics/> (last visited Jan. 19, 2020).

<sup>96</sup> Elizabeth Schulze, *EU Says Facebook, Google and Twitter Are Getting Faster at Removing Hate Speech Online*, CNBC (Feb. 4, 2019, 6:06 AM), <https://www.cnbc.com/2019/02/04/facebook-google-and-twitter-are-getting-faster-at-removing-hate-speech-online-eu-finds--.html>. Hate speech, according to the European Commission, is "the public incitement to violence or hatred directed to groups or individuals on the basis of certain characteristics, including race, color, religion, descent and national or ethnic origin." *Id.* (citation omitted).

<sup>97</sup> STROSSEN, *supra* note 26, at 32.

<sup>98</sup> Mehreen Khan, *More 'Hate Speech' Being Removed from Social Media*, FIN. TIMES (Feb. 4, 2019), <https://www.ft.com/content/868f9d96-27bc-11e9-a5ab-ff8ef2b976c7> (reporting that Facebook took action against 2.9 million pieces of content in the third quarter of 2018, a rate of nearly 1 million per month).

<sup>99</sup> *Community Standards Enforcement Report*, FACEBOOK, <https://transparency.facebook.com/community-standards-enforcement#hate-speech> (last visited May, 26 2020).

<sup>100</sup> Nick Perry & Mark Baker, *New Zealand Mosque Shooter Broadcast Slaughter on Facebook*, ASSOCIATED PRESS (Mar. 15, 2019), <https://apnews.com/ce9e1d267af149dab40e3e5391254530>.

<sup>101</sup> *A Further Update on New Zealand Terrorist Attack*, FACEBOOK (Mar. 20, 2019), <https://about.fb.com/news/2019/03/technical-update-on-new-zealand/>.

status.”<sup>102</sup> Indeed, YouTube’s policy goes beyond Germany’s incitement to hatred laws by prohibiting speech based on message alone and without reference to the potential for real-world impact.<sup>103</sup> Twitter’s Hateful Conduct Policy hews more closely to the First Amendment and censors tweets that “promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease.”<sup>104</sup> Except for hateful images or symbols in a profile image or header that express hate towards a person, group, or protected category, Twitter’s rules allow hate speech as long as it does not incite violence or attack someone on the basis of their membership in a group.<sup>105</sup>

Of the mainstream social media companies, Facebook’s definition of hate speech is the broadest and is implemented by the most aggressive procedures for moderating content. Facebook defines hate speech as “a direct attack on people based on what we call protected characteristics—race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability.”<sup>106</sup> Facebook also provides “some protections for immigration status.”<sup>107</sup> It defines an attack “as violent or dehumanizing speech, statements of inferiority, or calls for exclusion or segregation.”<sup>108</sup>

Facebook separates attacks into three tiers: Tier 1 attacks are those that target a person or group of people who share one of the above-listed characteristics or immigration status, when the attack is violent or dehumanizing speech.<sup>109</sup> Dehumanizing speech involves referring to or comparing the target to, among other things, insects, animals, filth, disease, or criminals.<sup>110</sup> Tier 2 attacks target persons or groups sharing the above-listed characteristics, when the attack includes statements of inferiority or an image implying a person’s or a group’s physical, mental, or moral deficiency or expressions of contempt or disgust.<sup>111</sup> Tier 3 attacks

---

<sup>102</sup> *Hate Speech Policy*, YOUTUBE, <https://support.google.com/youtube/answer/2801939?hl=en> (last visited Jan. 16, 2020).

<sup>103</sup> See, e.g., Strafgesetzbuch [StGB] [Penal Code], June 19, 2019, BGBL I at 844, § 130, no. 1 (Ger.), [https://www.gesetze-im-internet.de/englisch\\_stgb/englisch\\_stgb.html](https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html) (discussing penalties for “[w]hoever, in a manner which is suitable for causing a disturbance of the public peace, incites hatred against a national, racial, religious group or a group defined by their ethnic origin”).

<sup>104</sup> *Hateful Conduct Policy*, TWITTER, <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (last visited Jan. 16, 2020).

<sup>105</sup> *Id.*

<sup>106</sup> *Hate Speech*, FACEBOOK COMMUNITY STANDARDS, [https://www.facebook.com/communitystandards/hate\\_speech](https://www.facebook.com/communitystandards/hate_speech) (last visited Jan. 16, 2020).

<sup>107</sup> *Id.*

<sup>108</sup> *Id.*

<sup>109</sup> *Id.*

<sup>110</sup> *Id.*

<sup>111</sup> *Id.*

are calls to exclude or segregate a person or group of people based on the above-listed characteristics and content that describes or negatively targets people with slurs; slurs are defined as words commonly used as “insulting labels for the above-listed characteristics.”<sup>112</sup>

Platform law is enforced by automated content tools and human moderators who assess content that has been reported by users.<sup>113</sup> Facebook has spent half a billion dollars on creating a regulatory structure that outsources the bulk of its content moderation to approximately 15,000 third-party contractors who employ over 35,000 moderators.<sup>114</sup> Harmful online content is classified through both automated moderation tools (algorithms) as well as user flagging.<sup>115</sup> Algorithms identify particular classes of harmful material such as child pornography, extremist content, and hate speech.<sup>116</sup> Additionally, Facebook and Twitter rely on users to identify harmful content, and each has created a flagging mechanism for users to tag content that is then sent to a human reviewer.<sup>117</sup>

The processes for identifying problematic content vary between Facebook and Twitter: Facebook allows users to report abusive content through a “Find Support or Report” link.<sup>118</sup> Twitter outlines how to respond to offensive posts on its “About Online Abuse” page, which suggests unfollowing and blocking abusive accounts and reporting posts to Twitter.<sup>119</sup>

At Facebook, a three-tier structure guides human content evaluation.<sup>120</sup> The bottom tier<sup>121</sup> consists of tens of thousands of moderators situated in various geographical locations around the world, clicking through hundreds of thousands of posts per day that have been flagged by users.<sup>122</sup>

<sup>112</sup> *Id.*

<sup>113</sup> See GILLESPIE, *supra* note 91, at 4.

<sup>114</sup> *Id.*; SARAH T. ROBERTS, BEHIND THE SCREEN: CONTENT MODERATION IN THE SHADOWS OF SOCIAL MEDIA (2019); David Gilbert, *Facebook Is Forcing Its Moderators to Log Every Second of Their Days — Even When They’re in the Bathroom*, VICE (Jan 9, 2020, 10:55 AM), [https://www.vice.com/en\\_ca/article/z3beea/facebook-moderators-lawsuit-ptsd-trauma-tracking-bathroom-breaks](https://www.vice.com/en_ca/article/z3beea/facebook-moderators-lawsuit-ptsd-trauma-tracking-bathroom-breaks); *Can Facebook Really Rely on Artificial Intelligence to Spot Abuse?*, BUS. DAY (Nov. 11, 2019), <https://businessday.ng/financial-times/article/can-facebook-really-rely-on-artificial-intelligence-to-spot-abuse/>.

<sup>115</sup> Kaye, *supra* note 89, ¶¶ 32–34.

<sup>116</sup> Press Release, Mark Zuckerberg, A Blueprint for Content Governance and Enforcement (Nov. 15, 2018), <https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>.

<sup>117</sup> *Id.*

<sup>118</sup> *How to Report Things on Facebook*, FACEBOOK, <https://www.facebook.com/help/181495968648557> (last visited Feb. 6, 2020).

<sup>119</sup> *About Online Abuse*, TWITTER, <https://help.twitter.com/en/safety-and-security/cyber-bullying-and-online-abuse> (last visited Feb. 6, 2020).

<sup>120</sup> Klonick, *supra* note 91, at 1639.

<sup>121</sup> *Id.* at 1639–40.

<sup>122</sup> Andrew Arshnt & Daniel Etcovitch, Commentary, *The Human Cost of Online Content Moderation*, JOLT DIG. (Mar. 2, 2018), <https://jolt.law.harvard.edu/digest/the-human-cost-of-online->

They are the first reviewers of reported content, judging its admissibility using Facebook's content moderation policies, and reviewing everything from attacks on race or sexual orientation to violent or pornographic content.<sup>123</sup> Tens of thousands of content moderators work under reportedly grim conditions, and they typically have under a minute to make a decision about whether an item of content violates the platform's rules.<sup>124</sup>

Facebook's moderators are instructed to follow an ever-changing document they call "the bible," a 10,000-word document that has twenty-four different categories, broken down into three groups: harmful behavior, sensitive content, and legal violations.<sup>125</sup> If a moderator is unsure whether a post is appropriate, they will escalate it to a more senior moderator, who is usually the supervisor of a content moderation center and has more experience in the field.<sup>126</sup> Finally, if that moderator still cannot determine whether the post is appropriate, the matter is escalated again, often to lawyers or policy makers at the main Facebook headquarters in Palo Alto, California.<sup>127</sup>

Over time, the regulatory reach and capacity of platform law has ratcheted up inexorably. It is worth noting that no rules or mechanisms governing the regulation of hate speech have been reversed or dismantled after they were designed and implemented. Furthermore, in 2018, 2019, and 2020, platforms made a flurry of policy changes, leading to new and even more protective community guidelines on hate speech. Google is at the forefront of addressing hate on its YouTube platform and made more than thirty policy updates in 2018 alone.<sup>128</sup> YouTube's policy changes had tangible consequences, including the removal of Alex Jones' Infowars Channel for hateful posts about Muslims and transgender people.<sup>129</sup> In June 2019, YouTube announced plans to remove thousands of videos and channels that advocated neo-Nazism, and in December it banned videos that "maliciously insult" people based on race, gender, or sexual orientation.<sup>130</sup> Also in 2019, Facebook decided to remove all white

---

content-moderation.

<sup>123</sup> Klonick, *supra* note 91, at 1640.

<sup>124</sup> Gilbert, *supra* note 114.

<sup>125</sup> *Id.*

<sup>126</sup> Klonick, *supra* note 91, at 1640–41.

<sup>127</sup> *Id.*

<sup>128</sup> *Our Ongoing Work to Tackle Hate*, YOUTUBE OFFICIAL BLOG (June 5, 2019), <https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>.

<sup>129</sup> Janko Roettgers, *Following Apple and Facebook, YouTube Bans Alex Jones, Infowars Channels*, VARIETY (Aug. 6, 2018), <https://variety.com/2018/digital/news/youtube-alex-jones-ban-1202896074/>.

<sup>130</sup> Mark Di Stefano, *YouTube Will Ban Videos That "Maliciously Insult" People Based on Race, Gender, or Sexual Orientation*, BUZZFEED NEWS (Dec. 11, 2019, 9:00 AM), <https://www.buzzfeed.com/markdistefano/youtube-harassment-guidelines-crowder-maza>; Kevin Roose & Kate Conger, *YouTube to Remove Thousands of Videos Pushing Extreme Views*, N.Y. TIMES (June 5,

nationalist groups' pages, depriving them of the ability to promote white supremacist ideas on a mainstream platform.<sup>131</sup> At around the same time, Twitter banned speech that dehumanizes a group of people based on their religion.<sup>132</sup>

In 2020, bowing to pressure to create an impartial regulatory mechanism to review its content removal decisions, Facebook created an oversight board managed by an independent trust.<sup>133</sup> The board of forty members<sup>134</sup> is empowered to review individual cases of content removal referred by Facebook or brought by users who disagree with Facebook's decision to remove their post.<sup>135</sup>

Some commentators hailed the oversight board as a "Supreme Court"<sup>136</sup> that will be "completely autonomous to Facebook"<sup>137</sup> and "can overrule Zuckerberg."<sup>138</sup> However, its authority will be quite limited, and its autonomy is circumscribed in important ways. Specifically, there is no provision in the bylaws for the board to review material that remains posted on the platform, only content that Facebook has taken down.<sup>139</sup> Further, the board lacks the authority to review Facebook's content moderation policies or procedures that guide removal decisions, including the algorithm that curates users' feeds and prioritizes hateful content.<sup>140</sup> In addition, the board's rulings must comply with "Facebook's content policies and values,"<sup>141</sup> which limits its ability to apply alternative sources of guidance, such as international human rights law.

---

2019), <https://www.nytimes.com/2019/06/05/business/youtube-remove-extremist-videos.html>.

<sup>131</sup> Liam Stack, *Facebook Announces New Policy to Ban White Nationalist Content*, N.Y. TIMES, (Mar. 27, 2019), <https://www.nytimes.com/2019/03/27/business/facebook-white-nationalist-supremacist.html>. See *Community Standards: Hate Speech*, FACEBOOK, [https://www.facebook.com/communitystandards/hate\\_speech](https://www.facebook.com/communitystandards/hate_speech) (last visited April 30, 2020) (setting out the policy that prohibits statements of inferiority or calls for exclusion or segregation).

<sup>132</sup> Emily Birnbaum, *Twitter Moves to Ban 'Dehumanizing' Language Toward Religious Groups*, HILL (July 9, 2019, 2:09 PM), <https://thehill.com/policy/technology/452226-twitter-expands-hateful-conduct-policy-to-include-dehumanizing-language>; *Hateful Conduct Policy*, TWITTER, <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (last visited Jan. 20, 2020).

<sup>133</sup> Facebook Oversight Board Bylaws, art. 4 (2020) [hereinafter Bylaws], [https://about.fb.com/wp-content/uploads/2020/01/Bylaws\\_v6.pdf](https://about.fb.com/wp-content/uploads/2020/01/Bylaws_v6.pdf).

<sup>134</sup> *Id.* art. 1, § 1.4.

<sup>135</sup> *Id.* art. 1, § 3 (Case Review and Decisions).

<sup>136</sup> Rebecca Bellan, *Facebook Announces Members of Its Supreme Court*, FORBES (May 6, 2020, 3:54 PM), <https://www.forbes.com/sites/rebeccabellan/2020/05/06/facebook-announces-members-of-its-supreme-court/#3759e11a3f12>.

<sup>137</sup> *Id.*

<sup>138</sup> Elizabeth Culliford, *Facebook Names First Members of Oversight Board That Can Overrule Zuckerberg*, REUTERS (May 6, 2020, 1:12 PM), <https://www.reuters.com/article/us-facebook-oversight/facebook-names-first-members-of-oversight-board-that-can-overrule-zuckerberg-idUSKBN2212LQ>.

<sup>139</sup> Bylaws, art. 3.

<sup>140</sup> Müller & Schwarz, *supra* note 69, at 27, 42.

<sup>141</sup> Bylaws, art. 1, § 3.



The board's decisions are binding on just one case at a time and do not unambiguously establish a policy precedent for all similar cases, a power that would be essential to ensuring that the board's decisions would have policy impact. Facebook must implement a board decision within a week,<sup>142</sup> unless the decision would be illegal according to national law. It is unclear how the board and Facebook will navigate jurisdictional issues surrounding national laws on hate speech and disinformation.<sup>143</sup>

The board's authority is limited in both scope and speed. The board cannot review content on Facebook Dating or Marketplace, or messages sent through WhatsApp, Messenger, or Instagram Direct, all services owned by Facebook.<sup>144</sup> Since its review process can take ninety days,<sup>145</sup> the board cannot act expeditiously to tackle viral content contributing to imminent public violence, as occurred during the genocide of the Rohingya.<sup>146</sup> The bylaws make allowance for expedited review requests which might speed the process up, but "expedited" can still mean up to thirty days.<sup>147</sup> Finally, the board can amend certain sections of the bylaws, but on the fundamental question of case review procedures (i.e., which cases are selected and how they are reviewed), Facebook reserves its right to veto any board amendments it disagrees with.<sup>148</sup>

As a result of these substantial restrictions, some critics have written off the oversight board as "toothless"<sup>149</sup> and inherently unable to address the source of Facebook's content moderation problems, namely its "attention economy"<sup>150</sup> business model that incentivizes contentious and hateful speech to sell more advertising.<sup>151</sup> They have also portrayed the board as a half-measure to forestall government oversight and anti-trust investigations, an assertion that seemed vindicated in May 2020 when

---

<sup>142</sup> *Id.* art. 2, § 2.3.1.

<sup>143</sup> See Jennifer C. Daskal, *Speech Across Borders*, 105 VA. L. REV. 1605, 1607–08 (2019) (discussing how social media companies navigate conflicting speech regulation regimes across national borders).

<sup>144</sup> Bylaws, art. 2, § 1.2.1.

<sup>145</sup> *Id.* art. 1, § 3.1 (Case Review Procedures and Timeline).

<sup>146</sup> Myanmar Report, *supra* note 77.

<sup>147</sup> Bylaws, art. 2, § 2.1.2.

<sup>148</sup> *Id.*, art. 5, § 1.

<sup>149</sup> Josh Constine, *Toothless: Facebook Proposes a Weak Oversight Board*, TECHCRUNCH (Jan. 28, 2020, 4:33 PM), <https://techcrunch.com/2020/01/28/under-consideration/>.

<sup>150</sup> Andrew Keen, *The 'Attention Economy' Created by Silicon Valley Is Bankrupting Us*, TECHCRUNCH (July 30, 2017, 6:30 PM), <https://techcrunch.com/2017/07/30/the-attention-economy-created-by-silicon-valley-is-bankrupting-us/>.

<sup>151</sup> Dipayan Ghosh, *Facebook's Oversight Board Is Not Enough*, HARV. BUS. REV. (Oct. 16, 2019), <https://hbr.org/2019/10/facebooks-oversight-board-is-not-enough>.

Facebook launched an advocacy group called “American Edge” to lobby in Washington against regulation of the tech sector.<sup>152</sup>

### B. *The Limits of Platform Law*

Social media companies today prohibit hate speech on their platforms. If promulgated by federal, state, or municipal authorities in the United States, platform law would undoubtedly violate the First Amendment, which prevents suppression of speech based on its content or viewpoint.<sup>153</sup> The U.S. government could not ban speech that “maliciously insults”<sup>154</sup> a person based on their sexual orientation (YouTube),<sup>155</sup> or forbid speech that dehumanizes a group based on religion (Twitter),<sup>156</sup> since both would violate constitutional prohibitions on content or viewpoint discrimination. But despite Strossen’s call for platforms to refrain from removing “expression that the First Amendment shields from government censorship,”<sup>157</sup> the companies have been charting a different course guided more by social norms than law. Social media platforms have created systems for suppressing speech—including speech protected under national and international law—in what has functionally become our public square. Indeed, according to the platforms, approximately 90% of hate speech is currently removed by automated evaluation tools even before it is seen by viewers.<sup>158</sup>

Given that these platforms operate globally and a majority of their users are located outside the United States,<sup>159</sup> a U.S.-centric approach to regulation of speech would not be appropriate, effective, or legitimate.

---

<sup>152</sup> Tony Romm, *Facebook Is Quietly Helping to Set Up a New Pro-tech Advocacy Group to Battle Washington*, WASH. POST, (May 12, 2020, 6:00 AM), [https://www.washingtonpost.com/technology/2020/05/12/facebook-lobbying-american-edge/?mc\\_cid=](https://www.washingtonpost.com/technology/2020/05/12/facebook-lobbying-american-edge/?mc_cid=)

<sup>153</sup> See *Reed v. Town of Gilbert*, 135 S. Ct. 2218, 2222 (2015) (holding that content-based laws are presumptively unconstitutional); *R.A.V. v. City of St. Paul*, 505 U.S. 377, 391 (1992) (reversing the City of St. Paul’s censoring of speech on the basis of bias-motivated content because it only placed prohibitions on speech that discriminated on the basis of “race, color, creed, religion or gender”); *Bible Believers v. Wayne Cty.*, 805 F.3d 228, 247 (6th Cir. 2015) (considering Wayne County’s suppression of an evangelical group’s anti-Muslim speech).

<sup>154</sup> *Harassment and Cyberbullying Policy*, YOUTUBE, [https://support.google.com/youtube/answer/2802268?hl=en&ref\\_topic=9282436](https://support.google.com/youtube/answer/2802268?hl=en&ref_topic=9282436) (last visited July 12, 2020).

<sup>155</sup> *Hate Speech Policy*, YOUTUBE, <https://support.google.com/youtube/answer/2801939?hl=en> (last visited July 12, 2020).

<sup>156</sup> *Hateful Conduct Policy*, TWITTER, <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (last visited July 12, 2020).

<sup>157</sup> STROSSEN, *supra* note 26, at 31.

<sup>158</sup> *Community Standards Enforcement Report*, FACEBOOK, <https://transparency.facebook.com/community-standards-enforcement#hate-speech> (last visited May 26, 2020).

<sup>159</sup> *Leading Countries Based on Facebook Audience Size as of April 2020*, STATISTA, <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/> (last visited Apr. 30, 2020). Only 189 million out of 1.623 billion (or 11%) of Facebook’s daily active users are based in the United States and Canada. *Facebook Q3 2019 Results*, *supra* note 9.

Further, although in the United States there is no legal compulsion for companies to limit hate speech—under § 230 of the Communications Decency Act of 1996, platforms cannot be held liable for content posted on their platforms<sup>160</sup>—the laws of other jurisdictions are increasingly imposing such liability. In addition, there are a host of commercial reasons for companies to moderate. Moderation of uncomfortable speech—as much as First Amendment advocates may resist it—is part and parcel of the service that social media companies offer. As Tarleton Gillespie explains, “moderation, far from being occasional or ancillary, is in fact an essential, constant, and definitional part of what platforms do. Moderation is the essence of platforms. It is the commodity they offer. It is their central value proposition.”<sup>161</sup> Platforms remove hate speech to create the environment they believe their users want and to comply with legislation in the national jurisdictions in which they operate.

Moreover, this ever-increasing prohibition on hate speech in platform law coincides with shifts in public attitudes in the United States. Successive surveys indicate that U.S. users broadly endorse what the companies are doing. In a 2018 survey in the United States conducted by the Anti-Defamation League, fifty-nine percent of respondents believed that “online hate and harassment make hate crimes more common.”<sup>162</sup> A survey of 1500 undergraduate students at public and private universities in the United States found that fewer than half of respondents believed that the First Amendment protected “hate speech.”<sup>163</sup> There was little variation by political affiliation or type of college (private v. public) but significant gender variation, with only thirty-one percent of females versus fifty-one percent of males indicating a belief that hate speech was constitutionally protected.<sup>164</sup> These results have been confirmed by subsequent surveys of college students, with one finding that only thirty-five percent of students think hate speech should be protected by the First Amendment.<sup>165</sup> Almost

---

<sup>160</sup> See Citron & Wittes, *supra* note 93, at 402–04 (reviewing the implications of § 230).

<sup>161</sup> Tarleton Gillespie, *Platforms Are Not Intermediaries*, 2 GEO. L. TECH. REV. 198, 201 (2018).

<sup>162</sup> Zachary Laub, *Hate Speech on Social Media: Global Comparisons*, COUNCIL ON FOREIGN REL. (June 7, 2019), <https://www.cfr.org/backgroundunder/hate-speech-social-media-global-comparisons>.

<sup>163</sup> John Villasenor, *Views Among College Students Regarding the First Amendment: Results from a New Survey*, BROOKINGS INST. (Sept. 18, 2017), <https://www.brookings.edu/blog/fixgov/2017/09/18/views-among-college-students-regarding-the-first-amendment-results-from-a-new-survey/>. The survey indicates a margin of error of two to six percent. The survey was funded by the conservative Charles Koch Foundation and has been reviewed critically, for instance by Lois Beckett, “*Junk Science*”: *Experts Cast Doubt on Widely Cited College Free Speech Survey*, GUARDIAN (Sept. 22, 2017), <https://www.theguardian.com/us-news/2017/sep/22/college-free-speech-violence-survey-junk-science>. Whatever its flaws, the Villasenor survey is considered here because its results are mostly confirmed by the subsequent FIRE survey. See *infra* note 165.

<sup>164</sup> Villasenor, *supra* note 163.

<sup>165</sup> KELSEY ANN NAUGHTON ET AL., FOUND. FOR INDIVIDUAL RIGHTS IN EDUC. (FIRE), *SPEAKING FREELY: WHAT STUDENTS THINK ABOUT EXPRESSION AT AMERICAN COLLEGES* 5, 17 (2017).

one-half of students (forty-eight percent) think the First Amendment should not protect hate speech,<sup>166</sup> and a majority (fifty-six percent) “agree that there are times when a college or university should withdraw a guest speaker’s invitation after the event has been announced.”<sup>167</sup>

Thus, the emergence of robust hate speech regulation on social media is not an aberration but a product of broader cultural shifts. Platform law has been driven more by the social norms of users and the desire of the companies to bridge U.S. and European approaches to hate speech than by the constitutional jurisprudence of the First Amendment.<sup>168</sup> Hate speech regulation by platforms in turn shapes social norms by helping to normalize regulation in the minds of social media users. Legal theorists need to be aware of these broad social dynamics since changes in real-world speech regulation on social media do not exclusively emanate from the rarified domain of statutes and case law.

Nonetheless, one need not be a free speech absolutist to be concerned about the emergent system of platform law of hate speech. Two concerns will be discussed here. First, even those who have advocated regulation of hate speech have not proposed that this regulation be carried out by private actors unaccountable to the public for their decisions. Many observers have expressed unease with the fact that speech is regulated by unaccountable private companies whose primary obligations are neither to voters nor to society at large, but to shareholders.<sup>169</sup> Rather than constrain this exercise of power, national governments are compounding it further by outsourcing their speech regulating powers to private companies with little or no oversight.<sup>170</sup> For example, in 2017, Germany passed the NetzDG (Network Enforcement Law)<sup>171</sup> that established measures imposing fines up to €50 million if platforms fail to remove speech that is “manifestly unlawful”<sup>172</sup> under German law (including hate speech, defamation, and inciting speech) from their platform in as little as twenty-four hours.<sup>173</sup>

---

<sup>166</sup> *Id.* at 17.

<sup>167</sup> *Id.* at 15.

<sup>168</sup> Kaye, *supra* note 89.

<sup>169</sup> See, e.g., Luca Belli et al., *Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police*, in PLATFORM REGULATIONS: HOW PLATFORMS ARE REGULATED AND HOW THEY REGULATE US 41, 53 (Luca Belli & Nicolo Zingales eds., 2017) (arguing that social media companies have too much discretion in enforcing EU Code of Conduct for online hate speech); Molly K. Land, *Against Privatized Censorship: Proposals for Responsible Delegation*, 60 VA. J. INT’L L. 363, 399–415 (2020) (examining the legality of privatized censorship).

<sup>170</sup> Land, *supra* note 169, at 412–13.

<sup>171</sup> The full name of the NetzDG is the *Netzwerkdurchsetzungsgesetz*. For a German-language version of the statute, see *Beschlussesempfehlung und Bericht [Resolution and Report]*, Deutscher Bundestag: Drucksache [BT] 18/13013 (Ger.), <http://dip21.bundestag.de/dip21/btd/18/130/1813013.pdf>.

<sup>172</sup> *Id.* § 3(2).

<sup>173</sup> Soraya Sarhaddi Nelson, *With Huge Fines, German Law Pushes Social Networks to Delete Abusive Posts*, NPR (Oct. 31, 2017, 7:44 AM), <https://www.npr.org/sections/parallels/2017/10/31/>

Then there are legitimate concerns about the process of regulating hate speech itself. David Kaye states that content policies must be consistent with the principle of legality, in which “[t]he restriction must be provided by laws that are precise, public and transparent.”<sup>174</sup> Platform law regulating hate speech, however, is anything but transparent. Although social media companies openly post their hate speech policies, they do not explain how their policies are translated into lines of operational machine code that identify and remove posts before they even appear on the platform. The use of artificial intelligence (AI) by social media companies for content moderation remains a black box. This matters because it has been widely observed that AI is poor at taking linguistic, cultural, and societal context into account and its design often reproduces many of the social biases and discriminatory schema that it is meant to regulate.<sup>175</sup> As Tim Wu wryly observes, AI is “inherently inhuman, and prone, at least for the foreseeable future, to make absurd errors that can be funny, horrific, or both.”<sup>176</sup>

We know little about how companies make determinations in hard cases. For instance, during the COVID-19 pandemic, Facebook announced that its depleted pool of content reviewers would prioritize content “related to real-world harm,”<sup>177</sup> but we do not know what Facebook’s criteria are for determining speech that is most likely to cause harm, or how these criteria are operationalized through automated tools. The process of turning policy into an ex ante content removal application is currently opaque and not open to independent scrutiny. Without more information about how decisions are made—such as, for example, through Kaye’s suggestion that companies publish their “caselaw”—it is difficult for users to know what rules are actually governing their speech, and how automated tools enforce these rules to remove content before users even see it.<sup>178</sup>

A little more is known about human content evaluation systems, but it is clear that they are not a panacea to the ills of hate speech either. The working conditions of the tens of thousands of content moderators located in warehouses in the Philippines, Ukraine, Ireland, and other places have

---

561024666/with-huge-fines-german-law-pushes-social-networks-to-delete-abusive-posts/.

<sup>174</sup> Kaye, *supra* note 30, ¶ 6.

<sup>175</sup> See Tim Wu, *Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems*, 119 COLUM. L. REV. 2001, 2013–18 (2019) (reviewing the role of AI in content moderation); David Kaye, *Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 32–38, U.N. Doc. A/73/348 (Aug. 29, 2018) (discussing the propensity of AI to perpetuate bias and discrimination). See also Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PENN. L. REV. 1445, 1448–52 (2013) (reviewing the implications of AI for speech regulation more generally).

<sup>176</sup> Wu, *supra* note 175, at 2021.

<sup>177</sup> *Update on Content Review Work*, FACEBOOK (Apr. 21, 2020, 3:30 PM), <https://about.fb.com/news/2020/04/coronavirus/>.

<sup>178</sup> David Kaye, *Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 63, U.N. Doc. A/HRC/32/38 (May 11, 2016).

been studied by researchers, and the conditions are distressing.<sup>179</sup> Silicon Valley moderators, paid the federal minimum wage of fifteen dollars per hour without any health care or retirement benefits, are said to have the “worst job in technology.”<sup>180</sup> Sarah Roberts documents the deleterious psychological effects of viewing child pornography, hate speech against marginalised groups, and users encouraging other users to commit suicide for eight hours a day, as well as how in 2018 and 2019 a group of former content moderators filed lawsuits in California and Dublin, Ireland against Facebook.<sup>181</sup>

There are also legitimate concerns about the accuracy of content policy enforcement. According to a study by ProPublica, human moderators inconsistently apply the guidelines in “the bible” of Facebook content evaluation.<sup>182</sup> ProPublica conducted a crowd-sourced study of 900 posts flagged by users as potential violations of Facebook’s Community Guidelines and asked the company about a random sample of forty-nine posts.<sup>183</sup> The company admitted that moderators made mistakes in their review of nearly half of them (twenty-two).<sup>184</sup>

Further, governments can and do take advantage of the real problems presented by online hate speech and disinformation in order to control social media and, in so doing, stifle public dissent.<sup>185</sup> Digital authoritarianism is a real and growing phenomenon. There are more and more examples with each passing day, as outright authoritarian governments or illiberal democracies have realized that, in order to win the propaganda war, they need to control social media.<sup>186</sup> For example,

---

<sup>179</sup> See, e.g., GILLESPIE, *supra* note 91, at 122 (discussing “grim” work conditions for content moderators in the Philippines); ROBERTS, *supra* note 114, at 172–73 (discussing the demanding and fast-paced work for content moderators in the Philippines); Andrew Arshat & Daniel Etcovitich, *The Human Cost of Online Content Moderation*, JOLT DIG. (Mar. 2, 2018), <https://jolt.law.harvard.edu/digest/the-human-cost-of-online-content-moderation> (discussing the psychological risks involved in content moderation).

<sup>180</sup> Queenie Wong, *Facebook Content Moderation Is an Ugly Business. Here’s Who Does It*, CNET (June 19, 2019), <https://www.cnet.com/news/facebook-content-moderation-is-an-ugly-business-heres-who-does-it/>.

<sup>181</sup> ROBERTS, *supra* note 114, at 160–65. The California case is *Scola v. Facebook Inc.* See Complaint for Declaratory & Injunctive Relief, *Scola v. Facebook Inc.*, No. 18CIV05135 (Cal. Super. Ct. Sept. 21, 2018). Olivia Miller, *Former Content Moderators File a Lawsuit in Dublin Against Facebook*, TECHNOWIZE (Dec. 10, 2019), <https://www.technowize.com/former-content-moderators-file-a-lawsuit-in-dublin-against-facebook/>.

<sup>182</sup> Ariana Tobin et al., *Facebook’s Uneven Enforcement of Hate Speech Rules Allows Vile Posts to Stay Up*, PROPUBLICA (Dec. 28, 2017, 5:53 PM), <https://www.propublica.org/article/facebook-enforcement-hate-speech-rules-mistakes>.

<sup>183</sup> *Id.*

<sup>184</sup> *Id.*

<sup>185</sup> Kaye, *supra* note 30, ¶ 1 (opining that the vagueness and the lack of consensus around the meaning of hate speech can be abused by governments).

<sup>186</sup> The phrase “illiberal democracy” was coined by Fareed Zakaria to refer to countries that hold elections, but where citizens do not enjoy basic civil liberties such as freedom of expression and do not

Singapore's recent crackdown on fake news has garnered significant controversy. The Protection from Online Falsehoods and Manipulation Act came into effect in Singapore in late 2019.<sup>187</sup> One of its first official measures was to compel Facebook to include a disclaimer at the bottom of a post accusing the government of running rigged elections stating, "[Facebook] is legally required to tell you that the Singapore government says this post has false information."<sup>188</sup> The Act also bans the use of fake accounts or bots to spread "false statements," and imposes penalties of over \$700,000 and/or a prison sentence of up to six years.<sup>189</sup> In Nigeria, a bill called the "Protection from Internet Falsehoods and Manipulation and Other Related Matters Bill 2019," which is almost identical to the Singapore Act, has catalyzed significant opposition.<sup>190</sup> Around the world, broad definitions of hate speech and other prohibited content in platform law are used by governments to suppress the speech of activists and minority groups.<sup>191</sup>

The current regime of hate speech regulation by social media platforms is also flawed because it is deeply acontextual. This regime is more aggressive than the approach advocated by Critical Race Theorists such as Matsuda and Lawrence thirty years ago, when they defended municipal hate crimes ordinances.<sup>192</sup> Certainly, the basic outlines of platform hate speech law prohibit the kind of speech that these theorists were critiquing. For example, Facebook's platform law prohibits dehumanizing speech and statements of inferiority based on race.<sup>193</sup> Twitter's policy against dehumanizing speech prohibits users from "[d]irect[ing] hate against a

---

have access to the information that is required to make fully informed democratic choices. Fareed Zakaria, *The Rise of Illiberal Democracy*, 76 FOREIGN AFF. 22, 22–23 (1997).

<sup>187</sup> Protection from Online Falsehoods and Manipulation Act of 2019 (Bill No. 10/2019) (Sing.), <https://sso.agc.gov.sg/Bills-Supp/10-2019/Published/20190401?DocDate=20190401#pr8->.

<sup>188</sup> *Facebook Bows to Singapore's 'Fake News' Law with Post 'Correction'*, BBC NEWS (Nov. 30, 2019), <https://www.bbc.com/news/world-asia-50613341>.

<sup>189</sup> Protection from Online Falsehoods and Manipulation Act of 2019, sec. 8 (Sing.).

<sup>190</sup> POLICY & LEGAL ADVOCACY CTR., PROTECTION FROM INTERNET FALSEHOODS AND MANIPULATION AND OTHER RELATED MATTERS BILL, 2019 (SB 132) (2019), <https://placng.org/wp/wp-content/uploads/2019/11/Protection-from-Internet-Falsehood-Bill-Summary.pdf>; Alexander O. Onukwue, *Nigeria's Social Media Bill Suffers Its Biggest Defeat Yet*, TECHCABAL (Mar. 9, 2020), <https://techcabal.com/2020/03/09/social-media-bill-public-hearing/>; Emmanuel Paul, *Everything You Need to Know About Nigeria's Social Media Bill and What You Can Do About It*, TECHPOINT.AFRICA (Nov. 28, 2019), <https://techpoint.africa/2019/11/28/nigerias-social-media-bill/>.

<sup>191</sup> Casey Newton, *Why Activists Get Frustrated with Facebook: When Political Speech Gets Removed, They Have Little Recourse—And a Lot to Lose*, VERGE (Jan. 14, 2020, 6:00 AM), <https://www.theverge.com/interface/2020/1/14/21063887/activists-facebook-iran-free-speech-authoritarianism>.

<sup>192</sup> Matsuda, *supra* note 18, at 2360.

<sup>193</sup> *Hate Speech*, FACEBOOK COMMUNITY STANDARDS, [https://www.facebook.com/communitystandards/hate\\_speech](https://www.facebook.com/communitystandards/hate_speech) (last visited Jan. 16, 2020).

protected category (e.g., race, religion, gender, orientation, disability).<sup>194</sup> A Twitter user who encounters a toxic tweet “can report abusive or harmful content directly from a tweet or profile.”<sup>195</sup> These policies resonate with the arguments of Charles Lawrence, who urged regulation of “group defamation”<sup>196</sup>—“insulting words” that are “aimed at an entire group with the effect of causing significant harm to individual group members.”<sup>197</sup>

However, Facebook’s policies are far broader than anything suggested by Lawrence or Matsuda. A central idea of Critical Race Theory is “skepticism toward dominant legal claims of neutrality, objectivity, color blindness, and meritocracy.”<sup>198</sup> Critical race scholars reject false equivalences and seek to recognize and embed in regulation a recognition of power. Matsuda, for example, advocated prohibition of speech only when it meets three criteria: “1. The message is of racial inferiority[;] 2. The message is directed against a historically oppressed group[; and] 3. The message is persecutory, hateful, and degrading.”<sup>199</sup> The second element, that the speech be directed to a vulnerable group, “recogniz[es] the connection of racism to power and subordination.”<sup>200</sup> These limiting principles—treating as particularly harmful speech that is directed to a member of a vulnerable group as opposed to any speech that is hateful regardless of its target—“narrows the field of interference with speech.”<sup>201</sup>

In support of this limiting principle, Lawrence argued that the impact of speech crucially depends on the status of the listener to whom it is directed. He recounts the story of one of his students who was targeted by homophobic insults that left the student “in a state of semishock, nauseous, dizzy, unable to muster the witty, sarcastic, articulate rejoinder he was accustomed to making.”<sup>202</sup> When asked if he reacted similarly to insults directed against him on the basis of his membership in a less vulnerable group—for example, insults such as “honkey,” “chauvinist pig,” or “mick”—the student said that he had not experienced the same effects. Lawrence notes: “The question of power, of the context of the power

<sup>194</sup> *Someone on Twitter Is Engaging in Abusive or Harassing Behavior*, TWITTER, <https://help.twitter.com/forms/abusiveuser> (last visited Mar. 1, 2020).

<sup>195</sup> *Toxic Twitter—The Reporting Process*, AMNESTY INT’L (2018), <https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-4/>.

<sup>196</sup> Charles R. Lawrence III, *If He Hollers Let Him Go: Regulating Racist Speech on Campus*, in WORDS THAT WOUND: CRITICAL RACE THEORY, ASSAULTIVE SPEECH, AND THE FIRST AMENDMENT 53, 75 (Mari J. Matsuda et al. eds., 1993).

<sup>197</sup> *Id.* at 55.

<sup>198</sup> Lawrence et al., *supra* note 13, at 6.

<sup>199</sup> Mari J. Matsuda, *Public Response to Racist Speech: Considering the Victim’s Story*, in WORDS THAT WOUND: CRITICAL RACE THEORY, ASSAULTIVE SPEECH, AND THE FIRST AMENDMENT 17, 36 (Mari J. Matsuda et al. eds., 1993).

<sup>200</sup> *Id.*

<sup>201</sup> *Id.*

<sup>202</sup> Lawrence, *supra* note 196, at 70.



relationships within which speech takes place, and the connection to violence must be considered as we decide how best to foster the freest and fullest dialogue within our communities.”<sup>203</sup>

Social media prohibitions on hate speech, in contrast, prohibit speech based on content rather than context. There is no consideration of local context, country- or region-specific meanings, the identity of the speaker or the target, or the relationship between speaker and listener. For example, under Facebook’s hate speech policy, “white people” constitute a protected group. An attack on someone based on their membership in a group defined by whiteness triggers Facebook’s hate speech policy because, under this policy, it is an attack based on race.

This deliberate disregard of power and context leads to overbroad and even irrational results. For example, because of the unfortunate way in which Facebook operationalized its policy, this meant that at one point in time, posts about “white men” would violate its community standards, while posts about “black children” would not.<sup>204</sup> (In short, this was because “men” is a protected category but “children” is not—thus, a post about “white men” could be a slur about a racial category while a post about “black children,” a subset of an unprotected category, was not.) For those seeking to talk about racism—which may require discussing the actions of white people—Facebook’s approach is a real barrier.<sup>205</sup> Definitions of hate speech that try to treat all speech as equivalent without regard to context end up taking down speech that—while perhaps uncomfortable—is also essential for important issues of public discourse.

Given the extensive regulation that now applies to speech on platforms, the debate about whether or not to regulate hate speech in these domains is for all practical purposes over. Now, it is more a matter of how speech will be restricted and with what consequences. What is urgently needed, however, is a recalibration of platform hate speech policies to integrate context and thus narrow its scope to prohibit only the most harmful instances of hate speech and to protect robust freedom of expression. The next part discusses how this might be accomplished.

---

<sup>203</sup> *Id.*

<sup>204</sup> Julia Angwin & Hannes Grassegger, *Facebook’s Secret Censorship Rules Protect White Men from Hate Speech But Not Black Children*, PROPUBLICA (June 28, 2017, 5:00 AM), <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms> [<https://perma.cc/ZAA4-KQSF>].

<sup>205</sup> Jessica Gynn, *Facebook While Black: Users Call It Getting ‘Zucked,’ Say Talking About Racism Is Censored as Hate Speech*, USA TODAY (Apr. 30, 2019, 2:23 PM), <https://www.usatoday.com/story/news/2019/04/24/facebook-while-black-zucked-users-say-they-get-blocked-racism-discussion/2859593002/>.

#### IV. TOWARDS CONTEXT-SPECIFIC CONTENT MODERATION

Although limits on hate speech on social media platforms may be inevitable at this point in time, the approach that companies have adopted to implement these limits is simultaneously overbroad and under-inclusive. Because they don't consider the context of speech, SMCs end up restricting speech that is either benign or in the public interest (or both), while missing pervasive denigrating speech with corrosive environmental effects. The purpose of this Part is to discuss the importance of context in distinguishing between speech that is harmful and that which is merely uncomfortable; describe the ways in which platform law fails to consider context; and provide recommendations for contextual moderation.

##### A. *The Role of Context*

In understanding the meaning of a speech act, context is everything. This is evident in both U.S. and international approaches to freedom of expression. David Kaye has called on platforms to use international human rights law as the basis for their approach to content moderation.<sup>206</sup> Given that a majority of users are located outside the United States,<sup>207</sup> and recognizing the wide variation in national laws on freedom of expression, human rights law provides a helpful and practical baseline.

International law regarding speech regulation requires consideration of context. In his report to the UN General Assembly regarding online hate speech, Kaye recommends that “any enforcement of hate speech rules involves an evaluation of context and the harm that the content imposes on users and the public.”<sup>208</sup> In particular, context is essential to assessing the impact of speech. Human rights law requires that any limits on speech be proportional to the legitimate ends that they serve,<sup>209</sup> and determining

---

<sup>206</sup> Kaye, *supra* note 178, ¶¶ 5–14. See Evelyn Mary Aswad, *The Future of Freedom of Expression Online*, 17 DUKE L. & TECH. REV. 26, 42–57 (2018) (applying international human rights law to the content moderation policies of social media companies).

<sup>207</sup> *Leading Countries Based on Facebook Audience Size as of April 2020*, STATISTA, <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/> (last visited Apr. 30, 2020). Only 189 million out of 1.623 billion (or 11%) of Facebook's daily active users are based in the United States and Canada. *Facebook Q3 2019 Results*, *supra* note 9.

<sup>208</sup> Kaye, *supra* note 178, ¶ 58(d); see also Barrie Sander, *Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation*, 43 FORDHAM INT'L L.J. 939, 979 (2020) (explaining that context is important in assessing the necessity of the limitation); Michael Karanicolas, *Squaring the Circle Between Freedom of Expression and Platform Law*, 20 PITT. J. TECH. L. & POL'Y 175, 197 (2020) (“[I]nternational courts have repeatedly found that assessing whether something is hate speech is an inherently contextual determination, based on whether the material was released under circumstances where it is likely to actually generate harm.”).

<sup>209</sup> Frank La Rue (Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression), *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 24, U.N. Doc. A/HRC/17/27 (May 16, 2011).

proportionality requires assessing speech in its context.<sup>210</sup> We insist on this principle as critical to the conversation about online hate speech: the need to evaluate expression in the context of its usage. In our understanding, the full meaning and potential effect of any speech act lies with the intent of the speaker, the content of the expression, and the context in which it is uttered.

This view is consistent with the entire thrust of philosophy of language since the middle of the twentieth century. In the ordinary language philosophy of John L. Austin, Gilbert Ryle, and Ludwig Wittgenstein, meaning is not to be found in the semantic genealogy of an expression or its putative essence, but instead in its immediate and contextual usage in a community of speakers who follow the rules of particular “language-games.”<sup>211</sup> This theory of language is encapsulated in Wittgenstein’s statement that “the meaning of a word is its use in the language.”<sup>212</sup> Derrida extends the analysis further by noting that, like a signature, the meaning of a word or expression can never be fixed because it is dependent on the setting in which it is communicated, consumed, and (re)positioned, and its context constantly changes over time.<sup>213</sup>

Contextual analysis of an utterance has also been an essential component of First Amendment jurisprudence for the past one hundred years. In *Schenck*, Justice Oliver Wendell Holmes replaced the overly broad “bad tendency” standard with the more stringent clear and present danger test,<sup>214</sup> which asked “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.”<sup>215</sup> We understand Holmes’s emphasis on the circumstances of speech acts and their proximity to “substantial evils” as an instruction to consider speech in its context of usage.

In First Amendment jurisprudence, it is now received wisdom that “the character of every act depends upon the circumstances in which it is

---

<sup>210</sup> The Rabat Plan of Action, a report produced in 2013 by a group of human rights experts convened by the UN High Commissioner for Human Rights, emphasized in particular the role of context in determining whether speech was so harmful as to require criminalization; their factors required consideration of the “social and political context,” the “status of the speaker,” the “[e]xtent or reach of the speech act,” and the extent to which it was likely to result in harm. Kaye, *supra* note 185, ¶¶ 13–14.

<sup>211</sup> LUDWIG WITTGENSTEIN, *PHILOSOPHICAL INVESTIGATIONS* 14 (P.M.S. Hacker & Joachim Schulte eds., 4th ed. 2009).

<sup>212</sup> *Id.* at 25.

<sup>213</sup> Jacques Derrida, *Signature Event Context*, in *MARGINS OF PHILOSOPHY* 307 (Alan Bass trans., Univ. of Chi. Press 1982) (1971).

<sup>214</sup> *Schenck v. United States*, 249 U.S. 47, 52 (1919); *see also* *Abrams v. United States*, 250 U.S. 616, 627 (1919) (using the expression “clear and imminent danger”).

<sup>215</sup> *Schenck*, 249 U.S. at 52.

done.”<sup>216</sup> *Brandenburg v. Ohio* established that for an expression to constitute incitement, it had to be likely that it would cause imminent lawless action, thus requiring an analysis of the circumstances of inciting speech.<sup>217</sup> In a workplace milieu, courts have held that the determination of whether an employee’s speech is a matter of public concern depends on the “content, form, and context” of their statement.<sup>218</sup> Contextual evaluation has become increasingly relevant in recent years, as courts scramble to adapt First Amendment law to the fast-moving advertising, internet, and social media environment, and address the challenges raised by hate speech and terrorist propaganda.<sup>219</sup>

The reasons for a contextual analysis are plainly evident. The same words that are benign in a peaceful setting may constitute incitement in a violent situation: “[An] utterance in a context of violence can lose its significance as an appeal to reason and become part of an instrument of force.”<sup>220</sup> Because the impact of speech depends on the context in which it is uttered, it is not possible to respond to anti-Rohingya sentiment—in a country where an authoritarian military government is engaged in genocidal acts against that population—using the same framework that one adopts to tackle the use of racial slurs in conversations between individuals in an established democracy.<sup>221</sup>

## B. *Law Without Context*

Most platform policies readily acknowledge that context is relevant to content moderation. YouTube, for example, states that, “context matters, so some videos could remain up because they discuss topics like pending legislation, aim to condemn or expose hate, or provide analysis of current events.”<sup>222</sup> Twitter says that context is important when evaluating harassment and it needs to hear from the person being targeted to most effectively review this type of content.<sup>223</sup>

---

<sup>216</sup> *Id.*

<sup>217</sup> See *Brandenburg v. Ohio*, 395 U.S. 444, 447 (1969) (per curiam) (explaining that constitutional guarantees forbid states from prohibiting free speech except where such advocacy incites imminent lawless action); Wilson & Kiper, *supra* note 42, at 72–81 (discussing the contextual elements of incitement).

<sup>218</sup> *Connick v. Myers*, 461 U.S. 138, 147–48 (1983).

<sup>219</sup> Alexander Tsesis, *Social Media Accountability for Terrorist Propaganda*, 86 *FORDHAM L. REV.* 605, 628 (2017) (discussing the role of context in determining the liability of intermediaries for social media posts that incite terrorism).

<sup>220</sup> *Milk Wagon Drivers Union v. Meadowmoor Dairies, Inc.*, 312 U.S. 287, 293 (1941).

<sup>221</sup> Molly K. Land & Rebecca Hamilton, *Beyond Takedown: Expanding the Tool Kit for Responding to Online Hate*, in *PROPAGANDA AND INTERNATIONAL CRIMINAL LAW: FROM COGNITION TO CRIMINALITY* 143, 144 (Predrag Dojčinović ed., 2020).

<sup>222</sup> *Our Ongoing Work to Tackle Hate*, YOUTUBE OFFICIAL BLOG (June 5, 2019), <https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>.

<sup>223</sup> Twitter Rules, *Abusive Behavior*, TWITTER, <https://help.twitter.com/en/rules-and>

Practically, however, platform consideration of “context” is limited to what appears in the post itself. Thus, in distinguishing satire from a slur, Twitter’s rules focus on the words used in the post, but do not reference external information or events to which that post might pertain, the specific meaning those words might have in that region, or the local context in which the post originated. Such a limited definition of context renders the resulting assessment of harm vastly underdetermined.

In part, platform law is defined without reference to context because of the problem of scale. In their efforts to ensure consistent application of their standards, platforms have created ostensibly “objective” rules that are designed to be applied by any moderator, regardless of that moderator’s knowledge or expertise.<sup>224</sup> Under this approach, context can only be defined with reference to information that would be available to a moderator who does not have any pre-existing knowledge of the context. The companies admit how difficult it is to contextualize posts, and they have not yet found ways to more systematically integrate geographically localized meanings into their policies and practices.<sup>225</sup> Thus, contemporary content evaluation procedures, and especially their automated tools, are largely detached from their socio-cultural and political contexts. This is apparent in all of the main elements of content evaluation: user flagging, artificial intelligence, and the growing ranks of human content moderators.

One way that social media companies purport to respond to the context of posts is through user flagging of content, and Twitter and Facebook both rely on user flagging to identify violations of their terms of service.<sup>226</sup> In the past, Facebook relied almost completely on user flagging for content evaluation.<sup>227</sup> After the genocide of the Rohingya, Facebook defended its inaction by stating that not enough of the offensive and inciting posts were flagged by users, so it could not review and remove them.<sup>228</sup> And for

---

policies/abusive-behavior (last visited Jan. 27, 2020).

<sup>224</sup> Klonick, *supra* note 91, at 1635, 1642.

<sup>225</sup> On context-specific approaches to content, see Monika Bickert, *Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process*, FACEBOOK NEWSROOM (Apr. 24, 2018), <https://perma.cc/VH7B-QLT9>; Nathaniel Gleicher, *Removing Bad Actors from Facebook*, FACEBOOK NEWSROOM (June 26, 2018), <https://perma.cc/TA3U-6TY3>.

<sup>226</sup> Niam Yaraghi, *How Should Social Media Platforms Combat Misinformation and Hate Speech*, BROOKINGS (Apr. 9, 2019), <https://www.brookings.edu/blog/techtank/2019/04/09/how-should-social-media-platforms-combat-misinformation-and-hate-speech/>; Kurt Wagner, *Twitter Says It’s Getting Better at Detecting Abusive Tweets Without Your Help*, VOX (Apr. 16, 2019, 3:35 PM), <https://www.brookings.edu/blog/techtank/2019/04/09/how-should-social-media-platforms-combat-misinformation-and-hate-speech/>.

<sup>227</sup> Rajagopalan et al., *supra* note 86.

<sup>228</sup> *Id.*

Twitter, certain objectionable content is only considered for removal if the target of the post flags the concern.<sup>229</sup>

User flagging alone, however, provides very little information about context. When they flag a post, users are generally able to choose from a range of pre-determined choices about why they are objecting to a particular piece of content. On Facebook, for example, users can flag speech as “hate speech” based on particular grounds such as “race or ethnicity,” but they are not allowed to offer further explanation.<sup>230</sup> Further, user flagging is not a reliable mechanism for identifying harmful speech because users may not be equally able or motivated, or technologically literate enough, to flag harmful content consistently. The user flagging system may also be “gamed” by coordinated campaigns to silence individuals or viewpoints.<sup>231</sup> User flagging also shifts the burden of protection onto vulnerable groups, who may be wary of condemning a post by a high-ranking member of the military or government. In cases of coordinated or systematic harassment, the targets of abuse may be unable to report all incidences of threatening or abusive language that is being directed toward them. “Flagging fatigue” by individuals or groups “who are frequently attacked coupled with a process that depends on users reporting every instance of harmful behavior represents a troubling aspect of content moderation practices.”<sup>232</sup>

Content review through automated means is also devoid of context. Since 2018, companies have relied more heavily on automated content moderation tools to determine whether a post merits further evaluation by content moderators.<sup>233</sup> Facebook, the biggest regulator of speech in the world,<sup>234</sup> has nearly replaced user flagging with AI.<sup>235</sup> According to Facebook, in the first three months of 2020, 88.8% of hate speech was

---

<sup>229</sup> Twitter Rules, *Abusive Behavior*, TWITTER, <https://help.twitter.com/en/rules-and-policies/abusive-behavior> (last visited Apr. 27, 2020).

<sup>230</sup> These options were viewed on December 12, 2020, on Facebook’s mobile app.

<sup>231</sup> See CTR. FOR AM. PROGRESS ET AL., RECOMMENDED INTERNET COMPANY CORPORATE POLICIES AND TERMS OF SERVICE TO REDUCE HATEFUL ACTIVITIES 1, 4, [https://assets.website-files.com/5bba6f4828dfc3686095bf6b/5bd0e36186e28d35874f0909\\_Recommended%20Internet%20Company%20Corporate%20Policies%20of%20Terms%20of%20Service\\_final-10-24.pdf](https://assets.website-files.com/5bba6f4828dfc3686095bf6b/5bd0e36186e28d35874f0909_Recommended%20Internet%20Company%20Corporate%20Policies%20of%20Terms%20of%20Service_final-10-24.pdf) (discussing the insufficiency of SMC reliance on user flagging) (last visited May 26, 2020).

<sup>232</sup> MUNA ABBAS ET AL., INVISIBLE THREATS: MITIGATING THE RISK OF VIOLENCE FROM ONLINE HATE SPEECH AGAINST HUMAN RIGHTS DEFENDERS IN GUATEMALA 30 (2019).

<sup>233</sup> CAMBRIDGE CONSULTANTS, USE OF AI IN ONLINE CONTENT MODERATION 18 (2019).

<sup>234</sup> *Facebook Q3 2019 Results*, *supra* note 9. With 2.4 billion users, Facebook is the largest social media company worldwide. Therefore, Facebook regulates more speech than any social media company and any state, including the government of China, which censors the speech of a population approximately half the size.

<sup>235</sup> Jason Koebler & Joseph Cox, *The Impossible Job: Inside Facebook’s Struggle to Moderate Two Billion People*, VICE (Aug. 23, 2018), [https://www.vice.com/en\\_us/article/xwk9zd/how-facebook-content-moderation-works](https://www.vice.com/en_us/article/xwk9zd/how-facebook-content-moderation-works).

identified through automated means before it was seen by viewers.<sup>236</sup> Using AI is arguably indispensable given the scale of online content moderation requirements and because of its ability to identify coordinated inauthentic behavior (i.e., integrated networks of “bots”), yet there are obvious drawbacks to an over-reliance on automated tools. As David Kaye has noted, social media companies “use the power of artificial intelligence to drive these systems, but the systems are notoriously bad at evaluating context.”<sup>237</sup>

AI may be indispensable for moderating content at scale, but what is lost is the context of usage that confers meaning, including physical signals, cultural particularity, and the social or political situation. Algorithmic identification is imprecise and unlikely to allow for contextual cues that might be necessary to distinguish extremist speech from parody, documentary footage, or legitimate protest.<sup>238</sup> As a result, some speech may be removed that is intended to challenge or lampoon hate speech. As York and Zuckerman note, “content moderation decisions are often highly subjective when made by humans, a problem that algorithms—designed by humans—are unlikely to solve.”<sup>239</sup> Next, AI is only as effective as the data it analyzes, and AI is often reliant on large datasets that may incorporate information generated by biased methods, with the result that AI reproduces the bias against historically disadvantaged populations.<sup>240</sup> Deploying artificial intelligence to flag content comes with the risk of over-censorship.

There are also limits on the extent to which human moderators can consider context. Human moderators usually have less than a minute to review a post,<sup>241</sup> and such brevity does not permit a wider consideration of the circumstances. Moderators review only one isolated post at a time or a short thread of several posts and may not see the entire ecosystem of posts in a locale. In this model, moderators may not appreciate the fact that posts are repeated in great numbers and may have cumulative effects.

---

<sup>236</sup> *Community Standards Enforcement Report*, FACEBOOK, <https://transparency.facebook.com/community-standards-enforcement#hate-speech> (last visited May, 26 2020).

<sup>237</sup> Kaye, *supra* note 30, ¶ 50.

<sup>238</sup> Sophia Cope et al., *Industry Efforts to Censor Pro-Terrorism Online Content Pose Risks to Free Speech*, ELECTRONIC FRONTIER FOUND. (July 12, 2017), <https://perma.cc/NUM5-Q9HY>; Emma Llansó, *Platforms Want Centralized Censorship. That Should Scare You*, WIRED (Apr. 18, 2019, 9:00 AM), <https://perma.cc/M8QA-UFZH>.

<sup>239</sup> Jillian C. York and Ethan Zuckerman, *Moderating the Public Sphere*, in HUMAN RIGHTS IN THE AGE OF PLATFORMS 137, 153 (Rikke Frank Jørgensen ed., 2019).

<sup>240</sup> Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 673–74 (2016); Aylin Caliskan et al., *Semantics Derived Automatically from Language Corpora Contain Human-like Biases*, 356 SCIENCE 183, 185 (2017).

<sup>241</sup> Gilbert, *supra* note 114.

A lack of a diverse moderation workforce also inhibits consideration of context. Before 2018, there were simply not enough content moderators.<sup>242</sup> Although in the intervening period, platforms have added tens of thousands of human moderators, there are still not enough of them for minority languages.<sup>243</sup> Even if there are numerous moderators in a major language such as English, Spanish, or Hindi, content moderators are often too far removed, both in terms of physical location and the local information they are provided, from the speech contexts they moderate. Content moderators may know little or nothing about the social, cultural, and political dimensions of the context from which a post emanates. A content moderator located in a distant country may not be aware of widespread election unrest, outbursts of communal violence, or a pattern of violence against sexual minorities in a locale. Since each moderator only sees a small sliver of the total range of expression about a topic or person, campaigns of systemic harassment are harder to identify.<sup>244</sup>

Then there are the multiple challenges of the universal and the particular in language. All languages are highly nuanced, with each word having many shades of meaning that depend on a number of factors including the speaker's intention, the listener's expectations, and the context in which an utterance is made. Even if a content moderator is a native speaker of a language, they may not be familiar with idiomatic usage in a locale far removed from their own personal experience. This is especially true of world languages and regional languages with hundreds of millions of speakers. For instance, we would not expect a Spanish speaker from Madrid to comprehend slang used on the streets of Guatemala City. Calling someone a "communist" or "destabilizer" may be benign in Madrid but constitute incitement to violence in Guatemala as a result of the countries' different political histories.<sup>245</sup> Yet there are credible reports of

---

<sup>242</sup> Civil society groups in Myanmar have indicated that Facebook did not have enough content reviewers who could speak Burmese during the time of the Rohingya genocide. Rajagopalan et al., *supra* note 86.

<sup>243</sup> *Id.*

<sup>244</sup> The Twitter Rules indicate:

Some Tweets may seem to be abusive when viewed in isolation, but may not be when viewed in the context of a larger conversation. When we review this type of content, it may not be clear whether it is intended to harass an individual, or if it is part of a consensual conversation. To help our teams understand the context of a conversation, we may need to hear directly from the person being targeted, to ensure that we have the information needed prior to taking any enforcement action.

*Abusive Behavior*, TWITTER, <https://help.twitter.com/en/rules-and-policies/abusive-behavior> (last visited Jan. 27, 2020).

<sup>245</sup> ABBAS ET AL., *supra* note 232, at 31–32 (describing the inciting implications of terms such as "communist," "destabilizer," and "terrorist" in Guatemala).



content moderators using Google Translate to gauge the meaning of a post they are reviewing.<sup>246</sup>

Social media companies have global lists of words and expressions that are banned, and these lists are universal; that is, they generally do not vary according to country or region.<sup>247</sup> As Chinmayi Arun observes, however, “[h]armful speech can be hyper-localised, significant in just one village or district, such that even people from the same state might not understand. This is the sort of thing that local police and local journalists might notice and address but a global corporation might miss completely.”<sup>248</sup> Highly local terms, such as the anti-Muslim racial epithet “kalar” that was repeatedly posted on Facebook by members of the Myanmar military during their genocidal campaign against the Rohingya, are often missed by content evaluation procedures either because they are in minority languages or because they are not on the list of universal hate speech terms.<sup>249</sup>

Alongside localized slurs, there is the problem of coded incitement and implicit threats that do not present as hate speech or violate hate speech guidelines, but are understood as incitements or threats by those with the requisite cultural knowledge. The genocidal campaign against the Rohingya in Myanmar raised with particular urgency the importance of identifying coded or euphemistic speech threats. For example, the Independent International Fact-Finding Mission on Myanmar noted that it was difficult to identify online hate speech against the Rohingya because of “[s]ubtleties in the Myanmar language and the use of fables and allegories.”<sup>250</sup>

Coded speech is present in virtually all instances of genocide and mass atrocity. For instance, after the 1994 Rwandan genocide, the International Criminal Tribunal for Rwanda (ICTR) had to decide whether exhortations from leaders such as Jean-Paul Akayesu calling on Rwandans to go “to work” and to fight and kill “the Inkotanyi” (literally in Kinyarwanda, “warriors”) represented incitement to genocide against the Tutsi

---

<sup>246</sup> Max Fisher, *Inside Facebook's Secret Rulebook for Global Political Speech*, N.Y. TIMES (Dec. 27, 2018), <https://www.nytimes.com/2018/12/27/world/facebook-moderators.html>.

<sup>247</sup> Billy Perrigo, *Facebook Says It's Removing More Hate Speech Than Ever Before. But There's a Catch*, TIME (Nov. 27, 2019, 4:42 AM), <https://time.com/5739688/facebook-hate-speech-languages/> (“The algorithms Facebook currently uses to remove hate speech only work in certain languages.”).

<sup>248</sup> Chinmayi Arun, *Rebalancing Regulation of Speech: Hyper-Local Content on Global Web-Based Platforms*, MEDIUM (Mar. 28, 2018), <https://medium.com/berkman-klein-center/rebalancing-regulation-of-speech-hyper-local-content-on-global-web-based-platforms-1-386d65d86e32>.

<sup>249</sup> Perrigo, *supra* note 247; *see also* Application Instituting Proceedings and Request for Provisional Measures (Gam. v. Myan.), Application, ¶¶ 38–39, 44–45 (Nov. 11, 2019), <https://www.icj-cij.org/en/case/178/institution-proceedings>.

<sup>250</sup> Myanmar Report, *supra* note 77, ¶ 1311.

minority.<sup>251</sup> On the basis of expert testimony from Rwandan sociolinguist Mathias Ruzindana about the use of veiled incitement by Hutu Power extremists, the ICTR decided that both expressions did constitute direct and public incitement to exterminate Tutsis.<sup>252</sup> Social media posts by a modern-day Akayesu implicitly inciting harm by encouraging his followers to “go to work” would likely not be removed under current content moderation policies because implicit speech requires sufficient contextual information to render its meaning evident.<sup>253</sup> Current content moderation policies do not integrate contextual factors into the decision-making and thus may permit coded calls to violence.<sup>254</sup>

### C. Recommendations

There are a number of steps that companies could take to ensure that context is taken into account more systematically with the aim of enhancing accurate content evaluation. We recognize that meticulous contextual analysis is demanding logistically, as well as resource-intensive, and we do not minimize the challenges ahead. However, there are a few actions that might be taken to bring more context into account while regulating hate speech. Many of these recommendations are likely implemented most effectively at higher levels of review; as a result, context-based moderation may require an expansion of the workforce at those senior levels as well as higher rates of escalation from front-line moderators to more senior levels of review.

First, platforms could start by pausing the dramatic shift in recent years towards artificial intelligence and by conducting a thorough, transparent, and independent review of the implications of automation. Along with independent experts, social media companies should examine the degree to which their content-generating algorithms elevate hateful speech on user feeds or perpetuate unconscious biases against particular minority groups. Companies must disclose the rate of false positives in their flagging algorithms and investigate how much hate-filled content remains online because it is coded or implicitly stated in a way that avoids the algorithm. Companies should design an iterative process in which AI and human content moderation are integrated and in which both are fully informed and updated by “a sufficiently large, trained team of internet company employees who are cognizant of relevant social, political, and cultural history and context.”<sup>255</sup>

---

<sup>251</sup> RICHARD ASHBY WILSON, INCITEMENT ON TRIAL: PROSECUTING INTERNATIONAL SPEECH CRIMES 191 (2017).

<sup>252</sup> *Id.* at 192.

<sup>253</sup> *Id.* at 190–92.

<sup>254</sup> Hamilton & Land, *supra* note 221, at 144.

<sup>255</sup> CTR. FOR AM. PROGRESS ET AL., *supra* note 231, at 5.

Second, we encourage companies to move away from a one-size-fits-all model of hate speech regulation and to build a number of different models that are appropriate to the conditions on the ground in the countries concerned. The work of programmers must be informed by country specialists as they construct automated tools that are sensitive to and responsive to local circumstances. This integration of human and machine systems could lead to the kind of “hybrid social-ordering systems” that Wu has proposed.<sup>256</sup> In sum, we endorse a human rights approach to artificial intelligence that prioritizes procedural fairness, transparency, and disclosure, and implies a commitment to human rights audits and regular reviews of potential discrimination in AI systems.<sup>257</sup>

Next, companies could take steps towards a pluralistic model of content moderation, starting with sub-regional or country-specific lists of slurs and hateful expressions that are drawn up after consultation with local civil society organizations, sociolinguists, and other experts. Facebook, for example, maintains country-specific lists of slurs based on input from civil society and experts,<sup>258</sup> and we recommend that other platforms adopt the same approach. Platforms could also create teams of moderators that look not at a single post or thread, but at the total range of expression in a country or region about a person, group, or topical area. When reviewing content, moderators should be allowed to take into consideration external information about the political and cultural context of a country and in particular indicators of the risk of violence against an individual or group.

Platforms could employ moderators who speak the local dialects of the language and who are culturally and linguistically close enough to the sources of the posts to be able to detect coded hate speech, threats, and incitement. As David Kaye stated, human content evaluation

must be based on real learning from the communities in which hate speech may be found, that is, people who can understand the “code” that language sometimes deploys to hide incitement to violence, evaluate the speaker’s intent, consider the nature of the speaker and audience and evaluate the environment in which hate speech can lead to violent acts. None of these things are possible with artificial

---

<sup>256</sup> Wu, *supra* note 175, at 2021.

<sup>257</sup> Kaye, *supra* note 175, ¶ 47 (advancing a human rights approach to artificial intelligence).

<sup>258</sup> *Understanding Social Media and Conflict*, FACEBOOK, <https://about.fb.com/news/2019/06/social-media-and-conflict/amp/> (last updated May 21, 2020); *see also* Sara Rieger, *Facebook to Investigate Whether Anti-Indigenous Slur Should Be Added to Hate Speech Guidelines*, CBC (Oct. 24, 2018), <https://www.cbc.ca/news/canada/calgary/facebook-indigenous-slur-1.4877225>.

intelligence alone, and the definitions and strategies should reflect the nuances of the problem.<sup>259</sup>

There ought to be some flexibility in the categories of persons who are protected under platform law, and these categories should be determined in light of local circumstances. In some contexts, the list of categories currently protected by platform law is underinclusive. For example, human rights defenders, including judges and prosecutors involved in anti-corruption efforts, have been subjected to coordinated online hate campaigns in Guatemala, Colombia, and Mexico in recent years.<sup>260</sup> In these three countries alone, over 200 human rights defenders are killed every year, about half of the global total.<sup>261</sup> Where human rights work is such a hazardous undertaking, it makes sense to designate human rights defenders as a protected category until the attacks start to decline. If there are other occupational groups who are not presently protected, such as social workers or humanitarian aid workers, then they too may be brought temporarily into the scope of protection.

We recommend that companies run a small number of pilot projects in at-risk countries that have experienced, or are currently experiencing, violent clashes during an election cycle; where minority groups are being persecuted in large numbers; or where watchdog groups, governments, or international agencies have identified a risk of mass atrocities or genocide. For countries receiving heightened scrutiny, companies could create specialist content evaluation teams with a short (three to six month), renewable mandate. Content regulation policy would be the product of consultations between multiple stakeholders including companies, governments, the Office of the United Nations High Commissioner for Human Rights, and civil society organizations with relevant human rights expertise. Content moderation itself would be carried out by local and native speakers of the relevant language(s) who are advised by expert policy analysts who know the history, culture, and politics of the country.

Some companies such as Twitter have explored “trusted partner” programs where they consult with local users and prioritize their flagging of problematic posts.<sup>262</sup> Trusted individuals or organizations can serve to identify problematic posts and alert companies through their content

---

<sup>259</sup> Kaye, *supra* note 206, ¶ 50; *see also* Jessica Dheere, *Considering Countering Violent Extremism in Lebanon*, SMEX (May 1, 2016), <https://smex.org/considering-countering-violent-extremism-in-lebanon/> (recommending that companies hire and train teams of moderators “with cultural and linguistic sensitivity”).

<sup>260</sup> FRONT LINE DEFS., GLOBAL ANALYSIS 2018, *supra* note 65, at 16 (reporting incidents of threats, attacks, and intimidation in Guatemala, Columbia, and Mexico, and that WhatsApp has become a “powerful tool to disseminate hate messages”); *see also* FRONT LINE DEFS., STOP THE KILLINGS 26 (2018) (describing intimidation of judges and public prosecutors in Guatemala).

<sup>261</sup> FRONT LINE DEFS., GLOBAL ANALYSIS 2018, *supra* note 65, at 4.

<sup>262</sup> ABBAS ET AL., *supra* note 232, at 39.

moderation processes. This would also allow for the identification of harmful posts without having to rely on vulnerable targets, since the act of reporting can put a target in danger. Trusted flaggers are likely to be more attuned to problematic posts because they have greater knowledge of the area, understand the context of the language used, and know how to navigate the content moderation process in general.<sup>263</sup> For example, Waze uses data to identify which users are trustworthy content moderators on the basis of positive feedback that the information they provide is accurate and reliable.<sup>264</sup> A similar approach might allow Facebook and Twitter to more effectively flag harmful speech.

Special content evaluation teams in at-risk countries ought to scrutinize carefully the social media posts of senior political leaders. There has been recent controversy over whether hate speech posted by senior politicians ought to be tolerated on the grounds that it is newsworthy and has public-interest value. In 2019, Facebook indicated that it would implement a newsworthy exemption for politicians.<sup>265</sup> In 2020, Twitter began to label President Donald J. Trump's tweets with a fact-check label after he falsely claimed that mail-in ballots were fraudulent.<sup>266</sup>

We have serious concerns about a newsworthiness exception as a global default on the grounds that senior politicians have played a vital role in inciting mass violence in almost all prior instances of genocide and crimes against humanity.<sup>267</sup> The weight of social science evidence finds that "incitement is almost certainly more harmful when uttered by leaders than by other users."<sup>268</sup> For instance, political scientist Scott Straus analyzed twenty-four cases of mass violence in Africa and discovered that the most critical factor in each case was government authorities who encouraged violence and also coordinated it through their speech acts.<sup>269</sup> Other studies corroborate Straus's findings and show that the greater the authority of the speaker, the more likely verbal instructions to commit

---

<sup>263</sup> *Id.*

<sup>264</sup> *Id.* at 39–40; *Places*, WAZEOPEDIA, [https://wazeopedia.waze.com/wiki/Canada/Places/#What\\_is\\_a\\_trusted\\_user.3F](https://wazeopedia.waze.com/wiki/Canada/Places/#What_is_a_trusted_user.3F) (last visited Jan. 29, 2020).

<sup>265</sup> Taylor Hatmaker, *Facebook: Anything Politicians Say Is Exempt from Platform's Rules*, DAILY BEAST (Sept. 25, 2019, 4:15 AM), <https://www.thedailybeast.com/facebook-anything-politicians-say-is-newsworthy-and-exempt-from-platforms-rules>.

<sup>266</sup> Elizabeth Dwoskin, *Twitter Labels Trump's Tweets with a Fact Check for the First Time*, WASH. POST (MAY 27, 2020, 8:07 AM), <https://www.washingtonpost.com/technology/2020/05/26/trump-twitter-label-fact-check/>.

<sup>267</sup> See WILSON, *supra* note 251, at 267–68 (explaining Bosnian Serb leaders' role in inciting genocide).

<sup>268</sup> Kaye, *supra* note 206, ¶ 47(d).

<sup>269</sup> SCOTT STRAUS, MAKING AND UNMAKING NATIONS: WAR, LEADERSHIP, AND GENOCIDE IN MODERN AFRICA 99 (2015).

violent acts will be acted upon.<sup>270</sup> Given these findings, a newsworthiness exemption for politicians might be re-examined in at-risk country settings, even if only on a temporary basis.

Finally, and perhaps most ambitiously, we could envisage a reorganization of companies' organizational structures and business models from industrial to artisanal models of content moderation.<sup>271</sup> Social media companies generate their revenue from advertising, and internet advertising represented the largest and fastest-growing share of the advertising market by type of media, totaling \$107 billion in 2018.<sup>272</sup> In 2019, Facebook's third quarter results reported that it is on track to make over \$60 billion in advertising revenue.<sup>273</sup> Internet advertising is the only media type experiencing double digit (and nearly 20%) growth year after year.<sup>274</sup> No self-respecting bricks-and-mortar advertising agency would operate in a market of any size without a physical presence or connection with a local subsidiary,<sup>275</sup> but social media platforms dispensed with this model to establish, at lightning speed, global empires from their headquarters in San Francisco or Palo Alto.

The corollaries and shortcomings of this unprecedented growth are now evident. Platforms are de facto global advertising companies that have yet to fully come to terms with this reality. In order to become truly global companies, they need to genuinely engage with the world. Otherwise, their policies and procedures, even with plenty of multi-stakeholder consultation, will continue to appear to be a form of U.S. cultural imperialism to many users outside the United States.<sup>276</sup> This must, of course, be balanced against the safety needs of employees who may be targets for authoritarian governments. But decentralization of moderation

<sup>270</sup> SARAH SORIAL, *SEDITION AND THE ADVOCACY OF VIOLENCE: FREE SPEECH AND COUNTER-TERRORISM* 84–85, 115–16 (2012).

<sup>271</sup> See ROBYN CAPLAN, *DATA & SOC'Y, CONTENT OR CONTEXT MODERATION? ARTISANAL, COMMUNITY-RELIANT, AND INDUSTRIAL APPROACHES* 17 (2018), <https://datasociety.net/library/content-or-context-moderation/> (reviewing artisanal and industrial models of content moderation employed by internet companies).

<sup>272</sup> PRICEWATERHOUSECOOPERS, IAB/PWC, *INTERNET AD REVENUE REPORT, FY 2018*, at 21 (2019). See Rebecca Tushnet, *Content Moderation in an Age of Extremes*, 10 CASE W. RES. J.L. TECH. & INTERNET 1, 11–12 (2019) (reviewing the implications of the advertising revenue model on content moderation).

<sup>273</sup> *Facebook Q3 2019 Results*, *supra* note 9.

<sup>274</sup> PRICEWATERHOUSECOOPERS, *supra* note 272, at 22.

<sup>275</sup> See Emad Rahim, *How Global Marketing Is Different From International Marketing*, TWEAK YOUR BIZ, <https://tweakyourbiz.com/management/how-global-marketing-is-different-from-international-marketing> (last updated Sept. 18, 2018) (explaining how international marketing agencies normally opened a subsidiary in a country).

<sup>276</sup> PAZ PEÑA OCHOA, *REPORTE DE LA SITUACIÓN DE AMÉRICA LATINA SOBRE LA VIOLENCIA DE GÉNERO EJERCIDA POR MEDIOS ELECTRÓNICOS* 44 (2017), <https://www.codingrights.org/wp-content/uploads/2017/11/Latin-American-Report-on-Online-Gender-Violence-final.pdf> (last visited Apr. 30, 2020).

would be an important means toward greater democratization and contextualization of content governance.

An artisanal content moderation model could involve dispensing with a universal list of hate speech terms implemented by automated tools and content moderators far removed from the circumstances of the speech. Companies are advised to open local offices or bureaus in each of the countries in which they operate, staffed by well-trained and recompensed local employees, at least some of whom are experienced investigative journalists. Their role would be to design and implement social media policy and content moderation procedures relating to disinformation, hate speech, and other forms of problematic expression, tailored to that country or region's specific context. Review of escalated moderation cases would be conducted in-house and would draw on locally generated and continually revised lists of terms and expressions that are likely to harm. Local gatekeepers would consciously gatekeep and do so in a way that is transparent. The companies' headquarters could exercise oversight and capacity building and support and maintain the automated tools, but local offices would have substantial leeway in content moderation policy formulation and application. This implies an openness on the part of companies to the "geographical segmentation" of their content moderation policies, and an acceptance of global legal pluralism in platform law.<sup>277</sup>

The concerns stated earlier about Facebook's oversight board are justified, but the board's creation also presents opportunities for a more context-specific moderation policy. The list of the first twenty members is promising.<sup>278</sup> The board is geographically diverse and includes representation from Africa, Asia, and Latin America as well as North America and Europe.<sup>279</sup> It comprises a former editor of *The Guardian*, a Nobel Prize winner, a former UN Special Rapporteur on Freedom of Peaceful Assembly and Association, constitutional and human rights law professors, and others with a demonstrated commitment to freedom of expression and a willingness to denounce governments for human rights violations.<sup>280</sup> The board may shed light on the currently opaque process of content moderation, as Facebook must respond publicly to all board

---

<sup>277</sup> See Daskal, *supra* note 143, at 1655 (proposing geographical segmentation of takedown policies in content moderation, with limits). See generally Molly K. Land, *The Problem of Platform Law: Pluralistic Legal Ordering and Social Media*, in OXFORD HANDBOOK OF GLOBAL LEGAL PLURALISM 958 (Paul Schiff Berman ed., 2020) (discussing how global legal pluralism provides an analytical model for understanding the competition between private regulation, national law, and international law); Berman, *supra* note 14, at 1157–58 (same).

<sup>278</sup> See Steven Overly & Alexandra S. Levine, *Facebook Announces First 20 Picks for Global Oversight Board*, POLITICO (May 6, 2020), <https://www.politico.com/news/2020/05/06/facebook-global-oversight-board-picks-240150> (considering the geographic diversity and biographical details of the first twenty oversight board members).

<sup>279</sup> *Id.*

<sup>280</sup> *Id.*

decisions and provide regular updates.<sup>281</sup> The board can issue advisory statements on policy matters and Facebook must respond to those within thirty days.<sup>282</sup> As Kadri and Klonick have suggested in their critique of Facebook's "newsworthiness" exception,<sup>283</sup> the board may be able to provide greater transparency on how concepts such as these are defined and applied.

The critics are right that the board lacks the authority to address essential issues, and the board should consider amendments to the bylaws immediately and before it starts operations. The board's competence should be expanded to include review of account removals ("deplatforming") and content left up on the platform, as well as content taken down. The board should be able to consider content removed by Facebook in compliance with national laws restricting speech. The board should also have the authority to review and make binding decisions on the policies, practices, and procedures relating to content removal and political advertising. This could include policies that govern decisions about which terms violate platform law, such as key terms like "abusive" or "threatening." Additionally, it could encompass procedures such as Facebook's algorithm for curating users' personal feeds and news feeds. For the board to be effective and promote meaningful accountability, the bylaws must clearly state that its decisions constitute precedent, in the legal sense that they govern Facebook's handling of all similar cases. To enhance transparency, the board should report quarterly rather than annually.<sup>284</sup> Finally, the current bylaws do not take context sufficiently into account and need to be revised to require staff to provide the board with a summary analysis of the relevant political, social, and cultural context of each user post under review, including information on the range of posts associated with it.

## CONCLUSION

The time-honored hate speech debate has now been superseded by the content regulation policies and procedures of global platforms that remove tens of thousands of abusive and inciting posts every day. The question is now not whether hate speech should be regulated in the public sphere, but how. We urge companies to develop and enhance a contextual element of their analysis, particularly in countries with a history of intergroup violence and weak state institutions. There are a number of measures that platforms could take to improve their content moderation policies and

---

<sup>281</sup> Bylaws, art. 2, § 2.3.2.

<sup>282</sup> *Id.*

<sup>283</sup> Thomas E. Kadri & Kate Klonick, *Facebook v. Sullivan: Public Figures and Newsworthiness in Online Speech*, 93 S. CAL. L. REV. 37, 96–97 (2019).

<sup>284</sup> Bylaws, art. 1, § 1.3.1.



consider context in their evaluation of potential hate speech, including drawing up localized lists of banned expression, ensuring that content moderators are native speakers of languages, and developing procedures to evaluate coded speech. In at-risk countries, companies should follow a more intensive strategy in which they provide heightened scrutiny of potentially harmful content, employ content moderators who are native and local speakers of languages, engage the guidance and assistance of local personnel and country experts in developing context-specific content moderation policies, consider the status of the speaker in evaluating the effect of speech on their platforms, and improve flagging processes to facilitate the gathering of context-specific information. These measures are not designed to be more restrictive of hate speech, but to ensure that any restrictions that are implemented are more precise, informed, and accurate.

Some of these measures are easily implemented, others require a sea-change in company mission and philosophy. Platforms are unlikely to embark upon a radical restructuring on the merits of well-reasoned arguments alone. However, we believe the current approach to moderation will become increasingly untenable. Users may be unwilling to tolerate a toxic environment, but it is also unclear whether they will tolerate a private platform broadly censoring their speech. Governments are requiring an adherence to national hate speech policies with little guidance on what that requires. Furthermore, events like the genocide of the Rohingya in which a company badly mishandles a situation can cause severe reputational damage. Until companies revise their content moderation policies to integrate the context in which speech occurs, they will remain exposed to accusations of responsibility for real-world violence.

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### The Power of the Civic Mindset: A Conceptual Framework for Overcoming Political Polarization

DANIEL L. SHAPIRO

*This Article proposes a new conceptual paradigm for overcoming political polarization—the civic mindset. I argue that the primary psychological barrier to bridging political divides is an adversarial state of mind called the partisan mindset, and I explain its specific characteristics, fundamental operating principles, and triggers. To combat polarization, I introduce the civic mindset, elucidate its basic features and functions, and explain how societal embrace of this unique outlook can advance a vibrant political space within which partisan competition and national unity can thrive.*

## ARTICLE CONTENTS

INTRODUCTION.....	1079
I. WHAT’S YOUR POLITICAL MINDSET?.....	1080
II. THE PARTISAN MINDSET.....	1081
A. KEY ATTRIBUTES .....	1082
B. OPERATING PRINCIPLES .....	1083
C. WHAT FOSTERS THE PARTISAN MINDSET?.....	1084
III. THE CIVIC MINDSET.....	1086
A. KEY ATTRIBUTES .....	1087
B. OPERATING PRINCIPLES .....	1088
C. FOSTERING THE CIVIC MINDSET.....	1089
IN SUMMARY .....	1091



# The Power of the Civic Mindset: A Conceptual Framework for Overcoming Political Polarization

DANIEL L. SHAPIRO \*

## INTRODUCTION

By design, democracy generates political tension. Opposing advocacy groups lobby for legislation, competing political parties must write, debate, and pass laws, and government branches conflict due to intentional overlap in responsibilities.<sup>1</sup> But while tension is expected, contemporary politics is on fire.<sup>2</sup> Political elites humiliate one another, legislators strategize ways to “win” a policy dispute and showcase the other party’s loss, and community members refuse to interact with their counterpart—generating a kind of “political warfare” that risks uprooting democratic institutions and the belief in their efficacy.<sup>3</sup> Partisanship is natural, but intense polarization can jeopardize the system itself.

Understanding the factors that contribute to political polarization can direct us to strategies to combat them. Social scientists have uncovered a variety of reasons for political divisiveness including economic disparities, discriminatory laws, gerrymandering, and media bias.<sup>4</sup> While such structural

---

\* Daniel L. Shapiro, Ph.D., is Founder and Director of the Harvard International Negotiation Program, Associate Professor of Psychology at Harvard Medical School / McLean Hospital, and Faculty Affiliate at the Program on Negotiation at Harvard Law School. Special thanks to Adam Kuegler, Alexandria Madjeric, Hannah Kalichman, and the entire team at the *Connecticut Law Review* for outstanding feedback and editorial assistance. This Article benefitted from perspectives on civil discourse examined at the *Connecticut Law Review*’s symposium, *How We Argue: The Moral Foundations of Politics and Law*. I would like to thank fellow panelists Richard Wilson, David Gergen, Carol Anderson, and Leah Rigueur. I also am grateful to Mikhaila Fogel for her insights on how to sharpen the concept of the civic mindset.

<sup>1</sup> *Separation of Powers—An Overview*, NAT’L CONF. ST. LEGISLATURES (May 1, 2019), <https://www.ncsl.org/research/about-state-legislatures/separation-of-powers-an-overview.aspx>.

<sup>2</sup> An alternative metaphor may be that political sentiments are becoming colder. A Pew Research Center study found that between 2016 and 2019, the number of Republicans who gave Democrats a “cold” rating on a “feeling thermometer” from 1–100 rose by 14% and the number of Democrats who gave Republicans a cold rating grew by 16%. *Growing Shares in Both Parties Give “Cold” Ratings to Those in Opposing Party*, PEW RES. CTR. (Oct. 10, 2019), [https://www.people-press.org/2019/10/10/partisan-antipathy-more-intense-more-personal/pp\\_2019-10-10\\_state-of-parties\\_0-01/](https://www.people-press.org/2019/10/10/partisan-antipathy-more-intense-more-personal/pp_2019-10-10_state-of-parties_0-01/) (showing a “feeling thermometer” to demonstrate how partisans give members of the opposite party a “cold rating”).

<sup>3</sup> Sean Theriault, *Partisan Warfare Is the Problem*, in *POLITICAL POLARIZATION IN AMERICAN POLITICS* 11, 11–15 (Daniel J. Hopkins & John Sides eds., 2015).

<sup>4</sup> *See* 2 RED AND BLUE NATION? CONSEQUENCES AND CORRECTION OF AMERICA’S POLARIZED

forces predispose citizens to polarization, the ultimate arbiter of political behavior is the human heart.<sup>5</sup> Political action from policy negotiations to voting behavior is heavily informed by mood, emotion, and attitude.<sup>6</sup> In contemporary society, politics has become so emotionally charged that brutal partisan politics is the norm—resulting in gridlock, harsh character judgments, and, perhaps most ominously, threat to the survival of the democratic system.<sup>7</sup>

This Article offers a new conceptual approach for overcoming polarization. I argue that the major psychological barrier to political cooperation is a divisive outlook I term the *partisan mindset*—a state of mind with specific characteristics, fundamental operating principles, and triggers. To overcome this psychological obstacle, I conceptualize an alternative state of mind—the *civic mindset*—that motivates concern for our own political interests *and* the legitimate interests of the multitude of political groups within society, resulting in a vibrant political space within which partisan competition *and* national unity can thrive.

## I. WHAT'S YOUR POLITICAL MINDSET?

Mindset frames how we see the world and our place in it—including what we attend to, what we ignore, and how we make meaning of events. Politicians who view themselves as adversaries can spend years unsuccessfully debating legislation while political allies can solve the same issue in a single hallway conversation. The way they view their relationship has a big impact on the outcome of their exchange.

Drawing on relational identity theory,<sup>8</sup> I conceive of *political mindset* as the lens through which we make sense of the political landscape and orient relationally to political stakeholders. This mindset patterns our cognitive and emotional world, providing us with affectively tinged assumptions about whom to trust or doubt. In the public arena, the mindset we adopt acts as a pair of glasses that colors the way we perceive the entire political landscape and our place in it. We easily can detect the political mindset in a neighbor

---

POLITICS (Pietro S. Nivola & David W. Brady eds., 2008) (containing a variety of essays discussing the roots of political divisiveness in the United States).

<sup>5</sup> See JONATHAN HAIDT, *THE RIGHTEOUS MIND: WHY GOOD PEOPLE ARE DIVIDED BY POLITICS AND RELIGION* 34 (2012) (discussing the role human emotion plays in political behavior and noting “[t]he head can’t even do head stuff without the heart”).

<sup>6</sup> See *id.* at 152–53 (discussing how the two ends of the political spectrum rely on “moral foundations”).

<sup>7</sup> See Steven Levitsky & Daniel Ziblatt, *How a Democracy Dies*, *NEW REPUBLIC*, Jan.–Feb. 2018, at 17, 19 (discussing how modern politics poses a threat to the U.S. democratic system).

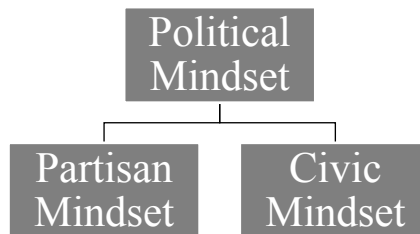
<sup>8</sup> See Daniel L. Shapiro, *Relational Identity Theory: A Systematic Approach for Transforming the Emotional Dimension of Conflict*, 65 *AM. PSYCHOLOGIST* 634, 634 (2010) (discussing relational identity theory).

who asks us, “Why on earth did you vote for *that* candidate?” This person holds an adversarial mindset and is primed to attack our views.

We tend to assume everyone wears distorting glasses but us. We think others see the world through glasses that bias their understanding of reality, whereas we see the world as it is, a glass-less truth. But this is obviously not true. Social cognition predisposes everyone toward a biased interpretation of reality.<sup>9</sup> Consequently, becoming aware of our own mindset frees us to decide whether to adopt a more constructive one.

In the following sections, I introduce two mindsets that offer us choice on how to view the political landscape. One fuels polarization; the other quells it.

*Chart 1. Two fundamental political mindsets.*



## II. THE PARTISAN MINDSET

The political mindset most responsible for fueling polarization is what I term the *partisan mindset*, a divisive outlook that pits us against another political entity.<sup>10</sup> This Section describes the nature and attributes of the partisan mindset, its operating principles, and the ways it gets fostered.

To understand the partisan mindset, we must appreciate the basic elements and function of a tribe. I define a tribe as any group whose members view themselves as like-kindred, kin-like in their relational connection, and emotionally invested in the group’s enhancement.<sup>11</sup> Being of like kind signifies that group members identify themselves as part of a shared political entity, whether a neighborhood organization, religious sect, or formal political party. Kin-like connection defines the nature of the relationship between members, who view themselves not as part of a loose affiliation or coalition but as members of an extended family, a deeply felt

---

<sup>9</sup> See DANIEL KAHNEMAN, THINKING, FAST AND SLOW 4 (2011) (describing how human impressions, intuitions, and decisions are not conscious choices).

<sup>10</sup> See DANIEL SHAPIRO, NEGOTIATING THE NONNEGOTIABLE: HOW TO RESOLVE YOUR MOST EMOTIONALLY CHARGED CONFLICTS, at xvii (2017). The partisan mindset is a politically nuanced subtype of the *tribes effect* described in NEGOTIATING THE NONNEGOTIABLE: HOW TO RESOLVE YOUR MOST EMOTIONALLY CHARGED CONFLICTS 26 (2017).

<sup>11</sup> Shapiro, *supra* note 8, at 638.

bond. Emotional investment refers to the motivation of members to enhance the tribe's status, power, and durability even in the face of personal sacrifice.

The partisan mindset activates when we feel a threat to our tribal identity. Specific social cognitive conditions stimulate this perspective shift. At a bare minimum, there must be a salient threat to our identity, we must prioritize our tribal identity over other social identities, and the mindset itself must be cognitively accessible.<sup>12</sup> The result is a relational outlook that predisposes us to experience divisive feelings, thoughts, and behavior toward the perceived source of threat.

### A. *Key Attributes*

The partisan mindset has three major elements that affect our political worldview:<sup>13</sup>

#### 1. *Adversarial*

We tend to view competing political parties through an adversarial lens that magnifies differences and minimizes similarities. A threat to our identity instigates a kind of relational amnesia, in which we ignore the countless personal and structural connections and fixate on political differences. Philosopher Martin Buber describes this movement toward disaffiliation as a shift from an "I-Thou" to an "I-It" relationship.<sup>14</sup> A state of emotional arousal, such as that triggered during times of political tension, reduces the cognitive complexity of our social perceptions and results in polarized evaluations of our counterpart.<sup>15</sup> At the extreme, daily confrontation with our own physical, mental, or collective mortality can heighten the perceived significance of our own group and lead us to devalue those who threaten our identity.<sup>16</sup>

#### 2. *Self-righteous*

The partisan mindset involves the self-serving conviction that our political views are not only right, but morally superior. We assume that morality exists on a single, exclusive moral plane and reject the idea that multiple perceptions of political truth can coexist.<sup>17</sup> Self-righteousness is

---

<sup>12</sup> SUSAN T. FISKE & SHELLY E. TAYLOR, *SOCIAL COGNITION: FROM BRAINS TO CULTURE* 342 (2007).

<sup>13</sup> SHAPIRO, *supra* note 10, at 27.

<sup>14</sup> MARTIN BUBER, *I AND THOU* 62–64 (Charles Scribner's Sons trans., Touchstone 1st ed. 1996) (1970).

<sup>15</sup> Delroy L. Paulhus & David T. K. Lim, *Arousal and Evaluative Extremity in Social Judgments: A Dynamic Complexity Model*, 24 *EUR. J. SOC. PSYCHOL.* 89, 90, 94 (1994).

<sup>16</sup> Jeff Greenberg et al., *Evidence for Terror Management Theory II: The Effects of Mortality Salience on Reactions to Those Who Threaten or Bolster the Cultural Worldview*, 58 *J. PERSONALITY & SOC. PSYCHOL.* 308, 317 (1990).

<sup>17</sup> The Pew Research Center found that in today's polarized society, a majority of Democrats and

founded on arrogance and is distinguishable from righteous indignation—the boiling anger that emerges when we confront injustice.

### 3. *Insular*

The partisan mindset locks us in a closed political system. We watch news shows that reinforce our political views, frequent social media sites that support our political values, build friendships with those who maintain similar perspectives, and resist talking politics with political counterparts. The more closely we follow public affairs, the more likely we are to express negative sentiments about the opposing party,<sup>18</sup> bolstering hostility toward our political counterpart.

#### B. *Operating Principles*

The partisan mindset operates on a few basic principles. First, loyalty to the tribe takes priority.<sup>19</sup> People who identify heavily with a political tribe will tend to make greater sacrifices for their own tribe than for other groups. Social evolutionists have discovered that the likelihood of our committing a costly altruistic action depends upon the genetic closeness to the recipient and the benefit to that person or group;<sup>20</sup> we sacrifice more for those who share our bloodlines. Human psychology extends this tendency to individuals who are connected through *perceived* kinship, thus imbuing political tribes with substantial influence over political action.

Second, partisan norms urge blind loyalty to the party platform—regardless of the degree to which policies serve personal interest.<sup>21</sup> This affords tribal leaders great personal power because they can rely on their political base to support their policies and block plans of political foes. But blind loyalty erodes merit-based decision making, because people care more about maintaining fidelity to their own political party than about finding ways to advance the legitimate interests of the multitude of political groups within society.

Third, cooperation with political outgroups is taboo. During times of polarization, the mere act of being seen talking with members of another political party, let alone negotiating in good faith, can fuel accusations of betrayal and result in political and social punishment.

---

Republicans view the other side as closed-minded, and a substantial set of partisans judged their counterpart as immoral. *Most Republicans and Democrats View Each Other as More Closed-Minded Than Other Americans*, PEW RES. CTR. (Oct. 10, 2019), [https://www.people-press.org/2019/10/10/how-partisans-view-each-other/pp\\_2019-10-10\\_state-of-parties\\_2-01/](https://www.people-press.org/2019/10/10/how-partisans-view-each-other/pp_2019-10-10_state-of-parties_2-01/); *Growing Shares in Both Parties Give “Cold” Ratings to Those in Opposing Party*, *supra* note 2.

<sup>18</sup> *Most Republicans and Democrats View Each Other as More Closed-Minded Than Other Americans*, *supra* note 17.

<sup>19</sup> Shapiro, *supra* note 8, at 638.

<sup>20</sup> W. D. Hamilton, *The Genetical Evolution of Social Behaviour*, 7 J. THEORETICAL BIOLOGY 1, 8 (1964).

<sup>21</sup> Shapiro, *supra* note 8, at 635, 639.



In sum, the partisan mindset prioritizes allegiance to the tribe, fosters blind loyalty to the party platform, and condemns cross-party engagement.

C. *What Fosters the Partisan Mindset?*

There are many roads to political polarization—but through my international work in conflict resolution over the past thirty years, I have observed a prevalent dynamic that I call the “Declare-Defend-Descend Model.”

*Declare.* This dynamic begins when a person or group communicates their political identity subtly or explicitly within a conversation or through a formal or informal platform. Consider an example between two attorneys, Leigh and Ron. As they enter the elevator to head home after work, Leigh says, “I can’t believe how awful that congressman’s speech was last night!” She implicitly declares aspects of her political identity through this statement. But Ron hears her words as an assault on his own sacred values, awakening the partisan mindset in him and impelling him to say, “I actually thought the congressman had a lot of guts to say what he did!” He declares *his* political identity.

*Defend.* Now these lawyers are experiencing a clash of identities, and their conversation moves toward self-defensive measures.<sup>22</sup> Ron frets over his working relationship with Leigh. Sharing opposing politics feels taboo, and he worries that if they delve too deeply into political conversation, they may never get out. Taboos protect their relationship and identities from harm.

In an attempt to understand Leigh’s perspective, Ron asks, “Why didn’t you like the congressman’s speech?” His intention is admirable, but his tone exudes self-righteous indignation. The more Leigh justifies her stance, the more compelled he feels to argue back. He fights within himself to resist turning the conversation into a fierce debate and musters willpower not to pick apart her arguments one-by-one, let alone to dismiss her entire character.

*Descend.* Finally, the time comes for Ron to share his own perspective. The moment he launches into his rationale, she attacks it with unexpected ferocity, and they get consumed in a vertigo-like swirl of exasperated anger. Fortunately, they temper their emotions and close the conversation on an amicable note. That night, Ron laments to a close friend, “How can Leigh—in her right mind—criticize the congressman’s effectiveness?” The friend supports Ron’s perspective, bolstering his belief in the legitimacy of his claims and vindicating him of intellectual and moral ineptitude. This is identity politics in action—enlisting his friend to affirm his political stance so he can feel “in the right”—despite that same confirmation fortifying the

---

<sup>22</sup> SHAPIRO, *supra* note 10, at 148.

partisan mindset. Unsurprisingly, the relationship between Ron and Leigh descends.

This experience illuminates a set of emotional dynamics that draws us into the partisan mindset. In my book, *Negotiating the Nonnegotiable*, I introduce these “lures,” which often affect our relationships outside of our conscious awareness:<sup>23</sup>

- *Vertigo* is a warped state of consciousness in which a conflictual relationship consumes our emotional energies. In a polarized society, we can become obsessed with conflict amid the twenty-four-hour news cycle of “breaking news,” the constant stream of on-line partisan criticism, and daily political smears by public officials and activists.
- An *assault on the sacred* is an attack on the most meaningful aspects of our identity, whether political values, views, or beliefs. Leaders easily can incite constituents to take political action by framing an issue as an assault on sacred beliefs, values, and allegiances. A politician, for example, may try to gain support for war by framing it as a critical means “to eliminate grave, imminent threats to the lives of our children here in our homeland.”
- *Taboos* are social prohibitions—actions we are not supposed to do, thoughts we are not supposed to think, and emotions we are not supposed to feel. During times of political polarization, a taboo on cross-party engagement can affect people at all levels of society—from senior leadership being accused of betrayal if they talk with the “enemy” to everyday citizens who avoid political or social conversation with colleagues holding opposing views.
- The *repetition compulsion* lures us to repeat a dysfunctional pattern of behavior, as when officials predictably reach political stalemate at the same time each year over the same policy issue. This lure afflicts interpersonal relations, too. When discussing heated political issues, we may reenact a destructive dispute resolution behavior that we learned in our younger years, such as treating every conflict as a confrontation.

---

<sup>23</sup> *Id.* at xvii. In *Negotiating the Nonnegotiable*, I call these forces the “five lures of the tribal mind.”

- *Identity politics* is the process of allying with a person or group in order to advance a political purpose. A political leader delivering a speech may state that “in the name of our lost soldiers and family members, in the name of our constitutional ideals, we must stick together to fight the evil enemy!” This leader builds affiliation with the audience through shared history and values, and uses those associations to garner political support for military action.

These five lures pull us toward the partisan mindset in conflicts of all sizes, from international divides to everyday political disputes. In fact, if you re-read the argument between Ron and Leigh, you will notice that each of the lures was present and drew the colleagues toward the partisan mindset. The two lawyers’ relationship survived, but a democracy that operates purely on political tribalism has no backbone and slowly, piece by piece, can fall.<sup>24</sup> It turns out, however, that the partisan mindset is not a *fait accompli*. I now introduce an alternative mindset that can be enlisted to increase societal cohesion and democratic ideals.

### III. THE CIVIC MINDSET

The political mindset most responsible for bridging partisan divides in democratic societies is what I term the *civic mindset*, a unifying outlook that connects people together via identification as fellow citizens who work together to address the legitimate interests of political groups, resolve differing interests through mutually acceptable processes, and take communal need into account. This Section introduces the mindset, its operating principles, and the ways in which it is fostered.

The origin of the word *civic* provides insight into its meaning. It derives from the French word *civique*, meaning citizen, and can be traced further back to the Latin phrase *corona civica*, a garland of oak leaves and acorns awarded to those who saved a fellow citizen from death.<sup>25</sup> *Civics* invokes a sense of connection, duty, and responsibility to one’s homeland and the people who reside within it.

I have chosen to describe the mindset as “civic,” not “civil.” While the two words are closely aligned and share a similar etymology, the word *civil* often implies respectful behavior, whereas the word *civic* emphasizes the overarching political identity that holds citizens together. My view is that a society with a strong enough civic foundation can withstand even fierce

---

<sup>24</sup> STEVEN LEVITSKY & DANIEL ZIBLATT, *HOW DEMOCRACIES DIE* 77 (2018).

<sup>25</sup> *Civic*, OXFORD ADVANCED LEARNER’S DICTIONARY, <https://www.oxfordlearnersdictionaries.com/us/definition/english/civic> (last visited Jan. 20, 2020).

political debate and, in the ideal, emerge with new collective understandings and mutually beneficial action ideas.<sup>26</sup>

The civic mindset embraces national identity—but not to the exclusion of tribal identity. In general, both dimensions are *equally* important to nurture within the democratic political sphere. A strong national identity encourages policies and political behavior that serve the common good, and inclusive politics empowers diverse groups to voice and address their unique concerns through the shared political system. While tension will inevitably emerge between national and tribal concerns, the civic mindset motivates the quest for a *pareto-optimal* system of political decision making that optimizes these two concerns to the extent possible.

### A. Key Attributes

The civic mindset has three key characteristics that shape our political worldview:

#### 1. *Cooperative – But Not Naïve*

The civic mindset motivates us to proactively seek opportunities for cross-party collaboration while recognizing the inherent competitiveness of politics. From a civic perspective, the purpose of the republic is to satisfy citizens' interests through a combination of competition *and* cooperation. Social scientists call this a mixed-motive context, because there is an incentive for citizens to compete and to collaborate.<sup>27</sup> Political parties must compete for votes and influence while the broader citizenry can cooperate on countless matters of common concern. Political tribes may battle over laws on abortion, but the entire society can work together on a public campaign to stop teen pregnancy. The civic mindset helps us see this kind of civic possibility. Additionally, this mindset reminds citizens from across political ideologies of the necessity to work together to strengthen core democratic institutions and procedures—the very structures in which political parties compete for power.

#### 2. *Pluralistic*

In valuing the concept of citizenry, the civic mindset encourages toleration of diversity and the notion that multiple perceptions of truth can coexist. This does not mean we must abandon our convictions or assume

---

<sup>26</sup> See, e.g., *Civic*, MERRIAM-WEBSTER DICTIONARY, <https://www.merriam-webster.com/dictionary/civic> (last visited Feb. 20, 2020) (defining “civic” as “of or relating to a citizen, a city, citizenship, or community affairs”); *Civil*, MERRIAM-WEBSTER DICTIONARY, <https://www.merriam-webster.com/dictionary/civil> (last visited Feb. 20, 2020) (defining “civil” as “adequate in courtesy and politeness”).

<sup>27</sup> See, e.g., Philip S. Gallo, Jr. & Charles G. McClintock, *Cooperative and Competitive Behavior in Mixed-Motive Games*, J. CONFLICT RESOL. 68, 68 (1965) (explaining that in mixed-motive situations, the players' goals are “partially coincident and partially in conflict”).

others' beliefs are true. The minimal threshold of pluralism is toleration of diverse perspectives of truth—as long as those worldviews do not impinge upon anyone's constitutional rights or dignity. While this creates space for all citizens' voices, society still must decide the limits of pluralism. Should people *tolerate* multiple perceptions of truth, accept them at a distance, or revere them? Pluralism is key to civic life in a democracy, and it is the people's responsibility to determine its bounds.

### 3. *Community-spirited*

The civic mindset emphasizes broadscale community welfare.<sup>28</sup> This conception of civic responsibility stretches beyond the words in the Constitution to the *spirit* of democracy. We serve the broader community not solely because we must do so by law, but because we *want to*: we internalize an emotional commitment to form a “more perfect union.”<sup>29</sup> This expansive identity creates political space for tribes of all types to feel emotionally included in society and to engage in the political process. Every citizen is an equal part of the national project. The mission of the United States Army, for example, is not to protect *some* citizens over others but to provide for “the defense of the United States, the Commonwealths and possessions and any areas occupied by the United States.”<sup>30</sup>

Chart 2 summarizes the qualities of the partisan mindset and civic mindset. The partisan mindset lures us toward polarization whereas the civic mindset opens political space for cooperation.

*Chart 2. The contrasting characteristics of the partisan mindset and civic mindset.*

Partisan Mindset		Civic Mindset	
1.	Adversarial	1.	Cooperative
2.	Self-righteous	2.	Pluralistic
3.	Insular	3.	Community-spirited

### B. *Operating Principles*

The civic mindset operates on a few basic principles. First, loyalty to the republic takes on deep importance. Even kindergarteners in most U.S. states pledge allegiance every day “to the flag of the United States of America, and to the Republic for which it stands.”<sup>31</sup> This does not mean we must abandon our tribal identity or view it as inferior, but that we locate it—in all its glory and wholeness—within the broader sphere of a civic identity, at least within

<sup>28</sup> The preamble of the U.S. Declaration of Independence states that “all men are created equal.” THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776).

<sup>29</sup> U.S. CONST. pmb1.

<sup>30</sup> 10 U.S.C. § 7062 (2018).

<sup>31</sup> 4 U.S.C. § 4 (2018).

the context of political decision making. Much like the relationship between an eggshell and yolk, there is a symbiotic relationship between national and tribal identification. Widespread affiliation with the nation enhances the felt connection between tribes, which increases trust in the broader political system and better enables individual tribes to address their concerns.

Second, wherever possible, political issues are tackled through cooperative decision making rather than through blind loyalty politics. Political tribes still compete for power and influence but also draw on cross-party cooperation to optimize societal decision making. Inter-party cooperation may sound like a wishful vision within a polarized society, but this civic-minded outlook has proven successful at various points in American political history. A classic example is President Kennedy's emphasis on public service. Other cultures also have leveraged the power of cross-group cooperation for societal benefit. Japanese economic scholars credit much of their country's industrial success to the embrace of the philosophy of its Omi merchants who, beginning in the medieval period, measured success by the degree to which their business was good for themselves, the customer, and society.<sup>32</sup>

Third, the civic mindset fosters the felt duty to negotiate across party lines to solve societal issues, and to feel ashamed if one fails to do so in good faith. If enough leaders and social influencers advocate for this approach, norms of political communication can shift.<sup>33</sup> Strikingly, interstate war joins citizens of opposing political persuasions in a united front against an external enemy, producing a civic mindset of sorts. As polarization tears at the seams of democracy, citizens would be wise to adopt that same mindset.

### C. *Fostering the Civic Mindset*

There are at least two major pathways to cultivate a civic mindset. The first is to adopt the role of a civic leader who acts cooperatively, thinks pluralistically, and engages with a community-minded spirit. Anyone can get involved in local politics, write an op-ed, start a political blog, or serve the community by volunteering at an eldercare facility or school. One also can encourage *others* to take up their civic duties. A simple example took place in my home last week. My fourteen-year old son Noah sometimes gets

---

<sup>32</sup> Kenzo Moriguchi, *Forum Holds Up Omi Feudal Merchants as Models of Corporate Responsibility*, JAPAN TIMES (Oct. 30, 2001), <https://www.japantimes.co.jp/community/2001/10/30/events/forum-holds-up-omi-feudal-merchants-as-models-of-corporate-responsibility/#.XiqG9mhKh3h>.

<sup>33</sup> What happens if one political party embraces a civic mindset and the other holds to a partisan outlook? One might assume that aggressive partisans would pressure civic-minded cooperators to accommodate to their demands. But this oversimplifies the mechanisms of politics. Within any political tribe, there are internal forces advocating for and against issues, there are backchannel negotiations affecting policy decisions, there are cross-party meetings between political advisors—and all of these forums offer the opportunity for the civic-minded leader to influence the decision-making process and to produce results that are better for each political tribe and for society as a whole.

into conflict with his eight-year-old brother Liam. I needed Noah to babysit Liam but was nervous that they might fight. I nearly told Noah not to bully his younger brother while I was gone but instead said, “While I’m out, can you help Liam make his breakfast?” Inviting Noah into the role of household helper fostered in him a kind of civic mindset that elevated his purpose and improved my sons’ dynamic that morning.

Another method to foster the civic mindset is via a method I call the “Connect-Respect-Transcend Model,” which provides an overarching framework for interacting constructively across tribal lines.

*Connect.* The first step is to build emotional connection.<sup>34</sup> Recall Ron and Leigh’s conversation about the congressman’s speech. Before launching into identity declarations, they could have spent a few minutes catching up on each other’s lives. Small talk has big importance, for people come to view each other as multifaceted human beings rather than as partisan stereotypes, and their emotional connection creates a “holding environment” that can sustain the relationship even in the face of acrimony.<sup>35</sup> Ron also could have been much more cautious in airing political differences. Rather than responding reflexively in political banter, he could have inquired about Leigh’s interest in talking politics. Though she initiated the conversation, she may have intended to make a simple declaration and not to engage in full-fledged political debate. With mutual consent, they could have entered the taboo territory of political dialogue with greater sensitivity.

*Respect.* Ron could have demonstrated greater respect for Leigh’s views by asking open-ended questions: What provoked her strong reaction to the congressman’s talk? What values felt assaulted? He could have communicated his understanding of her views, checked in with her to make sure he understood correctly, and shared which values of hers most resonated with him. By respecting her experience, he could have built greater emotional connection.

Resisting the repetition compulsion was paramount to the modest success of their conversation. Given the intensity of Ron’s political beliefs, he was hyperaware of the risk of their conversation becoming adversarial and sought to temper the expression of his strong views. Nevertheless, he could have suggested a simple process to guide their conversation, such as having them each share the *personal significance* of their views, turning the

---

<sup>34</sup> Emotional connections must be built at the national as well as regional and interpersonal levels. In examining ways to stem the tide of political tribalism in the United States, Amy Chua notes that “citizens will . . . need to collectively fashion a national identity capable of resonating with and holding together Americans of all sorts—old and young, immigrant and native born, urban and rural, rich and poor, descendants of slaves as well as descendants of slave owners.” Amy Chua, *Tribal World: Group Identity Is All*, 97 FOREIGN AFF. 25, 33 (2018).

<sup>35</sup> See D. W. Winnicott, *The Theory of the Parent-Infant Relationship*, 41 INT’L J. PSYCHO-ANALYSIS 585, 591 (1960) (indicating that the holding environment of infants is a form of loving).

debate over whose politics is “right” into an emotionally revealing exchange about the roots of their political identities.

*Transcend.* Ron and Leigh could have been more open to listening for sake of learning. Neither individual’s political views would have changed, but they could have expanded their understanding of the political terrain and important interests at stake for various constituencies.

This is the civic mindset coming to life: We approach differences jointly, commit to the “relentless we,” and refuse to fall prey to political tribalism.<sup>36</sup> The more we listen, learn, and share, the more we fall into a positive vertigo—a free-flowing conversation that stirs enthusiasm, curiosity, and fulfillment of our civic duty. We maintain our beliefs while transcending difference.

#### IN SUMMARY

When political polarization threatens to undermine the ideals and functionality of democracy, there is a societal imperative to embrace a civic mindset. While the partisan mindset pits groups against one another and reduces political space for collaboration, the civic mindset motivates national cohesion and expands political space for mutually beneficial decision making. In this frame of mind, citizens across interest groups cooperate on issues of shared concern, embrace pluralism, and foster a community-minded spirit. Political parties still compete for power and influence—but within a broader identity that binds them together in the quest for a more perfect union.

---

<sup>36</sup> SHAPIRO, *supra* note 10, at 128–29.





# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### The Role of Lawyers and Law Schools in Fostering Civil Public Debate

JENNIFER K. ROBBENOLT & VIKRAM D. AMAR

*Partisanship can make policy discussion and civil debate difficult. Partisan differences in how facts and policies are understood contribute to the escalation of conflict and a lack of cooperation. Lawyers are not immune from these human tendencies. But good lawyers have, and good law schools teach, values, knowledge, and skills that can aid in fostering and modeling more productive debate and resolution of conflict.*

*Lawyers are trained and socialized to internalize and safeguard the foundational tenets of our constitutional democracy, to uphold the law even when it does not reflect their own individual preferences. The professional rules of conduct encourage lawyers to separate the professional from the personal, and expect that vigorous debate, dissent, and zealous advocacy will be done in a professional manner. Lawyers are taught to think about issues, cases, or arguments from multiple sides and to value rational argument, the primacy of evidence and facts, and neutral processes in which cases are decided on their merits. The nuanced approaches to conflict that are required of lawyers—distinguishing productive and unproductive conflict, both creating and claiming value, and acting as both advisors and advocates—equip lawyers with abilities that help them generate and manage more productive debate.*

*Law schools, then, should strive to provide even better grounding in these values, knowledge, and skills. Lawyers should endeavor to highlight for themselves, their clients and colleagues, and their opponents nuanced approaches to conflict and debate. And law schools and lawyers should work to educate the broader citizenry about the values of our constitutional democracy and to model effective and civil dispute resolution strategies.*



# The Role of Lawyers and Law Schools in Fostering Civil Public Debate

JENNIFER K. ROBBENNOLT\* & VIKRAM D. AMAR\*\*

We are in a time in which political polarization is frequently in the headlines,<sup>1</sup> public opinion polls reveal a pervasive sense of division<sup>2</sup> and a sense that political discussions have become less grounded in facts,<sup>3</sup> and many are concerned about the civility (or lack thereof) with which we treat each other across political differences.<sup>4</sup> Partisanship can, indeed, make policy discussion and civil debate difficult. And there are many aspects of human psychology that can contribute to the hurdles. Take a prominent example: people interpret policies and information differently depending

---

\* Alice Curtis Campbell Professor of Law, Professor of Psychology, University of Illinois College of Law. Our thanks go to Jean Sternlight for her very helpful comments and suggestions.

\*\* Dean, Iwan Foundation Professor of Law, University of Illinois College of Law.

<sup>1</sup> See, e.g., Jaclyn Gallucci, *When It Comes to Politics, Americans Are Divided. Can Data Change That?*, FORTUNE (July 17, 2019, 1:18 PM), <https://fortune.com/2019/07/17/political-polarization-in-america-define/> (discussing “polarization unity” and how polarized voters and campaign donor influence furthers political division); Natalie Pattillo, *As Shutdown Pauses, Coverage Focuses on Partisan Polarization*, COLUM. JOURNALISM REV. (Jan. 25, 2019), <https://www.cjr.org/politics/shutdown-partisan-coverage.php> (describing focus of news coverage on polarization). See also Shanto Iyengar et al., *The Origins and Consequences of Affective Polarization in the United States*, 22 ANN. REV. POL. SCI. 129, 130–31 (2019) (describing affective polarization and how to mitigate it); Samara Klar et al., *Opinion, Is America Hopelessly Polarized, or Just Allergic to Politics?*, N.Y. TIMES (Apr. 12, 2019), <https://www.nytimes.com/2019/04/12/opinion/polarization-politics-democrats-republicans.html> (suggesting that we are less polarized than we think). See generally JOSHUA GREENE, MORAL TRIBES: EMOTION, REASON, AND THE GAP BETWEEN US AND THEM 5, 14–16 (2014) (describing “us v. them” conflicts and the role they play in moral decision making); JONATHAN HAIDT, THE RIGHTEOUS MIND: WHY GOOD PEOPLE ARE DIVIDED BY POLITICS AND RELIGION, at xi–xii (2012) (describing how a “righteous” mind enables both cooperation and moralistic conflict).

<sup>2</sup> See, e.g., ROBERT P. JONES & MAXINE NAJLE, PRRI, AMERICAN DEMOCRACY IN CRISIS: THE FATE OF PLURALISM IN A DIVIDED NATION 28 (2019), <https://www.prii.org/research/american-democracy-in-crisis-the-fate-of-pluralism-in-a-divided-nation/> (finding that “Americans are nearly unanimous in their belief that the country is divided over politics (91%), with 74% of Americans saying that the country is very divided”).

<sup>3</sup> *Most Americans Say Political Debate in the U.S. Has Become Less Respectful, Fact-Based, Substantive*, PEW RES. CTR. (July 18, 2019), [https://www.pewresearch.org/fact-tank/2019/07/18/americans-say-the-nations-political-debate-has-grown-more-toxic-and-heated-rhetoric-could-lead-to-violence/ft\\_19-07-18\\_toxicpolitics\\_most-americans-say-political-debate-us-less-respectful-fact-based-substantive/](https://www.pewresearch.org/fact-tank/2019/07/18/americans-say-the-nations-political-debate-has-grown-more-toxic-and-heated-rhetoric-could-lead-to-violence/ft_19-07-18_toxicpolitics_most-americans-say-political-debate-us-less-respectful-fact-based-substantive/) (finding that 76% of American adults believe political debate has grown less fact-based in recent years).

<sup>4</sup> WEBER SHANDWICK, POWELL TATE & KRC RESEARCH, CIVILITY IN AMERICA 2019: SOLUTIONS FOR TOMORROW 2, 8–9, 10–15 (2019), <https://www.webershandwick.com/news/civility-in-america-2019-solutions-for-tomorrow/>.

on their perspective and preferences,<sup>5</sup> and assess evidence in accordance with preferences and prior beliefs.<sup>6</sup> And recent research has found that people's ideological beliefs can even make it difficult to evaluate the basic logical validity of arguments.<sup>7</sup>

Despite these tendencies, we tend to believe that our own perceptions and experiences are objective and accurate, and often fail to realize the ways that our perceptions are influenced by our own perspective, knowledge, expectations, and desires—a phenomenon known as *naïve realism*.<sup>8</sup> This naïve realism creates the “feeling that [our] own take on the world enjoys particular authenticity, and that other actors will, or at least should, share that take, if they are attentive, rational, and objective perceivers of reality and open-minded seekers of truth.”<sup>9</sup> This feeling tends to make us confident that we should be able to persuade others of the rightness of our positions.<sup>10</sup> But when others persist in having different views, it can lead us to conclude that they are unreasonable, biased, or

<sup>5</sup> See, e.g., Albert H. Hastorf & Hadley Cantril, *They Saw a Game: A Case Study*, 49 J. ABNORMAL & SOC. PSYCHOL. 129, 132–34 (1954) (finding partisan interpretations by the fans of opposing sports teams); Dan M. Kahan et al., “*They Saw a Protest*”: *Cognitive Illiberalism and the Speech-Conduct Distinction*, 64 STAN. L. REV. 851, 854–55, 883–85 (2012) (finding that preferences and prior beliefs influenced interpretations of videotape of a political demonstration); David Tannenbaum et al., *On the Misplaced Politics of Behavioural Policy Interventions*, 1 NATURE HUM. BEHAV. 1, 5 (2017) (finding that people find behavioral interventions more ethical when the nature of the intervention matches their political beliefs and less ethical when it does not); Leaf Van Boven et al., *Psychological Barriers to Bipartisan Public Support for Climate Policy*, 13 PERSP. ON PSYCHOL. SCI. 492, 493, 496–500 (2018) (describing how partisans devalue policy proposals from an opposing party).

<sup>6</sup> Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098, 2108 (1979); Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175, 197 (1998).

<sup>7</sup> Anup Gampa et al., *(Ideo)Logical Reasoning: Ideology Impairs Sound Reasoning*, 10 SOC. PSYCHOL. & PERSONALITY SCI. 1075, 1082 (2019). See generally Peter H. Ditto et al., *At Least Bias Is Bipartisan: A Meta-Analytic Comparison of Partisan Bias in Liberals and Conservatives*, 14 PERSP. ON PSYCHOL. SCI. 273 (2019) (reporting a meta-analysis of the tendency “to evaluate otherwise identical information more favorably when it supports one’s political beliefs and allegiances”).

<sup>8</sup> Emily Pronin et al., *Understanding Misunderstanding: Social Psychological Perspectives*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 636, 646 (Thomas Gilovich et al. eds., Cambridge Univ. Press 2002).

<sup>9</sup> *Id.* Naïve realism also leads to the belief that we make more accurate assessments of other people than they make of us. This is the *illusion of asymmetric insight*. “We insist that our ‘outsider perspective’ affords us insights about our peers that they are denied by their defensiveness, egocentricity, or other sources of bias. By contrast, we rarely entertain the notion that others are seeing us more clearly and objectively than we see ourselves.” Emily Pronin et al., *You Don’t Know Me, But I Know You: The Illusion of Asymmetric Insight*, 81 J. PERSONALITY & SOC. PSYCHOL. 639, 639 (2001). This can mean that we are prone to “talk when we would do well to listen and to be less patient than we ought to be when others express the conviction that they are the ones who are being misunderstood or judged unfairly.” *Id.* at 652–53.

<sup>10</sup> Lee Ross & Andrew Ward, *Naïve Realism in Everyday Life: Implications for Social Conflict and Misunderstanding*, in VALUES AND KNOWLEDGE 103, 116 (Edward S. Reed et al. eds., 1996).

ill-motived.<sup>11</sup> Research has found that people commonly conclude that those who disagree with them are biased, simply because they disagree.<sup>12</sup> Once a person attributes bias to another, they tend to see their conflict as more pervasive, to expect cooperation to be less worthwhile, and to act more competitively. Not surprisingly, this tends to cause the other person to respond in kind, creating a spiral of conflict.<sup>13</sup> And incivility makes arguments seem less sound,<sup>14</sup> likely contributing to the escalation of conflict.

When differences in perspective are particularly focal—e.g., two people are on different sides of a contentious issue—we tend to overestimate those differences.<sup>15</sup> Similarly, we tend to overestimate the degree to which things like ideology and self-interest influence other people's views and behavior, believing that others are more motivated or influenced by these than we are ourselves.<sup>16</sup> One study asked people with varying views on an issue to express their own judgments and also to predict how their understandings would differ from those with other political views.<sup>17</sup> While there were, in fact, differences in how people with different political views perceived the case, these differences were relatively small compared to the large differences predicted by the participants.<sup>18</sup> These sorts of mispredictions can mean that people are overly doubtful and cynical about the potential fruits of collaboration or finding common ground.<sup>19</sup>

---

<sup>11</sup> *Id.* at 111; Emily Pronin et al., *Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self Versus Others*, 111 *PSYCHOL. REV.* 781, 793 (2004); Leigh Thompson & George Loewenstein, *Egocentric Interpretations of Fairness and Interpersonal Conflict*, 51 *ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES* 176, 193 (1992).

<sup>12</sup> Kathleen A. Kennedy & Emily Pronin, *When Disagreement Gets Ugly: Perceptions of Bias and the Escalation of Conflict*, 34 *PERSONALITY & SOC. PSYCHOL. BULL.* 833, 845 (2008).

<sup>13</sup> *Id.*

<sup>14</sup> Jason R. Popan et al., *Testing the Effects of Incivility During Internet Political Discussion on Perceptions of Rational Argument and Evaluations of a Political Outgroup*, 96 *COMPUTERS HUM. BEHAV.* 123, 130 (2019).

<sup>15</sup> Nicholas Epley & Eugene M. Caruso, *Perspective Taking: Misstepping into Others' Shoes*, in *HANDBOOK OF IMAGINATION AND MENTAL SIMULATION* 297, 304 (Keith D. Markman et al. eds., 2009).

<sup>16</sup> Chip Heath, *On the Social Psychology of Agency Relationships: Lay Theories of Motivation Overemphasize Extrinsic Incentives*, 78 *ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES* 25, 26 (1999); Dale T. Miller, *The Norm of Self-Interest*, 54 *AM. PSYCHOL.* 1053, 1053 (1999); Rebecca K. Ratner & Dale T. Miller, *The Norm of Self-Interest and Its Effects on Social Action*, 81 *J. PERSONALITY & SOC. PSYCHOL.* 5, 14 (2001); see also Justin Kruger & Thomas Gilovich, *"Naive Cynicism" in Everyday Theories of Responsibility Assessment: On Biased Assumptions of Bias*, 76 *J. PERSONALITY & SOC. PSYCHOL.* 743, 751 (1999) (finding that "people have cynical intuitions about how others assess responsibility").

<sup>17</sup> Robert J. Robinson et al., *Actual Versus Assumed Differences in Construal: "Naive Realism" in Intergroup Perception and Conflict*, 68 *J. PERSONALITY & SOC. PSYCHOL.* 404, 414 (1995).

<sup>18</sup> *Id.*

<sup>19</sup> *Id.* at 416.

Lawyers, of course, are not immune from these human tendencies. The adversarial nature of our legal system and the representative nature of legal practice means that lawyers must necessarily operate much of the time with a partisan perspective.<sup>20</sup> This partisan lens can contribute to the sorts of spirals just described.<sup>21</sup> And incivility in the profession has been a topic of concern.<sup>22</sup>

But good lawyers have, and good law schools teach, a range of values, knowledge, and skills that should be useful in fostering and modeling more productive debate and resolution of conflict.<sup>23</sup>

Importantly, lawyers are trained and socialized to internalize and safeguard the foundational tenets of our constitutional democracy.<sup>24</sup> The rule of law in the United States—and the prospect over time of formulating better policy that itself will be respected as legitimate—depends on notice and opportunity to be heard, the robust exercise of freedom of speech and a free press, substantive engagement of ideas, and confidence that dissenting viewpoints are engaged on their merits rather than merely overridden or

---

<sup>20</sup> See, e.g., George Loewenstein et al., *Self-Serving Assessments of Fairness and Pretrial Bargaining*, 22 J. LEGAL STUD. 135, 150–51 (1993) (finding that representation of a party on one side of a legal case influences perceptions of fairness). See also Dan M. Kahan et al., *Whose Eyes Are You Going to Believe? Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 HARV. L. REV. 837, 838, 887, 896–97 (2009) (critiquing the Supreme Court’s failure to recognize how different people might interpret a video of a police chase in *Scott v. Harris*).

<sup>21</sup> See, e.g., Stephen M. Garcia et al., *Morally Questionable Tactics: Negotiations Between District Attorneys and Public Defenders*, 27 PERSONALITY & SOC. PSYCHOL. BULL. 731, 737 (2001) (finding that attorneys viewed questionable negotiation tactics as more appropriate when used in response to the perceived use of questionable tactics by the other side).

<sup>22</sup> See NAT’L CTR. FOR PROF’L & RESEARCH ETHICS, SURVEY ON PROFESSIONALISM: A STUDY OF ILLINOIS LAWYERS 2014, at 5 (2014), <https://www.2civility.org/wp-content/uploads/2015/04/Study-of-Illinois-Lawyers-2014.pdf> (noting that more than eighty-five percent of lawyers surveyed reported experience with some kind of uncivil behavior in the past six months, such as sarcasm, condescension, misrepresentation, or negotiating in bad faith); LAUREN STILLER RIKLEEN, RIKLEEN INST. FOR STRATEGIC LEADERSHIP, SURVEY OF WORKPLACE CONDUCT AND BEHAVIORS IN LAW FIRMS 33, 38 (2018), <https://wbawbf.org/sites/WBAR-PR1/files/WBA%20Survey%20of%20Workplace%20Conduct%20and%20Behaviors%20in%20Law%20Firms%20FINAL.pdf> (describing inappropriate behavior at law firms); Sam Skolnik, *More Than Third of Female Lawyers Harassed at Work, Survey Shows*, BLOOMBERG L.: BIG L. BUS. (Nov. 29, 2018), <https://biglawbusiness.com/more-than-third-of-female-lawyers-harassed-at-work-survey-shows> (reporting that more than a third of female lawyers have been sexually harassed at work).

<sup>23</sup> See Carrie Menkel-Meadow, *The Lawyer’s Role(s) in Deliberative Democracy*, 5 NEV. L.J. 347, 349–51 (2004) (suggesting that lawyers, as neutral advocates, are well-suited to assist in democratic discourse).

<sup>24</sup> Interestingly, the President of the American Bar Association and the President of the Association of American Law Schools both recently highlighted the importance of lawyers and legal education in upholding and educating the public about the rule of law and the “pillars” of constitutional democracy. Judy Perry Martinez, *President’s Letter: Promise to a Nation*, A.B.A. J., Sept.–Oct. 2019, at 6; Vicki Jackson, President, Ass’n of Am. Law Sch., 2019 Presidential Address at the Second Meeting of the AALS House of Representatives: Pillars of Democracy: Law, Representation, and Knowledge (Jan. 4, 2019), <https://www.aals.org/about/publications/newsletters/winter-2019/pillars-of-democracy/>.

ignored.<sup>25</sup> Lawyers are the cultural custodians of this distinctive government by the people, for the people, and of the people.<sup>26</sup> When we say—as we often do—that we are a nation of laws, not people,<sup>27</sup> what we mean is that our highest obedience is to a set of principles of governance, not to the particular people who govern. This is why no person is above (or below) the law, and why lawyers are required and trained to uphold the law, even when it does not reflect their own individual preferences.<sup>28</sup> That does not, of course, mean that lawyers passively accept laws that they believe to be unjust. Indeed, a big part of a lawyer’s role is to work for legal reform<sup>29</sup> through the mechanisms of our constitutional democracy. That is why lawyers are permitted to take positions that are not supported by existing law, provided they are, in the words of one important ethics formulation, “warranted . . . by a nonfrivolous argument for the exten[sion], modif[ication], or revers[al of] existing law[,] or for [the] establish[ment] of new law.”<sup>30</sup>

In serving as institutional and cultural custodians, lawyers are required to assume particular roles. It is for this reason that professional rules of conduct encourage—and successful law schools teach—lawyers to separate the professional from the personal.<sup>31</sup> Vigorous debate, dissent, and zealous advocacy are all valued—and can all be done in a professional manner. As Shakespeare once said: “[D]o as adversaries do in law, strive

---

<sup>25</sup> *Overview – Rule of Law*, U.S. CTS., <https://www.uscourts.gov/educational-resources/educational-activities/overview-rule-law> (last visited Jan. 26, 2020).

<sup>26</sup> *Former Solicitor General: Government Lawyers Critical to Rule of Law in Troubling Times*, A.B.A. NEWS (Oct. 26, 2018), <https://www.americanbar.org/news/abanews/aba-news-archives/2018/10/former-solicitor-general--government-lawyers-critical-to-rule-of/>.

<sup>27</sup> David Davenport, *A Nation of Laws, Not Men*, HOOVER INSTITUTION (Sept. 2, 2013), <https://www.hoover.org/research/nation-laws-not-men>.

<sup>28</sup> MODEL RULES OF PROF’L CONDUCT r. 1.2 (AM. BAR ASS’N 1983) (stating the lawyer must abide by a client’s decision and that representation of a client does not mean the lawyer endorses the client’s political or moral views).

<sup>29</sup> The preamble to the Model Rules of Professional Conduct states, “[a]s a public citizen, a lawyer should seek improvement of the law, the administration of justice and the quality of service rendered by the legal profession. As a member of a learned profession, a lawyer should cultivate knowledge of the law beyond its use for clients, employ that knowledge in reform of the law and work to strengthen legal education.” MODEL RULES OF PROF’L CONDUCT pmb1. (AM. BAR ASS’N 1983).

<sup>30</sup> FED. R. CIV. P. 11(b)(2).

<sup>31</sup> *See, e.g.*, CODE OF PROF’L COURTESY no. 10 (KY. BAR ASS’N), <https://www.kybar.org/page/procourtesy> (“A lawyer should recognize that the conflicts within a legal matter are professional and not personal and should endeavor to maintain a friendly and professional relationship with other attorneys in the matter. In other words, ‘leave the matter in the courtroom.’”); OBA STANDARDS OF PROFESSIONALISM r. 2.7 (OKLA. BAR ASS’N 2006), <https://www.okbar.org/ec/standardsofprofessionalism/> (“We understand, and will impress upon our client, that reasonable people can disagree without being disagreeable; and that effective representation does not require, and in fact is impaired by, conduct which objectively can be characterized as uncivil, rude, abrasive, abusive, vulgar, antagonistic, obstructive or obnoxious.”). For a content analysis of state bar civility codes, see Donald E. Campbell, *Raise Your Right Hand and Swear to Be Civil: Defining Civility as an Obligation of Professional Responsibility*, 47 GONZ. L. REV. 99, 107–28 (2011).



mightily, but eat and drink as friends.”<sup>32</sup> In addition to the rules and norms of professionalism, lawyers also have their own interests in treating those on the opposing side of a particular issue, case, or dispute with respect. Lawyers are repeat-players and are likely to encounter each other in future cases. Reputation—for ethicality, competence, problem-solving ability, or civility—is an important asset that should not be squandered.<sup>33</sup> Moreover, someone who is on the opposing side in this case may be a partner or collaborator in the next.

Lawyers are trained to think about issues, cases, or arguments from multiple sides.<sup>34</sup> Lawyers cannot make good predictions if they have not thought about an issue in a complex, and multifaceted way. And lawyers would not be able to act as good advocates if they hadn’t at least anticipated the counterarguments. As a profession, moreover, we value principled analysis and rational argument, rather than foregone conclusions.<sup>35</sup> Our system is grounded in the primacy of evidence and facts<sup>36</sup> and the value of neutral processes in which cases are decided on their merits.<sup>37</sup>

---

<sup>32</sup> WILLIAM SHAKESPEARE, *THE TAMING OF THE SHREW* act 1, sc. 2, ll. 281–82 (Oxford, Clarendon Press 1921) (1594).

<sup>33</sup> See, e.g., Catherine H. Tinsley et al., *Tough Guys Finish Last: The Perils of a Distributive Reputation*, 88 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 621, 640 (2002) (finding that negotiator reputation influenced negotiator behavior). See also Nancy A. Welsh, *The Reputational Advantages of Demonstrating Trustworthiness: Using the Reputation Index with Law Students*, 28 NEGOT. J. 117, 120 (2012) (“Perhaps paradoxically, the negotiators who are most likely to have a reputation for effectiveness are those who acknowledge that legal negotiation is just as much about the other people who are involved and abiding by relevant professional norms as it is about the task of competing for a favorable share of apparently scarce resources.”). See Catherine H. Tinsley et al., *Reputation in Negotiation*, in *THE NEGOTIATOR’S DESK REFERENCE* 255, 256–58 (Chris Honeyman & Andrea Kupfer Schneider eds., 2017) (noting that a lawyer with an integrative reputation is perceived to be more effective by her negotiation counterpart).

<sup>34</sup> Charles G. Lord et al., *Considering the Opposite: A Corrective Strategy for Social Judgment*, 47 J. PERSONALITY & SOC. PSYCHOL. 1231, 1239–41 (1984) (exploring the effects of considering alternative outcomes); David McCraw, *Think Like a Libel Lawyer*, N.Y. TIMES (Mar. 9, 2019), <https://www.nytimes.com/2019/03/09/opinion/sunday/think-like-a-libel-lawyer.html>. See also Russell Korobkin, *Psychological Impediments to Mediation Success: Theory and Practice*, 21 OHIO ST. J. ON DISP. RESOL. 281, 296 (2006) (suggesting the utility of taking the perspective of a disagreeable adjudicator).

<sup>35</sup> Vikram David Amar, *The Helpful Role Lawyers Can Play in Rebuilding American Democracy*, JUSTIA: VERDICT (Jan. 12, 2018), <https://verdict.justia.com/2018/01/12/helpful-role-lawyers-can-play-rebuilding-american-democracy> (“Lawyers apply logic—and not preconceived notions or forgone conclusions—to the facts. Logic must be tempered by history and experience but at base relies on principled reasoning.”).

<sup>36</sup> *Id.* (“Lawyers deal in facts, grounded in evidence—they don’t trade in speculation, and certainly they do not create or promote fabricated falsehood.”). See also McCraw, *supra* note 34 (discussing the importance of facts to libel lawyers).

<sup>37</sup> See, e.g., Steven L. Blader & Tom R. Tyler, *A Four-Component Model of Procedural Justice: Defining the Meaning of a “Fair” Process*, 29 PERSONALITY & SOC. PSYCHOL. BULL. 747, 748 (2003) (describing the importance of neutrality to procedural justice).

Conflict resolution skills—often taught in dispute resolution courses or clinics—are also an important part of the lawyer’s toolkit.<sup>38</sup> Lawyers are routinely called upon to assist parties on opposite sides of a deal, case, or issue to come to mutual agreement. Listening actively, with curiosity, and for understanding—and listening with respect even in disagreement—helps lawyers understand the interests of the parties.<sup>39</sup> Lawyers can help bring clients along to agreement by counseling them to assess both their own interests and those of the other side.<sup>40</sup> Lawyers rely on empathy and creativity to craft or frame proposals that will satisfy the interests of both sides.<sup>41</sup> Lawyers know that differences in interests or values make it possible to create value through exchange.<sup>42</sup>

Of course, any serious and productive attempt to better align law school curricula and culture with modes of argumentation that might better serve individual law school graduates and society must reckon with the reality that conflict itself and lawyers’ roles are each varied. These variations, and the nuanced approaches to conflict that they require of lawyers, may themselves equip lawyers with abilities that can help them generate and manage more productive debate.

Conflict is often thought of as necessarily bad. But conflict theorists distinguish between constructive and destructive conflict.<sup>43</sup> And, indeed,

---

<sup>38</sup> See RANDALL KISER, *SOFT SKILLS FOR THE EFFECTIVE LAWYER* 96 (2017) (“Higher levels of self-control are correlated with . . . superior conflict resolution skills.”); JENNIFER K. ROBBENOLT & JEAN STERNLIGHT, *PSYCHOLOGY FOR LAWYERS: UNDERSTANDING THE HUMAN FACTORS IN NEGOTIATION, LITIGATION, AND DECISION MAKING* 1 (2012) (describing the skills that will help lawyers to be better negotiators and counselors); Menkel-Meadow, *supra* note 23, at 359–60 (describing conflict management and consensus building skills). See also Symposium, *ADR’s Place in Navigating a Polarized Era*, TEX. A&M L. REV. (forthcoming) (considering the pros and cons of applying alternative dispute resolution techniques to contentious issues in a polarized climate).

<sup>39</sup> See, e.g., RISKIN ET AL., *DISPUTE RESOLUTION AND LAWYERS* 76–90 (6th ed. 2019) (describing active listening); Jack Zenger & Joseph Folkman, *What Great Listeners Actually Do*, HARV. BUS. REV. (July 14, 2016), <https://hbr.org/2016/07/what-great-listeners-actually-do> (“[P]eople perceive the best listeners to be those who periodically ask questions that promote discovery and insight. . . . Good listeners may challenge assumptions and disagree, but the person being listened to feels the listener is trying to help.”). See also Jonathan R. Cohen, “Open-Minded Listening”, 5 CHARLOTTE L. REV. 139, 144 (2014) (discussing the importance of open-minded listening and the factors that hinder and promote it).

<sup>40</sup> Perspective taking is complicated. See, e.g., Tal Eyal et al., *Perspective Mistaking: Accurately Understanding the Mind of Another Requires Getting Perspective, Not Taking Perspective*, 114 J. PERSONALITY & SOC. PSYCHOL. 547, 547 (2018) (finding that perspective taking is difficult and that accuracy about another person is better aided by engaging in conversation with them).

<sup>41</sup> Andrea Kupfer Schneider, *Teaching a New Negotiation Skills Paradigm*, 39 WASH. U. J.L. & POL’Y 13, 27–37 (2012).

<sup>42</sup> See, e.g., Carrie Menkel-Meadow, *Why We Can’t “Just All Get Along”*: Dysfunction in the Polity and Conflict Resolution and What We Might Do About It, J. DISP. RESOL. 5, 9 (2018) (describing the importance of focusing on interests and values in conflict resolution).

<sup>43</sup> See LEWIS A. COSER, *FUNCTIONS OF SOCIAL CONFLICT* 16, 20–21, 47 (1956) (discussing both the dysfunctional and beneficial aspects of conflict); MORTON DEUTSCH, *THE RESOLUTION OF*

while conflict can sometimes be devastating, it can also be “the seedbed that nourishes social change”<sup>44</sup> or the impetus to engage in creative thinking about how to accommodate or reconcile legitimate, though differing, interests. The goal, therefore, is not to eliminate disagreement or to invariably compromise quickly. Instead, it is important to distinguish between conflict that is necessary or useful and conflict that is unnecessary or unproductive. A lawyer might, for example, reasonably choose to litigate a case rather than agree to a settlement that does not meet her clients’ interests or might pursue a strategy of litigation in the service of legal reform. But she might also readily agree to a request for a delay in the proceedings from opposing counsel, when doing so would not compromise her client’s interests.

Not only is conflict itself multifaceted, but lawyers are trained to operate in many different types of advocacy roles. In a transactional setting, getting to an agreement with other stakeholders (while preserving the things that are most important to one’s own client) is often the ultimate mark of success; lawyers who are unable to ultimately facilitate deal-making have a tough time earning a living in transactional practice areas. To the extent that naïve realism and other confirmation biases make it harder to appreciate—much less work to address or accommodate in relatively low-cost ways—the interests, perspectives, and proposals of other stakeholders in an agreement, classroom and skills training that helps students recognize and combat such subjective blind spots can only be to the good.

In litigation (or litigation-like) arenas, things might get even more complicated. Even in litigation, most cases end up being settled out of court<sup>45</sup> and, even in cases that go to trial, there are many procedural or substantive agreements to be made along the way.<sup>46</sup> Litigators, therefore, need to be dealmakers. But not all cases can or should be settled,<sup>47</sup> and our

---

CONFLICT: CONSTRUCTIVE AND DESTRUCTIVE PROCESSES 17 (1973) (differentiating constructive and destructive conflict).

<sup>44</sup> DEAN G. PRUITT & SUNG HEE KIM, *SOCIAL CONFLICT: ESCALATION, STALEMATE, AND SETTLEMENT* 10 (3d ed. 2004).

<sup>45</sup> Theodore Eisenberg & Charlotte Lanvers, *What Is the Settlement Rate and Why Should We Care?*, 6 J. EMPIRICAL LEGAL STUD. 111, 112 (2009); Marc Galanter & Mia Cahill, “Most Cases Settle”: *Judicial Promotion and Regulation of Settlements*, 46 STAN. L. REV. 1339, 1339 (1994). Litigation is really a process of “litigotiation.” Marc Galanter, *Worlds of Deals: Using Negotiation to Teach About Legal Process*, 34 J. LEGAL EDUC. 268, 268 (1984).

<sup>46</sup> See, e.g., J.J. Prescott & Kathryn E. Spier, *A Comprehensive Theory of Civil Settlement*, 91 N.Y.U. L. REV. 59, 62 (2016) (describing the range of agreements made by parties throughout the litigation process).

<sup>47</sup> The right balance is debated. See Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073, 1075 (1984) (discussing problems with settlement); Carrie Menkel-Meadow, *Whose Dispute Is It Anyway?: A Philosophical and Democratic Defense of Settlement (in Some Cases)*, 83 GEO. L.J. 2663, 2665–66 (1995) (responding to Fiss); Symposium, *Against Settlement: Twenty-Five Years Later*, 78 FORDHAM L. REV. 1117 (2009) (discussing various views and perspectives on settlements).

adversarial process is premised on the notion of vigorous advocacy. This means that litigators must be prepared to simultaneously cooperate and advocate. It also means that litigators must necessarily think about other audiences, including judges, juries, and arbitrators. To be sure, to the extent that these neutral finders of fact (or law) lament needless disagreement, bickering, and incivility by the participants in a dispute-resolution process, those lawyers who are best-trained in being reasonable—and appearing to be reasonable—will be rewarded. But there is nothing to guarantee that deciders of cases always react negatively to entrenched or combative presentation.<sup>48</sup> A lawyer who has reason to know (or even think), for example, that a particular “old-school” judge (or a particular jury panel) will view empathy for and acknowledgement of the plausibility of the other side’s positions as weaknesses or implicit doubts about the validity of one’s own arguments, is duty bound by her oath to take such information into account when framing her presentation.

Even within the roles of transactional dealmaker or litigator, the lawyer’s role has many facets. For example, to appropriately advise their clients as to the merits of a deal or lawsuit and the prospects for a better deal or settlement, lawyers must be able to objectively evaluate the deal or case. At the same time, to effectively promote clients’ interests, lawyers must act as advocates. It is not easy to wear these two hats—neutral observer and partisan advocate—at the same time,<sup>49</sup> as these roles require different skills. Similarly, lawyers “cannot steward . . . effective deal[s] without both minimizing and facilitating risk taking.”<sup>50</sup> As advisors, lawyers must often simultaneously seek creative solutions for clients while also ensuring their compliance with the law.<sup>51</sup>

---

<sup>48</sup> The research literature is sparse and somewhat mixed. See Margaret S. Gibbs et al., *Cross-Examination of the Expert Witness: Do Hostile Tactics Affect Impressions of a Simulated Jury?*, 7 BEHAV. SCI. & L. 275, 280 (1989) (finding a negative effect of hostile and leading cross examination tactics); Peter W. Hahn & Susan D. Clayton, *The Effects of Attorney Presentation Style, Attorney Gender, and Juror Gender on Juror Decisions*, 20 LAW & HUM. BEHAV. 533, 548 (1996) (finding that an aggressive presentation style was more effective than a passive style); William M. O’Barr & John M. Conley, *When a Juror Watches a Lawyer*, 3 BARRISTER 8, 11 (1976) (discussing the effects of language and presentation); Janet Sigal et al., *The Effect of Presentation Style and Sex of Lawyer on Jury Decision-Making Behavior*, 22 PSYCHOLOGY 13, 16 (1985) (finding that an aggressive presentation style was seen as more effective than a passive style). See generally Dominic A. Infante & Andrew S. Rancer, *Argumentativeness and Verbal Aggressiveness: A Review of Recent Theory and Research*, 19 ANN. INT’L COMM. ASS’N 319, 327–44 (1996) (reviewing research finding that argumentativeness is associated with higher credibility, but verbal aggression is associated with lower credibility). For recent research in a different context, see Popan et al., *supra* note 14, at 123.

<sup>49</sup> See Don A. Moore, Lloyd Tanlu & Max H. Bazerman, *Conflict of Interest and the Intrusion of Bias*, 5 JUDGMENT & DECISION MAKING 37, 43 (2010) (describing the difficulty of simultaneously enacting multiple roles).

<sup>50</sup> Susan P. Sturm, *Lawyering Paradoxes: Making Meaning of the Contradictions* 7 (Columbia Pub. L. Research, Working Paper No. 14-642, 2019).

<sup>51</sup> *Id.*

When engaging in negotiation—whether the negotiation of a contract, the settlement of a lawsuit, or any other negotiation—lawyers often find themselves with multiple, and conflicting, goals. In identifying the “Negotiator’s Dilemma,” scholars have recognized that negotiations involve “two separate but complementary negotiation tasks—claiming or distributing value, often described as ‘dividing the pie,’ and potentially creating new value from the opportunities that the negotiation presents, often described as ‘enlarging the pie.’”<sup>52</sup> The best lawyers draw on skills—some cooperative, some competitive—that allow them to be successful in both of these tasks.<sup>53</sup>

All of this suggests a few possible directions for reform. First, law schools can strive to provide (1) an even better grounding in establishing and critically evaluating facts<sup>54</sup> and a deeper understanding of empirical evidence,<sup>55</sup> (2) an understanding of the habits of mind that influence policy debate,<sup>56</sup> (3) more training in a wide range of approaches to dispute resolution and the relevant toolbox of skills,<sup>57</sup> (4) facility in navigating the multiplicity of roles and making the nuanced distinctions required of

---

<sup>52</sup> RISKIN ET AL., *supra* note 39, at 149–66. See also DAVID A. LAX & JAMES K. SEBENIUS, *THE MANAGER AS NEGOTIATOR* 29–45 (1986) (“There is a central, inescapable tension between cooperative moves to create value jointly and competitive moves to gain individual advantage.”). See Keith G. Allred, *Distinguishing Best and Strategic Practices: A Framework for Managing the Dilemma Between Claiming and Creating Value*, 16 NEGOT. J. 387 (2000) (discussing best practices and strategies to manage this tension).

<sup>53</sup> Andrea Kupfer Schneider, *Shattering Negotiation Myths: Empirical Evidence on the Effectiveness of Negotiation Style*, 7 HARV. NEGOT. L. REV. 143, 147–49 (2002); Tinsley et al., *supra* note 33, at 621–22, 624, 637.

<sup>54</sup> See Beryl Blaustone & Lisa Radtke Bliss, *The Role of the Lawyer and the Essential Skills to Teach Law Students in an Era of Fake News, “Alternative Facts,” and Governing by Disruption*, 19 LOY. J. PUB. INT. L. 139, 154–55 (2018) (discussing lawyers’ role in investigating, identifying, and establishing facts; challenging assumptions; looking behind alleged facts; and thinking about permissible inferences). Also important is the ability to distinguish the less than helpful concepts of “alternative facts” or “nonexistent truth” from the more useful notion of “constructive ambiguity” in dispute resolution. Noam Ebner, *Begun, The Trust War Has: Teaching Negotiation When Truth Isn’t Truth*, 35 NEGOT. J. 207, 208 (2019).

<sup>55</sup> See ROBERT M. LAWLESS ET AL., *EMPIRICAL METHODS IN LAW* 1–2 (2d ed. 2016) (explaining the importance for lawyers of understanding research design and statistics).

<sup>56</sup> See, e.g., JOHN COOK & STEPHAN LEWANDOWSKY, *THE DEBUNKING HANDBOOK* 1 (2011) (examining how attempts to debunk myths can reinforce those myths); ROBBENOLT & STERNLIGHT, *supra* note 38, at 1 (describing psychological phenomena that influence conflict and its resolution); Jiin Jung et al., *A Multidisciplinary Understanding of Polarization*, 74 AM. PSYCHOL. 301, 307–10 (2019) (discussing a multidisciplinary approach to understanding polarization); Sami R. Yousif et al., *The Illusion of Consensus: A Failure to Distinguish Between True and False Consensus*, 30 PSYCHOL. SCI. 1195, 1195 (2019) (discussing how false consensus can influence what information we trust); *supra* notes 5–18 and accompanying text (discussing how preferences and prior beliefs can lead to biases when evaluating arguments).

<sup>57</sup> See generally RISKIN ET AL., *supra* note 39 (surveying a range of skills and approaches relevant to dispute resolution).

lawyers;<sup>58</sup> and (5) a foundation of essential skills for making good decisions and working effectively with other people, including adversaries.<sup>59</sup>

Second, at the very least, a lawyer who is trained in civil advocacy could present to her client the choice between a no-holds barred approach that might yield a somewhat better financial outcome and a more enlightened strategy that may leave the client with a bit less money in her pocket but serve other interests the client may have, leaving her feeling better about herself, the opposing party (with whom the client may have an ongoing business or personal relationship), and the legal system in general.<sup>60</sup> Even more broadly, lawyers can work to distinguish the professional and the personal, to distinguish necessary and unnecessary conflict, to distinguish instances in which litigation is necessary from those in which a consensual solution is attainable, to distinguish their roles of advocate and advisor, and to distinguish the ways in which they present arguments to different audiences.<sup>61</sup> Educating clients and opponents about these nuances can open the door to more problem-solving and less needless conflict.

Third, law schools should—in the short and long term—look to educate would-be decision makers (current and future judges and the citizenry at large) about how the adversarial system can generate the most accurate and fair results in individual cases and for society at large.<sup>62</sup> Just

---

<sup>58</sup> See Sturm, *supra* note 50, at 5 (discussing the importance of “making sense of, and being able to forge constructive tension between [the] oppositional aspects of lawyering”).

<sup>59</sup> See KISER, *supra* note 38, at 4–5, 9–11 (describing the increasing importance of intrapersonal and personal competencies to the success of lawyers and other professionals); ROBBENOLT & STERNLIGHT, *supra* note 38, at 5 (noting the most important skills for lawyers); John M. Lande & Jean R. Sternlight, *The Potential Contribution of ADR to an Integrated Curriculum: Preparing Law Students for Real World Lawyering*, 25 OHIO ST. J. ON DISP. RESOL. 247, 251 (2010) (highlighting the importance for lawyers of understanding emotion, communication skills, and creativity); Jean R. Sternlight & Jennifer K. Robbennolt, *Psychology and Effective Lawyering: Insights for Legal Educators*, 64 J. LEGAL EDUC. 365, 365–73 (2015) (discussing the connection between psychology and the core competencies of working with other people and making good decisions).

<sup>60</sup> See MODEL RULES OF PROF'L CONDUCT r. 2.1 (AM. BAR ASS'N 1983) (“In rendering advice, a lawyer may refer not only to law but to other considerations such as moral, economic, social and political factors, that may be relevant to the client’s situation.”).

<sup>61</sup> Sturm, *supra* note 50, at 7 (noting that these sorts of tensions “lie at the heart of what makes lawyers distinctive, necessary, and effective. The most successful and impactful lawyers live in these tensions. This capacity to hold paradox may be what equips lawyers to exercise truly effective leadership”); Symposium, *ADR’s Place in Navigating a Polarized Era*, *supra* note 38 (considering “whether and how we can teach our students to be discerning in making appropriate use of these approaches and skills, both in their future representation of clients and in their future roles as leaders within their local, professional, religious, and political communities”).

<sup>62</sup> The preamble to the Model Rules of Professional Conduct notes that “a lawyer should further the public’s understanding of and confidence in the rule of law and the justice system because legal institutions in a constitutional democracy depend on popular participation and support to maintain their authority.” MODEL RULES OF PROF'L CONDUCT pmb1. 6 (AM. BAR ASS'N 1983). See Justin Sevier, *A*

as with democracy, there is no guarantee that America's distinctive contribution to formal dispute resolution (an adversarial system with robust lawyer ethical obligations and information protection, combined with mechanisms to help all kinds of clients secure competent and zealous representation) will be maintained in the next century as it has over the last two. We are at an historical moment in which sharp disagreements and seeming inability to appreciate the other side's points is causing policy leaders in Washington and on the campaign trail to threaten major reform of the country's highest legal institutions, so nothing can or should be taken for granted.

At the same time, we are not suggesting a world in which lawyers are trained to facilitate consensual resolution of all controversies. Indeed, one increasingly prominent critique of the nation's dispute-resolution system is that certain kinds of cases are settled too frequently, such that institutions that represent the public, like appellate courts and legislatures, are starved of fodder to reflect on and weigh in on major policy issues that should, because of their external and symbolic effects, not be left entirely to private ordering.<sup>63</sup> But even here, disputes that are best resolved by our government leaders among the three branches can be clarified, streamlined, and facilitated—not hindered—when lawyers better understand how to present arguments in a less histrionic and more balanced and data-informed way. Our adversarial system is most worth preserving when we keep firmly in mind that adversarial is not the same thing as belligerent, and certainly not the same thing as bellicose.

---

[*Relational*] *Theory of Procedure*, 104 MINN. L. REV. 1987, 1996–98 (2020) (describing establishing truth and providing justice as key objectives of dispute resolution). For empirical work on how U.S. laypeople view adversarial and inquisitorial systems, see Justin Sevier, *The Trust-Justice Tradeoff: Perceptions of Decisional Accuracy and Procedural Justice in Adversarial and Inquisitorial Legal Systems*, 20 PSYCHOL. PUB. POL'Y & L. 212, 213 (2014).

<sup>63</sup> See, e.g., Fiss, *supra* note 47, at 1075 (describing the potential costs of settlement); Menkel-Meadow, *supra* note 47, at 2663–67 (discussing “when, how, and under what circumstances” disputes should be settled); Symposium, *Against Settlement: Twenty-Five Years Later*, *supra* note 47 (presenting a range of views on settlement).

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### Political Polarization and Moral Outrage on Social Media

JORDAN CARPENTER, WILLIAM BRADY, MOLLY CROCKETT, RENÉ  
WEBER & WALTER SINNOTT-ARMSTRONG

*Many theorists claim that social media contribute to political polarization, but it is not clear how these effects occur. We propose and explain a theoretical model of this process that focuses on moral outrage. This combination of anger and disgust can emerge from a mismatch between evolved human nature and certain features of political discussions on the internet. We identify three specific types of socially negative behavior that moral outrage facilitates: aggression (behavior intended to harm others), sophistry (poor argumentation), and withdrawal (avoiding discussions of politics). We describe psychological mechanisms through which moral outrage can lead to these outcomes, specifically focusing on dehumanization and group antagonism. We discuss research justifying our proposed model and suggest new ways to empirically test its links. Our model should be useful for researchers exploring the question of when and how political discussions on social media go wrong as well as what to do about these problems.*



## ARTICLE CONTENTS

BACKGROUND .....	1109
I. THESIS.....	1110
II. MODEL.....	1111
III. LINKS.....	1113
A. DOES MORAL CONTENT INCREASE MORAL OUTRAGE?.....	1113
B. DO SOCIAL MEDIA AMPLIFY MORAL OUTRAGE? .....	1114
C. DOES DIGITAL OUTRAGE INCREASE GROUP ANTAGONISM?.....	1116
D. DOES DIGITAL OUTRAGE LEAD TO DEHUMANIZATION? .....	1116
E. HOW CAN WE TEST THESE LINKS? .....	1117
F. DOES ONLINE MORAL OUTRAGE LEAD TO AGGRESSION? .....	1118
G. DOES ONLINE MORAL OUTRAGE INCREASE SOPHISTRY? .....	1119
H. DOES ONLINE MORAL OUTRAGE MOTIVATE WITHDRAWAL? .....	1119
IV. IMPACT .....	1120



# Political Polarization and Moral Outrage on Social Media

JORDAN CARPENTER, WILLIAM BRADY, MOLLY CROCKETT, RENÉ WEBER  
& WALTER SINNOTT-ARMSTRONG \*

## BACKGROUND

Decades ago, experts hailed the internet as a grand, new opportunity for political enlightenment. It was thought that the web would provide a convenient and widely available way to remove limitations imposed by geography and resources, expanding access to information, increasing understanding and empathy among people, and making the world better. Today, this optimistic view is tempered by fears that certain aspects of internet use—most notably social media—have the potential to exacerbate threats to democracy, including political polarization.<sup>1</sup>

Political polarization is sometimes understood merely as ideological distance between political parties or homogeneity within parties.<sup>2</sup> However, group coherence and disagreement by themselves are not the main problems here. The more threatening kind of polarization, which is often described as affective group polarization, involves intense, negative attitudes toward the political outgroup.<sup>3</sup> According to Pew Research

---

\* William Brady is an NSF postdoctoral fellow in the psychology department at Yale University. Jordan Carpenter is a postdoctoral fellow in the Kenan Institute for Ethics at Duke University. Molly Crockett is an Assistant Professor of Psychology at Yale University and a Distinguished Research Fellow at the Oxford Centre for Neuroethics, University of Oxford. Walter Sinnott-Armstrong is the Chauncey Stillman Professor of Practical Ethics at Duke University in the Philosophy Department, the Kenan Institute for Ethics, the Psychology and Neuroscience Department, and the Law School. René Weber (M.D., University of Aachen, Germany; Ph.D. University of Technology Berlin, Germany) is a professor in the Department of Communication at the University of California in Santa Barbara, director of UCSB's Media Neuroscience Lab (<https://medianeuroscience.org>), and a Fellow of the International Communication Association.

<sup>1</sup> William J. Brady & M.J. Crockett, *How Effective Is Online Outrage?*, 23 *TRENDS COGNITIVE SCI.* 79, 79–80 (2019). See CASS R. SUNSTEIN, #REPUBLIC: DIVIDED DEMOCRACY IN THE AGE OF SOCIAL MEDIA 5–6 (2017) (advocating for “an architecture of serendipity” as the ultimate way to salvage democracy, as “[t]o the extent that social media allow us to create our very own feeds, and essentially live in them, they create serious problems. . . . Self-insulation and personalization . . . spread falsehoods, and promote polarization and fragmentation”); ANAMITRA DEB ET AL., IS SOCIAL MEDIA A THREAT TO DEMOCRACY? 3–4 (2017), <https://www.omidyargroup.com/wp-content/uploads/2017/10/Social-Media-and-Democracy-October-5-2017.pdf> (reporting the six key features of social media that challenge democratic principles).

<sup>2</sup> Christopher Hare & Keith T. Poole, *The Polarization of Contemporary American Politics*, 46 *POLITY* 411, 412 (2014).

<sup>3</sup> Shanto Iyengar & Sean J. Westwood, *Fear and Loathing Across Party Lines: New Evidence on*

surveys in 2014, deep antipathy toward one's political outgroup grew by 24% in the preceding decade among the American public, and nearly 32% of Americans saw the opposing party's policies as threats to the nation or its well-being.<sup>4</sup> These strong feelings have contributed to violent clashes, such as those between far-right political groups and liberals in New Orleans and Charlottesville.<sup>5</sup> More generally, increasing affective group polarization has led to a decline in the kind of civil discourse that many hold to be a cornerstone of democracy.<sup>6</sup>

While social media use is widely believed to contribute to growing polarization,<sup>7</sup> data directly addressing this claim are scarce and in part lead to controversial interpretations and conclusions. As a result, the processes through which social media might exacerbate polarization are not well understood. We need to figure out the processes behind polarization in order to figure out what to do about it. Solutions require understanding.

## I. THESIS

We propose here that moral outrage is central to understanding how social media use is related to affective group polarization. Moral outrage is an intense negative emotion combining anger and disgust triggered by a perception that someone violated a moral norm.<sup>8</sup> Messages that describe or evoke moral outrage are increasingly prevalent in contemporary political contexts, especially those accusing political opponents of moral norm

*Group Polarization*, 59 AM. J. POL. SCI. 690, 690 (2015); Matt Motyl, *Liberals and Conservatives Are (Geographically) Dividing*, in SOCIAL PSYCHOLOGY OF POLITICAL POLARIZATION 7, 21 (Piercarlo Valdesolo & Jesse Graham eds., 2016).

<sup>4</sup> *Political Polarization in the American Public: Section 2: Growing Partisan Antipathy*, PEW RES. CTR. (June 12, 2014), <https://www.people-press.org/2014/06/12/section-2-growing-partisan-antipathy/>.

<sup>5</sup> See Nicholas Bogel-Burroughs, *What Is Antifa? Explaining the Movement to Confront the Far Right*, N.Y. TIMES (July 2, 2019), <https://www.nytimes.com/2019/07/02/us/what-is-antifa.html> (reporting on the movements of "antifa," a contraction of the word "anti-fascist," including protests in Charlottesville that turned violent); Alan Feuer & Jeremy W. Peters, *Fringe Groups Revel as Protests Turn Violent*, N.Y. TIMES (June 2, 2017), <https://www.nytimes.com/2017/06/02/us/politics/white-nationalists-alt-knights-protests-colleges.html> (describing various groups' attempts to mobilize, including that of the Proud Boys—a clan of conservative nationalists—in New Orleans over the removal of Confederate monuments).

<sup>6</sup> Jürgen Habermas, *Three Normative Models of Democracy*, 1 CONSTELLATIONS 1, 7 (1994); WALTER SINNOTT-ARMSTRONG, THINK AGAIN: HOW TO REASON AND ARGUE 2–4 (2018) [hereinafter, SINNOTT-ARMSTRONG, THINK AGAIN].

<sup>7</sup> Levi Boxell et al., *Greater Internet Use Is Not Associated with Faster Growth in Political Polarization Among US Demographic Groups*, 114 PNAS 10,612, 10,612–16 (2017); JOSHUA A. TUCKER ET AL., SOCIAL MEDIA, POLITICAL POLARIZATION, AND POLITICAL DISINFORMATION: A REVIEW OF THE SCIENTIFIC LITERATURE 3–5 (2018), <https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf>.

<sup>8</sup> Jessica M. Salerno & Liana C. Peter-Hagene, *The Interactive Effect of Anger and Disgust on Moral Outrage and Judgments*, 24 PSYCHOL. SCI. 2069, 2074 (2013).

violations.<sup>9</sup> The moral nature of such messages makes them more likely to capture audiences' attention<sup>10</sup> and intensifies receivers' emotional involvement.<sup>11</sup> The resulting moral outrage is associated with especially stubborn political views<sup>12</sup> and can even facilitate political violence.<sup>13</sup>

Recent theorizing suggests that the design of social media platforms amplifies moral outrage by lowering the social costs associated with outrage and increasing its personal benefits,<sup>14</sup> especially when moral content interacts with moral sensitivities to shape exposure to social media and subsequent behavior.<sup>15</sup> Thus, moral outrage sparked by messages on social media and the internet more broadly is likely a crucial factor in explaining recent alarming trends in societal discourse and their consequences for increasing polarization and the decay of democratic norms.

## II. MODEL

To understand affective group polarization, we propose a model describing how a mismatch between our evolutionary past and current social media amplifies moral outrage in online contexts. This, among other factors, leads to affective group polarization, involving group antagonism and dehumanization, which subsequently motivates social behaviors that directly threaten democracy.

---

<sup>9</sup> Spassena P. Koleva et al., *Tracing the Threads: How Five Moral Concerns (Especially Purity) Help Explain Culture War Attitudes*, 46 J. RES. PERSONALITY 184, 191–93 (2012).

<sup>10</sup> William J. Brady et al., *Attentional Capture Helps Explain Why Moral and Emotional Content Go Viral*, J. EXPERIMENTAL PSYCHOL. GEN. 1, 4 (2019); Ana P. Gantman & Jay J. Van Bavel, *The Moral Pop-Out Effect: Enhanced Perceptual Awareness of Morally Relevant Stimuli*, 132 COGNITION 22, 28 (2014).

<sup>11</sup> William J. Brady et al., *Emotion Shapes the Diffusion of Moralized Content in Social Networks*, 114 PNAS 7313, 7316 (2017).

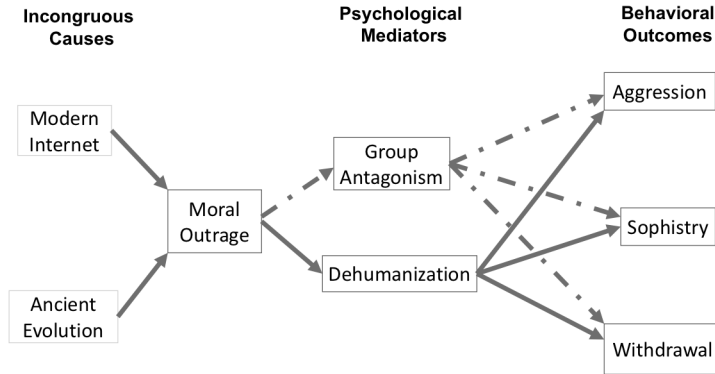
<sup>12</sup> Linda J. Skitka et al., *Moral Conviction: Another Contributor to Attitude Strength or Something More?*, 88 J. PERSONALITY & SOC. PSYCHOL. 895, 903 (2005) (testing “the degree that moral conviction was correlated” with political orientation).

<sup>13</sup> ALAN PAGE FISKE & TAGE SHAKTI RAI, *VIRTUOUS VIOLENCE: HURTING AND KILLING TO CREATE, SUSTAIN, END, AND HONOR SOCIAL RELATIONSHIPS 1–2* (2015) (discussing virtuous violence theory); Marlon Mooijman et al., *Moralization in Social Networks and the Emergence of Violence During Protests*, 2 NATURE HUM. BEHAV. 389, 389 (2018).

<sup>14</sup> M. J. Crockett, *Moral Outrage in the Digital Age*, 1 NATURE HUM. BEHAV. 769, 769–71 (2017).

<sup>15</sup> Richard Huskey et al., *Things We Know About Media and Morality*, 2 NATURE HUM. BEHAV. 315, 315 (2018).

Figure 1: Proposed model of online moral outrage.



Our model begins with some fundamental sources of online moral outrage. Human psychology developed over evolutionary time in small communities, where observing egregious acts was a rare and noteworthy event. By contrast, the modern world, particularly with the development of social media, supplies a near-constant barrage of material that evokes moral outrage when political discussions occur. Other features of online contexts that exacerbate moral outrage include the psychological distance between conversation partners and the rarity of punitive consequences for bad behavior,<sup>16</sup> as well as the predominantly written nature of online communication, which can intensify the emotional impact of messages.<sup>17</sup> It was also much harder and more dangerous to leave one's small community in evolutionary times than it is to drop out of online exchanges. This mismatch between the circumstances in which our ancestors evolved and the online worlds that many of us inhabit today plays a large role in instigating the problem of moral outrage online.

In the next stage of our model, online moral outrage leads to two psychological states that characterize affective group polarization: *group antagonism* (antipathy toward groups of political opponents)<sup>18</sup> and *dehumanization* (failure to recognize others' human mental attributes).<sup>19</sup> These psychological states then lead to three distinct social behaviors: *aggression* (behavior intended to harm another individual),<sup>20</sup> *sophistry*

<sup>16</sup> Crockett, *supra* note 14, at 769–71.

<sup>17</sup> Huskey et al., *supra* note 15, at 315; Joseph B. Walther, *Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction*, 23 COMM. RES. 3, 3–5, 7–8 (1996).

<sup>18</sup> Shanto Iyengar & Sean J. Westwood, *Fear and Loathing Across Party Lines: New Evidence on Group Polarization*, 59 AM. J. POL. SCI. 690, 690, 704 (2015).

<sup>19</sup> Lasana T. Harris & Susan T. Fiske, *Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups*, 17 PSYCHOL. SCI. 847, 847–48, 850 (2006); Nick Haslam, *Dehumanization: An Integrative Review*, 10 PERSONALITY & SOC. PSYCHOL. REV. 252, 252–53 (2006).

<sup>20</sup> ROBERT A. BARON & DEBORAH R. RICHARDSON, HUMAN AGGRESSION (PERSPECTIVES IN SOCIAL PSYCHOLOGY) 7 (2d ed. 1994).

(using empty, misleading, or irrelevant arguments),<sup>21</sup> and *withdrawal* (deliberately avoiding political participation, including voting, contribution, discussion, or even learning about political issues). These behaviors can threaten democracy by restricting communication, cooperation, civic participation, and the ability to react appropriately to political events.

### III. LINKS

In order to test each link in this model, we need to pose a variety of research questions. We cannot answer any of these questions yet, but asking them will illuminate the central claims in our model and will show why we think our model is at least plausible.

#### A. *Does moral content increase moral outrage?*

Because moral outrage is triggered when a perceiver of a message believes an important moral norm has been violated, messages (e.g., tweets) without moral information are less likely to elicit moral outrage than messages that contain information about moral wrongdoing or moral conflict. In addition, the model of intuitive morality and exemplars (MIME)<sup>22</sup> has shown that effects of social media messages are intensified when their content addresses violations or upholdings of moral norms that the audience endorses and sees as important.<sup>23</sup> Furthermore, evidence from communication diffusion models repeatedly suggests that media effects are a function of both stimulus prevalence and stimulus density over a given time interval (e.g., the number of communicators or the number of message repetitions).<sup>24</sup> Hence, the high prevalence and density of moral information and moral conflict in social media could help to explain why social media trigger such strong emotions.

---

<sup>21</sup> See SINNOTT-ARMSTRONG, *THINK AGAIN*, *supra* note 6, at 183–84 (discussing the sophistical fallacy of misleading others by jumping topics and avoiding the question asked).

<sup>22</sup> Ron Tamborini, *Model of Intuitive Morality and Exemplars*, in *MEDIA AND THE MORAL MIND* 43, 43 (Ron Tamborini ed., 2015); Ron Tamborini & René Weber, *Advancing the Model of Intuitive Morality and Exemplars*, in *THE HANDBOOK OF COMMUNICATION SCIENCE AND BIOLOGY* 456, 456 (Kory Floyd & René Weber eds., 2020).

<sup>23</sup> Tamborini, *Model of Intuitive Morality and Exemplars*, *supra* note 22, at 50–51; *see also* Graham J. Haidt et al., *Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism*, in *47 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY* 55, 82, 83 (James Olson ed., 2013) (discussing a study that showed individuals “were more likely . . . to favor those who personified virtues related to” ideals that are stereotypically valued by the side of the political spectrum with which the individual identified).

<sup>24</sup> *See* Ronald E. Rice, *Intermediality and the Diffusion of Innovations*, 43 *HUM. COMM. RES.* 531, 531 (2017) (discussing the communication diffusion perspective of innovation).

### B. *Do social media amplify moral outrage?*

Moral outrage has been singled out as especially likely to occur in online political discourse.<sup>25</sup> Although moral outrage can also occur in face-to-face interactions, several factors exacerbate its effects online. Offline, people rarely encounter egregious moral violations, but social media and other technologies allow people to become aware of others' worst behaviors much more easily. People are highly motivated to express outrage about immoral actions, which makes such information especially likely to go viral.<sup>26</sup> In addition, expressing outrage is easy online, because the target of the outrage need not be present, the potential for retaliation is minimal, and distant targets inspire less empathic concern.<sup>27</sup> It is also easy to express outrage and leave the website without waiting for any response and, hence, without knowledge of how harmful one's expression might have been.

Data from previous studies using experience sampling<sup>28</sup> suggest that people experience more intense outrage in response to immoral events that they encounter online compared to events that they encounter in person or via traditional media (e.g., TV, radio, newspaper).<sup>29</sup> Spending time on social media would therefore seem to increase the likelihood of experiencing strong moral outrage.

These effects might be moderated by other factors, including age and political ideology. Age is of particular interest in light of recent evidence that older Americans have shown the greatest increases in group antagonism in recent years, despite using the internet and social media the least;<sup>30</sup> however, other evidence suggests that older adults who do use social media are the most polarized.<sup>31</sup> Certain demographic groups, such as older people, may be more vulnerable to the exacerbation of online moral outrage than others, which may help to explain demographic asymmetries in polarization. Just as not everyone exposed to a virus will fall ill, not everyone exposed to partisan content online will be influenced in the same

---

<sup>25</sup> See Crockett, *supra* note 14, at 769 (explaining how the internet facilitates the spread of moral outrage).

<sup>26</sup> Brady et al., *supra* note 11, at 7316.

<sup>27</sup> Crockett, *supra* note 14, at 770.

<sup>28</sup> MIHALY CSIKSZENTMIHALYI, *FLOW AND THE FOUNDATIONS OF POSITIVE PSYCHOLOGY* 21 (2014) ("The Experience Sampling Method (ESM) is a research procedure for studying what people do, feel, and think during their daily lives."). For examples of studies using experience sampling, see Crockett, *supra* note 14, at 770 and Wilhelm Hofmann et al., *Morality in Everyday Life*, 345 *SCIENCE* 1340, 1340–41 (2014).

<sup>29</sup> Crockett, *supra* note 14, at 770.

<sup>30</sup> Boxell et al., *supra* note 7, at 10,612.

<sup>31</sup> *Id.*; see also *National Politics on Twitter: Small Share of U.S. Adults Produce Majority of Tweets*, PEW RES. CTR. (Oct. 23, 2019), <https://www.people-press.org/2019/10/23/national-politics-on-twitter-small-share-of-u-s-adults-produce-majority-of-tweets/> (concluding that older Americans are tweeting the most about national politics).

way. Because older adults show changes in brain systems related to processing social feedback,<sup>32</sup> older adults could be differentially susceptible to online amplification of moral outrage.

The specific mechanisms by which internet usage combines with evolved characteristics can be further specified. For instance, it is well documented in the literature that repeated exposure to media, including social media, influences emotions and behaviors by altering the *salience* of moral and political content.<sup>33</sup> In a 2010 study by Leidner, Castano, Zaiser, and Giner-Sorolla, content that emphasized in-group glorification reduced the demands for justice when a violent perpetrator was an in-group member.<sup>34</sup> Moreover, this effect was mediated by moral disengagement (de-emphasizing suffering by victims' families and dehumanizing victims), which in turn is linked to violence and terrorism.<sup>35</sup>

The MIME mentioned above suggests that, over time, exposure to a consistent communication diet emphasizing the superiority of one moral intuition over another will either increase the salience of the emphasized intuitions or maintain their salience in the face of opposing influences. According to the MIME, polarization is expected in relatively closed systems, where outside influence is limited or blocked (such as in fundamentalist religious or political groups); whereas self-regulation is more likely in relatively open systems where external factors exert opposing forces (as in social media networks with fast and inexpensive information). The MIME holds that more isolated communicative networks with insulation from value-inconsistent messages should foster polarized values within such groups, intensify responses to moral conflicts between groups, and reduce openness to divergent views. Several studies have found these predicted effects in media content produced for and consumed by sub-groups that differ by age, political interest and orientation, moral intuition salience, culture, location, and dosage of exposure.<sup>36</sup>

---

<sup>32</sup> See Lars Bäckman et al., *The Correlative Triad Among Aging, Dopamine, and Cognition: Current Status and Future Prospects*, 30 *NEUROSCIENCE & BIOBEHAVIORAL REV.* 791, 797 (2006); Jean-Claude Dreher et al., *Age-Related Changes in Midbrain Dopaminergic Regulation of the Human Reward System*, 105 *PNAS* 15,106, 15,109 (2008); Ben Eppinger et al., *Reduced Striatal Responses to Reward Prediction Errors in Older Compared with Younger Adults*, 33 *J. NEUROSCIENCE* 9905, 9908 (2013); Shu-Chen Li et al., *Dopaminergic Modulation of Cognition Across the Life Span*, 34 *NEUROSCIENCE & BIOBEHAVIORAL REV.* 625, 628 (2010).

<sup>33</sup> William J. Brady, Killian McLoughlin & Molly J. Crockett, *Theory-Driven Measurement of Emotion (Expressions) in Social Media Text*, in *THE ATLAS OF LANGUAGE ANALYSIS IN PSYCHOLOGY* (Morteza Dehghani & Ryan Boyd eds.) (forthcoming) (manuscript at 17).

<sup>34</sup> Bernhard Leidner et al., *Ingroup Glorification, Moral Disengagement, and Justice in the Context of Collective Violence*, 36 *PERSONALITY & SOC. PSYCHOL. BULL.* 1115, 1116 (2010).

<sup>35</sup> E.g., Alfred L. McAlister et al., *Mechanisms of Moral Disengagement in Support of Military Force: The Impact of September 11*, 25 *J. SOC. & CLINICAL PSYCHOL.* 141, 162–63 (2006).

<sup>36</sup> For an overview, see Tamborini & Weber, *supra* note 22, at 457–58.



### C. *Does digital outrage increase group antagonism?*

Antagonism is not mere partisan disagreement. It involves hatred of political opponents in contrast to civil, substantive disputes about values and policies.<sup>37</sup> Our model concerns antagonism rather than civil disagreement.

Antagonism is related to moral outrage in that they both involve intensely negative emotions. However, whereas moral outrage is usually a response to an *individual* person or behavior, political antagonism is often directed against groups. Such political group antagonism is characterized by feelings of hostility towards the other political party and by beliefs that the other party is dangerous or evil. What began as a negative feeling towards an individual person or act grows into antagonism to their entire group.

### D. *Does digital outrage lead to dehumanization?*

Dehumanization is a process of denying a person abilities and tendencies that are typical of human mental life.<sup>38</sup> It is distinct from antagonism in that it is possible to hate someone without dehumanizing them and vice versa.<sup>39</sup> However, we hypothesize that antagonism and dehumanization can feed one another and co-occur in the context of contentious political discourse online.

Dehumanization takes two distinct forms: a target can be denied *agency* (the ability to make reasonable decisions) or *feeling* (the ability to suffer).<sup>40</sup> People see the other side as “less than human”<sup>41</sup> either in their ability to reason or in their ability to feel pain. Both kinds of dehumanization can be a consequence of moral outrage, largely because of its emotional element of disgust, which is associated with

<sup>37</sup> Shanto Iyengar et al., *Affect, Not Ideology: A Social Identity Perspective on Polarization*, 76 PUB. OPINION Q. 405, 405, 408, 421 (2012).

<sup>38</sup> Nick Haslam, *Dehumanization: An Integrative Review*, 10 PERSONALITY & SOC. PSYCHOL. REV. 252, 252, 254 (2006); see also Lasana T. Harris & Susan T. Fiske, *Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups*, 17 PSYCHOL. SCI. 847, 847 (2006) (“[W]e present new social neuroscience data indicating that extreme forms of prejudice may deny their targets even full humanity.”).

<sup>39</sup> Taze S. Rai et al., *Dehumanization Increases Instrumental Violence, but Not Moral Violence*, 114 PNAS 8511, 8514–15 (2017).

<sup>40</sup> See Mengyao Li, Bernhard Leidner & Emanuele Castano, *Toward a Comprehensive Taxonomy of Dehumanization: Integrating Two Senses of Humanness, Mind Perception Theory, and Stereotype Content Model*, 21 TPM 285, 287 (2014) (defining “agency” as “the capacity for planning and acting” and defining “experience” as “the capacity for desires and feelings”). For further discussion of experience and agency, see Heather M. Gray et al., *Dimensions of Mind Perception*, 315 SCIENCE 619, 619 (2007).

<sup>41</sup> Madeleine Dalsklev & Jonas Rønningsdalen Kunst, *The Effect of Disgust-Eliciting Media Portrayals on Outgroup Dehumanization and Support of Deportation in a Norwegian Sample*, 47 INT’L J. INTERCULTURAL REL. 28, 29 (2015).

dehumanization.<sup>42</sup> Studies have also found that communication by text as opposed to voice leads to greater dehumanization,<sup>43</sup> so the fact that most online communication takes the form of writing might increase its contribution to dehumanization.

On the other hand, recent evidence suggests that people do not dehumanize the victims of violence when those victims are perceived as immoral, which is likely to be the case in the context of political conflict.<sup>44</sup> Perceiving someone as immoral in fact usually requires perceiving them as having nefarious or malicious intentions, which are human mental states. Merely humanizing opponents by ascribing some mental states to them is then not enough to forestall antagonism and aggression towards them.<sup>45</sup> Beliefs that they have bad intentions can instead make their suffering seem less aversive and increase antagonism and aggression towards them.<sup>46</sup> In this way, inaccurate perceptions of others' mental states can sometimes be just as pernicious as dehumanization.

The sources of group antagonism and dehumanization need to be determined in order to design remedies. Many proposed interventions on affective group polarization (such as those designed to increase empathy for political opponents) are predicated on the assumption that affective polarization leads people to spontaneously generate limited, simplistic theories about their opponents' motivations or emotions. These interventions are unlikely to succeed if their assumptions are inaccurate.<sup>47</sup>

#### E. *How can we test these links?*

To verify or falsify these assumptions, we need to measure relationships among outrage, group antagonism, and dehumanization among social media users. This task can now be approached with tools that have become available only recently, such as natural language processing<sup>48</sup>

---

<sup>42</sup> *Id.* at 29, 37–38; Katrina M. Fincher & Philip E. Tetlock, *Perceptual Dehumanization of Faces Is Activated by Norm Violations and Facilitates Norm Enforcement*, 145 J. EXPERIMENTAL PSYCHOL. 131, 132 (2016); Harris & Fiske, *supra* note 38, at 852.

<sup>43</sup> Juliana Schroeder et al., *The Humanizing Voice: Speech Reveals, and Text Conceals, a More Thoughtful Mind in the Midst of Disagreement*, 28 PSYCHOL. SCI. 1745, 1746, 1760 (2017).

<sup>44</sup> Rai et al., *supra* note 39, at 8513–14.

<sup>45</sup> *Id.* at 8512.

<sup>46</sup> *Id.* at 8511–12.

<sup>47</sup> See Scott Barry Kaufman, *Can Empathic Concern Actually Increase Political Polarization?*, SCI. AM. (Nov. 6, 2019), <https://blogs.scientificamerican.com/beautiful-minds/can-empathic-concern-actually-increase-political-polarization/> (discussing how biases are likely to increase hostility toward the “outgroup”).

<sup>48</sup> Frederic R. Hopp et al., *The Extended Moral Foundations Dictionary (eMFD): Development and Applications of a Crowd-Sourced Approach to Extracting Moral Intuitions from Text*, BEHAV. RES. METHODS 2 (2020), <https://doi.org/10.3758/s13428-020-01433-0>; CHRISTOPHER D. MANNING & HINRICH SCHÜTZE, FOUNDATIONS OF STATISTICAL NATURAL LANGUAGE PROCESSING 4 (1999); Eyal Sagi & Morteza Dehghani, *Measuring Moral Rhetoric in Text*, 32 SOC. SCI. COMPUTER REV. 132, 142

and supervised learning classification.<sup>49</sup> We predict that these tools can be used to uncover a positive relationship between expressions of moral outrage online and language that expresses antagonism towards groups and that dehumanizes opponents, such as by referring to them as animals.<sup>50</sup> We also predict that social reinforcement of expressions of moral outrage (in the form of “likes” and “retweets”) will increase subsequent use of antagonistic and dehumanizing language in online discourse so that participants who receive the greatest amount of positive social feedback when they express moral outrage in their social media posts will show the highest levels of antagonism and dehumanization. Evidence for these predictions would support the corresponding links in our model between online moral outrage and the two psychological mediators: group antagonism and dehumanization.

#### F. *Does online moral outrage lead to aggression?*

Our model’s next set of research questions asks whether moral outrage, through the mediators of both group antagonism and dehumanization, will lead to certain behaviors. Our model focuses on three actions: aggression, sophistry, and withdrawal.

To understand online *aggression*, recall that moral outrage begins as a negative emotional reaction to a single individual’s act,<sup>51</sup> whereas antagonism is directed towards a group.<sup>52</sup> The transition from moral outrage to antagonism thus involves the spreading of negative feeling from one person to their entire group.

Anger at outgroups is associated with prejudice<sup>53</sup> and has been shown to be related specifically to disliking political outgroups more and tolerating them less.<sup>54</sup> Therefore, higher levels of moral outrage tend to lead to higher levels of prejudice and intolerance towards groups

---

(2013); René Weber et al., *Extracting Latent Moral Information from Text Narratives: Relevance, Challenges, and Solutions*, 12 COMM. METHODS & MEASURES 119, 124, 137 (2018).

<sup>49</sup> Brady, McLoughlin & Crockett, *supra* note 33 (manuscript at 6–9).

<sup>50</sup> See, e.g., Florian Arendt & Narin Karadas, *Content Analysis of Mediated Associations: An Automated Text-Analytic Approach*, 11 COMM. METHODS & MEASURES 105, 112 (2017) (analyzing the use of animal-related terms over a four-month period to demonstrate the dehumanization of Muslims in German news coverage of Islam).

<sup>51</sup> See Salerno & Peter-Hagene, *supra* note 8, at 2069 (closely linking moral outrage with anger).

<sup>52</sup> See Giulia Evolvi, *#Islamexit: Inter-Group Antagonism on Twitter*, 22 INFO. COMM. & SOC’Y 386, 397 (2019) (studying group antagonism through anti-Muslim tweets during the Brexit debate in the United Kingdom).

<sup>53</sup> Nilanjana Dasgupta et al., *Fanning the Flames of Prejudice: The Influence of Specific Incidental Emotions on Implicit Prejudice*, 9 EMOTION 585, 589 (2009).

<sup>54</sup> Linda J. Skitka et al., *Political Tolerance and Coming to Psychological Closure Following the September 11, 2001, Terrorist Attacks: An Integrative Approach*, 30 PERSONALITY & SOC. PSYCHOL. BULL. 743, 754 (2004).

associated with the particular source of the moral outrage and then to representatives of those groups.

As a result, it seems likely that increased group antagonism will make people more willing to act aggressively to individual members of opposing groups based on group membership. Similarly, previous research has found that dehumanization is strongly associated with aggression, such that when a person perceives opponents as lacking feeling, they become more willing to inflict harm against opponents through bullying or harassment.<sup>55</sup> Therefore, dehumanization, particularly a lack of concern for the feelings of the target, would also seem to lead to aggression, at least in some cases.

### G. Does online moral outrage increase sophistry?

Moral outrage also seems to lead people to engage in *sophistry*, or bad arguments, partly because one component of outrage is anger, which impairs judgment and decision making.<sup>56</sup> Ideally, the purpose of presenting arguments is to increase understanding of opposing points of view (including why others hold those positions) as well as to influence beliefs and attitudes on both sides of a controversy.<sup>57</sup> However, for many people talking about politics in social media, the focus is instead on competition and provocation (beating opponents by embarrassing, exhausting, or bewildering them) or theater (appearing more intelligent to observers who are allies or potential allies).<sup>58</sup> Even when people intend to create good arguments against opposing positions (or for their own), they often miss their targets because of a simplistic understanding of their opponents. This tendency seems to be exacerbated by higher levels of antagonism and dehumanization, which leads people to adopt a competitive or theatrical mindset during political discussions online, resulting in sophistry.

### H. Does online moral outrage motivate withdrawal?

For the same reasons that moral outrage is galvanizing for some people, it leads others to *withdraw* from politics.<sup>59</sup> Intense animus can be overwhelming and unpleasant and will motivate many people to withdraw

---

<sup>55</sup> Albert Bandura et al., *Disinhibition of Aggression Through Diffusion of Responsibility and Dehumanization of Victims*, 9 J. RES. PERSONALITY 253, 266 (1975); Brock Bastian et al., *The Roles of Dehumanization and Moral Outrage in Retributive Justice*, 8 PLOS ONE 1, 9 (2013).

<sup>56</sup> Jennifer S. Lerner & Larissa Z. Tiedens, *Portrait of the Angry Decision Maker: How Appraisal Tendencies Shape Anger's Influence on Cognition*, 19 J. BEHAV. DECISION MAKING 115, 132 (2006).

<sup>57</sup> SINNOTT-ARMSTRONG, THINK AGAIN, *supra* note 6, at 56.

<sup>58</sup> See, e.g., Ashley A. Anderson & Heidi E. Huntington, *Social Media, Science, and Attack Discourse: How Twitter Discussions of Climate Change Use Sarcasm and Incivility*, 39 SCI. COMM. 598, 600 (2017) (analyzing the use of sarcastic or uncivilized rhetoric in online discourse surrounding climate change).

<sup>59</sup> Elizabeth A. Bennett et al., *Disavowing Politics: Civic Engagement in an Era of Political Skepticism*, 119 AM. J. SOC. 518–19 (2013).

in order to avoid or reduce associated negative emotions.<sup>60</sup> It is not clear what leads one person to become aggressive and another person to withdraw, but various individual differences might moderate these effects, such as the degree to which people experience moral outrage as unpleasant. In any case, antagonism and dehumanization by politically active people seem to lead some people into cynicism or apathy about politics. Just as many people effortfully avoid feeling sympathy for widespread suffering out of a desire to avoid emotional exhaustion,<sup>61</sup> so many people are motivated to avoid engaging with politics in order to keep from experiencing the hostility that characterizes contemporary partisan politics. They see politics as unpleasant, difficult, and exhausting; they foresee few compensating benefits for engaging in political activity, especially because of the sophistry and vicious attacks that characterize so much of political discourse online. For such reasons, both antagonism and dehumanization seem to lead many people to withdraw from politics—and understandably so.

#### IV. IMPACT

Many claims in our model remain speculations in need of further empirical support, but it could prove important and useful. If even approximately correct, our proposed model and its further specifications could illuminate the sources of many unpleasant psychological states and politically harmful behaviors on social media and elsewhere. It could help us understand an important social problem by providing a greater sense of the emotional and cognitive factors that lead people to behave badly when engaging in politics online. Because we need to understand a problem before we can solve it, our model could also potentially guide interventions that reduce political polarization and ensuing social problems.

All of this remains to be seen, because our model so far is only that: a hypothesized model—an educated guess. We would never claim to have established it as accurate. Much more research needs to be done to test it. All we can claim for now is that we find it plausible, promising, and potentially useful. We hope that others do, too.

---

<sup>60</sup> *Id.*

<sup>61</sup> C. Daryl Cameron & B. Keith Payne, *Escaping Affect: How Motivated Emotion Regulation Creates Insensitivity to Mass Suffering*, 100 J. PERSONALITY & SOC. PSYCHOL. 1, 2–3 (2011).

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### How People Judge Institutional Corruption

ELINOR AMIT, EUGY HAN, ANN-CHRISTIN POSTEN  
& STEVEN SLOMAN

*Institutional corruption refers to actions that are legal yet carry negative consequences for the greater good. Such legal yet harmful behaviors have been observed among politicians and donors who establish quid-pro-quo relationships in exchange for money and among doctors who receive gifts from pharmaceutical companies in return for recommending the companies' drugs. How does the general public reconcile the tension between the legal status of an action and its impact on the greater good and judge the action's moral acceptability? We explored this question empirically by comparing the relative weight people give to the legal status of actions and to the impact of actions when judging moral acceptability. Results show that people unequivocally rely on legal status and ignore the impact of the actions. We conclude that people outsource their moral judgments to the law. The law does not simply reflect people's sense of corruption but determines it. Together, our research suggests a surprising and ironic role for the law: that it diminishes independent, critical thinking.*

## ARTICLE CONTENTS

I. EMPIRICAL EVIDENCE .....	1127
A.    EXPERIMENT 1A .....	1128
B.    EXPERIMENT 1B .....	1131
C.    EXPERIMENT 2 .....	1133
II. GENERAL DISCUSSION: THE LAW AS A SHORTCUT FOR MORAL JUDGMENT .....	1135
III. IMPLICATIONS .....	1137



# How People Judge Institutional Corruption

ELINOR AMIT <sup>\*</sup>, EUGY HAN <sup>\*\*</sup>, ANN-CHRISTIN POSTEN <sup>\*\*\*</sup>  
& STEVEN SLOMAN <sup>\*\*\*\*</sup>

Former CEO of Turing Pharmaceuticals, Martin Shkreli, gained a high profile in 2015 when he purchased the rights to a lifesaving HIV drug, Daraprim, and hiked up its price by more than 5000%.<sup>1</sup> While such acts of price gouging can often be observed in business, Shkreli's act created controversy because Daraprim is on the World Health Organization's List of Essential Medicines.<sup>2</sup> Though Shkreli was heavily criticized, he justified his actions with the response: "[E]verything we've done is legal."<sup>3</sup> Shkreli was right. His actions, potentially ruining the lives of a large number of people, were legal. And he was not alone. Legal actions that result in terrible consequences are, in fact, quite common, from pharmaceutical companies partnering with doctors to prescribe their more expensive drugs,<sup>4</sup> to politicians who establish quid-pro-quo relationships with donors who anonymously give them unlimited amounts of money,<sup>5</sup> to the

---

<sup>\*</sup> Coller School of Management, Tel Aviv University, Israel.

<sup>\*\*</sup> Brown University.

<sup>\*\*\*</sup> University of Cologne; University of Limerick.

<sup>\*\*\*\*</sup> Brown University.

<sup>1</sup> Robert Mclean, *Turing Cuts Hospital Price for Drug It Hiked 5,000%*, CNNMONEY (Nov. 25, 2015, 8:17 AM), <https://money.cnn.com/2015/11/25/news/companies/turing-pharmaceuticals-daraprim-price-drop/>.

<sup>2</sup> *Id.*

<sup>3</sup> See ROBERT B. REICH, *THE COMMON GOOD* 10 (2018) (describing Shkreli's response to criticism of the Daraprim price hike); Richard Mark Kirkner, *Drug Pricing Regulation Pushed from Many Sides*, MANAGED CARE, Jan. 2016, at 14, 15 (noting criticism of Shkreli and Turing by 2016 presidential candidates and Merck CEO).

<sup>4</sup> See Lisa Cosgrove & Robert Whitaker, *Finding Solutions to Institutional Corruption: Lessons from Cognitive Dissonance Theory* 16–19 (Edmond J. Safra Ctr. for Ethics, Working Paper No. 9, 2013), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2261375](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2261375) (explaining how relationships between doctors and pharmaceutical companies create conflicts of interest and worsen quality of care); BEN GOLDACRE, *BAD PHARMA: HOW DRUG COMPANIES MISLEAD DOCTORS AND HARM PATIENTS* 316 (2012) (describing how senior doctors were paid by pharmaceutical companies to promote their drugs, changing the prescribing behavior of other doctors).

<sup>5</sup> See Elinor Amit et al., *Institutional Corruption Revisited: Exploring Open Questions Within the Institutional Corruption Literature*, 26 S. CAL. INTERDISC. L.J. 447, 453 (2017) (explaining how the current United States campaign finance system results in the "donor class" exerting a disproportionate political influence); Girish J. Gulati, *Super PACs and Financing the 2012 Presidential Election*, 49 SOCIETY 409, 409 (2012) (explaining that the relationship between candidates and donors threatens democratic governance and creates opportunities for corruption); Jonathan Mendilow & Michael Brogan, *Perceptions of Corruption and Trust in Government in the United States*, in *CORRUPTION AND GOVERNMENTAL LEGITIMACY: A TWENTY-FIRST CENTURY PERSPECTIVE* 59, 62 (Jonathan Mendilow



President of the United States who does not shy away from conflicts of interest by maintaining control over his private business while in office.<sup>6</sup> All of these examples have one thing in common: the action is legal yet carries negative consequences for the greater good.

Tension between the legal status of an action and its impact on the greater good poses a particular type of challenge when judging its moral acceptability. In a democratic society, laws are expected to represent the interests of its people, protect the greater good from individual misconducts, or at the very least not harm the society. While this is usually true, lawmakers cannot foresee each and every incident in which the law will be relevant. As a result, some legal behaviors may end up harming the greater good. Such legal yet harmful behaviors have been observed in research, design, manufacturing, and marketing processes of new products and policies released by various institutions.<sup>7</sup> For instance, gaming companies use legal yet socially irresponsible marketing strategies and tactics to target vulnerable populations and encourage gambling.<sup>8</sup>

In the farming industry, hog, poultry, and cattle farms often use legal antibiotics as a feed additive to stimulate artificial growth in their livestock.<sup>9</sup> Through overuse, these antibiotics eventually become ineffective for bacterial infections, which people often contract through meat consumption and drinking infected water.<sup>10</sup> As more and more antibiotics are becoming ineffective, new, usually more expensive, drugs have to be developed, leaving many individuals with no viable methods of treatment.<sup>11</sup> This is one of many deadly but legal<sup>12</sup> practices exercised by the farming industry. Big factory farming industries often lobby for

& Ilan Peleg eds., 2016) (noting that a quid pro quo between donors and politicians is the “most flagrant form of political corruption”).

<sup>6</sup> Steve Reilly et al., *Did Trump Keep His 19 Promises to Insulate Himself from His Business? Only He Knows.*, USA TODAY, <https://www.usatoday.com/in-depth/news/politics/2019/03/18/president-donald-trumps-promises-didnt-end-business-entanglements/3030377002/> (last updated Mar. 18, 2019, 9:08 PM).

<sup>7</sup> See Nikos Passas, *Lawful but Awful: ‘Legal corporate crimes’*, 34 J. SOCIO-ECON. 771, 777 (2005) (noting that the activities of several industries generate “hidden costs” that are mainly borne by the least privileged); John Warren Kindt, *The Costs of Legalized Gambling: An Economic Approach*, in IT’S LEGAL BUT IT AIN’T RIGHT: HARMFUL SOCIAL CONSEQUENCES OF LEGAL INDUSTRIES 115, 124 (Nikos Passas & Neva Goodwin eds., 2004) (explaining how the gambling industry targets lower income individuals and encourages excessive gambling by “consistent gamblers”).

<sup>8</sup> June Buchanan et al., *The Marketing of Legal but Potentially Harmful Products and Corporate Social Responsibility: The Gaming Industry View*, 4 INT’L J. INTERDISC. SOC. SCI. 81, 84 (2009).

<sup>9</sup> Mallory Russo, *Food for Thought: Analyzing the Impacts of Livestock Factory Farming in the United States* 40 (May 15, 2017) (unpublished B.A. thesis, Fordham University) (on file with Fordham University Libraries); Mark Ritchie, *The High Price of Cheap Food*, in IT’S LEGAL BUT IT AIN’T RIGHT: HARMFUL SOCIAL CONSEQUENCES OF LEGAL INDUSTRIES 178, 180 (Nikos Passas & Neva Goodwin eds., 2010).

<sup>10</sup> Ritchie, *supra* note 9, at 185.

<sup>11</sup> *Id.*

<sup>12</sup> *Id.*

government regulations and laws that can profit their operations, allowing them to sustain inadequate facilities and exploit environmental resources.<sup>13</sup> These actions lead to irreversible consequences such as extinction of species, deadly working conditions for workers (e.g., farm workers are exposed to harmful chemicals and may die from asphyxiation), land and soil degradation, water contamination, and exacerbation or climate change.<sup>14</sup> While such practices adhere to legal requirements, they bring harm of various degrees to society.

How do people resolve this conflict and judge the moral acceptability of legal yet harmful actions? This question is important, because people's moral judgments determine a wide variety of behaviors, from purchasing products,<sup>15</sup> to protesting against companies,<sup>16</sup> to voting for or against politicians in public elections.<sup>17</sup>

In recent years there has been a growing interest in behaviors which are legal yet harm the greater good, collectively labeled as "institutional corruption."<sup>18</sup> According to the classic definition by Lessig (2013):

*Institutional corruption* is manifest when there is a systemic and strategic influence which is legal, or even currently ethical, that undermines the institution's effectiveness by diverting it from its purpose or weakening its ability to achieve its purpose, including, to the extent relevant to its purpose, weakening either the public's trust in that institution or the institution's inherent trustworthiness.<sup>19</sup>

A surge of research on institutional corruption in the past decade has examined its manifestations in various domains, including: the pharmaceutical industry,<sup>20</sup> psychiatry,<sup>21</sup> food production and distribution

<sup>13</sup> Russo, *supra* note 9, at 35.

<sup>14</sup> *Id.* at 16, 43.

<sup>15</sup> See Johannes Brinkmann, *Looking at Consumer Behavior in a Moral Perspective*, 51 J. BUS. ETHICS 129, 129 (2004) (claiming that "business ethics and consumer behavior could profit from further development of their overlap"); Oliver M. Freestone & Peter J. McGoldrick, *Motivations of the Ethical Consumer*, 79 J. BUS. ETHICS 445, 445 (2008) (discussing the rise of consumer action "in the form of boycott activity, pressure groups and other forms of consumer activism").

<sup>16</sup> Norman E. Bowie & Thomas W. Dunfee, *Confronting Morality in Markets*, 38 J. BUS. ETHICS 381, 385, 389 (2002).

<sup>17</sup> See JASON BRENNAN, *THE ETHICS OF VOTING* 1–2 (2011) (arguing the moral significance and ethical implications of voting).

<sup>18</sup> See Amit et al., *supra* note 5, at 448 (describing the increased scholarly attention to institutional corruption). Lessig provides a definition for institutional corruption in Lawrence Lessig, *Foreword: "Institutional Corruption" Defined*, 41 J.L. MED. & ETHICS 2, 2 (2013).

<sup>19</sup> Lessig, *supra* note 18, at 553 (emphasis added).

<sup>20</sup> See sources cited *supra* note 4 (illustrating the influence pharmaceutical companies have over doctors); see also Garry C. Gray, *The Ethics of Pharmaceutical Research Funding: A Social Organization Approach*, 41 J.L. MED. & ETHICS 629, 629 (2013) (illustrating the subtle ways pharmaceutical industry funding influences medical research); Donald W. Light, *From Institutional*

companies,<sup>22</sup> and political campaigns.<sup>23</sup> For example, research on the pharmaceutical industry has revealed that funding from drug companies influences medical researchers to produce favorable reviews of the company's products, thus enabling a significant number of harmful drugs to reach the market.<sup>24</sup> This undermines pharmaceutical companies' mission of improving patient care and public health and promoting drug safety.<sup>25</sup> Similarly, in the field of psychiatry, financial conflicts of interest have compromised biomedical research, teaching, and practice.<sup>26</sup> Organizations such as the American Psychiatric Association (APA) receive substantial funding from the drug industry, and many of the individuals who serve as diagnostic panel members have ties with the drug industry.<sup>27</sup> "Industry financial relationships can . . . affect researchers' and clinicians' behavior in subtle ways" potentially influencing "decisions about the criteria for and measurement of diagnoses."<sup>28</sup>

Institutional corruption has also been observed in political campaigns.<sup>29</sup> For instance, the creation of super political action committees (super PACs) in 2010 sparked significant controversy, as committees were allowed to receive and spend unlimited sums of money on independent

---

*Corruption to Pharmageddon?*, 1 LAB DISPATCHES 69, 69–70 (2013) (reviewing David Healy's observation that the FDA and pharmaceutical companies have major influence over the medical industry); Marc A. Rodwin, *Conflicts of Interest, Institutional Corruption, and Pharma: An Agenda for Reform*, 40 J.L. MED. & ETHICS 511, 511 (2012) ("Physician relations with pharmaceutical firms are a source of conflicts of interest that can bias their prescriptions and advice.").

<sup>21</sup> See ROBERT WHITAKER & LISA COSGROVE, *PSYCHIATRY UNDER THE INFLUENCE: INSTITUTIONAL CORRUPTION, SOCIAL INJURY, AND PRESCRIPTIONS FOR REFORM* 4 (2015) (examining the societal impacts of the psychiatry field); Lisa Cosgrove & Emily E. Wheeler, *Industry's Colonization of Psychiatry: Ethical and Practical Implications of Financial Conflicts of Interest in the DSM-5*, 23 FEMINISM & PSYCHOL. 93, 93 (2013) (arguing that modern psychiatry has been captured by the pharmaceutical industry); Marc-André Gagnon, *Corruption of Pharmaceutical Markets: Addressing the Misalignment of Financial Incentives and Public Health*, 41 J.L. MED. & ETHICS 571, 574 (2013) (highlighting how new anti-psychotic drugs dominate the market despite being more expensive and less effective than older drugs).

<sup>22</sup> See Amit et al., *supra* note 5, at 450 (discussing the case of Del Monte Foods); see also Sylvia Rowe et al., *Funding Food Science and Nutrition Research: Financial Conflicts and Scientific Integrity*, 67 NUTRITION REVIEWS 264, 264–65 (2009) (describing issues of conflict and scientific bias in the food industry); Wendy Wagner & David Michaels, *Equal Treatment for Regulatory Science: Extending the Controls Governing the Quality of Public Research to Private Research*, 30 AM. J.L. & MED. 119, 142 (2004) (discussing the Food and Drug Administration's conflict policy).

<sup>23</sup> See sources cited *supra* note 5; see also Malcolm S. Salter, *Lawful but Corrupt: Gaming and the Problem of Institutional Corruption in the Private Sector* 2, 24 (Harvard Bus. Sch., Working Paper 11-060, 2010) (describing the scope and size of private money influencing Congress).

<sup>24</sup> See sources cited *supra* note 4 (illustrating the influence pharmaceutical companies have over doctors).

<sup>25</sup> See sources cited *supra* note 18 (discussing institutional corruption).

<sup>26</sup> Cosgrove & Wheeler, *supra* note 21, at 94.

<sup>27</sup> *Id.* at 102.

<sup>28</sup> *Id.* at 97.

<sup>29</sup> Gulati, *supra* note 5, at 409.

campaigning in support of or in opposition to political candidates.<sup>30</sup> Politicians receiving unlimited anonymous donations from super PACs suggested that there could be a disproportionate influence on policy and electoral outcomes.<sup>31</sup>

Despite its importance, determining which behaviors fall under the category of institutional corruption is challenging. Even amongst experts there is disagreement over what counts as institutional corruption. In a survey we conducted, a set of ten scenarios that represent different facets of institutional corruption was presented to members of the Edmond J. Safra Center for Ethics at Harvard University—all of whom were experts on the topic of institutional corruption research. These scenarios covered topics that are relevant for the concept of institutional corruption in various settings and were selected based on the research of the fellows of the center. In this survey, the experts were asked to judge whether the scenarios represent institutional corruption. For half of the scenarios we presented, 30% of the subjects or more thought that they do not represent instances of institutional corruption. The results illustrate the challenge in clearly identifying instances of institutional corruption, even by experts.

The current paper is concerned with how the public morally judge instances of institutional corruption. We bring evidence that institutional corruption is perceived as more morally acceptable than criminal behavior, even when the portrayed action is identical except for its legal status. Moreover, although people's moral judgments are sensitive to information about the legal status of the action—what we henceforth deem the distinguishing characteristic between institutional corruption and criminal action—people are not sensitive to information about the magnitude of the harm caused by the action. Thus, cases of institutional corruption with horrible consequences are judged as more morally acceptable than criminal actions with benign consequences. Together, our research suggests a surprising and ironic role for the law: that it diminishes independent, critical thinking. While criminal actions that have mildly negative consequences can be construed as immoral, institutional corruption will be seen as moral despite having terrible consequences.

## I. EMPIRICAL EVIDENCE

The following experiments were designed to investigate how people judge institutional corruption versus criminal actions. Moral scenarios were presented, and people judged the moral acceptability of the action of a protagonist. We varied whether an action was institutional corruption or criminal by manipulating the legal status of the action: for the “institutional

---

<sup>30</sup> *Id.* at 410.

<sup>31</sup> *Id.* at 409.

corruption” group the action was presented as legal, and for the “criminal behavior” group the action was presented as illegal. All the presented actions had a negative impact on the greater good. Additionally, we varied the severity of the direct consequences of the actions. Research shows that actions with bad consequences are judged as less morally acceptable than those with less bad consequences.<sup>32</sup> We therefore wanted to rule out the possibility that people’s judgments are contingent on a particular outcome.

## A. *Experiment 1A*

### 1. *Participants and Design*

Participants were 131 MTurkers (75 females; Mage = 34.06, SD = 11.15) who participated in exchange for payment. The participants were randomly assigned to one of the four conditions of a 2 (Legality: Legal vs. Illegal) x 2 (Impact: Small vs. Large) between-subjects design. In this and all subsequent studies: (A) we set a sample size based on previous research that had used similar methods and stimuli<sup>33</sup>; (B) our only a-priori selection criteria was recruiting subjects residing in the United States; and (C) we did not exclude any subject from the analysis. Any gap between the number of subjects recruited and reported in the “Participants” section and the final analysis are due to incompleteness of the questionnaire by subjects (less than 0.5% of the recruited subjects in each experiment).

### 2. *Procedure*

Subjects were asked to read three scenarios that described actions of a protagonist. Half of the subjects were informed that this action is legal, and the other half that it is illegal. In each of these groups, half of the subjects were informed that the action is slightly harmful and the other half that it is significantly harmful. After reading each scenario, the subjects were asked whether the actions of the protagonist are morally permissible on a scale that ranged between 1 (completely impermissible) to 7 (completely permissible). The subjects were subsequently asked whether the action of the protagonist is legal on a scale that ranged from 1 (completely legal) to 7 (completely illegal) and how would they rate the impact of the action on a scale that ranged from 1 (negligible impact) to 7 (strong impact). Finally, in this and all subsequent experiments, subjects were asked several demographic questions, including gender, age, whether they live in the

---

<sup>32</sup> See, e.g., Elinor Amit & Joshua D. Greene, *You See, the Ends Don’t Justify the Means: Visual Imagery and Moral Judgment*, 23 PSYCHOL. SCI. 861, 861 (2012) (discussing how a consequentialist judgment that “favor[s] the greater good” is often deemed to be more morally salient); Jonathan Baron, *Nonconsequentialist Decisions*, 17 BEHAV. & BRAIN SCI. 1, 3 (1994).

<sup>33</sup> Elinor Amit, Jonathan Koralnik & Ann-Christin Posten, *Mental Imagery of Institutional Corruption* (unpublished manuscript) (on file with authors).

United States, what is their native language, level of education, income (range), whether they support federal regulation, how liberal versus conservative they are with regards to economic and social issues, and who they voted for in the last election.

The three scenarios are presented below. In the following scenarios, legal actions appear in the text; illegal in parentheses. Big impact appears in the text; small impact in parentheses.

i. Scenario 1 (hospital)

Suppose there is a head of a hospital network who receives money from a pharmaceutical company to give a talk. The head of the network then contemplates changing the hospital network's electronic drug prescribing system from prescribing generic drugs as the default to prescribing brand name drugs as the default. Changing the default option from generic to brand name drugs has been shown to increase the percentage of brand name drugs prescribed. While the brand name drug is effectively identical to the generic, the brand name costs patients and insurance companies more, and earns the drug company more. The hospital director decides to change the electronic drug prescribing system to make brand name drugs the default. The increase in the percentage of brand name drugs prescribed causes patients and insurance companies to pay significantly (slightly) more for medications. According to the State Medical Board, it is legal for doctors to accept payments from pharmaceutical companies of up to \$10,000 (\$4000). The pharmaceutical company paid the doctor \$7000.

ii. Scenario 2 (coal)

Under the U.S. Clean Air Act, the Environmental Protection Agency (EPA) sets limits on certain pollutants. State governments can choose to make the state laws stricter than those of the EPA, though no state can have more lenient laws than those of the EPA. Suppose there is a CEO of a coal factory who is deciding whether to increase the energy produced by the plant. In order to produce energy, the factory burns fossil fuels, which emit air pollution into the surrounding towns. Notably, the percentage of residents in the closest town that have asthma is significantly (slightly) greater than the national average. According to the state law, it is legal for the coal factory to emit up to 40,000 (20,000) tons per year. The CEO decides to increase the factory's air pollution emission to approximately 30,000 tons per year to meet his energy goals.

iii. Scenario 3 (army)

Suppose there is a retired army general who serves as a director at a defense contractor and also advises the Department of Defense (DoD). As a director of a defense contractor, his job is to advocate for the weapons his

company produces. However, as an advisor, his job is to help find and purchase the most effective weaponry for the United States Army. The American military relies on the Department of Defense to choose top quality weapons. The general advises the DoD in favor of choosing his company's products and the Department decides to follow his advice. Notably, some of the weapons sold by his company may not be the most cost-effective choices. Because of the cost of these weapons, the Army will have to significantly (slightly) decrease the budgets of other agencies. According to federal law, it is legal for DoD advisors to withhold information about conflicts of interest on purchases under \$30 (\$20) billion. The general advises the DoD to purchase his weapons from his defense contractor for \$25 billion.

### 3. *Analyses and Results*

We conducted a multivariate analysis of variance (MANOVA), with legality (legal, illegal) and impact (small, large) as factors, and moral judgment, legality judgment and impact judgment as dependent measures. The results show that the legal status of the action affected moral judgments such that actions that were legal were judged to be more morally permissible than actions that are illegal ( $M_s = 3.5$  and  $2.08$ , respectively);  $F(1, 127) = 35.38, p < .001, \eta^2 = .22$ . In contrast, the severity of the consequences of the action did not affect moral judgment ( $M_s = 2.8$  and  $2.8$ , respectively);  $F(1, 127) = 0.02, p = .877$ . There was no interaction between legality and consequences  $F(1, 127) = 0.19, p = .663$ . Legal actions were perceived as more legal than illegal actions ( $M_s = 3.2$  and  $6.02$ , respectively);  $F(1, 127) = 87.23, p < .001, \eta^2 = .4$ . Legality also affected perceived impact: legal actions were perceived to have smaller impact than illegal actions ( $M_s = 5.09$  and  $4.6$ , respectively);  $F(1, 127) = 3.94, p < .049, \eta^2 = .03$ .

In order to further explore the results, we investigated whether political orientation affects the relative weight people give to legality versus morality. Among our 131 subjects, 48 mentioned that in the 2016 presidential elections they voted for Hillary Clinton, and 33 for Donald Trump (the rest either did not disclose or gave other names). For the ease of interpretation of the effect of political orientation on moral judgments, we analyzed only subjects who endorsed Clinton or Trump. The results show that legal actions were judged as more morally acceptable than illegal actions ( $M_s = 3.43$  and  $1.89$ , respectively);  $F(1, 73) = 22.98, p < .0001, \eta^2 = .23$ . The rest of the effects were not significant. Thus, political orientation did not affect moral judgments nor interact with legality or impact.

## B. *Experiment 1B*

### 1. *Overview*

In Experiment 1B, we replicated the procedure of Experiment 1A using a different set of scenarios.

### 2. *Participants and Design*

Participants were 130 MTurkers (79 females; Mage = 36.98, SD = 11.90) who participated for a payment in the experiment. The participants were randomly assigned to one of the four conditions of a 2 (Legality: Legal vs. Illegal) x 2 (Impact: Small vs. Big) between-subjects design.

### 3. *Procedure*

The method of Experiment 1B was identical to that of 1A, with a different set of scenarios. The scenarios used in this experiment appear below. In the following scenarios legal actions appear in the text; illegal in parentheses. Big impact appears in the text; small impact in parentheses.

#### i. Scenario 1 (CEO)

Suppose there is a CEO of a large food chain who must decide whether to offer products containing ABA, a chemical compound that is often used in plastic containers and canned food. This chemical is cost-effective; using it minimizes costs and maximizes profit margins. The CEO decides to offer the products that contain ABA. Notably, ABA consumption is significantly (slightly) correlated with an increased health risk of headaches and nausea. According to the U.S. Food and Drug Administration (FDA), it is legal to sell foods containing up to 2 mg (1.5 mg) of ABA. The products the CEO decides to offer contain up to 1.75 mg of ABA.

#### ii. Scenario 2 (scientist)

Suppose a scientist who works for a pharmaceutical company runs several studies to investigate the safety and efficacy of a new medicine. The scientist obtains mixed results. Some show that the drug is safe and effective. Others show no effect of the drug. Some further studies also reveal a significant (slight) correlation between using this drug and experiencing various side effects, such as increased blood pressure. The scientist decides to re-run the studies that had the best results. These results confirm the positive findings from before. The scientist only publishes the positive results in the medical literature. The FDA approves the drug. According to the FDA, it is legal for scientists investigating the safety of prescription drugs to exclude up to 30% (20%) of their results. The scientist excluded 25% of their results.



iii. Scenario 3 (prison)

Suppose there is a director of a private prison who is trying to maximize the prison's earnings. The greatest profits can be achieved when the prison is fully booked. For the highest degree of capacity utilization, it is better to transfer prisoners from one prison to another. The prison director decides to transfer prisoners to maximize occupancy of the prison. Notably, research shows that moving prisoners away from their support system of family and friends leads to a significant (slight) increase in their rates of recidivism. According to federal law, it is legal for private prisons to transfer inmates up to 200 miles (100 miles). The director decides to send some to prisons up to 150 miles away.

4. *Analyses and Results*

As in Experiment 1A, we conducted a multivariate analysis of variance (MANOVA), with legality (legal, illegal) and impact (small, large) as factors, and moral judgment, legality judgment, and impact judgment as dependent measures. The results show that the legal status of the action affected moral judgments such that legal actions were perceived as more morally permissible than illegal actions ( $M_s = 3.48$  and  $1.76$ , respectively);  $F(1, 126) = 50.91, p < .001, \eta^2 = .29$ . In contrast, the severity of the consequences of the action did not affect moral judgment ( $M_s = 2.53$  and  $2.71$ , respectively);  $F(1, 126) = 0.60, p = .44$ . There was no interaction between legality and consequences  $F(1, 126) = 0.38, p = .54$ . Legality affected perceived legality, such that legal actions were perceived as more legal than illegal actions ( $M_s = 3.19$  and  $6.22$ , respectively);  $F(1, 126) = 93.2, p < .0001, \eta^2 = .42$ . No other effect was significant.

As in Experiment 1A, in order to further explore the results, we investigated whether political orientation affects the relative weight people give to legality versus morality. Among our 130 subjects, 58 mentioned that in the 2016 presidential elections they voted for Hillary Clinton, and 29 voted for Donald Trump (the rest either did not disclose or gave other names). As in Experiment 1A, we analyzed only subjects who endorsed Clinton or Trump. The results show that legal actions were judged as more morally acceptable than illegal actions ( $M_s = 4.02$  and  $1.79$ , respectively);  $F(1, 79) = 57.14, p < .0001, \eta^2 = .42$ . Subjects who endorsed Clinton judged the behavior as less morally acceptable than subjects who endorsed Trump ( $M_s = 2.54$  and  $3.28$ , respectively);  $F(1, 79) = 6.36, p < .014, \eta^2 = .075$ . We also found a significant interaction between voting and outcomes. For Clinton voters, when the outcome was small, the behavior was judged as more morally acceptable than when the outcome was large ( $M_s = 3.09$  and  $1.98$ , respectively,  $p < .001$ ). For Trump voters, there was no significant difference between small and large outcomes ( $M_s = 1.18$  and  $3.38$ , respectively,  $p < .68$ ). The rest of the effects were not significant.

### C. *Experiment 2*

In Experiment 2 we used the same basic design of Experiment 1, and manipulated legality (legal, illegal) and impact (big, small). The main difference was that in Experiment 2 we described impacts as distinct outcomes (e.g., diabetes versus skin rash).

#### 1. *Participants and Design*

Participants were 132 MTurkers (85 females; Mage = 36.75, SD = 12.10) who were paid for participation. The participants were randomly assigned to one of the four conditions of a 2 (Legality: Legal vs. Illegal) x 2 (Impact: Small vs. Big) between-subjects design.

#### 2. *Procedure*

The method of Experiment 2 was identical to that of 1B, except that impact was described as distinct outcomes. In the following scenarios legal actions appear in the text; illegal actions appear in parentheses. Big impact appears in the text; small impact appears in parentheses.

##### i. Scenario 1 (CEO)

Suppose there is a CEO of a large food chain who must decide whether to offer products containing ABA, a chemical compound that is often used in plastic containers and canned food. This chemical is cost-effective; using it minimizes costs and maximizes profit margins. The CEO decides to offer the products that contain ABA. Notably, ABA consumption is correlated with an increased health risk of diabetes (skin rashes). According to the FDA, it is legal to sell food containing up to 2 mg (1.5 mg) of ABA. The products the CEO decides to offer contain up to 1.75 mg of ABA.

##### ii. Scenario 2 (scientist)

Suppose a scientist who works for a pharmaceutical company runs several studies to investigate the safety and efficacy of a new medicine. The scientist obtains mixed results. Some show that the drug is safe and effective. Others show no effect of the drug. Some further studies also reveal a correlation between using this drug and experiencing side effects such as irregular heart rhythms (headaches). The scientist decides to re-run the studies that had the best results. These results confirm the positive findings from before. The scientist only publishes the positive results in the medical literature. The FDA approves the drug. According to the FDA, it is legal for scientists investigating the safety of prescription drugs to exclude up to 30% (20%) of their results. The scientist excluded 25% of their results.

## iii. Scenario 3 (prison)

Suppose there is a director of a private prison who is trying to maximize the prison's earnings. The greatest profits can be achieved when the prison is fully booked. For the highest degree of capacity utilization, it is better to transfer prisoners from one prison to another. The prison director decides to transfer prisoners to maximize occupancy of the prison. Notably, research shows that moving prisoners away from their support system of family and friends leads to an increase in the rate of prisoners returning to committing felonies such as burglaries (misdemeanors such as shoplifting). According to federal law, it is legal for private prisons to transfer inmates up to 200 miles (100 miles). The director decides to send some to prisons up to 150 miles away.

3. *Analyses and Results*

As in Experiment 1, we conducted a multivariate analysis of variance (MANOVA), with legality (legal, illegal) and impact (small, large) as factors, and moral judgment, legality judgment, and impact judgment as dependent measures. The results show that subjects used only legality to determine the moral permissibility of the actions, such that they judged legal actions to be more morally permissible than illegal actions ( $M_s = 3.1$  and  $1.69$ , respectively);  $F(1, 128) = 44.04, p < .001, \eta^2 = .26$ . There was no effect of consequences ( $M_s = 2.39$  and  $2.48$ ),  $F(1, 128) = 0.13, p = .715$ , and impact did not interact with legality,  $F(1, 128) = 0.07, p = .932$ . Legality affected the perceived legality of the action, such that legal actions were perceived as more legal than illegal actions ( $M_s = 3.08$  and  $6.23$ , respectively);  $F(1, 128) = 98.42, p < .001, \eta^2 = .43$ . Finally, when asked about the severity of consequences, subjects estimated illegal actions to be more severe than legal actions ( $M_s = 5.58$  and  $4.82$ , respectively);  $F(1, 128) = 8.76, p = .004, \eta^2 = .06$ . Unlike the pretest, here, where subjects were given information about legality, the consequences did not affect the judged severity of 10 consequences,  $F(1, 128) = 0.32, p = .58$ , nor was there an interaction,  $F(1, 128) = 0.44, p = .51$ .

As in Experiment 1, in order to further explore the results, we investigated whether political orientation affects the relative weight people give to legality versus morality. Among our 132 subjects, 45 mentioned that in the 2016 presidential elections they voted for Hillary Clinton, and 36 voted for Donald Trump (the rest either did not disclose or gave other names). The results show that legal actions were judged as more morally acceptable than illegal actions ( $M_s = 3.41$  and  $1.89$ , respectively);  $F(1, 73) = 26.16, p < .0001, \eta^2 = .26$ . Subjects who endorsed Clinton judged the behavior as less morally acceptable than subjects who endorsed Trump ( $M_s = 3.01$  and  $2.29$ , respectively);  $F(1, 73) = 5.91, p < .018, \eta^2 = .075$ . The remaining effects were not significant.

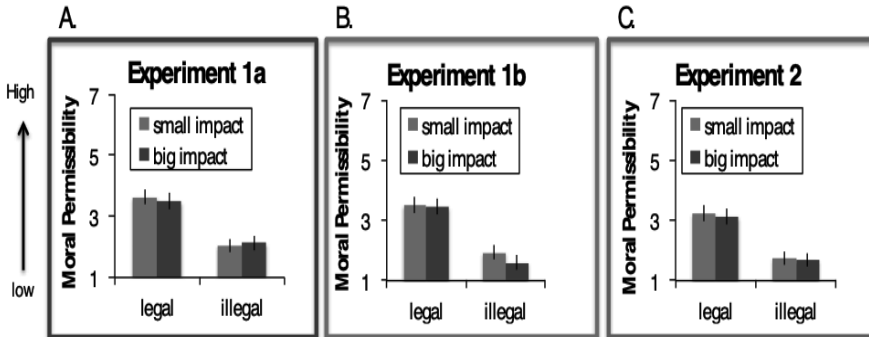


Figure 1. People rely on legality when judging moral permissibility of actions. Subjects judged the moral permissibility of actions that are either legal or illegal. In Experiment 1a (Fig 1A) and 1b (Fig 1B), the impact of the action was defined as having “slightly” or “significantly” negative consequences. In Experiment 2 (Fig 1C) the impact was manipulated as different outcomes that a pretest showed were considered severe or not (e.g., diabetes vs. skin rashes). In all experiments, subjects judged the legal actions as more morally permissible than the illegal actions, suggesting outsourcing of moral judgment to the Community of Knowledge.

## II. GENERAL DISCUSSION: THE LAW AS A SHORTCUT FOR MORAL JUDGMENT

The current research shows that people judge institutional corruption as more morally acceptable than criminal actions, even when the portrayed events are identical. Moreover, our data suggest that the severity of outcomes of the portrayed action does not play a role in people’s judgments. Thus, institutionally corrupt actions were judged as more acceptable than criminal actions, regardless of whether the action had severe or benign consequences. Finally, we observed that political orientation (operationalized as voting for Trump or Clinton in the 2016 presidential elections) did not affect the relative weight of legality and outcomes, nor did it diminish any of those effects.

Why is institutional corruption judged as more morally acceptable than criminal actions? One explanation for our findings is that people use the law as a shortcut to judge whether an action is morally right or wrong. Evaluating the morality of actions is not easy. Moral judgments of real-life events involve numerous considerations, including the intentions and beliefs of actors, outcomes of actions,<sup>34</sup> protected values,<sup>35</sup> and one’s emotional reactions.<sup>36</sup>

<sup>34</sup> Amit & Greene, *supra* note 32, at 861; see Fiery Cushman, Victor Kumar & Peter Railton,

Research from cognitive science and psychology suggest that, when faced with complex assessment tasks, people use simplifying strategies to make decisions.<sup>37</sup> It is therefore possible that when it comes to moral judgments, one simplifying strategy that people adopt is to not consider the moral dilemma at all. Instead, they outsource the question to the law.<sup>38</sup> According to this logic, the law does not simply reflect people's sense of corruption, but actually determines it. This means that framing actions in terms of their legality is more than just a signal of the actions' morality; the framing shapes their morality.

An appeal to the law is an appeal to the moral reasoning of a community of thought leaders (such as legislators and policy makers) who have determined the legality of a class of actions that includes the one under consideration. Outsourcing cognitive tasks to the community is necessary when problems are too complex for individual reasoning alone. The many factors that are taken into account when judging moral dilemmas, including inferences about intentions, outcomes,<sup>39</sup> the need to

*Moral Learning: Psychological and Philosophical Perspectives*, 167 COGNITION 1, 3 (2017) (discussing the role empathy plays in moral responses to outcomes of actions); Liane Young & Rebecca Saxe, *The Neural Basis of Belief Encoding and Integration in Moral Judgment*, 40 NEUROIMAGE 1912, 1913 (2008) (discussing how children base their moral judgments primarily on an action's consequence).

<sup>35</sup> See Jonathan Baron & Mark Spranca, *Protected Values*, 70 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1, 1 (1997) (defining protected values as those that resist trade-offs with other values, particularly economic values); Philip E. Tetlock, *Thinking the Unthinkable: Sacred Values and Taboo Cognitions*, 7 TRENDS COGNITIVE SCI. 320, 320–21 (2003) (discussing how moral outrage may be triggered by discovering that community members have compromised sacred values).

<sup>36</sup> See Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCIENCE 2105, 2106 (2001) (“Some moral dilemmas . . . engage emotional processing to a greater extent than others . . . and these differences in emotional engagement affect people's judgments.”); Jonathan Haidt, *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 PSYCHOL. REV. 814, 814 (2001) (discussing how, under a social intuitionist model, “moral intuitions (including moral emotions) come first and directly cause moral judgments”).

<sup>37</sup> See Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 49, 81 (Thomas Gilovich et al. eds., 2002) (“The substitution of one question for another, the representation of categories by prototypes, the view of erroneous intuitions are easy to override but almost impossible to eradicate . . . We show here that the same ideas apply to a diverse class of difficult judgments . . .”); JOHN W. PAYNE ET AL., *THE ADAPTIVE DECISION MAKER 2* (1993) (“When faced with more complex choice problems involving many alternatives, people often adopt simplifying (heuristic) strategies that are much more selective in the use of information.”); Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5 COGNITIVE PSYCHOL. 207, 207 (1973) (“We propose that when faced with the difficult task of judging probability or frequency, people employ a limited number of heuristics which reduce these judgments to simpler ones.”).

<sup>38</sup> See MASS. INST. TECH., *HANDBOOK OF COLLECTIVE INTELLIGENCE* (Thomas W. Malone & Michael S. Bernstein eds., 2015).

<sup>39</sup> See Cushman et al., *supra* note 34, at 3 (discussing the ways in which learning the outcomes of their actions affects children's future decisions); Young & Saxe, *supra* note 34, at 1913 (“Even though they can represent beliefs, these children continue to base their moral judgments primarily on the action's consequences. . .”).

protect sacred values,<sup>40</sup> and the need to integrate emotional reactions to events,<sup>41</sup> make them complex enough to be strong candidates for outsourcing. For example, an assessment of big pharma's culpability for aggressively selling opioids requires an analysis of the various companies' prior knowledge and goals when selling opioids, the extent of their responsibility for the multiple traumatic consequences of the opioid crisis, as well as the amount of suffering that was reduced by the administration of opioids, and this must be weighed against each of one's basic moral values while controlling one's sense of outrage.

There are alternative explanations for our results. One is that people hold the law in such high esteem that breaking the law is itself morally wrong and hence illegal actions are by definition immoral. To the extent the law is a set of rules that serves to protect people and their rights, it should play a substantial role in maintaining an ordered society, and an act that breaks the law should be viewed as inherently immoral. This may be, but it does not explain why consequences had no influence on judgment. Presumably consequences should be an additional contributor to our assessments of an act's morality. Another alternative explanation is that legality is easier to evaluate than consequences<sup>42</sup> because it is binary (legal versus illegal) whereas consequences are complex. This is possible although the consequences in our scenarios were not really complex and the differences between the bad and very bad consequences were actually quite stark. It is also possible that legality is more salient than consequences. This seems unlikely and does not explain why consequences had no effect whatsoever. If it is a matter of salience, one would expect a less salient dimension to have a smaller effect, but not no effect.

### III. IMPLICATIONS

Overall, these data show that people judge institutional corruption as more morally acceptable than criminal actions. Two otherwise identical actions were given different moral appraisals by virtue of their legal status. Actions that carry severely negative consequences may pass in the public

---

<sup>40</sup> See Baron & Spranca, *supra* note 35, at 1 (defining protected values as those that resist trade-offs with other values, particularly economic values); Tetlock, *supra* note 35, at 320–21 (discussing how moral outrage may be triggered by discovering that community members have compromised sacred values).

<sup>41</sup> See Greene et al., *supra* note 36, at 2106 (“Some moral dilemmas . . . engage emotional processing to a greater extent than others . . . and these differences in emotional engagement affect people’s judgments.”); Haidt, *supra* note 36, at 814 (discussing how, under a social intuitionist model, “moral intuitions (including moral emotions) come first and directly cause moral judgments”).

<sup>42</sup> See Christopher K. Hsee, *The Evaluability Hypothesis: An Explanation for Preference Reversals Between Joint and Separate Evaluations of Alternatives*, 67 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 247, 249 (1996) (discussing the way in which individuals will choose one decision-making process over another because it is easier).

as morally acceptable if they are legal, while criminal actions with relatively minor consequences may be perceived as morally unacceptable.

Regardless of the explanation for the effect of legality and the neglect of consequences for judgments of morality, the results suggest that people rely on the law not only to prescribe and prohibit actions, but to actually determine their sense of what is moral. In other words, the law not only shapes how people act, but it also shapes what they believe. These findings have implications for both legislation and how we assess our society and institutions.

With regard to legislation, lawmakers can take advantage of our findings by using them as justification to use the law to nudge people in socially beneficial ways, such as discouraging people from smoking in public spaces, protecting endangered species, and pushing people to avoid sugary drinks. As makers of the law—and consequently, shapers of morality—lawmakers have the power to redefine the relationship between institutions and people who are affected both directly and indirectly by those institutions. Through legislative measures, institutional behavior can be restructured, both internally (e.g., eliminating conflicts of interest, restructuring guidelines followed by the institution) and externally (e.g., evaluating the impact on the greater good).

But the findings also suggest that we should be aware that our evaluations of existing institutional actions may be influenced in ways that we are not entirely aware of. Laws may carry information about whether actions are morally acceptable or not, but laws may reflect values that are out-of-date, they may have been inspired by technology or other facts that are no longer relevant, and they may be influenced by political and economic interests. In other words, there are a variety of reasons to be skeptical of the information carried by current law. Thus, it is important to have an independent means to judge the morality of action, not to rely only on the law as it is written. The foreseen consequences of an action seem a worthwhile basis for such judgment. Without considering such consequences, the legal status quo will sustain itself without adequate rationale. Pharmaceutical companies will continue to gouge innocent victims and politicians will continue to bend the law in favor of themselves and their supporters.

Our findings thus provide additional reason to critically evaluate both the law and the morality of our institutions. We need to evaluate the consequences of the products we purchase and the policies and politicians we support. Society cannot rely only on existing law to make moral assessments; it needs to constantly refresh its justifications for the actions it considers right and wrong.

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### Minimizing the Impact of Cognitive Bias in Transactional Legal Education

ALINA BALL

*This Article explores methods law professors can employ to address the cognitive biases their law students possess. This Article provides concrete thoughts on how transactional law clinics can utilize the social, political, and neuroscience research included in this symposium edition.*





# Minimizing the Impact of Cognitive Bias in Transactional Legal Education

ALINA BALL \*

The partisan divide that evolved over appropriate measures to prevent the spread of COVID-19 has provided yet another public display of political tribalism, populace polarization, and epistemic authority.<sup>1</sup> During a presidential term pledged with brazen politics and mutual disdain for the other side,<sup>2</sup> even the global crisis of the pandemic has not been sufficient to close the chasm.<sup>3</sup> There seems to be no better moment than the present to take seriously the themes raised in this symposium volume and reflect on the role law schools can have in making law students sensitive to the complexity of human decision making. Society relies on lawyers to reconcile conflicting interests, ensure flow of reliable information, minimize opportunism that might otherwise exist between opposing parties, and marshal evidence that facilitates problem solving in the midst of ambiguity. But lawyers, judges, and politicians are all themselves susceptible to the same cognitive vulnerabilities that breed the current political polarization and, more generally, exacerbate conflict. This Article explores methods law professors can employ to align these inconsistencies between the role of the lawyer and the humanity of lawyers. Clearly, law professors can make an impact on the lawyers their students become. This

---

\* Professor of Law, UC Hastings College of the Law.

<sup>1</sup> See Cary Funk & Alec Tyson, *Partisan Differences Over the Pandemic Response Are Growing*, PEW RES. CTR. (June 3, 2020), <https://www.pewresearch.org/science/2020/06/03/partisan-differences-over-the-pandemic-response-are-growing/>; Marc J. Hetherington & Isaac D. Mehlhaff, *American Attitudes Toward Covid-19 Are Divided by Party. The Pandemic Itself Might Undo That.*, WASH. POST (Aug. 18, 2020), <https://www.washingtonpost.com/politics/2020/08/18/american-attitudes-toward-covid-19-are-divided-by-party-pandemic-itself-might-undo-that/>.

<sup>2</sup> Emma Green, *Americans Hate One Another. Impeachment Isn't Helping*, ATLANTIC (Nov. 2, 2019), <https://www.theatlantic.com/politics/archive/2019/11/impeachment-democrats-republicans-polarization/601264/>; Ryan Warner, *What Impeachment Says About Tribalism; Pushback on Polis*, CPR NEWS (Jan. 17, 2020), <https://www.cpr.org/show-episode/jan-17-2020-what-impeachment-says-about-tribalism-pushback-on-polis/> (recording at 24:10); Joel Rose, *Americans Increasingly Polarized When It Comes To Racial Justice Protests, Poll Finds*, NPR (Sept. 3, 2020), <https://www.npr.org/2020/09/03/908878610/americans-increasingly-polarized-when-it-comes-to-racial-justice-protests-poll-f/>; and Grace Sparks, *CNN Poll: Americans Are Divided Over Amy Coney Barrett*, CNN (Oct. 7, 2020), <https://www.cnn.com/2020/10/07/politics/cnn-poll-scotus-aca-october/index.html>.

<sup>3</sup> See Alex Fitzpatrick, *Why the U.S. Is Losing the War on COVID-19*, TIME (Aug. 13 2020), <https://time.com/5879086/us-covid-19/> (describing a “disturbing partisan trend” regarding views on the pandemic).

Article provides some concrete thoughts on how transactional law clinics can utilize the social, political, and neuroscience research included in this symposium edition.

In her article, *Winning Isn't Everything*, Professor Carrie Menkel-Meadow argues that even though it is often rare that everyone in a legal matter will “win,” lawyers must recognize that they could do better for their clients by practicing problem-solving lawyering.<sup>4</sup> She goes on to present problem-solving lawyering in opposition to conventional, adversarial lawyering.<sup>5</sup> Problem solving, she argues, produces better solutions not by simply compromising<sup>6</sup> or seeking to avoid conflict. Instead, problem solving requires collaboration and coordination between counterparties and adversaries.<sup>7</sup> It also requires the lawyer to exercise creativity<sup>8</sup> and engage in robust fact investigation to better understand context and consequences.

The narrative example of problem-solving negotiation is a story Professor Menkel-Meadow shares about her and her brother regularly fighting over a piece of chocolate cake.<sup>9</sup> Their mom would let their fighting go on for a while and then eventually intervene by cutting the piece of cake down the middle.<sup>10</sup> While this would stop the argument, it didn't resolve the issue between the siblings.<sup>11</sup> Both the young Professor Menkel-Meadow and her brother would walk away aggrieved<sup>12</sup> because she wanted the icing while her brother preferred the cake.<sup>13</sup> If her mom had asked her and her brother what their underlying motives and goals were, her mom could have uncovered this difference in interests.<sup>14</sup> Professor Menkel-Meadow uses this narrative as a tangible example of how, with problem-solving negotiation, even with finite resources, it is possible for opposing parties to get 100 percent of what they want if the lawyers understand each party's underlying interests.<sup>15</sup> It is the role of the problem-solving lawyer to conceive the multiple possibilities that even a single piece of cake presents. Even when “the pie” cannot be expanded, it

---

<sup>4</sup> Carrie J. Menkel-Meadow, *When Winning Isn't Everything: The Lawyer as Problem Solver*, 28 HOFSTRA L. REV. 905, 906 (2000).

<sup>5</sup> *Id.* at 907.

<sup>6</sup> *Id.* at 906, 911.

<sup>7</sup> *Id.* at 910–11.

<sup>8</sup> *Id.* at 912, 915.

<sup>9</sup> *Id.* at 911.

<sup>10</sup> *Id.*

<sup>11</sup> *Id.* at 911–12.

<sup>12</sup> *Id.* at 912.

<sup>13</sup> *Id.*

<sup>14</sup> *Id.*

<sup>15</sup> *Id.*

can be possible for the lawyer to creatively divide it to optimize client objectives.<sup>16</sup>

All lawyers have an ethical obligation to act in the best interest of their client.<sup>17</sup> If problem-solving lawyering is in the best interest of the client, it is surprising that it is not the conventional form of lawyering. Far from the norm, Professor Menkel-Meadow explains that conventional lawyering is centered around a culture of adversarialism.<sup>18</sup> As she explains, the hallmarks of the legal profession are argument, selective marshaling of facts, debate, competition, and performance of toughness, all of which often escalate conflict and lead to a stalemate.<sup>19</sup> Conventional lawyering adopts the zero-sum mentality,<sup>20</sup> described by Professor Daniel Shapiro in *Overcoming Political Polarization*,<sup>21</sup> which animates the dominant legal culture and performance of lawyers. Ultimately, Professor Menkel-Meadow argues it is a lack of creativity leading many lawyers to view scarcity or limited resources as zero-sum. “[A]dversarialism . . . leads us to argue in oppositional modes, to see black or white, to resist nuance and complexity . . . .”<sup>22</sup> In other words, adversarialism fosters cognitive biases that impede a lawyer’s ability to represent their client. Too often, lawyers are failing to reach their minimum goals for their represented cases.<sup>23</sup> This phenomenon of dominant lawyering culture has serious consequences in the legal profession—consequences that attorneys have an ethical obligation to be mindful of—because this affects their clients. In light of this, law professors should attempt to mitigate dominant lawyering culture in their classrooms.

Professor Menkel-Meadow provides steps to becoming problem-solving lawyers. First, she advises that there are several questions problem-solving lawyers must ask: (1) what are the client’s needs and goals; (2) what are the motivations of the counterparty; (3) what are the underlying interests of the counterparty; (4) what is at stake in this dispute;

---

<sup>16</sup> *Id.* at 916.

<sup>17</sup> See MODEL RULES OF PROF’L CONDUCT r. 1.3 cmt. (AM. BAR ASS’N 1983) (“A lawyer should act with commitment and dedication to the interests of the client and with zeal in advocacy upon the client’s behalf.”); *id.* r. 1.7 cmt. (stating that a lawyer’s duty is to prioritize the client’s interest above the lawyer’s own).

<sup>18</sup> Menkel-Meadow, *supra* note 4, at 907.

<sup>19</sup> *Id.*

<sup>20</sup> *Id.* at 906.

<sup>21</sup> See Daniel L. Shapiro, *Overcoming Political Polarization: The Power of the Civic Mindset*, 52 CONN. L. REV. \_ (forthcoming 2020) (describing the tribal mindset and its emotional dynamics that foster political polarization).

<sup>22</sup> Menkel-Meadow, *supra* note 4, at 909.

<sup>23</sup> James H. Stark & Maxim Milyavsky, *Towards a Better Understanding of Lawyers’ Judgmental Biases in Client Representation: The Role of Need for Cognitive Closure*, 59 WASH. U. J.L. & POL’Y 173, 176–78 (2019).

and (5) what are the outcomes produced by a given process.<sup>24</sup> In the conventional model of lawyering, the lawyer is likely to stop their inquiry after contemplating the motivations of the counterparty. Professor Menkel-Meadow explains that moving beyond motivation to understanding the underlying interests of the counterparty, can provide meaningful information for the lawyer. The example of Ann Atwater and C.P. Ellis's relationship demonstrates why this can be a powerful tactic in moving towards problem solving.<sup>25</sup> Research demonstrates that lawyers should be skeptical of their ability to effectively ascertain the counterparty's underlying interests in an adversarial context.<sup>26</sup> Given that lawyers may not be readily adept at obtaining or understanding a counterparty's underlying interests, there is even more justification for why law schools should prioritize teaching and allowing students to cultivate this skill. Law school should be an opportunity for budding attorneys to develop this metacognitive awareness and development. Practice utilizing the problem-solving questions Professor Menkel-Meadow proposes, would better prepare law students to achieve the goals of their future clients and make them more likely to resolve what might otherwise lead to a stalemate in a negotiation.

Dominant lawyering culture is not the only hurdle that prevents lawyers from effective problem solving. Poor problem solving is also a result of cognitive biases, which impede one's ability to effectively problem solve for their clients. Professor Menkel-Meadow addresses this in her analysis of problem-solving lawyering. Reactive devaluation, where one cannot hear something because it is coming from the other side, is a cognitive bias that occurs when a proposal is devalued if it appears to originate from an antagonist.<sup>27</sup>

Lawyers must acknowledge and overcome cognitive biases to effectively represent clients and for the legitimacy of the legal profession. Well-developed psychological evidence demonstrates that implicit bias is a strong cognitive bias that impedes human ability to effectively problem solve in social situations.<sup>28</sup> In the last decade, implicit bias has become the primary frame for contemporary discussion on social injustice, with

---

<sup>24</sup> Menkel-Meadow, *supra* note 4, at 916.

<sup>25</sup> *An Unlikely Friendship* (PBS television broadcast Feb. 1, 2019), <https://www.pbs.org/video/an-unlikely-friendship-tbnri0/>.

<sup>26</sup> See Jean R. Sternlight & Jennifer Robbennolt, *Good Lawyers Should Be Good Psychologists: Insights for Interviewing and Counseling Clients*, 23 OHIO ST. J. ON DISP. RESOL. 437, 448–73 (2008) (describing research from social and cognitive psychology that explores how people perceive and interpret information and make judgments).

<sup>27</sup> Menkel-Meadow, *supra* note 4, at 914.

<sup>28</sup> See, e.g., CHERYL STAATS ET AL., KIRWAN INST. FOR THE STUDY OF RACE & ETHNICITY, STATE OF THE SCIENCE: IMPLICIT BIAS REVIEW 19–20 (2017) (noting the effects of implicit bias on police encounters).

implicit bias on one end of the spectrum and deliberate, explicit racism, sexism, homophobia on the other end.<sup>29</sup> The idea of implicit bias has become ingrained in popular culture.<sup>30</sup> As Secretary Clinton mentioned during the 2016 presidential election, “[i]mplicit bias is a problem for everyone.”<sup>31</sup> When invoked in popular conversation, implicit bias is often used to describe the attitudes that are beyond conscious awareness or control. As the narrative goes, individuals manifest implicit bias without necessarily knowing it. In other words, implicit biases reflect the unconscious biases that individuals are unaware of<sup>32</sup> and, thus, do not control.

While the developing neuroscience does not support this popular concept of implicit or unconscious biases,<sup>33</sup> it is helpful to recognize that individuals possess unacknowledged biases. These biases are pervasive, deep seeded, and contrary to purported notions of fairness and justice. But if everyone is responsible, the old saying goes, then no one is responsible. Implicit bias can become the way that individuals absolve themselves of the consequences of their actions. Individuals can easily find themselves feeling comfortable because these biases are so pervasive there is little they can do about them. How does a law professor access what occurs beyond the conscious reach of the decision maker, which their introspection is not likely to reveal? Can a law professor do anything to bridge the gap between the dissociation of what students believe about themselves and what their biases reflect? These are the questions I continue to wrestle with because there are no easy answers or universal remedies.

I teach a corporate law clinic at UC Hastings College of the Law. My course prepares upper division law students to be corporate lawyers by having them represent social enterprise clients, businesses that use market-based strategies to achieve a social mission. For example, the clinic may represent a for-profit LLC marketing and design firm with the social

---

<sup>29</sup> Jonathan Kahn, *The 911 Covenant: Policing Black Bodies in White Spaces and the Limits of Implicit Bias as a Tool of Racial Justice*, 15 STAN. J. C.R. & C.L. 1, 2 (2019).

<sup>30</sup> See Shankar Vedantam et al., *Radio Replay: The Mind of the Village*, NPR: HIDDEN BRAIN (Mar. 9, 2018, 5:38 PM), <https://www.npr.org/2018/03/09/591895426/the-mind-of-the-village-understanding-our-implicit-biases> (discussing research about implicit bias and how it affects people’s perceptions).

<sup>31</sup> John A. Powell, *Implicit Bias in the Presidential Debate*, OTHER & BELONGING INST., <https://belonging.berkeley.edu/implicit-bias-presidential-debate> (last visited Jan. 26, 2020); *Clinton on Implicit Bias in Policing*, WASH. POST (Sept. 26, 2016, 9:00 PM), [https://www.washingtonpost.com/video/politics/clinton-on-implicit-bias-inpolicing/2016/09/26/46e1e88c-8441-11e6-b57d-dd49277af02f\\_video.html](https://www.washingtonpost.com/video/politics/clinton-on-implicit-bias-inpolicing/2016/09/26/46e1e88c-8441-11e6-b57d-dd49277af02f_video.html).

<sup>32</sup> See *Implicit Bias Module Series*, KIRWAN INST. FOR STUDY RACE & ETHNICITY, <http://kirwaninstitute.osu.edu/implicit-bias-training/> (defining implicit bias as “the attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner”).

<sup>33</sup> LASANA T. HARRIS, *INVISIBLE MIND: FLEXIBLE SOCIAL COGNITION AND DEHUMANIZATION* 159–60 (2017).

mission to hire and promote nontraditionally trained graphic designers who have criminal records. The company mission serves as a means for not only providing living wage employment to individuals often excluded in the formal economy but also elevating the voices and perspectives of their employees. The legal matters the clinic might take on for this client could range from drafting and analyzing business-to-business contracts, to renegotiating its commercial lease agreement, to amending and restating its operating agreement over the course of the representation. Law students in my course gain and refine their substantive corporate law knowledge in a context where they are applying their technical lawyering skills and critical thinking to pressing social and environmental issues central to their client's corporate purpose. The course provides a rare opportunity to future corporate lawyers to meaningfully engage in issues of social and economic justice.

Moreover, this is a class that affords me the bandwidth to help students recognize and confront their biases in the process of becoming better problem solvers.<sup>34</sup> In early iterations of my course, I began by asking students early in the semester (as part of a due diligence memo) to identify any biases that they may have going into this representation. But what I found is that students—like most of us—often do not have the tools to self-assess and be meaningfully reflective regarding cognitive biases, especially those that are contrary to their ideas of fairness, justice, and equity. I maintained the question as a part of the assignment for years, but was rarely satisfied with the responses it elicited from students. Regularly a student would write that they had no biases going into the representation. Once that response was provided, it would effectively end the conversation on their biases and provide me no basis to help them dig deeper. It occurred to me the assignment may be more harmful than productive in regards to bias reduction.

Recently, I adopted another method to the assignment with promising results. In my new iteration of the same due diligence memo, I have removed the explicit question about biases altogether. Instead, I now ask students to map out the client and its counterparty, or other key parties in the transaction, preferably in a sketch or diagram.<sup>35</sup> The prompt asks them to draw the relational dimensions of the client and other parties involved. Lastly, the prompt asks the student to identify the role of the lawyer and opportunities for the lawyer to influence the outcome of the transaction

---

<sup>34</sup> See generally Mark Neal Aaronson & Stefan H. Krieger, *Teaching Problem-Solving Lawyering: An Exchange of Ideas*, 11 CLINICAL L. REV. 485 (2004) (discussing the interaction between substantive legal knowledge and critical thinking skills in legal education).

<sup>35</sup> See JAY A. MITCHELL, *SKETCH PAD AS LEGAL PAD: PICTURING CORPORATE PRACTICE* 3 (2015) (discussing the benefit and practice of transactional lawyers sketching clients and deals).

consistent with the client's goals.<sup>36</sup> Transactional lawyers commonly sketch transactions and client matters. In the context of the due diligence memo assignment, the benefit of this diagramming and sketching is that it invites law students to memorialize their assumptions about distributions of power within the transaction or legal matter. Completing this assignment often requires the law students to make assumptions and draw conclusions outside of the facts of the client file.

When I debrief with the student teams, I then ask them to identify where their assumptions and cognitive biases are at play in the depiction. This is an attempt to not only identify, but begin to disrupt their cognitive biases.<sup>37</sup> During our conversation, I also invite the law students to think about the biases of the other actors in the representation. Some of the questions I ask them are, "What assumptions are other parties likely to have about your client? About you?" In this way, I am helping students become more aware and self-reflective of their cognitive biases and how biases impact their ability to represent their clients. This exercise is one concrete example of an attempt to debias my law students during their time in my course.

Professor Menkel-Meadow's model of problem-solving lawyering asks lawyers to answer a series of questions before attempting to problem solve.<sup>38</sup> This slowing down of the problem-solving process is intended to help the lawyers better identify the underlying interests of the counterparty, what is at stake in this dispute, and the outcomes produced by a given process.<sup>39</sup> These questions are also a useful debiasing technique. Her model of problem-solving lawyering allows law students to reflect on their own biases towards their clients, and how those might be contributing to the marginalization of their client in the representation. Lawyers have an opportunity to be architects and engineers of justice.<sup>40</sup> Lawyers have the capability to leave the world a better place. I believe that wholeheartedly and try to instill this notion in my students each semester.

As future transactional lawyers, I want my law students to understand they have the ability and the ethical responsibility to think outside the box, be creative, and engineer solutions that are structured around justice. By helping them acknowledge their cognitive biases early in the client representation, I believe I am facilitating them in becoming the better

---

<sup>36</sup> See Alicia Plerhoples, *Risks, Goals, and Pictographs: Lawyering to the Social Entrepreneur*, 19 LEWIS & CLARK L. REV. 301, 313 (2015) (explaining the utility of pictographs to memorialize client goals).

<sup>37</sup> See Joan C. Williams & Sky Mihaylo, *How the Best Bosses Interrupt Bias on Their Teams*, HARV. BUS. REV. (Nov.-Dec. 2019), <https://hbr.org/2019/11/how-the-best-bosses-interrupt-bias-on-their-teams>.

<sup>38</sup> Menkel-Meadow, *supra* note 4, at 909–10.

<sup>39</sup> *Id.*

<sup>40</sup> *Id.* at 910.



problem-solving lawyers that society desperately needs. More importantly, I am heartened by the experiences in my clinical course that there is even more transactional courses can do to make law students sensitive to the complexity of decision making and better prepared to defuse conflicts. If law professors are willing to engage with their students in the struggle of their ideological becoming, I believe the next generation of lawyers will be the change-makers this moment requires.

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### Semantic Descent: More Trouble for Civility

ROBERT B. TALISSE

*Civility is widely regarded as a duty of democratic citizenship. This Article identifies a difficulty inherent within the enterprise of developing an adequate conception of civility. Challenging the idea civility is the requirement to remain calm, peaceable, or dispassionate in political debate, it is argued that that civility is instead the requirement to address one's political arguments to one's interlocutors. In this way, civility is a second-order requirement, a norm governing our conduct in political disagreement. From there, a conceptual problem for civility so understood is raised, the problem of semantic descent. It is argued that any plausible conception of civility is prone to being "weaponized," transformed into a partisan device for incivility. The general upshot is that as important as civility is for a well-functioning democracy, its usefulness as a diagnostic tool for repairing political dysfunctions is limited.*

## ARTICLE CONTENTS

INTRODUCTION.....	1151
I. THE PUZZLE OF CIVILITY.....	1153
II. CIVILITY AS CIVIC ADDRESS.....	1157
III. CIVILITY AS A RECIPROCAL DUTY.....	1160
IV. CIVILITY AS A SECOND-ORDER NORM.....	1162
V. THE PROBLEM OF SEMANTIC DESCENT.....	1165
CONCLUSION.....	1167



# Semantic Descent: More Trouble for Civility

ROBERT B. TALISSE \*

## INTRODUCTION

It is commonly observed by commentators from across the political spectrum in the United States that the past decade has seen a marked intensification of political antagonism.<sup>1</sup> Office holders and party officials seem to have lost the capacity to cooperate with those who do not share their political affiliation, resulting in political deadlock and stagnation.<sup>2</sup> In fact, the trouble runs deeper than that. Whereas partisan animosity has long prevailed among politicians, it only recently has trickled down to, and intensified among, the U.S. citizenry.<sup>3</sup> Citizens struggle these days to communicate productively across partisan divides. Thus, in addition to the intransigence within government, popular political discourse is beset by a range of dysfunctions that are commonly referred to collectively as *incivility*.<sup>4</sup> A common thought is that we can begin to repair our democracy only by first restoring a commitment to civility.

That incivility poses trouble for democracy is acknowledged by the American citizenry. A 2019 Pew study finds that most Americans believe that political debate has become dangerously toxic and disrespectful in recent years.<sup>5</sup> A prior study finds that Americans disapprove of the incivility that has taken hold of democracy and want greater comity and cooperation among politicians.<sup>6</sup> These findings offer some hope that the remedy for current political dysfunctions is not beyond reach. If citizens

---

\* W. Alton Jones Professor of Philosophy and Professor of Political Science, Vanderbilt University. Thanks to Scott Aikin, Jody Azzouni, and Lisa Madura for comments on an earlier version of this Article.

<sup>1</sup> Shanto Iyengar & Sean J Westwood, *Fear and Loathing Across Party Lines: New Evidence on Group Polarization*, 59 AM. J. POL. SCI. 690, 692 (2014).

<sup>2</sup> *The Partisan Divide on Political Values Grows Even Wider: Sharp Shifts Among Democrats on Aid to Needy, Race, Immigration*, PEW RES. CTR. (Oct. 5, 2017), <https://www.people-press.org/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/>.

<sup>3</sup> *Id.*

<sup>4</sup> Kathleen Hall Jamieson et al., *The Political Uses and Abuses of Civility and Incivility*, in THE OXFORD HANDBOOK OF POLITICAL COMMUNICATION 205, 205–19 (Kate Kenski & Kathleen Hall Jamieson eds., 2017).

<sup>5</sup> Bruce Drake & Jocelyn Kiley, *Americans Say the Nation's Political Debate Has Grown More Toxic and 'Heated' Rhetoric Could Lead to Violence*, PEW RES. CTR. (July 18, 2019), <https://www.pewresearch.org/fact-tank/2019/07/18/americans-say-the-nations-political-debate-has-grown-more-toxic-and-heated-rhetoric-could-lead-to-violence/>.

<sup>6</sup> See *supra* note 2.

dislike over-the-top political rancor and want a more civil mode of politics, then the challenge is simply that of incentivizing politicians and citizens to turn down the heat.

Alas, this optimism fades in light of a further finding of the same study.<sup>7</sup> Many who disapprove of the uncivil streak in contemporary U.S. politics also believe that the blame for incivility lies squarely with their political opponents; moreover, they think the solution is to force capitulation from the other side.<sup>8</sup> In calling for more civility, many citizens seek *submission* or *acquiescence* from their political rivals. Surely civility cannot require capitulation from one's political opposition. Something is amiss.

The problem rests with the concept of civility itself. We all identify it as a political good, and incivility a political bad, yet we do not share a sufficiently nuanced understanding of what civility is. Similarly, we agree that democratic citizens and office holders have a duty to be civil and are therefore blameworthy when they behave uncivilly; nonetheless, in the absence of a clearer account of what civility requires, the duty of civility lacks definite content. Finally, as it is popularly understood, civility is closely related to politeness, courtesy, and generosity; while satisfying these norms may be admirable, the duty of civility arguably involves something else, as it is not clear that citizens have a duty to be polite, courteous, or generous. In short, there is the conceptual work to be done. If incivility is to play a central diagnostic and prescriptive role in our analysis of current political dysfunctions, and if, as is commonly held, civility is to be understood as a duty of democratic citizenship, then we need to work from a suitably detailed conception of what civility is and why it is valuable.

To anticipate a point that will feature in a later part of this Article, in pursuing this conceptual task, it is crucial that we develop a conception of civility that is itself nonpartisan. That is, in order to perform its prescriptive and diagnostic job, our understanding of civility must not be rigged in favor of our own favored political stances. Whatever civility is, it must be satisfiable by citizens from across the full spectrum of democratically valid political opinion.<sup>9</sup> In other words, a conception of civility that renders out of bounds all advocacy for political positions that we oppose is no

---

<sup>7</sup> *Id.*

<sup>8</sup> *Political Polarization in the American Public: How Increasing Ideological Uniformity and Partisan Antipathy Affect Politics, Compromise and Everyday Life*, PEW RES. CTR. (June 12, 2014), <https://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/>.

<sup>9</sup> This formulation leaves open the possibility that some political viewpoints are *intrinsically* uncivil, such that their expression is necessarily a violation of the duty of civility. Certain overtly racist, sexist, and nationalistic views can be plausibly regarded as uncivil. The point, which will be discussed further below, is that our conception of civility must allow for civil expressions of political viewpoints that we find erroneous, perhaps severely so.

conception of civility at all. We might say, then, that a proper conception of civility must itself be endorsable across the spectrum of democratically acceptable political perspectives. It must be in this sense *public*.

This Article identifies a difficulty inherent within the enterprise of developing a public conception of civility. I begin by raising a puzzle about civility as popularly understood, which in turn reveals that the duty of civility cannot be the requirement to remain calm, peaceable, or dispassionate in political debate. In fact, the initial puzzle suggests that civility cannot be a norm governing one's *tone* or *demeanor* in political disputation at all. It will be proposed in Section Two that civility rather is a norm of discursive engagement that requires one to *address* one's contributions to political argument to one's interlocutors; civility is a norm of *engagement*, pertaining to how well one's contributions to debate track the contributions of one's interlocutors. It will then be argued, in Section Three, that whatever the more precise details of civility might be, the duty to be civil is a *reciprocal* requirement; that is, we are bound by the norm of civility only when there is sufficient reason to expect that one's interlocutors will abide by it as well. In Section Four, I further specify conditions that a viable conception of civility must meet. In particular, I draw upon some terminology introduced by Willard Van Orman Quine to show that civility must be understood as a second-order norm, a norm governing our conduct in contexts of disagreement. In Section Five, I spell out a conceptual problem for civility so understood, what will be called the *problem of semantic descent*. It will be argued that any plausible conception of civility is prone to being "weaponized," transformed into a partisan device for incivility. The general upshot is that as important as civility is for a well-functioning democracy, its usefulness as a diagnostic tool for repairing political dysfunctions is limited.

## I. THE PUZZLE OF CIVILITY

Begin by assuming a popular conception of civility. According to it, civility involves remaining calm, reserved, cooperative, courteous, and dispassionate in political debate. This conception of civility has many merits, and perhaps chief among them is that it comports well with common usage of the term, including use in contexts outside of politics. Thus, in assuming the popular conception of civility, one is claiming that in politics one ought to abide by norms that govern other familiar discursive contexts where disputation might arise, from the family dinner table to the workplace, classroom, and queue at the supermarket. In short, the popular conception of civility in politics has the virtue of making civility a *unitary* concept, a set of norms applicable to interpersonal affairs across the board, as it were.

Of course, the nastiness of politics may be regrettable without thereby being democratically pathological. Therefore, laments over the incivility of

contemporary politics are often accompanied by an additional (though often tacit) claim, namely that democratic citizens have a duty to be civil. This additional contention is needed if incivility is to be regarded as a democratic failure or political pathology. That is, in order for incivility to register as a democratic dysfunction, we need to make a case for thinking that democratic citizens and office holders have a duty to be civil.

The case for the duty of civility runs roughly as follows: If it is to flourish, a democracy needs to get things done; this means that political rivals—especially office holders, but citizens too—must cooperate and sometimes even compromise. When discourse is overly antagonistic, participants are incentivized to dig in their heels and adopt a stance of intransigence. This not only leads to deadlock, it also encourages a mode of politics that abandons the democratic ideal of self-government among equals, replacing it with a strictly strategic conception that values only winning and so regards political opponents as merely obstacles to be neutralized. Yet democracy is a partnership among political equals; thus, democratic citizens, including office holders, have a duty of civility.<sup>10</sup> Consequently, when they fail to be civil, citizens have fallen short of proper citizenship; and when incivility is rampant within a polity, democracy is failing.

This reasoning is solid, but it occasions a puzzle. Understood as self-government among political equals, democracy calls upon citizens to take responsibility for their collective political life; they must in some sense *take ownership* of their political order. This means that citizens must be *participants* in the tasks of democracy. Accordingly, they have duties to be informed, competent, and reflective as citizens.<sup>11</sup> In addition, they have a duty to *exercise their judgment* about public affairs. They are expected to think through the political issues of the day and to form their own political opinions. What is more, as citizens are one another's equals, the style of judgment required by citizenship involves a kind of perspective-taking; democratic citizens are expected to reason from a collective point of view,

---

<sup>10</sup> See RONALD DWORKIN, IS DEMOCRACY POSSIBLE HERE?: PRINCIPLES FOR A NEW POLITICAL DEBATE 131–34 (2006) (detailing the “partnership” view of democracy).

<sup>11</sup> There is a massive literature focusing on whether citizens are up to the tasks of democratic citizenship. Though this question is important, it is not my focus here. For a range of views, see generally CHRISTOPHER H. ACHEN & LARRY M. BARTELS, DEMOCRACY FOR REALISTS: WHY ELECTIONS DO NOT PRODUCE RESPONSIVE GOVERNMENT (2017) (detailing how the idea of thoughtful citizens guiding a democratic state from the voting booth is fundamentally misguided); ROBERT GOODIN, REFLECTIVE DEMOCRACY (2003) (offering a solution when political disputes in a democracy invariably mix facts with values); HÉLÈNE LANDEMORE, DEMOCRATIC REASON: POLITICS, COLLECTIVE INTELLIGENCE, AND THE RULE OF THE MANY (2013) (arguing that democracy is legitimate, just, and smart because the decision by the many will be more accurate than decision by the few); ILYA SOMIN, DEMOCRACY AND POLITICAL IGNORANCE: WHY SMALLER GOVERNMENT IS SMARTER (2016) (revealing how the depths of political ignorance in America cause a major challenge for democracy).

rather than from a narrowly first-personal perspective. That is, in thinking through political issues, each citizen must look beyond their individual interests and attempt to discern the collective or *public* good.<sup>12</sup> Many democratic theorists claim that participating in public political discussion is an indispensable step in adopting the collective point of view; hence these theorists hold that democratic citizens are also obligated to engage in forms of public deliberation and discussion.<sup>13</sup>

The puzzle emerges from the fact that under conditions of intellectual freedom and equality, citizens' judgments over normative matters are unlikely to converge. As John Rawls observed, the freedoms secured in a constitutional democracy give rise to a "reasonable pluralism" of political ideas, doctrines, perspectives, and opinions.<sup>14</sup> Accordingly, the aim of public participation and deliberation is not *consensus* but, as indicated above, the achievement of a properly *public* perspective from which to form one's political judgements and conduct political debate. Rawls's insight, well-confirmed in democratic practice, is that properly public judgments about complex political issues nonetheless diverge.<sup>15</sup> Hence, political disagreement and disputation are inexorable from democracy.

The task of much of democratic theory is to envision and design processes and institutions that can *manage* our political divides. One might go as far as to say that constitutions simply are rulebooks for containing our divisions within the navigable boundaries of fair rules and procedures. But when it comes to interactions among the disputants, this managerial function is fraught. Although some political disagreements concern relatively nominal matters, many invoke citizens' sense of justice. Though the familiar debates over healthcare, poverty, taxation, immigration, and the environment are sometimes articulated as the strictly managerial disputes of what the country can afford, they are ultimately disputes over justice, and disagreements about what a morally acceptable society is required to do. Hence, heat frequently accompanies these disputes. When arguing about matters invoking our sense of justice, we tend to see our interlocutors as being not merely on the *incorrect* side of the question, but on the *unjust* side. And when their side prevails politically, we are bound

---

<sup>12</sup> See Joshua Cohen, *Procedure and Substance in Deliberative Democracy*, in DEMOCRACY AND DIFFERENCE: CONTESTING THE BOUNDARIES OF THE POLITICAL 95, 95–96 (Seyla Benhabib ed., 1996) (explaining the importance of decisions being made collectively within a democracy, even when there are no shared comprehensive moral or religious views); Seyla Benhabib, *Deliberative Rationality and Models of Democratic Legitimacy*, 1 CONSTELLATIONS 26, 28–29 (1994) (distinguishing between collective good and individual interests in the democratic process).

<sup>13</sup> The deliberative democrats are perhaps the most obvious proponents for this kind of view. But the idea that citizens have a duty to discuss politics is not limited to deliberativists.

<sup>14</sup> JOHN RAWLS, POLITICAL LIBERALISM 36 (2005).

<sup>15</sup> GERALD GAUS, THE TYRANNY OF THE IDEAL: JUSTICE IN A DIVERSE SOCIETY ch. 1 (2016).



to regard that outcome as not merely suboptimal or regrettable, but morally unacceptable.

Here, then, is the puzzle occasioned by the popular conception of civility and its corresponding duty. Heated tones and antagonistic demeanors are precisely what one should expect in disputes where the interlocutors take themselves to be defending against injustice. Moreover, as was noted above, democratic citizens have a duty to take responsibility for their collective political life. Accordingly, when matters of justice are at stake, they are thus likely to bring to political disagreements the emotional and normative investments that tend to manifest in heat and hostility. In short, antagonism, animosity, resentment, and indignation are what we should expect from democratically engaged citizens when arguing sincerely over important public matters. The features of political discourse that are commonly cited as constitutive of incivility appear to be the natural byproducts of responsible citizenship. The duty of civility, taken in its popular conception, looks inconsistent with proper democratic citizenship.

A defender of the popular view might retort that the reasoning above shows not that civility is incompatible with proper citizenship, but only that it is difficult. She may contend that the duty of civility is the requirement to *contain* or *suppress* the tendency, natural though it may be, to be exercised in political discussions when matters of justice are at stake. The duty of civility, it may be claimed, is the requirement to not allow our passions to get the better of us in political disputation.

This line of defense prompts a formidable critique of the idea that the duty of civility involves the requirement to be soft-spoken, collected, and reserved in political debate. Theorists working in various feminist idioms have argued—correctly, in my view—that, when civility is understood broadly as politeness, the duty of civility gives unwarranted advantage to the status quo and the existing balance of political power.<sup>16</sup> One of the ways in which the duty of civility so understood accomplishes this is by creating opportunities for the politically powerful to defuse criticisms and objections by fixing on the manner in which they are expressed. Those for whom the status quo is most objectionable tend also to be those who are most passionate and indignant. Consequently, the duty of civility, again

---

<sup>16</sup> See, e.g., Virginia Held, *Non-Contractual Society: A Feminist View*, 13 CAN. J. PHIL. SUPPLEMENTARY VOLUME 111–37 (1987), reprinted in SCIENCE, MORALITY & FEMINIST THEORY 209, 213–14 (Marsha Hanen & Kai Nielsen eds., 1987) (noting that women were not “expected to demand equal rights”); Iris Marion Young, *Communication and the Other: Beyond Deliberative Democracy*, in DEMOCRACY AND DIFFERENCE: CONTESTING THE BOUNDARIES OF THE POLITICAL 120, 122–24 (Seyla Benhabib ed., 1996) (arguing that “differences of speech privilege correlate with other differences of social privilege”); Lynn M. Sanders, *Against Deliberation*, 25 POL. THEORY 347, 348–49 (1997) (arguing that “appeals to deliberation do nothing to challenge an undesirable status quo” and ultimately disadvantage individuals “who are already underrepresented in formal political institutions”).

assuming the popular understanding of civility that we have been exploring, is in effect an instrument by which democratic responsibility is deflected and democratic judgment disabled. Once again, given the popular understanding of civility, the duty of civility seems incompatible with responsible democratic citizenship.

This provides a clue as to what civility and its corresponding duty cannot be. Whatever value there might be in remaining calm and collected in the midst of political disagreement, it cannot be a requirement of democratic citizenship to do so. It may be an effective strategy to manifest a reserved and dispassionate demeanor in political discussion, but it is not a democratic duty to do so. The heat, tone, and animosity of contemporary political discourse might be regrettable for a variety of reasons, but it does not itself constitute a failure of citizenship. Consequently, the duty of civility requires something else of us.

To be clear, civility in the popular sense that we have thus far adopted nonetheless captures a common collection of *pro tanto* norms which I do not take the arguments above to impugn. Moreover, I take it that there is a *pro tanto* duty that applies to persons to be civil in their interactions, and perhaps especially their disputes. The argument thus far has fixed only on the idea that civility, as it is popularly construed, is the core of a duty of citizenship. Accordingly, the upshot thus far may be formulated as follows: Civility is not a unitary concept. Rather, there is a distinct sense of civility that applies strictly to contexts of democratic politics, and consequently, a duty of civility that applies to persons specifically in their role as democratic citizens.

## II. CIVILITY AS CIVIC ADDRESS

What, then, does this distinctively democratic concept of civility and its corresponding duty amount to? In pursuing this, we can draw further a lesson from the feminist line of criticism introduced above. Recall that the target of that critique was a conception of the duty of civility that too easily enabled powerholders and advocates of the status quo to deflect objections from those who find the status quo unacceptable. Widening the focus slightly, we can see the feminist critique as emphasizing a crucial feature of democratic responsibility: in order to realize the ideal of self-government among equals who take ownership of their collective political order, citizens must be *accountable* to one another. When it comes to citizens who are also holders of public office, this means that they must render themselves *vulnerable* to the criticisms of their fellow citizens. Broadly, democratic citizens owe to one another actual *engagement*; in conducting their political disputes, they must *address* one another in a way that is consistent with their standing as democratic citizens, that is, as political equals.

Addressing another person in a way that is consistent with their standing as an equal citizen is not easy, especially when matters of justice are at stake. Just as it is easy to manifest hostility towards those who one perceives to be advocating injustice, it is easy to dismiss one's political opponents as political inferiors. However, unlike expressions of political hostility, attitudes of political superiority are obviously at odds with the core democratic ideal of self-government among equals. In other words, the political disputation that is inexorable from—because it is constitutive of— democracy is always disputation among equals. I want to propose, then, that the core of civility (now taken in its distinctive democratic sense) is the norm of *addressing* political interlocutors as one's fellow citizens, and thus as equal partners in self-government. In short, I propose that we think of civility as a norm of *civic address*, and the corresponding duty as requiring that we address our fellow citizens in a way that is consistent with their standing as our equals.

Central to civility as *civic address* are three broad requirements. First, there is *responsiveness*. In political discourse, including of course argumentation, citizens must strive to be *responsive* to their interlocutors' actually stated views and reasons, rather than with strawmen or other opportunistically constructed fabrications. Second, is the requirement of *connection*. That is, in political discourse, citizens must strive to address their contributions to one another, rather than to onlooking audiences or a choir of sympathetic partisans; in argumentative contexts, interlocutors must not use one another as mere props, foils against which to mug to their allies in the audience. Third, citizens must endeavor to conduct their political discourse by means of reasons and considerations that they sincerely believe that their interlocutors will appreciate the force of. Call this the *mutuality* requirement for civic address. In short, in communicative and argumentative exchanges, we manifest a due recognition for our interlocutor's status as an equal when we strive to address them in a way that responds to their actual views, connects with them directly, and attempts sincerely to offer reasons and considerations that they will accept as such. Insofar as democratic citizenship involves a standing requirement to acknowledge our fellow citizens as our political equals, there is a duty of civility.

Readers will have noticed that the requirements of civic address have been formulated as requirements to *strive* and *endeavor* to engage with one's fellow citizens in a particular way. The formulation as requirements to *try* is necessary if we are to take due account of the fact that argumentative discourse is one of the ways in which citizens come to learn about their oppositions' perspectives and arguments. Consequently, civility must be consistent with a certain degree of sincere misunderstanding of others' views. Consider that it should not count as uncivil when a citizen fails to offer her interlocutors reasons that they could accept due to her

being honestly mistaken about the substance of their position. Borrowing a term proposed in a different context by Christopher Eberle, we can say that the resulting duty of civility is a duty of “*conscientious* engagement.”<sup>17</sup>

It is not difficult to discern some of the leading merits of this approach to conceptualizing civility and its corresponding duty. For example, on the civic address account, the mark of incivility does not lie in tone, heat, or demeanor, but rather in opportunistic refusals to take one’s critics or opponents as nonetheless one’s equal citizens. And this view enables us to identify popular modes of political discourse as especially uncivil. Note how often political argumentation involves strategic mischaracterization of the oppositions’ actual views, the mere pantomime of answering objections while simply restating one’s views for the sake of rallying one’s allies, and the tactic of offering as decisive reasons in favor of one’s views claims that are the very ones being called into question by one’s critics.<sup>18</sup> That these tactics are compatible with maintaining a soft, concessive tone and polite demeanor indicates the limitations of the popular conception of political civility.

The conception of civility as civic address is also able to accommodate the thought that hostility and animosity are generally *regrettable* features of our politics. After all, heat and attitude are tactics for *escalating* conflict, and when they are employed by those who are already unduly advantaged, they serve to diminish critics and smother criticisms. One could go so far as to say that the marks of incivility as popularly understood (aggression, name-calling, shouting, impoliteness, and the like) are reliable *signals* that civility in the sense of civic address is being breached. What matters in assessing a mode of discourse with respect to civility, then, is how well the interlocutors succeed in actually addressing one another in the relevant ways. Heat, animosity, and tone are consistent with civic address, even though they are frequently manifestations of its violation. Consequently, democratic citizens have a *pro tanto* duty to be civil in the popular sense, but this is parasitic on the duty to civically address their fellow citizens. This means that a failure to be civil in the popular sense does not itself constitute a failure of citizenship.

I realize that I have barely sketched my conception of civility as civic address. A complete view of civility would need to include far more detail than can be provided here. But recall that my objective in this Article is not to develop a conception of civility, but rather to identify a problem that any viable conception of civility will confront. My contention at this juncture is simply that the conception of civility as civil address, along with its

---

<sup>17</sup> CHRISTOPHER J. EBERLE, RELIGIOUS CONVICTION IN LIBERAL POLITICS 104 (2002).

<sup>18</sup> See SCOTT F. AIKIN & ROBERT B. TALISSE, WHY WE ARGUE (AND HOW WE SHOULD): A GUIDE TO POLITICAL DISAGREEMENT IN AN AGE OF UNREASON 10–12 (2d ed. 2019) (offering detailed analyses of these pathologies).

corresponding duty, is both commonsensical and noncontroversial enough to serve as a template from which more detailed articulations of civility would proceed. Spelling out the details of civility as civic address is undertaken in other work.<sup>19</sup>

### III. CIVILITY AS A RECIPROCAL DUTY

Any conception of civility that takes something like civic address as its foundation will conceive of civility as a norm of *engagement* with one's fellow citizens, and this will naturally entail that the duty of civility is what may be called a *reciprocal* duty.<sup>20</sup>

To explain, some duties are *first-personal* while others are *reciprocal*.<sup>21</sup> An analogy with garden-variety moral virtue will be helpful. Consider a virtue like *moderation*. This virtue establishes a standard of conduct that requires an individual's temperance in the pursuit of enjoyment. This standard is *first-personal*.<sup>22</sup> What it requires is not contingent on the presence of other temperate people; the virtue of temperance *applies* to individuals as individuals, and demands of them individual moderation, even in the presence of immoderate company. Another example of a first-personal virtue is *courage*. The courageous person must stand firm in fearful situations, even when surrounded by cowards. To be sure, precisely *what* course of action courage requires might depend on one's company and what they are currently doing; nonetheless, that others are cowards does not license anything less than courage from the courageous person. Again, courage, as a first-personal virtue, applies to the individual.

Now contrast these first-personal virtues with requirements of a different kind. These do not primarily attach to individuals, but instead govern *groups* of individuals or are exhibited in relations between them. That is, they establish a standard of conduct for *us* rather than simply for *me* and *you*. For example, we teach our children the policy "keep your hands to yourself." But notice that the policy of keeping one's hands to oneself establishes a standard of conduct for *those on the playground*; more importantly, it is in virtue of its *collective application* that individuals are bound to comply with its requirements. Consequently, when Billy violates the norm by grabbing Danny, and Danny retaliates, it would be absurd to criticize Danny for failing to keep his hands to himself. With Billy's violation, the collective norm is suspended, and in extricating himself by pushing Billy away, Danny does not himself *break* the rule. To

---

<sup>19</sup> SCOTT F. AIKIN & ROBERT B. TALISSE, POLITICAL ARGUMENT IN A POLARIZED AGE 53–56 (forthcoming 2020).

<sup>20</sup> The discussion that follows draws heavily from AIKIN & TALISSE, *supra* note 18, ch. 9.

<sup>21</sup> *Id.*

<sup>22</sup> *Id.*

better capture this, notice that the norm “keep your hands to yourself” is an abbreviated version of the more complex norm “keep your hands to yourself on the condition that others are keeping their hands to themselves.” We see, then, that the norm and its corresponding duty are *reciprocal*; they establish a standard of conduct that applies to groups, and individuals are required to abide by the norm, as long as others generally do so as well.<sup>23</sup>

Notice that in this playground case, the norm does not indicate what one is permitted to do in response to its violation. Surely there are certain retaliatory acts that Danny could perform against Billy that would be inappropriate or even impermissible. That Billy’s violation suspends the collective norm does not afford to Danny moral *carte blanche* to respond however he wishes. Though his retaliatory response does not itself constitute a violation of the “keep your hands to yourself” norm, Danny may still retaliate in ways that render him worthy of criticism, perhaps even punishment.

Return now to politics. It is clear that some political duties are first-personal. As a citizen, one’s engagements with others must manifest the requirements of honesty and evenhandedness. That one’s fellow citizens are inveterate dissemblers does not license one to be dishonest or biased. In fact, when dishonesty is widespread, honesty and evenhandedness are all the more important. However, other political duties are reciprocal. They prescribe modes of conduct to us—*collectively*, so to speak. Accordingly, individuals are required to abide by these requirements only when they are embraced and generally practiced by the group. Where the norm corresponding to a reciprocal duty is commonly disregarded within a group, the duty itself is rendered inactive, as it establishes a standard of behavior *only under the conditions where the norm is collectively embraced*.

Given that the duty of civility follows from the democratic aspiration of maintaining responsibility for the citizens’ shared political order, it looks clearly reciprocal. We are required to hold ourselves politically accountable to our fellow citizens, and hence to render ourselves vulnerable to their criticisms, as a way of manifesting our recognition of their status as equal partners in democratic self-government. When others are disposed to incivility, they decline to hold themselves politically accountable to us; under such conditions, it would be perverse to take ourselves to nonetheless be bound by the duty of civility. Indeed, abiding by the duty of civility when it is generally flouted is tantamount to abetting our own political subordination, as it places constraints on our political activity and puts our views and objectives at a relative disadvantage. One

---

<sup>23</sup> *Id.*

might say, then, that abiding by the duty of civility under conditions where it is not reciprocally practiced helps to empower and entrench incivility, which in turn serves to further deteriorate democratic conditions. In abiding by the duty of civility when dealing with the uncivil, we contribute to the strategic effectiveness of incivility. Under certain extreme circumstances, there may be a duty of incivility.<sup>24</sup>

#### IV. CIVILITY AS A SECOND-ORDER NORM

The next stage of the argument begins from a brief digression into what might seem to be exceedingly distant territory—the views of the twentieth century American empiricist philosopher Willard Van Orman Quine, who wrote nothing in political philosophy and only one curious essay in moral philosophy.<sup>25</sup> Our present interest in Quine is limited to a particular bit of nomenclature—specifically, the term *semantic ascent*<sup>26</sup>—that he introduced as he was theorizing the structure of philosophical disagreements over what might be called *ontology*, or disagreements over what exists.<sup>27</sup> I suspect readers are familiar with certain longstanding philosophical disputes that belong to the subfield of metaphysics. In particular, there is a debate going back to at least the Presocratics about whether everything that exists is physical.<sup>28</sup> This is closely related to a debate over whether everything that exists has *causal* properties. Typically, these debates focus on the seemingly indispensable nature of certain *abstracta* (that is, nonphysical items), such as numbers, sets, classes, and kinds, whose role in the causal order is on anyone's view mysterious.<sup>29</sup> Quine was concerned with whether disputes over the existence of such (purported) items could be well-ordered. It would seem that between the physicalist (who denies that there are *abstracta*) and the nonphysicalist (who affirms that some *abstracta* exist), there could be no proper

---

<sup>24</sup> See generally CANDICE DELMAS, A DUTY TO RESIST: WHEN DISOBEDIENCE SHOULD BE UNCIVIL (2018) for a similar discussion in the context of the duty to obey the law. Delmas argues that in cases where the political order is significantly unjust, there is not only no duty to obey the law, but possibly a duty to disobey, perhaps even in ways that overtly violate norms of civil disobedience. *Id.* at 224–25.

<sup>25</sup> For a reliable overview of Quine's philosophy, see, for example, CHRISTOPHER HOOKWAY, QUINE: LANGUAGE, EXPERIENCE AND REALITY 50–53 (1988) (describing Quine's views on pragmatism, relativism, and realism).

<sup>26</sup> WILLARD VAN ORMAN QUINE, WORD AND OBJECT 270–74 (1960) [hereinafter QUINE, WORD AND OBJECT].

<sup>27</sup> See HOOKWAY, *supra* note 25, at 20 (“Quine speaks of the set of objects which a speaker takes to exist as his ‘ontology’: an assertion ‘ontologically commits’ its author to objects of a certain kind if the assertion would only be true if objects of that kind existed.”).

<sup>28</sup> See JONATHAN BARNES, THE PRESOCRATIC PHILOSOPHERS (1979) (describing various ideas of the Presocratics).

<sup>29</sup> For an example, see generally, Kathrin Koslicki, *On the Substantive Nature of Disagreements in Ontology*, 71 PHIL. & PHENOMENOLOGICAL RES. 85 (2007).

engagement, as the argumentative maneuvers from the one side seem destined to appear question-begging to the other.<sup>30</sup> Indeed, it seems that among these interlocutors, their disagreement is likely to extend to the question over what should count even as a relevant consideration with respect to the dispute.

Quine gave the name *semantic ascent* to the strategy of dealing with disputes of this kind by shifting the site of the dispute, at least momentarily, towards the terms each side employs in formulating their position; interlocutors should try ascending from talking about the matter in dispute to talking about the way they talk about the matter in dispute. Quine writes that semantic ascent “is the shift from talking in certain terms to talking about them.”<sup>31</sup> He reasons that “[w]ords . . . unlike [e.g.] classes . . . are tangible objects of the size so popular in the marketplace, where men of unlike conceptual schemes communicate at their best.”<sup>32</sup>

The strategy of semantic ascent, then, is that of looking for relatively uncontested common ground, in the hope that disputants might find some foothold from which to more productively conduct their more fundamental disagreements. Importantly, the strategy involves not simply a change in the topic of the disputants’ conversation, but a shift in the *level* of the topic of the conversation. They are to shift from arguing about, e.g., whether classes exist to talking about how they talk about whether classes exist. This *second-order* conversation will focus on hopefully more tractable questions such as what each takes to be adequate *definitions* or *conceptualizations* of the disputed items, what renders those definitions and conceptualizations adequate, and so on.

Again, the hope is that by ascending to semantic level, interlocutors might discover that there is enough difference in their nomenclature as to render their dispute over classes merely verbal. (In that case, they may nonetheless have to confront a prolonged debate at the semantic level.) Or they could discover that they are largely in agreement over the semantics, which itself might constitute a kind of progress in their first-order dispute about classes. Semantic ascent, then, is not a way to resolve disputes, but rather a means for making disputes more orderly.

Semantic ascent is of course a commonsense strategy that we intuitively employ in disputes of all kinds. What is of particular interest at present is that the strategy calls attention to a general fact of disagreements, namely that they run simultaneously along two evaluative tracks that may be distinguished by referring to *first-order* and *second-order* levels of evaluation. In this way, argumentation functions a

---

<sup>30</sup> See generally Jody Azzouni, *On “On What There Is”*, 79 PACIFIC PHIL. Q. 1 (1998) (discussing “debates over the Quine-Putnam Indispensability thesis”).

<sup>31</sup> QUINE, WORD AND OBJECT, *supra* note 26, at 271.

<sup>32</sup> *Id.* at 272.



lot like communication more generally. More specifically, in order to carry on a conversation, communicators must simultaneously track two distinct levels of language use, namely the meanings of the words being used (the semantics, the first order) and the grammatical structure in which they are used (the syntax, the second order). Missteps in tracking either distinct level typically result in communicative failure.<sup>33</sup> What is more, communicators need a common second-order idiom in order to diagnose and correct certain kinds of communicative failure.

Similarly, in order to successfully engage in argumentation, disputants must track and evaluate each other's claims (the first order) as well as what might be called their *argumentative performance* (the second order). For example, when engaged in argumentation about, say, gun ownership laws, disputants must be able to track one another's first-order claims about guns, ownership, and the law, while also attending to the ways in which those claims figure into the evolving dialectic among them. Minimally, in order to engage in argumentation, each interlocutor's contribution must: take account of the relevant prior statements introduced into the discussion; avoid contradicting their own prior statements; refrain from needlessly repeating claims that are already manifestly agreed upon; decline to repeatedly assert as a premise that which is disputed among them; and so on. In other words, interlocutors must perpetually be "keeping score" of the state of play in their disagreement.<sup>34</sup> This is a second-order evaluative site that is accordingly distinct from the first-order assessment of the reasons offered about gun ownership laws. Accordingly, just as an unsound argument can have a true conclusion, an impeccable first-order case for stricter gun laws can be a dialectical failure.

We are now able to tie the hanging threads together. As a collection of norms governing our engagement with fellow citizens in political disagreements, civility as civic address has largely to do with what has just been called our *argumentative performance*. We fail to be civil not simply

---

<sup>33</sup> This is not intended as a complete analysis of communication, but only as identifying two necessary conditions for communicative success. Command of the relevant semantics and syntax does not suffice for successful communication; according to many views, one must also develop an adequate conception of the interlocutor's communicative *intention* in talking as she does. See PAUL GRICE, *STUDIES IN THE WAY OF WORDS* 26 (1989) (formulating "the Cooperative Principle" for participants in conversation: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged").

<sup>34</sup> David Lewis, *Scorekeeping in a Language Game*, 8 J. PHIL. LOGIC 339, 344–46 (1979). See also ROBERT B. BRANDOM, *MAKING IT EXPLICIT: REASONING, REPRESENTING, AND DISCURSIVE COMMITMENT* 6 (1994) (explaining an aim of the work is "to address the question, How should the relation between representation . . . and the discursive concepts of reason and truth be understood?"); MARY KATE MCGOWAN, *JUST WORDS: ON SPEECH AND HIDDEN HARM* 2 (2019) (identifying "a previously overlooked mechanism by which ordinary speech by ordinary speakers under ordinary circumstances enacts harmful norms and thus constitutes, rather than merely causes, harm. Harm is constituted when the harm is brought about via adherence to norms enacted").

in virtue of the position we hold, nor simply in virtue of the heat or tone with which we express our position, but centrally in virtue of the ways in which our presentation of our position fails to properly address our interlocutor. In this sense, civility is a second-order norm, and the corresponding duty is similarly second order—it is a duty to engage political disagreement in a particular way. In short, the duty of civility has to do with our dialectical conduct.

The second-order nature of civility may seem so obvious as to not require extended comment. However, it is worth emphasizing this feature of the duty in order to properly understand its role in democratic life. The concept of civility is needed to articulate a feasible understanding of what citizens owe to one another, given that political disagreement is an inevitable product of democracy itself. That is, democratic citizens need to be able to evaluate one another's dialectical conduct as a way of managing their first-order disputes. As with the ontological debates that concerned Quine, citizens need to be able to talk about the way they talk about the issues that divide them, as this second-order talk can help them to arrive at mutually agreeable understandings of where their disagreements lie and the precise shape they take. Perhaps the ability to ascend to talk in the second order about how well one's arguments address one's interlocutors can sometimes assist in making progress with first-order disputes over policy. At the very least, the ability to avail ourselves of a second-order idiom for evaluating discursive conduct may be able to help prevent undue escalation in contexts of disagreement.

And here is where a point made at the beginning of this Article comes into its own. Note that the idiom of civility can perform these democratic tasks only if it is second order. That is, a conception of civility can play its proper role in democracy only if it specifies requirements that can be met by those who hold erroneous political views, and also violated by those who hold correct political views. We might say that the requirements specified by the duty of civility must be nonpartisan, that is, *impartial* with respect to political error and correctness. Or, put another way, a viable conception of civility and its corresponding duty must enable us to countenance cases where those with whom we agree at the first order are nonetheless failing at civility, and those with whom we disagree at the first order are nonetheless civil.

## V. THE PROBLEM OF SEMANTIC DESCENT

My hope is that nearly everything that has been said thus far strikes the reader as unobjectionable, perhaps even commonsensical. Still, the elements are in place for formulating what I see as a serious difficulty confronting any plausible conception of civility. When it comes to political disagreement, there are no analogues to Quine's "tangible objects" that allow people "of unlike conceptual schemes" to "communicate at their

best.”<sup>35</sup> To put the point bluntly, what goes up can come down, too. And our second-order idiom for evaluating each other’s discursive conduct is subject to what can be called *semantic descent*—the shift by which our second-order evaluative terms are conscripted into playing a combat role in first-order skirmishes.

To get the flavor of what I have in mind, consider a term used in talking about online discourse: “troll.” This term entered into our vernacular as a way to identify a mode of online conduct that is objectionable in virtue of being provocative and haranguing simply for the sake of disrupting discussion threads.<sup>36</sup> In this original usage, being a “troll” has little to do with one’s substantive commitments; “trolling” is the second-order phenomenon of engaging in a way designed to derail conversation. Once the term gained currency, however, it quickly became vulnerable to semantic descent: the term is now widely deployed as a term of abuse to characterize those with whom one disagrees over issues at the first order. “Trolls” are simply one’s political opponents, and “trolling” is what trolls (so understood) do in articulating their views.

Another example of the phenomenon of semantic descent is the term “fake news.” The term was introduced to describe a source that poses as journalistic, but in fact is not.<sup>37</sup> The term thus initially denoted a second-order feature of various web sites, television programs, and print media. However, “fake news” is now widely used as a term to deride the *content* that is reported by a journalist.<sup>38</sup> At its worst, “fake news” is simply what one calls reportage that is favorable to one’s political rivals.<sup>39</sup> In this way, the term has descended from its second-order function to our first-order debates. In other words, it initially served an *umpiring* function, but is now just another player in the political game.

Finally, consider the verb to “politicize.” This term entered the vernacular as a way of identifying cases where political actors attempt to gain politically from a high-profile event (typically a tragedy) that arguably ought to stand above the fray of politics. But now the term is itself

---

<sup>35</sup> QUINE, WORD AND OBJECT, *supra* note 26, at 272.

<sup>36</sup> ALICE MARWICK & REBECCA LEWIS, DATA & SOC’Y RES. INST., MEDIA MANIPULATION AND DISINFORMATION ONLINE 4–5 (2017), [https://datasociety.net/wp-content/uploads/2017/05/DataAndSociety\\_MediaManipulationAndDisinformationOnline-1.pdf](https://datasociety.net/wp-content/uploads/2017/05/DataAndSociety_MediaManipulationAndDisinformationOnline-1.pdf).

<sup>37</sup> JULIE POSETTI ET AL., JOURNALISM, ‘FAKE NEWS’ & DISINFORMATION: HANDBOOK FOR JOURNALISM EDUCATION AND TRAINING 14 (2018). This book references the “growing international concern about a ‘disinformation war’ in which journalism and journalists are prime targets,” and offers as evidence the joint statement issued by the United Nations Special Rapporteur for Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe’s Representative on Freedom of the Media, the Organization of American States’ Special Rapporteur on Freedom of Expression, and the African Commission on Human and People’s Rights Special Rapporteur on Freedom of Expression and Access to Information. *Id.*

<sup>38</sup> *Id.* at 66.

<sup>39</sup> *Id.* at 14–15.

deployed as a means for gaining politically under circumstances of that kind. More specifically, the term functions as an accusation that one wields against one's opposition as a way to claim for oneself the political high ground by disparaging one's rivals. But, of course, vying for the political high ground is itself an attempt to gain politically.

The trouble is that any conceptualization of civility, and indeed the very concept of civility itself, is subject to semantic descent. That is, although civility is useful to us as a way of thinking about how we conduct ourselves in political argumentation, it can always be opportunistically wielded as just another way of carrying out our first-order debates. When *civility* has descended, it functions as a term of approval for those who express positions that we find favorable, and *incivility* serves as a term for expressing disapprobation at the first-order positions that we find objectionable.

The matter grows bleaker still once we recall that civility norms are best understood as requirements to *try* to engage in ways that successfully address one's fellow citizens. This means that the question of whether a citizen is being civil in a particular dispute is largely a matter of judgment: we must assess whether our interlocutor is endeavoring adequately to address us as an equal and communicating in good faith. A significant body of empirical literature demonstrates that evaluations of this kind are highly responsive to exogeneous factors. In short, we are very likely to regard those with whom we disagree politically as ill-motivated, untrustworthy, and unreliable. Accordingly, our assessments of their dialectical conduct will be responsive to our evaluations of their first-order views in ways that will lead us to assess our opponents as uncivil.

Recall the argument from Section Three that the duty of civility is a reciprocal requirement. This means that individual citizens are bound by the requirement only provided that their interlocutors seem prepared to reciprocate. Once civility has descended from its second-order function, we grow increasingly unable to regard our political rivals as capable of reciprocating. We therefore take ourselves to not be bound by civility's norms. Crucially, our opposition reasons similarly. The result is democratically degenerative—a condition where citizens are able to interact civilly only with those who are politically just like themselves, and, with some justification, see those who are politically different as undeserving of civility (because unwilling or unable to reciprocate).

#### CONCLUSION

Return to the Pew results mentioned at the beginning of this Article. Citizens say they want a more civil mode of politics, but they also see their political opposition as the source of incivility, and accordingly want their rivals simply to stand down and acquiesce. This collection of attitudes is precisely what we should expect when civility has semantically descended,

when its second-order content has been degraded to the point where it can no longer serve as a basis for impartial evaluation of discursive conduct. The upshot, I think, is that the concept of civility is at this point probably unable to play a useful role in diagnosing and repairing our democracy. To put the point in a different way, no suitably detailed conception of civility will be sufficiently public and impartial to serve as a tool for addressing our political dysfunctions.

Where does that leave us? Although I cannot develop the thought here, my sense is that the partisanship, animosity, polarization, and distrust that have taken hold of our democracy cannot be remedied with prescriptions for better politics. Any such prescription is likely to be received by a large portion of those to whom it is addressed as yet another partisan maneuver. What is needed instead is the recognition that these dysfunctions are possibly the natural result of allowing politics to play too great a role in our collective life. As other scholars have noted, the intensification of partisan animosity and polarization has occurred alongside sociological shifts in the country that have served to place individuals' partisan affiliations at the center of their social identity. In short, we now tend to see ourselves as *fundamentally* defined by our political loyalties.<sup>40</sup> Along with this shift in our self-understanding has come a transformation of the social spaces we inhabit in our day-to-day lives: they have become increasingly segregated along partisan divides.<sup>41</sup> Perhaps the right response, then, is to devise channels by which citizens can come to see in one another something beyond partisan identities. As paradoxical as it might sound, the way to repair our democracy is to keep ourselves mindful of the fact that, as important as democracy is, we cannot live well together as equals solely by means of politics.<sup>42</sup>

---

<sup>40</sup> See LILLIANA MASON, *UNCIVIL AGREEMENT: HOW POLITICS BECAME OUR IDENTITY* 5–6 (2018) (“Parties simplify the voting decision. . . . [M]ost voters have a sense of party loyalty. They know, either through a lifetime of learning, from parental socialization, from news media, or through some combination thereof, that one party is better suited to them.”).

<sup>41</sup> See BILL BISHOP, *THE BIG SORT: WHY THE CLUSTERING OF LIKE-MINDED AMERICA IS TEARING US APART* 5 (2009) (“As Americans have moved over the past three decades, they have clustered in communities of sameness, among people with similar ways of life, beliefs, and, in the end, politics. . . . When people move, they also make choices about who their neighbors will be and who will share their new lives. Those are now political decisions, and they are having a profound effect on the nation’s public life.”).

<sup>42</sup> See ROBERT B. TALISSE, *OVERDOING DEMOCRACY: WHY WE MUST PUT POLITICS IN ITS PLACE* 4–5 (2019) (discussing “the *ubiquity* of democratic politics, the saturation of social life with activities and projects that are overtly organized around the categories and divisions of current politics”).

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Article

### The Long Con of Civility

LYNN MIE ITAGAKI

*Civility has been much on the minds of pundits in local and national political discussions since the 1990s. Periods of civil unrest or irreconcilable divisions in governance intensify concerns about civility. While its more archaic definitions refer to citizenry and civilization, civility is often promoted as the foundation or goal of deliberative democracies. However, less acknowledged is its disciplinary, repressive effects in maintaining or deepening racial, gendered, heteronormative, and ableist hierarchies that distinguish some populations for full citizenship and others for partial rights and protections.*

*In Part I, I examine a recent series of civility polls, their contradictory results, and how these contradictions can importantly expose the fissures of our contemporary moment and our body politic. In Part II, I describe the historical background of civility around race, gender, and sexuality and the unacknowledged difficulty in defining civility and incivility. In Part III, I extend this discussion to address the recent cases before the Supreme Court concerning LGBTQ+ employment discrimination and lack of accessibility. In conclusion, I identify what it would mean to analyze civility in terms of dignity on the basis of these cases about the equal rights and protections of their LGBTQ+ and disabled plaintiffs. We should be deeply suspicious with demands for civility that are often deployed to quell dissent from marginalized populations and to dampen democratic practices.*

## ARTICLE CONTENTS

INTRODUCTION.....	1171
I. THE PARADOX OF THE POLLS .....	1172
II. THE CULTURE WARS WERE CIVILITY WARS.....	1175
III. INCIVILITIES, INEQUALITIES, INDIGNITIES, INHUMANITIES	1181
CONCLUSION .....	1185



# The Long Con of Civility

LYNN MIE ITAGAKI\*

## INTRODUCTION

Civility, now more than ever, is a concern of our everyday conversations and political debates. Civility is a flexible concept which has referred to everything from polite conversation to an entire civilization, and its very expansiveness undermines its effectivity. Most often, civility denotes orderly turn-taking or polite manners exhibited by a “civilized” citizenry. It also includes internal contradictions, depending on the situation and its practitioners. Civility is the way in which we *should* talk about important things; civility is also the way in which we *do not talk* about important things. One person’s civility is another’s incivility. From the smallest rude remark to the institutionalized exclusion of groups, the flexibility of the term “civility” encompasses a tremendous range of customs, behaviors, practices, and traditions. Moreover, when people advocate for civility, they often neglect to disclose their implicit assumptions about to whom they owe civility, which may not include those with whom they disagree or disparage. Civility is often understood as the foundation or goal of deliberative democracies; however, it also has disciplinary, repressive effects in maintaining or exacerbating racial, gendered, heteronormative, and ableist hierarchies that distinguish some populations for full citizenship and others for partial rights and protections.

In Part I, I examine a recent series of civility polls, their contradictory results, and how these contradictions can importantly expose the fissures of our contemporary moment and our body politic. In Part II, I describe the historical background of civility around race, gender, and sexuality and the unacknowledged difficulty in defining civility and incivility. In Part III, I extend this discussion to address the recent cases before the Supreme Court concerning LGBTQ+ employment discrimination and lack of accessibility. In conclusion, I identify what it would mean to analyze civility in terms of dignity on the basis of these cases about the equal rights and protections of their LGBTQ+ and disabled plaintiffs. We should be deeply suspicious

---

\* Associate Professor of English and Women’s and Gender Studies, University of Missouri, Columbia. I sincerely thank Richard Wilson, the *Law Review* organizers, and editors of the 2019 symposium “How We Argue Now: The Moral Foundations of Politics and Law.” Thank you to Devin Fergus, Jennifer M. Gully, Srirupa Prasad, Linda Reeder, Leslie Wingard, and my colleagues at the University of Missouri for their invaluable feedback.



with demands for civility that are often deployed to quell dissent from marginalized populations and to dampen democratic practices.

### I. THE PARADOX OF THE POLLS

The consensus among Americans is that incivility seems to have gotten worse. Poll after poll appears to represent respondents' increasing frustration with incivility in our media and political debates. A late October 2019 poll<sup>1</sup> about civility in politics seemed to affirm with numbers the political polarization and profound partisanship that most people felt. In response to the statement "[t]he political, racial, and class divisions in this country are getting worse and our national dialogue is breaking down," seventy-seven percent agreed, a six point swing from April 2019's eighty-three percent (albeit within the  $\pm 3.1$  margin of error).<sup>2</sup> Other similar overwhelming majorities agreed with variations of these statements such as "I am frustrated by the uncivil and rude behavior of many politicians" (eighty-seven percent) and "[b]ehavior that used to be seen as unacceptable is now accepted as normal behavior" (eighty-four percent).<sup>3</sup>

However, polling data for this Civility Poll revealed an apparent contradiction expressed by a majority of respondents: while eighty-seven percent (with sixty-four percent strongly agreeing) agreed that "compromise and common ground should be the goal for political leaders," eighty-four percent (with sixty-three percent strongly agreeing) also agreed with the statement, "I'm tired of leaders compromising my values and ideals. I want leaders who will stand up to the other side."<sup>4</sup> Taking cues from the press release, media coverage repeated and amplified this apparent contradiction between respondents simultaneously wanting

---

<sup>1</sup> *October 2019 Civility Poll*, GEO. U. INST. POL. & PUB. SERV. (Oct. 2019), <http://politics.georgetown.edu/october-2019-civility-poll-2/>. This poll is run semi-annually by the Institute of Politics at Georgetown University with Republican-leaning and Democrat-leaning polling outfits.

<sup>2</sup> *Compare THE TARRANCE GRP. & LAKE RESEARCH PARTNERS, BATTLEGROUND 65 FINAL STUDY #16440*, at 1, 3 (2019), <http://politics.georgetown.edu/full-cross-tabs-and-questionnaire-october-2019-2/>, with *THE TARRANCE GRP. & LAKE RESEARCH PARTNERS, BATTLEGROUND 64 STUDY # 16354*, at 1, 3 (2020), <http://politics.georgetown.edu/civility-poll-tables-cross-tabs-questionnaire-april-2019/>.

<sup>3</sup> *THE TARRANCE GRP. & LAKE RESEARCH PARTNERS, BATTLEGROUND 65 FINAL STUDY #16440*, at 1, 3 (2019), <http://politics.georgetown.edu/full-cross-tabs-and-questionnaire-october-2019-2/>. The Georgetown Institute of Politics and Public Service Battleground Civility Poll from April 4, 2019 showed a three percent decrease and one percent increase, respectively, again within the margin of error. See *THE TARRANCE GRP. & LAKE RESEARCH PARTNERS, BATTLEGROUND 64 STUDY # 16354*, at 1, 3 (2020), <http://politics.georgetown.edu/civility-poll-tables-cross-tabs-questionnaire-april-2019/> (displaying results of the April 2019 poll).

<sup>4</sup> GU Politics, *New Poll: Voters Find Political Divisions So Bad, Believe U.S. Is Two-Thirds of the Way to "Edge of a Civil War"*, GEO. U. INST. POL. & PUB. SERV. (Oct. 23, 2019), <http://politics.georgetown.edu/press-releases/civility-press-release-oct-2019/>.

compromise and not wanting one's political leaders to compromise.<sup>5</sup> Mo Elleithe, Executive Director of the Georgetown Institute of Politics and Public Service, says that this contradiction "creates mixed messages for even the most skilled political leader trying to decide whether to be a fighter or a dealmaker."<sup>6</sup> But how could at least two-thirds, if not as many as three-quarters, of the same respondents say that they wanted these two contradictory things?<sup>7</sup>

This contradiction points to the benefits and drawbacks of civility, especially in terms of the promise of civil discourse and respectful debate it offers to many. The ostensible inconsistency with these two statements reflects one of the fundamental problems of civility and its "meaning clusters"<sup>8</sup> that go far beyond mere politeness, but rather activate its more obscure definitions of the citizenry and civilization. Assertions of civility rely on certain assumptions about to whom one owes civility.<sup>9</sup> For example, whether one should be civil to another individual or member of another group can depend on if one perceives these individuals or group members to share values and commitments to the nation.<sup>10</sup> The practices, tactics, and strategies for how we might achieve civility can require strong advocacy in and of themselves, as opposed to the most common demand to just be polite.

However, the increasing polarization of the U.S. electorate can lead to a more restrictive notion of civility, summoning the interrelated force of its archaic definitions of citizenry and civilization by limiting those to whom one should show civility. Consolidating this in-group and excluding others determine who is valued or disdained, a full citizen or second-class one, patriot or enemy.<sup>11</sup> Under these conditions, civility's flexible and wide-ranging definition may stretch to mean even its opposite. For example, Civility Poll respondents who wanted compromise and common ground might actually imply for respondents that those on the other side of the issue are responsible for reaching a compromise and finding common ground with their stated position. In other words, those on the other side of the debate should listen to me and to views like mine, expressed by my

---

<sup>5</sup> Paul Bedard, *Battleground: 7 in 10 Say US 'On the edge of Civil War'*, WASH. EXAMINER (Oct. 23, 2019, 1:27 PM), <https://www.washingtonexaminer.com/washington-secrets/battleground-7-in-10-say-us-on-the-edge-of-civil-war>; Victor Westerkamp, *Americans Believe We're Two-Thirds of the Way to a Civil War*, NTD (Oct. 24, 2019), [https://www.ntd.com/americans-believe-were-two-thirds-of-the-way-to-a-civil-war\\_396411.html](https://www.ntd.com/americans-believe-were-two-thirds-of-the-way-to-a-civil-war_396411.html).

<sup>6</sup> GU Politics, *supra* note 4 (alteration in original).

<sup>7</sup> *Id.* Given the margin of error, the number of respondents who agreed with both statements could be as high as seventy-seven, and as low as sixty-five, albeit the range clearly signals a distinct majority.

<sup>8</sup> Virginia Sapiro, *Considering Political Civility Historically: A Case Study of the United States 2* (1999) (unpublished manuscript), <http://blogs.bu.edu/vsapiro/files/2018/06/Considering-Civility.pdf>.

<sup>9</sup> *Id.* at 5–6.

<sup>10</sup> *Id.* at 7.

<sup>11</sup> *Id.* at 9.

elected representative, but I and people who agree with me shouldn't have to. Compromise describes what others do, therefore I may listen to those other perspectives but I will not change my position, as that would compromise my own values. I could justify this redefinition of compromise to mean not compromising by identifying the other side as not worthy of compromise by labeling these antagonists as criminal, unpatriotic, evil, subhuman—i.e., not worthy of civility. Thus, I only need to find compromise and common ground among those I consider citizens, patriots, voters, humans—essentially people who think like me. Moreover, I categorize as “humans” only those who think like me, or those with whom I would not have to compromise my values.

Respondents could be reflecting a social desirability bias as “shy” anti-compromise and anti-common-ground citizens and voters who might not admit certain opinions to pollsters. Whites might be especially fearful of appearing racist in ways that might parallel women's and racial minorities' fears of being stereotyped,<sup>12</sup> and Moss-Racusin et al. posit a “compensatory egalitarianism” in which “individuals may reassure themselves that although they did not select a minority candidate, they ‘repaid’ this candidate with implicit antiprejudice attitudes on their behalf.”<sup>13</sup> I posit that respondents might be sensitive to being perceived as not supporting democratic values or egalitarianism by being unwilling to find a compromise or common ground

This bias skews polls by small amounts that can alter expectations of close runs and flip anticipated results, as occurred in the 2016 presidential election outcome.<sup>14</sup> Known as the “Bradley effect,”<sup>15</sup> scholars and pundits named this phenomenon after the 1982 gubernatorial race of California in which African American mayor Tom Bradley led by a comfortable margin in the pre-election polls but eventually lost by fewer than 100,000 votes, or 1.2%.<sup>16</sup> Researchers identified respondents' concerns that they would be perceived as racist for not voting for Bradley, a candidate who could have

---

<sup>12</sup> Phillip Atiba Goff, Claude M. Steele & Paul G. Davies, *The Space Between Us: Stereotype Threat and Distance in Interracial Contexts*, 94 J. PERSONALITY & SOC. PSYCHOL. 91, 92 (2008).

<sup>13</sup> Corinne Moss-Racusin, Julie Phelan & Laurie Rudman, “I’m Not Prejudiced, but . . .”: *Compensatory Egalitarianism in the 2008 Democratic Presidential Primary*, 31 POL. PSYCHOL. 543, 546 (2010).

<sup>14</sup> Andrew Mercer et al., *Why 2016 Election Polls Missed Their Mark*, PEW RES. CTR.: FACT TANK (Nov. 9, 2016), <https://www.pewresearch.org/fact-tank/2016/11/09/why-2016-election-polls-missed-their-mark/>; Joseph P. Williams, *The Trump Effect*, U.S. NEWS: REP. (July 1, 2016, 3:30 PM), <https://www.usnews.com/news/articles/2016-07-01/are-voters-too-embarrassed-to-say-they-support-trump>.

<sup>15</sup> Ronald J. Vogel & Phillip Ardoin, *Ask Me No Questions, I’ll Tell You No Lies: Does the Bradley Effect Still Exist?*, 15 RACE GENDER & CLASS 65, 66 (2008).

<sup>16</sup> *Id.*; John Canalis, *Is the Bradley Effect Real? Deukmejian Says No*, LONG BEACH PRESS TELEGRAM (Oct. 28, 2008, 12:00 AM), <https://www.presstelegram.com/2008/10/28/is-the-bradley-effect-real-deukmejian-says-no/>.

been the first Black governor of California.<sup>17</sup> This statistical issue of how researchers and pollsters should most accurately factor in bias to avoid tipping the scales results from the phenomena of “covering,” as Kenji Yoshino calls what he takes up from Erving Goffman’s notion of hiding or downplaying a disfavored, stigmatized identity, in this case, an anti-Black racist one.<sup>18</sup> The generally unanticipated election of Donald J. Trump in 2016 and the 2015 surprising majority vote of the British to withdraw from the European Union are only the most recent spectacular examples of covering: the Shy Trump voter or the Shy Brexit voter whose answers reflected this bias, and for whom pollsters did not adequately account.<sup>19</sup>

Pollsters and jury selection consultants, among other experts, can attempt to ascertain more accurate opinions with indirect, rather than direct, questions. For example, in 2008 pollsters were confronted with the unprecedented question of whether registered Democrats would have a problem voting for a Black or female candidate, as was likely with Barack Obama and Hillary Clinton as the presidential nominee frontrunners. Pollsters had to ask indirect questions that more accurately predicted voting preferences. They knew they would get skewed results by asking participants directly if they were racist or sexist or even if they would support a qualified Black or female candidate, so instead pollsters had to develop versions of these questions that asked if “most of my neighbors”<sup>20</sup> or “most people they knew” would support these presidential candidates and if “America was ready” for them.<sup>21</sup> Pollsters need to work around a perceived norm, or what respondents might see as the more socially desirable or civil answer, because respondents are covering (hiding or toning down) what they think are disfavored political views and identities and ones that might generate hostility and conflict.

## II. THE CULTURE WARS WERE CIVILITY WARS

I posit civility as not only a confidence trick played on the U.S. public for decades, but also as referencing the general confidence in civility as a universal solution for many contemporary political ills. This use of civility usually appears as a default strategy and non-legal remedy for leveling hierarchies and inequalities.<sup>22</sup> In this context, interpreting the above polls

---

<sup>17</sup> Vogel & Ardoin, *supra* note 15, at 67.

<sup>18</sup> KENJI YOSHINO, COVERING: THE HIDDEN ASSAULT ON OUR CIVIL RIGHTS 18 (2007).

<sup>19</sup> Nick O’Malley, *Donald Trump, Brexit, and the Shy Voter Theory*, SYDNEY MORNING HERALD (Oct. 26, 2016, 8:57 AM), <https://www.smh.com.au/world/north-america/donald-trump-brexit-and-the-shy-voter-theory-20161026-gsao0a.html>.

<sup>20</sup> Susan Page, *Call Her Madame President*, USA TODAY, Oct. 10, 2005, at C1.

<sup>21</sup> Kathy Frankovic, *Race, Gender and Bias in the Electorate*, CBS NEWS (Mar. 17, 2008, 3:00 PM), <https://www.cbsnews.com/news/race-gender-and-bias-in-the-electorate/>.

<sup>22</sup> LYNN MIE ITAGAKI, CIVIL RACISM: THE 1992 LOS ANGELES REBELLION AND THE CRISIS OF RACIAL BURNOUT 19–22 (2016).

assumes that voters' frustrations with decreasing civility might change how they are voting or will vote.

Civility has multiple purposes evoked by its range of definitions—in short: citizenry, civil society, and civilization.<sup>23</sup> Political scientist Virginia Sapiro notes that “civility has been traditionally defined as civic virtues, community-building, and deliberative practices of good citizens.”<sup>24</sup> Civility is often considered to be the foundation of deliberative democracies.<sup>25</sup> It is thus “a concept, political goal, and measure.”<sup>26</sup> Civility is a strategy and may also be an obstacle to maintaining a vibrant democracy as it can be a tool of exclusion used against those who are not deemed “civil.”<sup>27</sup>

All these definitions of civility are at play or are in crisis in the all-capitals topline of the GU Politics Civility Poll press release: “VOTERS FIND POLITICAL DIVISIONS SO BAD, BELIEVE U.S. IS TWO-THIRDS OF THE WAY TO ‘EDGE OF A CIVIL WAR.’”<sup>28</sup> While polarization deepened during the Obama and Trump Administrations, the concern of an imminent civil war alludes to the cultural and historical backdrop of another era, of another allegedly impending race war. Following the 1989 fall of the Berlin Wall and 1991 collapse of the Soviet Union, “political upheavals and civil wars . . . instigated an anxious outpouring of U.S. research on civil society and its contribution to the durability and integrity of democratic states.”<sup>29</sup> The post-cold war era saw that “new social and political forces were at play, both domestically and internationally” and largely superseded what legal scholar Derrick Bell and historian Mary Dudziak have argued is “the ‘interest convergence’ of cold war politics and the modern civil rights movement [that] contributed to more democratic racial practices.”<sup>30</sup> And although capitalism had ostensibly triumphed over communism,<sup>31</sup> the 1992 violence in Los Angeles “punctured the illusion of [U.S.] exceptionalism”<sup>32</sup> and its invincibility from civil unrest—in this case, over police brutality and racial inequalities.

Paleoconservative Pat Buchanan referenced the 1992 Los Angeles “riots” as he concluded his serious challenge to the Republican presidential

---

<sup>23</sup> *Id.* at 15.

<sup>24</sup> *Id.* at 28–29.

<sup>25</sup> Sapiro, *supra* note 8, at 3.

<sup>26</sup> ITAGAKI, *supra* note 22, at xv.

<sup>27</sup> *Id.* at 19–20.

<sup>28</sup> GU Politics, *supra* note 4.

<sup>29</sup> ITAGAKI, *supra* note 22, at 14.

<sup>30</sup> ITAGAKI, *supra* note 22, at 14. See generally Derrick A. Bell, Jr., Comment, *Brown v. Board of Education and the Interest-Convergence Dilemma*, 93 HARV. L. REV. 518, 518–33 (1980) (outlining the scholar's interest convergence argument). See also MARY L. DUDZIAK, *COLD WAR CIVIL RIGHTS: RACE AND THE IMAGE OF AMERICAN DEMOCRACY* 14, 258 (2000) (supporting Bell's interest convergence argument).

<sup>31</sup> *Id.*

<sup>32</sup> *Id.* at 15.

incumbent by conceding to President George H.W. Bush at the 1992 Republican National Convention in his famous “Culture War” speech:

My friends, this election is about more than who gets what. It is about who we are. It is about what we believe, and what we stand for as Americans. There is a religious war going on in this country. It is a cultural war, as critical to the kind of nation we shall be as was the Cold War itself, for this war is for the soul of America.<sup>33</sup>

Buchanan arrayed both Bill and Hillary Clinton against Bush, who “is on our side.”<sup>34</sup> Addressing his primary voters who posed a significant challenge within the Republican Party to then-President Bush’s re-election campaign, Buchanan both describes and directs them in martial language: “And so, to the Buchanan Brigades out there, we have to come home and stand beside George Bush.”<sup>35</sup> Emphasizing that this culture war is a physical one, and one “for the soul of America,” Buchanan closes his speech with the martial image of teenaged National Guard troops who have returned from the 1991 Persian Gulf War to fight against “a mob” during the 1992 Los Angeles “riots”: “And as those boys took back the streets of Los Angeles, block by block, my friends, we must take back our cities, and take back our culture, and take back our country.”<sup>36</sup>

Civility was fueled by the master narrative of American progress, the exceptionalist *Bildungsroman* of the maturing nation adopting increasingly robust notions of democracy and equality. By calling that dominant perception of the nation into question, the 1992 Los Angeles Rebellion had a profound, though seldom acknowledged, influence on the way Americans thought of their country in the post-cold war, post-civil rights eras. For example, political scientist Robert Putnam’s popular essay and bestselling book of the same title, *Bowling Alone*, warned of a civic culture in this country declining at the very moment when newly democratic governments in Eastern Europe and elsewhere were looking to the United States as a model:

Many students of the new democracies that have emerged over the past decade and a half have emphasized the importance of a strong and active civil society to the consolidation of democracy. Especially with regard to the postcommunist countries, scholars and democratic activists alike have lamented the absence or obliteration of traditions

---

<sup>33</sup> Patrick Joseph Buchanan, Culture War Speech: Address to the Republican National Convention (Aug. 17, 1992), <https://voicesofdemocracy.umd.edu/buchanan-culture-war-speech-speech-text/>.

<sup>34</sup> *Id.*

<sup>35</sup> *Id.*

<sup>36</sup> *Id.*

of independent civic engagement and a widespread tendency toward passive reliance on the state. To those concerned with the weakness of civil societies in the developing or postcommunist world, the advanced Western democracies and above all the United States have typically been taken as models to be emulated. There is striking evidence, however, that the vibrancy of American civil society has notably declined over the past several decades.<sup>37</sup>

In the mid-1990s, the civility debates reached their peak in public discourse among academics, politicians, journalists, and pundits; these “unresolved differences of civic values and virtues have been blamed for fracturing U.S. society, resulting in a so-called civility crisis.”<sup>38</sup> The civility crisis had its roots in the legacy of affirmative action, the promises of President Johnson’s Great Society policies, and the sociopolitical tumult of the 1960s anti-war, anti-poverty, decolonial, lesbian and gay, feminist, and anti-racist movements.<sup>39</sup> Historian Daryl Michael Scott identifies a racial neoconservative perspective that perceived liberals as:

too apologetic for what they viewed as the riotous behavior of urban blacks, and [neoconservatives] emphasized the need for law and order. More important here, [neoconservatives] also tended to have serious reservations about preferential programs such as affirmative action and efforts to promote integration such as school busing. They reasserted the traditional racial liberal call for a color-blind state, which would protect only the civil rights of individuals.<sup>40</sup>

Scott notes the perception of uncivil behavior, “riotous behavior,” and even the “serious reservations” over supporting equal opportunity as outcomes rather than as processes.<sup>41</sup> The title of his study, *Contempt and Pity*, taken from W.E.B. Du Bois’s *The Souls of Black Folk* in its trenchant explanation of anti-Black attitudes, identifies how the prevalent notion of “pity” was Whites’ implicit reason for changing public sentiment during major civil rights landmarks, such as *Brown v. Board*, Great Society

---

<sup>37</sup> Robert D. Putnam, *Bowling Alone: America’s Declining Social Capital*, 6 J. DEMOCRACY 65, 65 (1995); see also ROBERT D. PUTNAM, *BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY* (2000).

<sup>38</sup> ITAGAKI, *supra* note 22, at 16.

<sup>39</sup> *Id.* at 9.

<sup>40</sup> DARYL MICHAEL SCOTT, *CONTEMPT AND PITY: SOCIAL POLICY AND THE IMAGE OF THE DAMAGED BLACK PSYCHE, 1880–1996* xiv (1997).

<sup>41</sup> *Id.*

reforms, and civil rights acts. However, this feeling masked contempt for Black people, in what I posit was the civil face of hateful, racist feelings.<sup>42</sup>

Inequality will be more indirectly channeled through ostensibly race-neutral governmental and institutional practices as historian Carol Anderson documents.<sup>43</sup> The incivility of “white rage” is now diffused through structural discrimination, indirect rather than direct, hidden instead of overt:

White rage is not about visible violence, but rather it works its way through the courts, the legislatures, and a range of government bureaucracies. It wreaks havoc subtly, almost imperceptibly. Too imperceptibly, certainly, for a nation consistently drawn to the spectacular—to what it can *see*. It’s not the Klan. White rage doesn’t have to wear sheets, burn crosses, or take to the streets. Working the halls of power, it can achieve its ends far more effectively, far more destructively.<sup>44</sup>

Anderson identifies the political logics of process and procedure that perpetuates inequalities more efficiently and pervasively. Critical ethnic studies scholars Jodi Byrd et al. echo Anderson’s arguments in relation to economic inequality, exposing the presumed “apolitical” nature of bureaucratic, procedural, and administrative acts to facilitate “dispossession”; and this façade of being apolitical flourishes despite these acts’ origins and maintenance “as the result of law and policy.”<sup>45</sup> Moreover, non-enforcement of constitutional rights and legislation and the selective protections available to some but not all persons create further inequities. The failure of constitutional amendments such as the Fifth, Fourteenth, and Fifteenth exemplifies the failure of enforcement since the nation’s founding alongside African chattel slavery, Indigenous genocide, and broken treaties. Civil rights acts of the 1960s have helped to close employment, educational, and voting gaps, among other things; however, it was the backlash against the laws that served to curtail their effectiveness and had a chilling effect on consistent enforcement. The racist destruction wrought by “halls of power” that Anderson references are not only political institutions, but also “still suffuse our social and economic system,

---

<sup>42</sup> *Id.* at xi–xii, xviii. See also W.E.B. DU BOIS, *Of Our Spiritual Strivings*, in *THE SOULS OF BLACK FOLK* 37, 38 (David W. Blight & Robert Gooding-Williams eds., Bedford Books 1997) (1903) (“It is a peculiar sensation, this double-consciousness, this sense of always looking at one’s self through the eyes of others, of measuring one’s soul by the tape of a world that looks on in amused contempt and pity.” (footnote omitted)).

<sup>43</sup> CAROL ANDERSON, *WHITE RAGE: THE UNSPOKEN TRUTH OF OUR RACIAL DIVIDE* 3–4 (2016).

<sup>44</sup> *Id.* at 3.

<sup>45</sup> Jodi A. Byrd et al., *Predatory Value: Economies of Dispossession and Disturbed Relationalities*, *SOC. TEXT*, June 2018, at 10.



buttressed by the informal modes of social control even more powerful than the law,” as historian William Chafe diagnoses the ultimately limited effects of the civil rights acts of the 1960s.<sup>46</sup> These informalities are often just as pernicious: “Although the means of keeping blacks in their place may now be implicit rather than explicit, they too often are just as effective as in the past.”<sup>47</sup>

Philosopher Jonathan Schonsheck identifies a typology that helpfully disambiguates incivilities among individuals (rudeness) and groups (rasp), as well as the basic rejection of what Schonsheck posits are the “metavalues” (repudiation): toleration and mutual respect, the foundation of a democracy.<sup>48</sup>

*Rasp* is the friction of jostling political, moral, religious, and ethnic groups that is inevitable in any multicultural “liberal democracy”—a system, or theory, or philosophy of government that cherishes the values of toleration and mutual respect. Not everyone, however, subscribes to toleration and mutual respect; the *repudiation* of these values generates the third, and most serious, category of incivility.<sup>49</sup>

Civility is also an appealing paradigm because of its admonitions of individual behavior rather than institutional changes that shape behavior through legal mandates such as affirmative action or antidiscrimination statutes. Civility can also normalize the violence of “repudiation”: the “friction” and “jostling” of *intolerance* and *disrespect*. In their “Eleven Theses on Civility,” performance scholars Kyla Wanzana Tompkins and Tavia Nyong’o expose how this repudiation is hidden: “Civility discourse enforces a false equation between incivility and violence that works to mask everyday violence as a civic norm.”<sup>50</sup> The problems with promoting civility as a panacea for political impasses are made more visible through an often unacknowledged subset of incivility—bullying—its normalization, and recent critiques of the research on this phenomenon.

In *Bully Nation: How the American Establishment Created a Bullying Society*, Charles Derber and Yale Magrass begin their wide-ranging study of the economic and militaristic institutions that implicitly and explicitly

---

<sup>46</sup> WILLIAM H. CHAFE, *CIVILITIES AND CIVIL RIGHTS: GREENSBORO, NORTH CAROLINA, AND THE BLACK STRUGGLE FOR FREEDOM*, at vii (1981).

<sup>47</sup> *Id.*

<sup>48</sup> Jonathan Schonsheck, *Rudeness, Rasp, and Repudiation*, in *CIVILITY AND ITS DISCONTENTS: ESSAYS ON CIVIC VIRTUE, TOLERATION, AND CULTURAL FRAGMENTATION* 169, 169–72 (Christine T. Sistare ed., 2004).

<sup>49</sup> *Id.* at 169.

<sup>50</sup> Tavia Nyong’o & Kyla Wanzana Tompkins, *Eleven Theses on Civility*, *SOCIAL TEXT* (July 11, 2018), <https://socialtextjournal.org/eleven-theses-on-civility/>.

promote bullying behavior among Americans.<sup>51</sup> Derber and Magrass provide critiques of anti-bullying programs in their emphasis on what they call the “psychiatric paradigm.”<sup>52</sup> I am not only interested in Derber and Magrass’s research on bullying for its relationship to civility but also as a means of making civility’s shortcomings also visible. Derber and Magrass are critical of how the psychiatric paradigm gives rise to the following popular beliefs: bullies are often misjudged as individuals acting alone and who are spontaneously maladjusted or troubled rather than organized and strategic in their harassment; the social ill of bullying is relegated to the realm of children, juvenile behavior, and immature development; bullying is not what adults do nor what venerable state apparatuses such as “the economy, military, culture, family, schools” facilitate which “all help create and perpetuate the bullying milieu.”<sup>53</sup> Incivility is similarly simplified as performed by individuals who are bad or criminal rather than strategic in deploying incivility, and incivility is generally condemned rather than examined for what its manifestation might expose about the issue or institution affected.

### III. INCIVILITIES, INEQUALITIES, INDIGNITIES, INHUMANITIES

I have more fully developed the concept of “civil racism” elsewhere in the context of the mid-1990s “civility debates” which promoted individual behaviors as leading to more effective community building, or, on the extreme other hand, the death of nations.<sup>54</sup> Civil racism acknowledges the growing field of civility studies that assesses the development or decline of our civil society in terms of social capital.<sup>55</sup> I argue that *civil racism* marks the cultural and legal shift from the modern civil rights movement to the post-civil rights era; “the goal of equality is portrayed as an ethical choice between the formal (legal) processes and empirical evaluations of the[ir] outcomes,” so equal process versus equal outcomes.<sup>56</sup> As a marker that distinguishes the post-civil rights era from earlier times, racism’s effects are “minimized” both in terms of its pervasive harm and its institutional purview: “Civil racism marks the ways in which racial discrimination has been allowed sanctuary in the private realms of individual, isolated behaviors.”<sup>57</sup> Moreover, “discussions of remedies for racial discrimination

---

<sup>51</sup> CHARLES DERBER & YALE R. MAGRASS, *BULLY NATION: HOW THE AMERICAN ESTABLISHMENT CREATED A BULLYING SOCIETY* 2 (2012).

<sup>52</sup> *Id.*

<sup>53</sup> *Id.*

<sup>54</sup> See ITAGAKI, *supra* note 22, at 15–16 (discussing the benefits of civility in dialogue and social behavior in deterring social anomie).

<sup>55</sup> *Id.* at 6.

<sup>56</sup> *Id.*

<sup>57</sup> *Id.* For “minimization of racism” as a “central frame of color-blind racism,” see EDUARDO BONILLA-SILVA, *WHITE SUPREMACY AND RACISM IN THE POST-CIVIL RIGHTS ERA* 142 (2001).

are marked as impolite and uncivil.”<sup>58</sup> Civil racism emphasizes the racial character of civility: the hierarchy depends on what is considered appropriate racial behavior in life, at work, and in leisure.<sup>59</sup> Given civil racism, one’s entitlement to rights and protections are based on one’s perceived good or bad behaviors rather than *inalienable* or *natural* as inherent to one’s humanity.<sup>60</sup>

I define “civil racism” as the preservation of civility at the expense of racial equality.<sup>61</sup> As an expression of structural racism, civil racism exhibits the active—though often unintentional—perpetuation of discrimination through one’s everyday engagement with the state and society.<sup>62</sup> Thus, I am skeptical of the investment of time and resources in civility. I posit calls for civility as a heuristic that can often expose an individual’s or group’s investment in thwarting more equitable processes and outcomes. Oppressed individuals and groups call for full equality, dignity, and humanity—rarely civility.

However, civility might entice some resistant stakeholders, possibly racial conservatives, centrists, and moderates into conversation with others holding different views. This rationale is often promoted, and marginalized speakers are often exhorted to behave civilly. Philosopher Hannah Arendt has famously commented on the need for one’s civil rights to be recognized before one has human rights, or the paradoxical right to have rights.<sup>63</sup> We can read this in a variety of ways. If one doesn’t have equality, then there is no true civility (mutual respect or cooperation). In cultivating or exercising citizenship, civility only functions between and among equals. Otherwise it is condescension for the powerful and submission for the less powerful. Being civil has been used as a gatekeeper to equality: if one isn’t civil then one doesn’t *deserve* equality.

My concern with a racist civility that is pursued at the personal or even institutional level was that it did not cost much if anything at all. In other words, there would be no material redistribution or even worse, a redistribution that flowed upward, just a friendlier face or personable interactions whatever one’s racist beliefs or society’s racist inequities. The emotional labor of civility was more often extracted from vulnerable populations who had to repeatedly overlook microaggressions or be forced

---

<sup>58</sup> *Id.*

<sup>59</sup> *See id.* at 24 (explaining how invisible forms of discrimination influence the structure of civil society).

<sup>60</sup> *See id.* at 21 (comparing proponents of civility and how they inadequately address existing social hierarchies).

<sup>61</sup> *See id.* at 6 (providing various definitions of civil racism).

<sup>62</sup> *Id.*

<sup>63</sup> *See* HANNAH ARENDT, *THE ORIGINS OF TOTALITARIANISM* 296–97 (1976) (delineating the right to have rights and the right to belong to an organized community).

to “cover.”<sup>64</sup> Kenji Yoshino interprets mainstream society, and indeed a number of court rulings, as failing to protect a mutable behavior that is not a protected civil right such as wearing no makeup, public affection between same-sex couples, or speaking Spanish at work when one is bilingual and hired for this fluency but the owners are not.<sup>65</sup> And given the three cases just taken up by the U.S. Supreme Court regarding employment discrimination against LGBTQ+ workers, these unprotected mutable behaviors in a number of states would also apply to gender performances not corresponding to the sex on one’s birth certificate—as in the case of Aimee Stephens, who lived and dressed as a woman—or in actions such as joining a gay softball league or loving someone of the same gender and disclosing that identity or relationship in a conversation with a work client.<sup>66</sup>

Returning to Arendt’s important description of cause-and-effect,<sup>67</sup> if one is not seen as having civil rights, then one does not have human rights.<sup>68</sup> If one is not perceived as being equal, then one can never be perceived as being civil. If one is not valued, then one can never be recognized or have a full self-expression that is respected and heard.<sup>69</sup> This tension between civility and equality appears in two Supreme Court cases in attempts to limit the extent to whom civility applies.

In *Robles v. Domino’s Pizza LLC*, Guillermo Robles has a visual disability and must use a screen reader for his computer.<sup>70</sup> However, he cannot engage with some parts of Domino’s Pizza’s online interface since screen readers cannot translate all of its website.<sup>71</sup> The case questioned whether the accessibility of public spaces under the Americans with Disabilities Act applied to commercial websites.<sup>72</sup> To extrapolate this, one cannot fully participate in online exchanges, whether economic, social, or political in a cyberspace that is inaccessible because websites and platforms cannot be translated by a screen reader.<sup>73</sup> This structural inequality devalues Robles and others with disabilities by implying that Domino’s does not consider this group of customers worth communicating with despite the relatively low cost of revamping its website to be fully

---

<sup>64</sup> YOSHINO, *supra* note 18, at 90, 130–31, 191.

<sup>65</sup> *Id.*

<sup>66</sup> Nina Totenberg, *Showdown Over LGBTQ Employment Rights Hits Supreme Court*, NPR (Oct. 8, 2019, 5:00 AM ET), <https://www.npr.org/2019/10/08/766386729/showdown-over-lgbtq-employment-rights-hits-supreme-court>.

<sup>67</sup> ARENDT, *supra* note 63, at 3, 131, 156.

<sup>68</sup> *See supra* note 63 and accompanying text.

<sup>69</sup> *Id.*

<sup>70</sup> *Robles v. Domino’s Pizza LLC*, 913 F.3d 898, 902 (9th Cir. 2019).

<sup>71</sup> *Id.*

<sup>72</sup> *Id.* at 904.

<sup>73</sup> Reply Brief for Plaintiff-Appellant at 16–17, 23, *Robles v. Domino’s Pizza LLC*, 913 F.3d 898 (9th Cir. 2019) (No. 17-55504).

accessible—tens of thousands of dollars for a multi-billion dollar company.<sup>74</sup>

In *Equal Employment Opportunity Commission v. R.G. & G.R. Harris Funeral Homes*, Harris Funeral Homes was probably quite civil when it fired Aimee Stephens after she explained she would live and work as a transgender woman although she had been hired initially as a cisgender man.<sup>75</sup> In fact, the civility and previous rapport between her and her employer, and the subsequent feelings of her employer betraying their civil relationship, was what encouraged her to pursue her employment discrimination case in the first place: “I’d given quite a few years to them, I had good reviews, we got along good—then all of a sudden it’s ‘We don’t need you anymore.’ I got mad enough to do something about it.”<sup>76</sup> At trial, the defense for Harris Funeral Homes referred to Stephens’s dead personal pronouns of “he” and “him.”<sup>77</sup> On appeal, the defense insisted that the owner would have fired a woman for violating the company’s dress code if she had worn pants instead of a skirt: one of the requirements of persuading the justices that the sex discrimination clause of Section VII of the 1964 Civil Rights Act does not apply.<sup>78</sup>

Trans+ activist and lawyer Richard Juang writes on the importance of recognition for trans+ individuals: “Being recognized within a liberal democracy means being valued, having one’s dignity protected, and possessing some access to public self-expression. The struggle for recognition’s key components—value, dignity, and self-expression—is a cornerstone of modern U.S. political, social, and cultural activity.”<sup>79</sup> Connecting Juang’s insights on trans+ rights to racial ones, “civil racism” that preserves civility over racial equality can also preclude dignity. The preservation of civility is also at the expense of human dignity. Given

<sup>74</sup> Petition for a Writ of Certiorari at 28, *Robles v. Domino’s Pizza LLC*, 913 F.3d 898 (9th Cir. 2019) (No. 17-55504); Kate Cox, *Accessibility, the Future, and Why Domino’s Matters*, ARS TECHNICA (Oct. 20, 2019, 10:15 AM), <https://arstechnica.com/tech-policy/2019/10/accessibility-the-future-and-why-dominos-matters/>.

<sup>75</sup> *EEOC v. R.G. & G.R. Harris Funeral Homes*, 884 F.3d 560, 568–69 (6th Cir. 2018); Totenberg, *supra* note 66.

<sup>76</sup> Katelyn Burns, *The Supreme Court Is Finally Taking on Trans Rights. Here’s the Woman Who Started It All*, VOX (Oct. 7, 2019, 4:40 PM), <https://www.vox.com/latest-news/2019/10/7/20903503/trans-supreme-court-decision-employment-discrimination-aimee-stephens>.

<sup>77</sup> See Defendant R.G. & G.R. Harris Funeral Homes, Inc.’s Motion to Dismiss at 11, *EEOC v. R.G. & G.R. Harris Funeral Homes*, 201 F. Supp. 3d 837 (E.D. Mich. 2016) (No. 2:14-cv-13710), 2014 WL 6805999 (“First, Anthony is not, in fact, a woman. As discussed above, *he* is biologically, anatomically and legally, a man. *He* may assert -- against all objective evidence -- that *he* is a woman, but there is no medical or legal authority that would support *him* in that assertion. *He* may intend to undergo therapy and surgery that might to some extent change *his* physical appearance to resemble a female. But doing so would not make *him* a female and, in any event, *he* has not done so yet.” (emphasis added)).

<sup>78</sup> *R.G. & G.R. Harris Funeral Homes*, 884 F.3d at 567.

<sup>79</sup> Richard M. Juang, *Transgendering the Politics of Recognition*, in *THE TRANSGENDER STUDIES READER* 706, 706 (Susan Stryker & Stephen Whittle eds., 2006).

recent frustrations with the limited universality and applicability of human rights and even equality, secular political theories have renewed interest in the concept of dignity. Through Black freedom struggles as Vincent Lloyd posits, dignity, in this case Black dignity, is achieved through collective struggle against White supremacy.<sup>80</sup> Thinking civility, humanity, and dignity together reveals that dignity can be present or absent in contesting or conforming, respectively, to multiple vectors of oppression (racist, ableist, gender, cis-heteronormative, economic, religious, colonialist) in what Patricia Hill Collins constructs as the “matrix of domination.”<sup>81</sup> In conceptualizing dignity, there is dignity through struggle against oppression, or where there is no struggle, there is no dignity. The presence or absence of dignity through struggle is unlike the hierarchies of human, in which even what constitutes the human is questioned from new materialist critiques, and has been historically hierarchized into what philosopher Sylvia Wynter terms “genres of being human,”<sup>82</sup> or value which also can be graduated or measured into more or less than.

#### CONCLUSION

Looking at violence—whether racist, gendered, queer/transphobic, or ableist—in relation to civility transforms notions of justice and fairness and compels us to reevaluate the responsibilities of the individual and the state to protect rights and foster democracy. Examining perceptions of civility in relation to protest often exposes the limits of inclusion and participation promised by the Constitution. Analyzing civility can change how we might try to align our behaviors with our understanding of democratic practices. The types of occasion in which we promote civility reveal the ways in which we think society and the state function best, as well as how individuals and institutions encourage or coerce us to facilitate or thwart structures that maintain civil society.

Civility is useful, but in a vastly limited sense that is little acknowledged by its proponents. If acts or people are called out for their incivility, then the incivilities reference—politely, of course—both the ugly acts of white supremacy, ableism, misogyny, or compulsory heterosexuality and the protest against them. Or the term incivility is used to weaken the claims of those accused of being uncivil. Civility can signal the demand to quell or ignore protest. Civility can also signal the obfuscation of an intent to silence or exclude. Thus, as popular terms in

---

<sup>80</sup> Vincent Lloyd, *Black Dignity*, 68 *CROSSCURRENTS* 73, 76, 78, 89–90 (2018).

<sup>81</sup> PATRICIA HILL COLLINS, *BLACK FEMINIST THOUGHT: KNOWLEDGE, CONSCIOUSNESS, AND THE POLITICS OF EMPOWERMENT* 1, 299 (2d ed. 2000).

<sup>82</sup> Sylvia Wynter, *Unsettling the Coloniality of Being/Power/Truth/Freedom: Towards the Human, After Man, Its Overrepresentation—An Argument*, 3 *CR: NEW CENTENNIAL REV.* 257, 269, 316–17 (2003).

wide use, civility and incivility are mere codewords, tools, or instruments that mask a political agenda separate from movements towards full citizenship, equality, dignity, and humanity.

# CONNECTICUT LAW REVIEW

---

VOLUME 52

FEBRUARY 2021

NUMBER 3

---

## Note

### The Tragedy of Central European University: Theorizing Hungarian Illiberal Democracy and Its Threat to Academic Freedom

JESSICA M. ZACCAGNINO

*The global proliferation of radical right political movements and the decline of democracy are defining features of our current moment. Authoritarian leaders ascend to power through the ballot box, but at once, they systematically consolidate control over the state and civil society. Hungarian Prime Minister Viktor Orbán and the Fidesz party is emblematic of illiberal democracy, a term originally coined by Fareed Zakaria. This Note applies Zakaria's illiberal democracy to Hungary while adjusting the contours of his theory to better account for the role of anti-intellectualism and nationalism in the illiberal toolkit. This Note also investigates the Orbán government's targeting of Central European University, one of the most notorious struggles between a university and an illiberal democracy for academic freedom. Central European University's situation illuminates the ways in which illiberal régimes attempt to smother spaces of resistance, using ethno-nationalist rhetoric to characterize universities and intellectuals as outsider threats to the illiberal nation-state.*



## NOTE CONTENTS

INTRODUCTION.....	1189
I. THEORETICAL UNDERPINNINGS: THE THIRD REVERSE WAVE, ILLIBERAL DEMOCRACY, AND THE NATION-STATE .....	1191
A.    DEMOCRACY’S THIRD REVERSE WAVE .....	1191
B.    THE RISE OF ILLIBERAL DEMOCRACY .....	1196
C.    THE ILLIBERAL TOOLKIT: NATIONALISM, ANTI-INTELLECTUALISM, AND OTHER ILLIBERAL VALUES .....	1201
II. THE HUNGARIAN SITUATION.....	1209
A.    WHAT IS CENTRAL EUROPE? .....	1209
B.    HUNGARIAN POLITICAL HISTORY (HAPSBURG—1989) .....	1212
C.    FLEDGLING DEMOCRACY: THE INTERIM YEARS (1989–2010).....	1213
D.    RISE OF HUNGARIAN ILLIBERAL DEMOCRACY (2010–PRESENT) ..	1218
E.    EMERGENCY AND TEMPORAL UNCERTAINTY IN HUNGARIAN ILLIBERAL DEMOCRACY .....	1227
III. THE INTERSECTION OF ANTI-INTELLECTUALISM, ILLIBERAL DEMOCRACY, AND ACADEMIC FREEDOM .....	1231
A.    LEX CEU: A BACKGROUND .....	1231
B.    ANTI-INTELLECTUALISM AND ILLIBERAL DEMOCRACY: A HUNGARIAN REPRISE.....	1233
C.    ILLIBERAL LEGAL CHALLENGES TO ACADEMIC FREEDOM.....	1235
CONCLUSION .....	1240



# The Tragedy of Central European University: Theorizing Hungarian Illiberal Democracy and Its Threat to Academic Freedom

JESSICA M. ZACCAGNINO \*

Every age has its own Fascism, and we see the warning signs wherever the concentration of power denies citizens the possibility and the means of expressing and acting on their own free will.

*Ogni tempo ha il suo fascismo: se ne notano i segni premonitori dovunque la concentrazione di potere nega al cittadino la possibilità e la capacità di esprimere ed attuare la sua volontà.*

—Primo Levi<sup>1</sup>

## INTRODUCTION

Far-right political parties have grown their influence globally by winning major, divisive elections in both recently democratized states and states that have long been considered cornerstones of liberal democracy.<sup>2</sup> Many of these parties emerged as a reaction to international unrest, including sweeping recessions and refugee crises. Guided by nationalism and populism, the far-right has shaken democracy to its core. Political parties such as Hungary's Fidesz and Jobbik, Poland's Law and Justice,

---

\* University of Connecticut School of Law, LL.M. Candidate, Human Rights & Social Justice; University of Connecticut School of Law, J.D. 2020; Bard College, B.A. 2017. I am incredibly grateful to my advisor, Professor Kiel Brennan-Marquez, without whom I could not have properly theorized this Note. I would also like to thank Professor Michael Fischl for his invaluable advice throughout the drafting process. I am forever indebted to the editors of the *Connecticut Law Review*, and like to thank Jillian Chambers, Hannah Kalichman, Adam Kuegler, Alexandria Madjeric, Carolyn Rennie, and Mallori Thompson in particular for their meticulous work, extraordinarily helpful feedback, and friendship. Most importantly, I would like to thank my parents (Debra and Robert Zaccagnino), grandparents (Anne and Robert Blackburn), sister (Melissa Zaccagnino), and partner (James Ninia) for their never-ending support.

<sup>1</sup> PRIMO LEVI, *A Past We Thought Would Never Return*, in *THE BLACK HOLE OF AUSCHWITZ* 31, 34 (Marco Belpoliti ed., Sharon Wood trans., 2005); PRIMO LEVI, *Un passato che credevamo non dovesse ritornare più*, in *L'ASIMMETRIA E LA VITA: ARTICOLI E SAGGI 1955-1987*, at 47, 50 (Marco Belpoliti ed., 2002).

<sup>2</sup> See, e.g., *Europe and Right-Wing Nationalism: A Country-By-Country Guide*, BBC (Nov. 13, 2019), <https://www.bbc.com/news/world-europe-36130006> (surveying electoral gains made by far-right parties in Europe).

and Italy's Lega Nord have quickly risen to prominence.<sup>3</sup> Likewise, far-right heads of government—including Donald Trump in the United States of America, Viktor Orbán in Hungary, and Recep Tayyip Erdoğan in Turkey—have swiftly gained power and begun to curb safeguards to democracy, like free speech and electoral access.

In many of these cases, the current political climate can be explained by the proliferation of illiberal democracy. Although illiberal democracies may at first appear to be a functioning “democracy,” upon closer look, it quickly becomes apparent that illiberal democracies are shells of their “liberal” counterparts.<sup>4</sup> Illiberal democracies may retain some democratic political liberties, like free elections or separation of powers, but they lack strong protections of the civil liberties provided by constitutional liberalism in liberal democracies. The rise of illiberal democracy in Hungary, the focus of this Note, is perfectly illustrated by Fidesz's ongoing attempt to push Central European University into exile and other unilateral attacks on academic freedom. In order to fully understand this complex situation, one must look not only to democratic theory, but to the history of Central Europe, Hungarian nationalism, and Fidesz's manipulation of the rule of law. This Note situates the Hungarian case as a cautionary tale and attempts to comprehend how the trend of illiberal democracy can flourish in the most prosperous age for democracy to date, using academic freedom as a point of focus. Part I situates this global phenomenon within a theoretical framework of democratic wave theory and illiberal democracy. Part II applies Part I to Prime Minister Viktor Orbán's rise to power in Hungary. Finally, Part III analyzes how Hungarian illiberal democracy has impacted academic freedom in the country.

---

<sup>3</sup> See, e.g., Jens Becker, *The Rise of Right-Wing Populism in Hungary*, 13 SEER: J. FOR LAB. & SOC. AFF. E. EUR. 29, 33 (2010) (“The European elections in 2009 had seen a writing on the wall regarding the future crash of the left, with FIDESZ obtaining 56.4 per cent of the votes against MSZP's 17.4 per cent. The elections of 2010 on 11 and 25 [of] April made this definitively clear.”); Michał Słowikowski & Michał Pierzgałski, *The Party System and Voting Behavior in Poland*, in CIVIC AND UNCIVIC VALUES IN POLAND: VALUE TRANSFORMATION, EDUCATION, AND CULTURE 41, 61 (Sabrina P. Ramet et al. eds., 2019) (“After the elections of 2015, the largest party in the parliament is now Law and Justice (PiS), which won an absolute majority of seats in both houses of the Polish parliament.”); CATHERINE FIESCHI, *POPULOCRACY: THE TYRANNY OF AUTHENTICITY AND THE RISE OF POPULISM* 101 (2019) (“The year 2008 marks the beginning of populism's full ideological development in Italy: [Movimento Cinque Stelle] began to capitalize on the deep transformation of the voters through its use of the Web, and the promise of a different, transparent and authentic bottom-up movement; while the Lega began to transcend its geographical limits and move southward with the aim of conquering Berlusconi strongholds through a discourse of common sense in the face of Italy's main challenges . . .”).

<sup>4</sup> *Infra* Part I.B.

## I. THEORETICAL UNDERPINNINGS: THE THIRD REVERSE WAVE, ILLIBERAL DEMOCRACY, AND THE NATION-STATE

### A. *Democracy's Third Reverse Wave*

In 1991, Samuel P. Huntington posited that democratization occurs in the form of waves. Huntington theorized that “[a] wave of democratization is a group of transitions from nondemocratic to democratic regimes that occur within a specified period of time that significantly outnumber transitions in the opposite direction during that period of time.”<sup>5</sup> Under Huntington’s wave theory, he also identifies the trend of the reverse wave. After each wave of democratization, there is a reverse wave, under which nation-states that “had previously made the transition to democracy reverted to nondemocratic rule.”<sup>6</sup> Huntington acknowledges that it would be arbitrary to prescribe a rigid date range pinpointing each wave, but nonetheless poses an approximate era for each wave.<sup>7</sup> Huntington subsequently proposes the following structure to describe the modern situation of democracy:

First, long wave of democratization

First reverse wave

Second, short wave of democratization

Second reverse wave

Third wave of democratization.<sup>8</sup>

Huntington’s first wave of democratization was indeed long, spanning between approximately 1828 to 1926.<sup>9</sup> This first wave was influenced by the American and French revolutions that took place nearly a century prior and was defined by a substantial widening of suffrage, reduced plural voting, and the secret ballot.<sup>10</sup> Under this first wave, twenty-nine democracies emerged.<sup>11</sup> The first reverse wave arrived in 1922 with Mussolini’s (democratic) ascension to power and ended with the defeat of the Axis forces in the Second World War.<sup>12</sup> The first reverse wave was characterized by “the shift away from democracy and either the return to

<sup>5</sup> SAMUEL P. HUNTINGTON, *THE THIRD WAVE: DEMOCRATIZATION IN THE LATE TWENTIETH CENTURY* 15 (1991).

<sup>6</sup> *Id.* at 16.

<sup>7</sup> *Id.* (“It is also arbitrary to attempt to specify precisely the dates of democratization waves and reverse waves. It is, nonetheless, often useful to be arbitrary . . .”).

<sup>8</sup> *Id.*

<sup>9</sup> *Id.* at 16–17.

<sup>10</sup> *Id.* at 16.

<sup>11</sup> Samuel P. Huntington, *Democracy's First Wave*, 2 J. DEMOCRACY 12, 12 (1991).

<sup>12</sup> HUNTINGTON, *supra* note 5, at 17–18.

traditional forms of authoritarian rule or the introduction of new mass-based, more brutal and pervasive forms of totalitarianism,”<sup>13</sup> and tended to occur in countries that had adopted democratic forms of governance either before the First World War or during the interwar period.<sup>14</sup> Countries that experienced nondemocratic régime change “reflected the rise of communist, fascist, and militaristic ideologies.”<sup>15</sup> Even in countries where democracy remained in place, antidemocratic movements gained strength.<sup>16</sup>

Huntington’s second wave of democratization followed the Allied victory and extended until the early 1960s, where “Allied occupation promoted inauguration of democratic institutions,”<sup>17</sup> and fledgling democracies emerged during the beginning of the end of colonialism.<sup>18</sup> Contrary to Europe’s democracies reemerging under Allied occupation, “no real effort was made to introduce democratic institutions” during decolonization in Africa and South Asia.<sup>19</sup> This led to mixed results: while some new states, such as Nigeria and India, established democracies that were maintained for at least a decade, in other states, democracy was tenuous and the institutions supporting it were shaky at best.<sup>20</sup> By the late 1950s and early 1960s, the second wave of democracy had ebbed and “political development[s] and regime transitions were taking on a heavily authoritarian cast.”<sup>21</sup> Latin America experienced numerous *coups d’état*, primarily led by military régimes, which established bureaucratic authoritarianism<sup>22</sup> throughout the region.<sup>23</sup> Similar military *coups d’état*

<sup>13</sup> *Id.* at 17.

<sup>14</sup> *Id.*

<sup>15</sup> *Id.* at 18.

<sup>16</sup> *Id.* (“In France, Britain, and other countries where democratic regimes remained in place, antidemocratic movements gained strength from the alienation of the 1920s and the depression of the 1930s.”).

<sup>17</sup> *Id.*

<sup>18</sup> *Id.* at 19.

<sup>19</sup> *Id.*

<sup>20</sup> *Id.*

<sup>21</sup> *Id.*

<sup>22</sup> The term “bureaucratic authoritarianism” was first coined by Guillermo O’Donnell in 1973 to explain the novel type of military rule in Latin America.

This form of rule has been interpreted as distinctively bureaucratic because national leadership was dominated by individuals who had risen to prominence not through political careers but through bureaucratic careers in large public and private organizations. . . . Decision-making styles among these leaders were commonly technocratic. This bureaucratic, technocratic orientation was generally accompanied by intense repression, which in most of the cases reached levels unprecedented in the region.

THE OXFORD COMPANION TO POLITICS OF THE WORLD 93 (Joel Krieger ed., 2d ed. 2001). The military-led coups in Brazil (1964), Chile (1973), and Argentina (1976) are all examples of bureaucratic authoritarian military régimes. See, e.g., *Remembering Brazil’s Military Coup 50 Years*

occurred throughout nations in Asia and the Mediterranean region.<sup>24</sup> At the same time, rapid decolonization caused democratic instability in new nations: “Thirty-three other African countries that became independent between 1956 and 1970 became authoritarian with independence or very shortly after independence.”<sup>25</sup> This era of decolonization led to the largest multiplication of authoritarian régimes in history,<sup>26</sup> accompanied by a worldwide decline in democratic governments. One study estimates that one third of the thirty-two functioning democracies in the world in 1958 had become authoritarian by the mid 1970s.<sup>27</sup> Huntington argues that this reverse wave was especially notable due to the fact that some nations undergoing nondemocratic régime changes had sustained democracy for over a quarter century.<sup>28</sup>

Huntington’s third and final democratic wave began with the Portuguese Carnation Revolution of 1974 and extended through *The Third Wave*’s publication in 1991.<sup>29</sup> During this time period, approximately thirty countries in Europe, Asia, and Latin America replaced their authoritarian régimes with democracies.<sup>30</sup> The wave began in Southern European nations—Portugal, Spain, Greece—and spread to the bureaucratic authoritarian régimes in Latin America and the military dictatorships in Asia throughout the 1970s and 1980s.<sup>31</sup> By the end of the 1980s, “the democratic wave engulfed the communist world.”<sup>32</sup> After forty-five years of Soviet occupation, Hungary transitioned to a multiparty system in

---

*Later*, NACLA (Apr. 1, 2014), <https://nacla.org/news/2014/4/1/remembering-brazils-military-coup-50-years-later> (recounting Brazil’s military dictatorship); Daniel Sheehy, *An Eyewitness Account of Pinochet’s Coup 45 Years Ago*, SMITHSONIAN MAG. (Sept. 10, 2018), <https://www.smithsonianmag.com/smithsonian-institution/eyewitness-account-pinochets-coup-45-years-ago-180970241/> (discussing a scholar’s daily life under Pinochet’s régime); Uki Goñi, *The Long Shadow of Argentina’s Dictatorship*, N.Y. TIMES (Mar. 21, 2016), <https://www.nytimes.com/2016/03/23/opinion/the-long-shadow-of-argentinas-dictatorship.html> (outlining the human rights abuses committed by Argentina’s military dictatorship).

<sup>23</sup> HUNTINGTON, *supra* note 5, at 19.

<sup>24</sup> *See id.* at 19–20 (discussing democratic backsliding and authoritarian régimes in Asian and Mediterranean countries in the 1950s through the 1980s).

<sup>25</sup> *Id.* at 20.

<sup>26</sup> *Id.* at 20–21.

<sup>27</sup> *Id.* at 21. Another report found that in 1962, thirteen governments were produced via *coups d’état*, and by 1975, thirty-eight were. *Id.* In 1960, nine out of ten South American nations that were former Iberian colonies had democratically elected governments. *Id.* But by 1973, only Venezuela and Colombia were left. *Id.*

<sup>28</sup> *See id.* (“This wave of transitions away from democracy was even more striking because it involved several countries, such as Chile, Uruguay (‘the Switzerland of South America’), India, and the Philippines, that had sustained democratic régimes for a quarter century or more.”)

<sup>29</sup> *Id.* at 21–27.

<sup>30</sup> *Id.* at 21.

<sup>31</sup> *Id.* at 21–23.

<sup>32</sup> *Id.* at 23.

1988.<sup>33</sup> Likewise, Hungary's Central and Eastern European neighbors began their own democratic transitions.<sup>34</sup> The number of democratic states rose from thirty in 1973 to fifty-eight by 1990, increasing the percentage of democratic states from 24.6 percent to 45 percent.<sup>35</sup> Soviet occupation was the principal obstacle to democratization for countries east of the Iron Curtain, and once removed, the region swiftly adopted democratic governments.<sup>36</sup>

By 1990, many of the catalysts for the third wave of democracy had stalled; “[n]either the White House, the Kremlin, the Vatican, nor the European Community were in a strong position to promote democracy. . . .”<sup>37</sup> At the same time, at least two of the new third wave democracies had already shifted back towards authoritarianism.<sup>38</sup> While it is difficult to definitively predict the duration of the third wave and what conditions may give rise to the next reverse wave, Huntington draws three generalizations from prior reverse waves to aid in comprehending the possible form of the third reverse wave. First, Huntington argues that “the causes of shifts from democratic to authoritarian political systems were at least as varied as and in part overlap with the causes of shifts from authoritarianism to democracy.”<sup>39</sup> Huntington provides a useful rubric of factors that contributed to the first and second reverse waves.<sup>40</sup> The factors are as follows:

- (1) the weakness of democratic values among key elite groups and the general public;
- (2) economic crisis or collapse that intensified social conflict and enhanced the popularity of remedies that could only be imposed by authoritarian governments;
- (3) social and political polarization often produced by leftist governments attempting to introduce or appearing to introduce too many major socioeconomic reforms too quickly;
- (4) the determination of conservative middle- and upper-class groups to exclude populist and leftist movements and lower-class groups from political power;

---

<sup>33</sup> *Id.*

<sup>34</sup> *Id.* (noting how such democratic transition included the Baltic republics and Poland).

<sup>35</sup> *Id.* at 26 tbl.1.1. Note that Huntington's figures exclude nations with populations under one million. *Id.*

<sup>36</sup> *Id.* at 288–89.

<sup>37</sup> *Id.* at 289.

<sup>38</sup> *Id.* at 290.

<sup>39</sup> *Id.*

<sup>40</sup> *Id.*

- (5) the breakdown of law and order resulting from terrorism or insurgency;
- (6) intervention or conquest by a nondemocratic foreign government;
- (7) snowballing in the form of the demonstration effects of the collapse or overthrow of democratic systems in other countries.<sup>41</sup>

Second, “transitions from democracy to authoritarianism were almost always produced by those in power or close to power in the democratic system.”<sup>42</sup> The vast majority of these previous transitions away from democracy occurred as either military *coups d’état* under which democratically elected leaders were ousted or through “executive coups,”<sup>43</sup> where democratically elected heads of government concentrated power in the executive by declaring a state of emergency or instituting martial law.<sup>44</sup> Finally, in each reverse wave, “democratic systems were replaced by historically new forms of authoritarian rule.”<sup>45</sup> Under the first wave, fascism differed from prior models of authoritarianism due to “its mass base, ideology, party organization, and efforts to penetrate and control most of society.”<sup>46</sup> Likewise, bureaucratic authoritarianism can be distinguished from other forms of authoritarian military rule by its institutional character.<sup>47</sup> Therefore, the authoritarianism set to emerge under the reverse wave theory should be expected to reinvent itself.

Currently, the world is in the throes of Huntington’s third reverse wave. Democracy has statistically entered an international era of decline.<sup>48</sup> Freedom House has documented “global declines in political rights and civil liberties” from 2005 to 2018 in their annual *Freedom in the World*

<sup>41</sup> *Id.* at 290–91.

<sup>42</sup> *Id.* at 291. This claim exempts régime changes that were produced by foreign actors. *Id.*

<sup>43</sup> *Id.*

<sup>44</sup> *Id.*

<sup>45</sup> *Id.* at 292.

<sup>46</sup> *Id.*

<sup>47</sup> *Id.*

<sup>48</sup> See, e.g., FREEDOM HOUSE, FREEDOM IN THE WORLD 2019: DEMOCRACY IN RETREAT 3–5 (2019), [https://freedomhouse.org/sites/default/files/Feb2019\\_FH\\_FITW\\_2019\\_Report\\_ForWeb-compressed.pdf](https://freedomhouse.org/sites/default/files/Feb2019_FH_FITW_2019_Report_ForWeb-compressed.pdf) (“*Freedom in the World* has recorded global declines in political rights and civil liberties for an alarming 13 consecutive years, from 2005 to 2018. The global average score has declined each year, and countries with net score declines have consistently outnumbered those with net improvements.”); *A Global Report on the Decline of Democracy*, FOREIGN AFF. (Apr. 17, 2018), <https://www.foreignaffairs.com/press/2018-04-17/global-report-decline-democracy> (summarizing articles published in the May/June 2018 *Foreign Affairs* issue on democratic decline); VARIETIES DEMOCRACY INST., DEMOCRACY FOR ALL?: V-DEM ANNUAL DEMOCRACY REPORT 2018, at 6 (2018), [https://www.v-dem.net/media/filer\\_public/3f/19/3f19efc9-e25f-4356-b159-b5c0ec894115/v-dem\\_democracy\\_report\\_2018.pdf](https://www.v-dem.net/media/filer_public/3f/19/3f19efc9-e25f-4356-b159-b5c0ec894115/v-dem_democracy_report_2018.pdf) (finding that autocratization has affected 2.5 billion people, or a third of the world).



rankings.<sup>49</sup> The international rise of the far-right satisfies and builds upon Huntington's three generalizations: the factors that may lead to democratic decline, the modes of régime transition, and the shifting image of authoritarianism. First, many of the issues contributing to the international rise of the far-right reflect the factors laid out by Huntington. These factors will be employed throughout the rest of this Note to analyze the preconditions for illiberal democracy. Huntington's second generalization regarding the mode of transition is the least apt to analyze the rise of the far-right vis-à-vis illiberal democracy. Under the majority of illiberal democracies, entire far-right political parties have seized power not through *coups d'état* or executive coups, but through the democratic process. In Hungary, this reverse wave is more party-centric than executive-centric. And finally, illiberal democracy is the answer to Huntington's claim that each reverse wave brings a novel form of authoritarianism. The following section will theorize illiberal democracy and the conditions that fomented its development with a focus on Orbán's Hungary.

### B. *The Rise of Illiberal Democracy*

The term "illiberal democracy" was originally coined by Fareed Zakaria in 1997<sup>50</sup> and was then later appropriated by Prime Minister Viktor Orbán as an ideological image of Fidesz's Hungary.<sup>51</sup> The emergence of illiberal democracy in the twenty-first century is the modern form of authoritarianism necessary for the third reverse wave of democracy.<sup>52</sup>

---

<sup>49</sup> FREEDOM HOUSE, *supra* note 48, at 4.

<sup>50</sup> Fareed Zakaria, *The Rise of Illiberal Democracy*, 76 FOREIGN AFF. 22, 22 (1997).

<sup>51</sup> Csaba Tóth, *Full Text of Viktor Orbán's Speech at Băile Tuşnad (Tusnádfürdő) of 26 July 2014*, BUDAPEST BEACON (July 29, 2014), <https://budapestbeacon.com/full-text-of-viktor-orbans-speech-at-baile-tusnad-tusnadfurdo-of-26-july-2014/> ("Meaning, that Hungarian nation is not a simple sum of individuals, but a community that needs to be organized, strengthened and developed, and in this sense, the new state we are building is an illiberal state, a non-liberal state. It does not deny foundational values of liberalism, as freedom, [et cetera]. But it does not make this ideology a central element of state organization, but applies a specific, national, particular approach in its stead.")

<sup>52</sup> Illiberal democracy does have a nomenclatural weakness—its name does not on its face convey the fascistic tendencies of many illiberal leaders. And, as a result, other terminologies also attempt to conceptualize this phenomenon. For example, Gáspár Miklós Tamás developed the concept of post-fascism to describe a "cluster" of behavior:

[P]olicies, practices, routines, and ideologies that can be observed everywhere in the contemporary world; that have little or nothing to do, except in Central Europe, with the legacy of Nazism; that are not totalitarian; that are not at all revolutionary; and that are not based on violent mass movements and irrationalist, voluntaristic philosophies, nor are they toying, even in jest, with anti-capitalism.

G.M. Tamás, *On Post-Fascism*, BOS. REV. (June 1, 2000), <http://bostonreview.net/world/g-m-tamas-post-fascism>. Post-fascism in ideology bears resemblance to classical fascism because of its open hostility to universal citizenship embraced by the Enlightenment, instead believing that some classes of

Unlike its predecessors, illiberal democracy arises more subtly than traditional forms of authoritarian rule. As opposed to traditional authoritarianism<sup>53</sup>—defined by closed society,<sup>54</sup> use of force, and formal censorship—illiberal democracies feign compliance with the general principles of democracy. Many of the national parties promoting illiberal democracy gain power through legitimate democratic means.<sup>55</sup> Once elected, “they use the levers of democratic institutions to consolidate control, all while claiming popular support from the people to protect the nation from foreign or domestic threats.”<sup>56</sup> While in power, these régimes parasitically sap the strength out of the democratic institutions through which they were elected. This Section explores the theoretical contours of illiberal democracy, contrasted with liberal democracy, to illustrate the individual case of Hungary.

To properly define illiberal democracy, one must first examine the characteristics of a *liberal* democracy. Zakaria’s article and subsequent book describes a liberal democracy as “a political system marked not only by free and fair elections, but also by the rule of law, a separation of powers, and the protection of basic liberties of speech, assembly, religion,

---

people do not deserve citizenship to the nation and the civic rights associated with membership. *Id.* This is particularly important as Enlightenment citizenship was equated with human dignity and by recognizing universal citizenship we, in turn, recognize those citizens as human. *Id.* When classical fascists and the post-fascists of today reject the citizenship of classes of people within the nation-state, fascists also reject their humanity. *Id.* Under classical fascism, “civic death was necessarily followed by natural death, that is, violent death, or death *tout court.*” *Id.* Post-fascism, however, replaces literal death with figurative death in an anti-Enlightenment illiberal democracy. The sovereign simultaneously grants citizenship to *some* residents of the nation-state while also refusing the humanity of others. Also, unlike classical fascism, “[p]ost-fascism finds its niche easily in the new world of global capitalism without upsetting the dominant political forms of electoral democracy and representative government.” *Id.*

<sup>53</sup> See HANNAH ARENDT, *THE ORIGINS OF TOTALITARIANISM* 408–09 (1968) (“As techniques of government, the totalitarian devices appear simple and ingeniously effective. They assure not only an absolute power monopoly, but unparalleled certainty that all commands will always be carried out; the multiplicity of the transmission belts, the confusion of the hierarchy, secure the dictator’s complete independence of all his inferiors and make possible the swift and surprising changes in policy for which totalitarianism has become famous.”).

<sup>54</sup> See K. R. POPPER, *THE OPEN SOCIETY AND ITS ENEMIES* 49 (1945) (“It is one of the characteristic features of the magical attitude of a primitive tribal or ‘closed’ society that it lives in a charmed circle of unchanging taboos, of laws and customs which are felt to be as inevitable as the rising of the sun, or the cycle of the seasons, or similar obvious regularities of nature.”); cf. HENRI BERGSON, *THE TWO SOURCES OF MORALITY AND RELIGION* 229 (R. Ashley Audra & Cloudesley Brereton trans., 1935) (“The closed society is that whose members hold together, caring nothing for the rest of humanity, on the alert for attack or defence, bound in fact, to a perpetual readiness for battle. . . . Man was made for this society, as the ant was made for the ant-heap.”).

<sup>55</sup> ALINA POLYAKOVA ET AL., *BROOKINGS, THE ANATOMY OF ILLIBERAL STATES: ASSESSING AND RESPONDING TO DEMOCRATIC DECLINE IN TURKEY AND CENTRAL EUROPE* 2 (2019), <https://www.brookings.edu/wp-content/uploads/2019/02/illiberal-states-web.pdf>.

<sup>56</sup> *Id.*

and property.”<sup>57</sup> Liberal democracies are also open societies<sup>58</sup> that tend to value civic nationalism.<sup>59</sup> Similarly, Freedom House’s annual *Freedom in the World* report splits its criteria for freedom into two separate rankings: political rights and civil liberties.<sup>60</sup> States with the highest political rights rankings “enjoy a wide range of political rights, including free and fair elections. Candidates who are elected actually rule, political parties are competitive, the opposition plays an important role and enjoys real power, and the interests of minority groups are well represented in politics and government.”<sup>61</sup> In order to attain a full forty-point ranking for political rights, states must score positively on criteria concerning the electoral process, political pluralism and participation, and the functioning of government.<sup>62</sup> In the realm of civil liberties, state treatment of the freedom of expression and belief, associational and organizational rights, rule of law, and personal autonomy and individual rights are considered criteria for a free state.<sup>63</sup> These two indicators, political liberties and civil liberties, correspond to democracy and constitutional liberalism, respectively.<sup>64</sup>

Liberal democracy can be divided into two elements: democracy and constitutional liberalism.<sup>65</sup> While democracy primarily governs political

---

<sup>57</sup> Zakaria, *supra* note 50, at 22.

<sup>58</sup> Henri Bergson first described the concept of open society in 1935 and it was later expanded upon by Karl Popper’s 1945 critique of totalitarianism, *The Open Society and Its Enemies*. As opposed to a closed society, an open society is one in which “individuals are confronted with personal decisions,” base their decisions on intelligence, and are critical of the taboos readily embraced in closed societies. POPPER, *supra* note 54, at 152, 178. Popper argues that totalitarianism is a type of “reactionary movements which have tried, and still try, to overthrow civilization and return to tribalism.” *Id.* at 1. Under totalitarianism and closed societies, critical thinking becomes impossible because these societies rely on “the suppression of reason and truth” and the “brutal and violent destruction of all that is human.” *Id.* at 177. For Popper, this is the danger of totalitarianism: a return to closed societies that threaten humanity. Open societies, therefore, must support freedom of thought and expression and protect them through the rule of law.

<sup>59</sup> Many scholars of nationalism tend to “distinguish[] ‘civic’ and ‘ethnic,’ western and eastern, liberal and illiberal forms of nationalism.” ROGERS BRUBAKER, *ETHNICITY WITHOUT GROUPS* 5 (2004). Like many topics in nationalism, the distinction between civic and ethnic nationalism is difficult to define and, according to Brubaker, “normatively problematic,” but it is nonetheless worth discussing in the context of liberal democracy. *Id.* Under civic nationalism, “nationhood and nationalism have been linked to democracy, self-determination, political legitimacy, social integration, civil religion, solidarity, dignity, identity, cultural survival, citizenship, patriotism, and liberation from alien rule.” *Id.* at 132.

<sup>60</sup> MICHAEL J. ABRAMOWITZ, *FREEDOM HOUSE, FREEDOM IN THE WORLD 2018: DEMOCRACY IN CRISIS* 2 (2018), <https://freedomhouse.org/report/freedom-world/2018/democracy-crisis>.

<sup>61</sup> *Id.*

<sup>62</sup> *Id.*

<sup>63</sup> *Id.*

<sup>64</sup> Zakaria, *supra* note 50, at 22–24.

<sup>65</sup> *Id.* at 22–23; see also POLYAKOVA ET AL., *supra* note 55, at 2 (“*Liberal principles*—political ideas that espouse the importance of individual liberties, minority rights, and the separation of power across levers of government—and *democratic institutions*—processes that translate popular will into public policy through legitimate elections . . .”).

rights, and is thus more process-oriented, constitutional liberalism's focus on civil rights is goal-oriented.<sup>66</sup> Zakaria explains that the term "constitutional liberalism" is a marriage between two interrelated concepts: "It is *liberal* because it draws on the philosophical strain, beginning with the Greeks, that emphasizes individual liberties. It is *constitutional* because it rests on the tradition, beginning with the Romans, of the rule of law."<sup>67</sup> Constitutional liberalism developed in Western Europe and the United States under thinkers including William Blackstone, Baron de Montesquieu, John Locke, John Stuart Mill, and Thomas Jefferson.<sup>68</sup> The general concept of constitutional liberalism, despite some variance, boils down to the argument "that human beings have certain natural (or 'inalienable') rights and that governments must accept a basic law, limiting its own powers, that secures them."<sup>69</sup> While the existence of constitutional liberalism in countries has led to the emergence of democracy, democracy does not necessarily give rise to constitutional liberalism.<sup>70</sup> Merely arranging free elections and protecting other political rights does not guarantee that those who democratically come to power will protect the civil liberties enshrined by constitutional liberalism.

Although democracy and constitutional liberalism are often associated as conjoined, the two are frequently in tension.<sup>71</sup> In particular, democracy and constitutional liberalism tend to conflict on the scope of government authority: "Constitutional liberalism is about the limitation of power, [and] democracy about its accumulation and use."<sup>72</sup> Democracy can undermine liberty without substantial safeguards for minority rights and liberties. John Stuart Mill warned of "the tyranny of the majority," under which the democratically-elected majority could subvert the liberties protected by constitutional liberalism.<sup>73</sup> Illiberal democracy, then, is symptomatic of this schism between democracy and constitutional liberalism.

---

<sup>66</sup> Zakaria, *supra* note 50, at 25 ("Constitutional liberalism, on the other hand, is not about the procedures for selecting government, but rather government's goals.").

<sup>67</sup> *Id.* at 26.

<sup>68</sup> *Id.* ("[Constitutional liberalism's] canonical figures include the poet John Milton, the jurist William Blackstone, statesmen such as Thomas Jefferson and James Madison, and philosophers such as Thomas Hobbes, John Locke, Adam Smith, Baron de Montesquieu, John Stewart Mill, and Isaiah Berlin.").

<sup>69</sup> *Id.*

<sup>70</sup> *Id.* at 28 ("Constitutional liberalism has led to democracy, but democracy does not seem to bring constitutional liberalism.").

<sup>71</sup> *Id.* at 30; *see also* POLYAKOVA ET AL., *supra* note 55, at 2 ("The rise of illiberal political parties and leaders within electoral democratic systems illustrates the schism between the foundational principles and institutions of liberal democracies.").

<sup>72</sup> Zakaria, *supra* note 50, at 30.

<sup>73</sup> JOHN STUART MILL, ON LIBERTY 7 (1956) ("[I]n political speculations 'the tyranny of the majority' is now generally included among the evils against which society requires to be on its guard."); *id.* at 3 ("By liberty was meant protection against the tyranny of the political rulers."); *id.* at 4 ("To prevent the weaker members of the community from being preyed upon by innumerable vultures,

Illiberal democracy, as is linguistically implied, is a form of faux-democracy whereby some legitimate democratic processes are present but without substantial safeguards of civil liberties as guaranteed by constitutional liberalism. This modern form of authoritarianism enshrines majoritarianism<sup>74</sup> and absolute sovereignty<sup>75</sup> as fundamental political values, which defy liberal democratic norms. Illiberal governments centralize authority and usurp power in a way that is “both horizontal (from other branches of the national government) and vertical (from regional and local authorities as well as private businesses and other nongovernmental groups).”<sup>76</sup> The claim that “unchecked centralization has been the enemy of liberal democracy”<sup>77</sup> is self-evident—Mussolini, for example, was a democratically elected fascist who quickly centralized power into totalitarian control.<sup>78</sup> Likewise, “[i]lliberal democracies gain legitimacy, and thus strength, from the fact that they are reasonably democratic. Conversely, the greatest danger that illiberal democracy poses—other than to its own people—is that it will discredit liberal democracy itself, casting a shadow on democratic governance.”<sup>79</sup> Put another way, illiberal democracy “is democratic because it respects the will of the majority; illiberal because it disregards the concerns of minorities.”<sup>80</sup> Emerging illiberal democracies can be identified by their *modus operandi* that enshrines values of nationalism, majoritarianism, dictatorship of law, absolute sovereignty, and anti-intellectualism.

---

it was needful that there should be an animal of prey stronger than the rest, commissioned to keep them down. . . . The aim, therefore, of patriots, was to set limits to the power which the ruler should be suffered to exercise over the community; and this limitation was what they meant by liberty.”)

<sup>74</sup> ARCH PUDDINGTON, FREEDOM HOUSE, *BREAKING DOWN DEMOCRACY: GOALS, STRATEGIES, AND METHODS OF MODERN AUTHORITARIANS* 7 (2017), [https://freedomhouse.org/sites/default/files/June2017\\_FH\\_Report\\_Breaking\\_Down\\_Democracy.pdf](https://freedomhouse.org/sites/default/files/June2017_FH_Report_Breaking_Down_Democracy.pdf)

(“A single idea of many authoritarians is the proposition that elections are winner-take-all affairs in which the victor has an absolute mandate, with little or no interference, from institutional checks and balances. . . . The Hungarian prime minister, Viktor Orbán, has instituted a thorough overhaul of the country’s constitution and national legislation with an eye toward measures that will insulate his party from future defeat.”)

<sup>75</sup> *Id.* (“A number of governments have invoked the doctrine of absolute sovereignty to rebuff international criticism of restrictions on the press, the smothering of civil society, the persecution of the political opposition, and the repression of minority groups. They claim that the enforcement of universal human rights standards or judgments from transnational legal bodies represent undue interference in their domestic affairs and a violation of national prerogatives.”)

<sup>76</sup> Zakaria, *supra* note 50, at 30.

<sup>77</sup> *Id.* at 32.

<sup>78</sup> John Foot & Christopher Hibbert, *Benito Mussolini*, *ENCYCLOPEDIA BRITANNICA*, <https://www.britannica.com/biography/Benito-Mussolini> (last updated Jan. 10, 2020).

<sup>79</sup> Zakaria, *supra* note 50, at 42.

<sup>80</sup> MADELEINE ALBRIGHT, *FASCISM: A WARNING* 172 (2018).

### C. *The Illiberal Toolkit: Nationalism, Anti-Intellectualism, and Other Illiberal Values*

#### 1. *Nationalism*

Modern nationalism guides much of the behavior exhibited by illiberal democracies and deeper exploration is necessary to fully understand its global rise. Illiberal democratic governments enshrine the needs of the community over individual rights.<sup>81</sup> These governments are wont to conceive “the community” as not citizens of the polity, but instead appeal to nationalistic conceptions of “the community” that are demarcated by bloodline.<sup>82</sup> Illiberal democracies differ from their liberal counterparts by subverting civil liberties traditionally delegated to “the people” by constitutional liberalism that is inclusive of all citizens, regardless of ethnic, religious, political, or other identities. Instead, these leaders narrowly tailor “the people” to mean those supportive of the illiberal government that belong to certain ethnic groups, while otherizing the rest.<sup>83</sup> This anti-pluralism stokes the flames of ethnic nationalism while bulldozing the civic nationalism traditionally associated with liberal democracies. Ethnic nationalism, a manifestation of nationalism in which the nation-state is defined on the basis of ethnicity,<sup>84</sup> is a core feature of illiberal democracies.

Nationalism as an ideology is a modern phenomenon and is deeply prevalent throughout contemporary societies. Nationalism is “notoriously difficult to define.”<sup>85</sup> In *Imagined Communities*, Benedict Anderson’s magnum opus, “the nation” is defined as “an imagined political community—and imagined as both inherently limited and sovereign.”<sup>86</sup> This definition builds upon Ernest Gellner’s conception of nationalism as “not the awakening of nations to self-consciousness: it invents nations where they do not exist.”<sup>87</sup> As such, the nation is a socially constructed

---

<sup>81</sup> See *id.* (“An illiberal democracy is centered on the supposed needs of the community rather than the inalienable rights of the individual. It is democratic because it respects the will of the majority; illiberal because it disregards the concerns of minorities.”).

<sup>82</sup> See *id.* (“[T]he togetherness [Orbán] envisions is defined by bloodlines, not borderlines.”).

<sup>83</sup> See *id.* (discussing how illiberal democracies disregard the needs and rights of minorities).

<sup>84</sup> See, e.g., BRUBAKER, *supra* note 59, at 132 (“[Ethnic] nationalism has been associated with militarism, war, irrationalism, chauvinism, intolerance, homogenization, forced assimilation, authoritarianism, parochialism, xenophobia, ethnocentrism, ethnic cleansing, even genocide; it has been characterized as the ‘starkest political shame of the twentieth century.’”).

<sup>85</sup> BENEDICT ANDERSON, *IMAGINED COMMUNITIES: REFLECTIONS ON THE ORIGIN AND SPREAD OF NATIONALISM* 3 (rev. ed. 2006).

<sup>86</sup> *Id.* at 6.

<sup>87</sup> ERNEST GELLNER, *THOUGHT AND CHANGE* 168 (1964). Gellner later expanded on his theories of nationalism in *NATIONS AND NATIONALISM* (1983).

community that is imagined<sup>88</sup> by those that identify as members of the group. The nation is imagined as a limited sovereign community: limited because its borders are finite,<sup>89</sup> sovereign because of the influence of Enlightenment ideals on political values,<sup>90</sup> and a community because of the deep comradeship prevalent amongst citizens of the nation-state.<sup>91</sup> Modernist theorists like Anderson contend that nationalism as an ideology was able to arise due to technological and socio-economic advances brought about by the Industrial Revolution.<sup>92</sup> Anderson singles out print-capitalism as a precursor to the development of national consciousness—a shared sense of national identity.<sup>93</sup> Print-capitalism unified local dialects into a language that members of a nation-state could all understand vis-à-vis mechanical reproduction<sup>94</sup> made possible by the printing press and the proliferation of capitalism.<sup>95</sup> Print-capitalism allowed nations to consolidate numerous vernaculars into a unified language representative of the nation, as will be seen with the Magyarization of Hungary.<sup>96</sup>

Nation-states are socially constructed imagined communities that purport to unify groups of peoples based on shared identity within the confines of their finite borders. Under this framework of nationalism,

<sup>88</sup> ANDERSON, *supra* note 85, at 6 (“It is *imagined* because the members of even the smallest nation will never know most of their fellow-members, meet them, or even hear of them, yet in the minds of each lives the image of their communion.”).

<sup>89</sup> *Id.* at 7 (“The nation is imagined as *limited* because even the largest of them, encompassing perhaps a billion living human beings, has finite, if elastic, boundaries, beyond which lie other nations.”).

<sup>90</sup> *Id.* (“It is imagined as *sovereign* because the concept was born in an age in which Enlightenment and Revolution were destroying the legitimacy of the divinely-ordained, hierarchical dynastic realm.”).

<sup>91</sup> *Id.* (“Finally, it is imagined as a *community*, because, regardless of the actual inequality and exploitation that may prevail in each, the nation is always conceived as a deep, horizontal comradeship. Ultimately it is this fraternity that makes it possible, over the past two centuries, for so many millions of people, not so much to kill, as willingly to die for such limited imaginings.”).

<sup>92</sup> *Id.* at 46.

<sup>93</sup> *Id.* at 44–45 (“These print-languages laid the bases for national consciousness in three distinct ways. First and foremost, they created unified fields of exchange and communication below Latin and above spoken vernaculars. . . . Second, print-capitalism gave a new fixity to language, which in the long run helped to build that image of antiquity so central to the subjective idea of the nation. . . . Third, print-capitalism created languages of power of a kind different from the older administrative vernaculars. Certain dialects inevitably were ‘closer’ to each print-language and dominated their final forms.”).

<sup>94</sup> See also WALTER BENJAMIN, *The Work of Art in the Age of Mechanical Reproduction*, in ILLUMINATIONS 217, 219–20 (Harry Zohn trans., 1969) (“Around 1900 technical reproduction had reached a standard that not only permitted it to reproduce all transmitted works of art and thus to cause the most profound change in their impact upon the public; it also had captured a place of its own among the artistic processes.”).

<sup>95</sup> See ANDERSON, *supra* note 85, at 37–44 (tracing the development of mass and mechanical reproduction and its effect on the consolidation of languages).

<sup>96</sup> See *infra* note 113 (discussing the history of Magyarization in nineteenth century Hungary).

“[t]he formulation of the nation thus appears as the fulfillment of a ‘project’ stretching over centuries.”<sup>97</sup> As illiberal democracies narrow the perception of who belongs in their imagined community, two questions are raised. First, “[w]hat makes the nation a ‘community’?”<sup>98</sup> And second, who are “the people” that are accepted into the community? The nation-as-community conceived of by Anderson is inherently tied to the *fraternité* expressed by those within the nation-state. Étienne Balibar<sup>99</sup> writes:

*Every social community reproduced by the functioning of institutions is imaginary, that is to say, it is based on the projection of individual existence into the weft of a collective narrative, on the recognition of a common name and on traditions lived as the trace of an immemorial past (even when they have been fabricated and inculcated in the recent past). But this comes down to accepting that, under certain conditions, only imaginary communities are real.*<sup>100</sup>

Therefore, the socially-constructed “community” is reified by the people that further the collective narratives and mythologies of the nation-state. Balibar argues that theoretically, this community of people recognizes itself as an entity or group distinct from other states *prior* to the foundation of the institutional state.<sup>101</sup> But this is clearly contradictory and impossible to actualize. Therefore, the nation cannot precede the state:

---

<sup>97</sup> ÉTIENNE BALIBAR & IMMANUEL WALLERSTEIN, RACE, NATION, CLASS: AMBIGUOUS IDENTITIES 86 (Chris Turner trans., 1991).

<sup>98</sup> *Id.* at 93.

<sup>99</sup> Étienne Balibar wrote, “Thinking about racism led us back to nationalism, and nationalism to uncertainty about the historical realities and categorization of the nation.” Étienne Balibar, *The Nation Form: History and Ideology*, 13 REV. (FERNAND BRAUDEL CTR.) 329, 329 (1990). Balibar sought to understand how the nation-state arose and found current explanations to be unsatisfactory. The state refers to the institutional apparatuses of a country, while the nation connotes some sort of identity. Balibar argues that logically, states must have predated the nation-state and could not have arose simultaneously. Therefore, he held, “it was by becoming ‘national’ that the states transformed themselves, more or less completely, into what we call the modern state.” *Id.* at 330. Balibar theorized that nation-states are created in one of three ways:

Either the states came into existence “endogenously,” seemingly autonomously, in tandem with a process of nationalizing the state that was already located in that territory, or they came into existence via “nationalist” (or “national liberation”) movements, by struggling against national states that already existed or were being created, or against “non-national” sovereign states (such as “multinational” empires, which thereby came to seem anachronistic).

*Id.* at 331.

<sup>100</sup> BALIBAR, *supra* note 97, at 93.

<sup>101</sup> *Id.*



In reality, the idea of nations without a state, or nations “before” the state, is thus a contradiction in terms, because the state always is implied in the historic framework of a national formation . . . . But this contradiction is masked by the fact that national states . . . project beneath their political existence to a preexisting “ethnic” or “popular” unity.<sup>102</sup>

As such, the nation must be a product of the state, which adopts the “nation” as its identity. By “becoming ‘national,’”<sup>103</sup> the state transforms itself into “the modern state”:<sup>104</sup> the nation-state.<sup>105</sup> In order to become “national,” the nation-state must manufacture an identity to define the confines of the community: “The fundamental problem is therefore to produce the people. More exactly, it is to make the people produce itself continually as a national community.”<sup>106</sup>

For Balibar, who was investigating the nation-state to understand “the causes and ‘deep’ structures of contemporary racism,”<sup>107</sup> that manufactured identity is “fictive ethnicity.”<sup>108</sup> An imagined community can become a nation-state only if it is made up of persons that embrace the *fraternité* that binds together the community, which does not exist naturally within any state.<sup>109</sup> A nation-state requires “the people” to be more than a mere abstraction and to share a common bond that legitimizes the “national” aspect of the nation-state.<sup>110</sup> This is instituted through fictive ethnicity, especially in the case of ethnic nationalism. By constructing a fictive ethnicity that appeals to an imagined community and convinces community members of their shared “ethnicity,” the nation-state and the

<sup>102</sup> Balibar, *supra* note 99, at 331.

<sup>103</sup> *Id.* at 330.

<sup>104</sup> *Id.*

<sup>105</sup> Balibar argues that the modern nation-state can be identified by: “its ideology and collective sovereignty; its juridical and administrative rationality; its particular mode of regulating social conflicts, especially class conflicts; and its ‘strategic’ objective of managing its territorial resources and population to enhance its economic and military power.” *Id.* at 330–31.

<sup>106</sup> BALIBAR, *supra* note 97, at 93.

<sup>107</sup> Balibar, *supra* note 99, at 329.

<sup>108</sup> BALIBAR, *supra* note 97, at 96.

<sup>109</sup> *See id.* at 93 (“A social formation only reproduces itself as a nation to the extent that, through a network of apparatuses and daily practices, the individual is instituted as *homo nationalis* from cradle to grave, at the same times he or she is instituted as *homo œconomicus, politicus, religious*. . . . [S]uch a people does not exist naturally, and even when it is tendentially constituted, it does not exist all the time. No modern nation possesses a given ‘ethnic’ basis, even when it arises out of a national independence struggle.”).

<sup>110</sup> *See id.* (“In the case of national formations, the imaginary which inscribes itself in the real in this way is that of the ‘people’. It is that of a community which recognizes itself in advance in the institution of the state, which recognizes that the state as ‘its own’ in opposition to other states and, in particular, inscribes its political struggles within the horizon of that state—by, for example, formulating its aspirations for reform and social revolution as projects for the transformation of ‘its national state’.”).

patriotism that accompanies it is reified.<sup>111</sup> In the metamorphosis of the state into the nation-state, the population within the borders of the new nation-state is “ethnicized” in a chrysalis that treats the group “as if they formed a natural community, possessing of itself an identity of origins, culture and interests which transcends individuals and social conditions.”<sup>112</sup> Ethnicity is produced through two modes—language<sup>113</sup> and race—and is established through various state institutions, including schooling.<sup>114</sup> Fictive ethnicity is weaponized by nationalistic illiberal democracies and is instrumental in understanding who are “the people” that are admitted to illiberal régimes and the compatriots that are excluded.

## 2. Majoritarianism, Dictatorship of Law, and Absolute Sovereignty

Despite its resemblance to democracy, illiberal democracies, informed by ethnic nationalism, have mobilized alternative values that challenge post-Cold War democratic norms. First, illiberal leaders tend to embrace majoritarianism, the notion that the majority of a population should be granted primacy when determining the outcome of a decision.<sup>115</sup> Majoritarianism has long been rebuked by foundational democratic

---

<sup>111</sup> *Id.* at 96 (“Fictive ethnicity is not purely and simply identical with the *ideal nation* which is the object of patriotism, but it is indispensable to it, for, without it, the nation would appear precisely only as an idea or an arbitrary abstraction; patriotism’s appeal would be addressed to no one. It is fictive identity which makes it possible for the expression of a preexisting unity to be seen in the state, and continually to measure the state against its ‘historic mission’ in the service of the nation, and as a consequence, to idealize politics.”).

<sup>112</sup> *Id.*

<sup>113</sup> In Hungary, the fictive ethnicity of the Magyar is deeply rooted in language. Hungarian is an extremely distinct language in the Finno-Ugric family whose closest relatives are Finnish and Estonian. Hungarian has no relation to the Slavic languages spoken in the nations surrounding Hungary. *Finno-Ugric Languages*, ENCYCLOPÆDIA BRITANNICA, <https://www.britannica.com/topic/Finno-Ugric-languages> (last visited Mar. 24, 2019). In the late nineteenth century, Hungary began the process of Magyarization, under which non-Magyar minorities in Hungary were forced to assimilate by adopting the Hungarian language and culture. See ANDERSON, *supra* note 85, at 101–07 (detailing the process of Magyarization). Linguistic nationalism still exists in Hungary, and can be seen as a common theme in Orbán’s speeches. See, e.g., Prime Minister Viktor Orbán, Orbán Viktor’s Ceremonial Speech on the 170th Anniversary of the Hungarian Revolution of 1848 (Mar. 15, 2018) (“They want us to hand it over to foreigners coming from other continents, who do not speak our language, and who do not respect our culture, our laws or our way of life: people who want to replace what is ours with what is theirs.”); Prime Minister Viktor Orbán, Prime Minister Viktor Orbán’s Address After Swearing the Prime-Ministerial Oath of Office (May 10, 2018) (“We are a unique species. We have a language that is unique to us. There is a world which we alone see and which we alone render through the prism of Hungarian language and culture. Without us human civilisation would certainly be deprived of a language, a view and a characterisation of the world. . . . I believe that Hungary, and the Hungarian language and culture, exert an enormous magnetic power, which will attract those Hungarians whom the wind has blown from the Carpathian Basin.”).

<sup>114</sup> BALIBAR, *supra* note 97, at 96 (“History shows us that there are two great competing routes to this: language and race.”).

<sup>115</sup> Nicholas Capaldi, *Majoritarianism*, ENCYCLOPÆDIA BRITANNICA, <https://www.britannica.com/topic/majoritarianism> (last visited Mar. 18, 2019).

political philosophers and is exemplified in Mill’s “tyranny of the majority.”<sup>116</sup> Majoritarianism is prevalent when illiberal democracies approach elections as a “winner-take-all affair[] in which the victor has an absolute mandate, with little or no interference from institutional checks and balances.”<sup>117</sup> These governments will often disregard the existence of institutional checks on their authority—particularly with the judiciary—or actually dismantle the democratic checks that were already in place, as seen with Hungary.<sup>118</sup> These actions delegitimize national legal systems and endanger pluralism both within governance and the country. Second, illiberal democracies will employ the “dictatorship of law,” originally coined by Vladimir Putin<sup>119</sup> to describe “the adoption of laws that are so vaguely written as to give the authorities wide discretion in applying them to regime opponents.”<sup>120</sup> Generally, these vague laws are paired with a weakened court system saturated with régime supporters that manipulate the legal system to carry out the régime’s political agenda.<sup>121</sup> Finally, illiberal democracies frequently invoke the doctrine of absolute sovereignty in order to insulate the state from international obligations and “criticism of restrictions on the press, the smothering of civil society, the persecution of political opposition, and the repression of minority groups.”<sup>122</sup> Sovereignty rhetoric is also deployed against international organizations, such as the United Nations or the European Union, that challenge state actions that run counter to international law and legal norms.<sup>123</sup>

### 3. *Anti-Intellectualism*

The final characteristic of illiberal democracies is anti-intellectualism. The term “anti-intellectualism” was coined in Richard Hofstadter’s

<sup>116</sup> MILL, *supra* note 73, at 7.

<sup>117</sup> PUDDINGTON, *supra* note 74, at 7.

<sup>118</sup> See *infra* Part II.D (discussing the dismantling of the Constitutional Court’s jurisdiction by the Fidesz party).

<sup>119</sup> *Russia’s Dictatorship of Law*, N.Y. TIMES (Nov. 20, 2010), <https://www.nytimes.com/2010/11/21/opinion/21sun2.html>. Putin’s “dictatorship of law” is exemplified in the case of Boris Berezovsky, a Russian oligarch who had helped finance Yeltsen’s reelection. In 2003, Berezovsky was granted political asylum by the United Kingdom and in 2013 was found dead in his home. His death remains an open investigation. Mary Dejevsky, *The Weird World of Boris Berezovsky: Alexander Litvinenko’s Inquest Has Provided an Intriguing Insight into the Dead Tycoon*, INDEPENDENT (Mar. 18, 2015), <https://www.independent.co.uk/news/world/europe/the-weird-world-of-boris-berezovsky-alexander-litvinenko-s-inquest-has-provided-an-intriguing-insight-10117927.html>.

<sup>120</sup> PUDDINGTON, *supra* note 74, at 7–8.

<sup>121</sup> *Id.* at 8 (“Such measures are typically paired with a court system that uses law merely to justify political instructions from the executive branch, making a mockery of due process and international conceptions of the rule of law.”).

<sup>122</sup> *Id.* at 7.

<sup>123</sup> *Id.*

*Anti-Intellectualism in American Life* to describe the “resentment and suspicion of the life of the mind and of those who are considered to represent it; and a disposition constantly to minimize that value of life.”<sup>124</sup> Anti-intellectuals that operate within populist political structures espouse anti-elitist and anti-rationalist<sup>125</sup> attitudes. Rational discourse is inextricably linked to values protected by constitutional liberalism, including those of free speech, assembly, and academic freedom.<sup>126</sup> Although Hofstadter’s book was written in response to McCarthyism in the United States,<sup>127</sup> his general framework of anti-intellectual thought can be transferred to the disdain for the intelligentsia expressed by authoritarians around the world.

Authoritarian governments instrumentalize anti-intellectualism to suppress political dissent by systematically removing the intelligentsia from power and public life.<sup>128</sup> Critical discourse and free thought, core components of that intellectual life, can undermine authoritarian projects by vocalizing opposition. Hannah Arendt reflects: “Intellectual, spiritual, and artistic initiative is as dangerous to totalitarianism as the gangster initiative of the mob, and both are more dangerous than mere political opposition.”<sup>129</sup> Authoritarian governments “seek[] to undermine public discourse by attacking and devaluing education, expertise, and language.”<sup>130</sup> By restricting access to education and spheres of critical

---

<sup>124</sup> RICHARD HOFSTADTER, *ANTI-INTELLECTUALISM IN AMERICAN LIFE* 27 (1963). Hofstadter also recognizes the difficulty in defining the term: “One reason that anti-intellectualism has not even been clearly defined is that its very vagueness makes it more serviceable in controversy as an epithet.” *Id.* at 26.

<sup>125</sup> In this context, anti-rationalism is a refusal of the “commitment to the value of critical thought and reasoned discourse in general,” not to be confused with anti-rationalist philosophical doctrine. Daniel Rigney, *Three Kinds of Anti-Intellectualism: Rethinking Hofstadter*, 61 *SOC. INQUIRY* 434, 436 (1991).

<sup>126</sup> *Id.* at 440 (“Rational discourse has its social bases in a constitutional system that protects free speech and assembly, in an adversarial system of political and judicial decision-making, in the institutions of scientific and scholarly inquiry and academic freedom, and in an emerging class of intelligentsia for whom the ‘culture of critical discourse’ is a shared ideology.”).

<sup>127</sup> HOFSTADTER, *supra* note 124, at 17–18 (“Primarily it was McCarthyism which aroused the fear that the critical mind was at a ruinous discount in this country. Of course, intellectuals were not the only targets of McCarthy’s constant denotations—he was after bigger game—but intellectuals were in the line of fire, and it seemed to give special rejoicing to his followers when they were hit.”).

<sup>128</sup> For example, the Czechoslovak Socialist Republic imprisoned philosopher and dissident, Václav Havel, numerous times between 1977 and 1989 before he became the first Czechoslovak president following the collapse of the communist régime. *Vaclav Havel: Timeline of the Former Czech President*, TELEGRAPH (Dec. 18, 2011, 1:46 PM), <https://www.telegraph.co.uk/news/worldnews/europe/czechrepublic/8964070/Vaclav-Havel-timeline-of-the-former-Czech-president.html>. Another example: Antonio Gramsci, an Italian Marxist philosopher, was sentenced to twenty years imprisonment under the Mussolini régime where he died in a medical clinic while serving his sentence. *Antonio Gramsci*, ENCYCLOPEDIA BRITANNICA, <https://www.britannica.com/biography/Antonio-Gramsci> (last visited Mar. 23, 2019).

<sup>129</sup> ARENDT, *supra* note 53, at 339.

<sup>130</sup> JASON STANLEY, *HOW FASCISM WORKS: THE POLITICS OF US AND THEM* 36 (2018).

debate, the capacity for intelligent discourse is limited, allowing the government to promote their single, “legitimate” ideology.<sup>131</sup> The very presence of critical academic discourse threatens the collective narratives and mythologies constructed to support the nation-state, especially when these narratives are recently constructed and do not conform to actual history.

When respected, free intellectual criticism may hold more weight than other forms of opposition, because of its ability to destabilize the flimsy theoretical grounds that authoritarian régimes use to legitimize their rule. Régimes, however, can also use a perversion of the university to their advantage to prop up fascistic ideology.<sup>132</sup> The university campus has become a battleground site for illiberal attacks on free thought, both at home and abroad.<sup>133</sup> Professors, students, disciplines, and universities are frequently targeted by illiberal régimes as dangerous voices that work to “indoctrinate” the nation’s children.<sup>134</sup> At the same time, régimes suppress critical viewpoints and manipulate the education system to reify mythic narratives of the nation-state as fact. Under these hyper-nationalist régimes, “the function of the education system is to glorify the mythic past, elevating the achievements of members of the nation and obscuring the perspectives and histories of those who do not belong.”<sup>135</sup> Disciplines such as gender studies are frequently attacked by far-right nationalist movements as undermining the traditions of the nation and its patriarchal ideology<sup>136</sup> and instead, disciplines that indoctrinate “hierarchal norms and national tradition”<sup>137</sup> are exalted by the régime. Higher education generally is depicted as an elitist institution symbolic of excess.<sup>138</sup> By rejecting and

---

<sup>131</sup> See *id.* (“In fascist ideology, there is only one legitimate viewpoint, that of the dominant nation.”).

<sup>132</sup> See *id.* (“Education therefore either poses a grave threat to fascism or becomes a pillar of support for the mythical nation.”).

<sup>133</sup> See *infra* note 367 (describing recent attempts by governments to limit academic freedom).

<sup>134</sup> Marxist thought tends to become the academic bogeyman for the far-right. Take, for example, “dangerous” university, course, and professor watchlists, like those promulgated by David Horowitz. See STANLEY, *supra* note 130, at 38–39 (“In 2006, Horowitz published a book, *The Professors*, naming the ‘101 most dangerous professors in America,’ a list of leftist and liberal professors, many of whom were supporters of Palestinian rights. In 2009, he published another book, *One-Party Classroom*, with a list of the ‘150 most dangerous courses in America.’”).

<sup>135</sup> *Id.* at 47.

<sup>136</sup> See *id.* at 42–43 (analyzing motives of the far-right in attacking gender studies). Particular anti-intellectual attacks on feminism and gender studies date back to Nazi-peddled myths that “feminism was a Jewish conspiracy to destroy fertility among Aryan women.” *Id.* at 43–44.

<sup>137</sup> *Id.* at 48.

<sup>138</sup> See *id.* at 37, 56 (“The media largely ignored these motivations [of the Black Lives Matter movement] and, representing protesting black students as an angry mob, used the situation as an opportunity to foment rage against the supposed liberal political excesses of the university. . . . In fascist politics, universities are debased in public discourse, and academics are undermined as legitimate sources of knowledge and expertise, represented as radical ‘Marxists’ or ‘feminists’ spreading a leftist ideological agenda under the guise of research.”); see also HOFSTADTER, *supra* note

mocking the value of academic expertise, the régime disrupts reality and inserts its own one-dimensional “reality.”<sup>139</sup> When the régime “is successful, its audience is left with a destabilized sense of loss, and a well of mistrust and anger against those who it has been told are responsible for this loss.”<sup>140</sup> By delegitimizing and forcibly targeting academics and universities, illiberal régimes aim to stall the possibility for robust debate and silence their critics. As such, anti-intellectualism remains in the toolkit of modern authoritarian leaders and their illiberal democracies. This is exemplified in the targeting of higher education institutions by modern authoritarian leaders in Central European countries, such as Hungary and Poland, as well as all over the world.

## II. THE HUNGARIAN SITUATION

### A. *What is Central Europe?*

This Paper investigates how illiberal democracy has developed in Hungary. To analyze how illiberal democracy has proliferated in Hungary, it is imperative to understand the historical conditions under which norms of governance and national image have developed in the region. But first, this begs the question: What is Central Europe?

Hungary is located in Central Europe—a region of small nation-states whose very existence is constantly under threat from larger surrounding powers.<sup>141</sup> Since the inception of the Cold War, Europe is often viewed as a dichotomy between Western and Eastern Europe, leaving little regard for the nations that lie somewhere in between.<sup>142</sup> Following the Second World War, nations whose cultures were traditionally associated with Western European values, such as Hungary and Czechoslovakia, found themselves on the Eastern side of the Iron Curtain. Czech author Milan Kundera described this experience in “The Tragedy of Central Europe.”<sup>143</sup> As a result of the sudden partition of some Central European nations, “three

---

124, at 53 (“Intellectuals, it may be held [by anti-intellectuals], are pretentious, conceited, effeminate, and snobbish; and very likely immoral, dangerous, and subversive.”).

<sup>139</sup> *Id.* at 57.

<sup>140</sup> *Id.*

<sup>141</sup> See Milan Kundera, *Die Weltliteratur: How We Read One Another*, NEW YORKER (Jan. 1, 2007), <https://www.newyorker.com/magazine/2007/01/08/die-weltliteratur> (“What distinguishes the small nations from the large is not the quantitative criterion of the number of their inhabitants; it is something deeper. For the small nations, existence is not a self-evident certainty but always a question, a wager, a risk; they are on the defensive against History, that force which is bigger than they, which does not take them into account, which does not even notice them.”).

<sup>142</sup> Stephen Shulman, *Challenging the Civic/Ethnic and West/East Dichotomies in the Study of Nationalism*, 35 COMP. POL. STUD. 554, 582–83 (2002) (“[T]he traditional civic-West/ethnic-East argument is a gross simplification of concepts of nationhood in the West, Central Europe, and Eastern Europe.”).

<sup>143</sup> Milan Kundera, *The Tragedy of Central Europe*, N.Y. REV. BOOKS, Apr. 26, 1984, at 33.

fundamental situations developed in Europe after [the Second World War]: that of Western Europe, that of Eastern Europe, and, most complicated, that of the part of Europe situated geographically in the center—culturally in the West and politically in the East.<sup>144</sup> Kundera characterized Central Europe as being “[b]oxed in by the Germans on one side and the Russians on the other”<sup>145</sup> and as “the least known and the most fragile part of the West,”<sup>146</sup> despite Central Europe’s wide contributions to the “European canon” from intellectuals such as Sigmund Freud, Béla Bartok, and Franz Kafka. Following the First World War, Central Europe was “transformed into a region of small, weak states, whose vulnerability ensured first Hitler’s conquest and ultimately Stalin’s triumph.”<sup>147</sup> Kundera argues that Central Europe is defined not by political power, but by its culture: “the great common situations that reassemble peoples, regroup them in ever new ways along the imaginary and ever-changing boundaries that mark a realm inhabited by the same memories, the same problems and conflicts, the same common tradition.”<sup>148</sup> Despite the vast cultural contributions made to “Western European culture” by Central Europe, as soon as the Iron Curtain closed around Central Europe, Western Europe was incapable of understanding the region as anything more than its politics, which was decidedly Eastern European.<sup>149</sup> But, Soviet usurpation was far more than a political struggle—it was also an attack on Central European civilization itself. The revolts in response to these existential attacks on Central European civilization were led by the intelligentsia<sup>150</sup> as a “struggle to preserve [Central European] identity—or, to put it another way, to preserve their Westernness.”<sup>151</sup> But, because the region became regarded as an Eastern political régime, “Europe [had not] noticed the disappearance of its cultural home because Europe no longer perceive[d] its unity as a cultural unity.”<sup>152</sup> This struggle for identity was completely ignored by Western Europe and, yet again, Central Europe was forgotten by its cultural brethren.<sup>153</sup> Kundera’s tragedy—that Central European nations had all but

---

<sup>144</sup> *Id.*

<sup>145</sup> *Id.* at 34.

<sup>146</sup> *Id.*

<sup>147</sup> *Id.*

<sup>148</sup> *Id.* at 35.

<sup>149</sup> *Id.* at 37.

<sup>150</sup> *Id.* These revolts “were prepared, shaped, realized by novels, poetry, theater, cinema, historiography, literary reviews, popular comedy and cabaret, philosophical discussions—that is, by culture.” *Id.*

<sup>151</sup> *Id.* at 34.

<sup>152</sup> *Id.* at 36.

<sup>153</sup> *Id.*

vanished from the map of Western Europe—rests not on the Soviet Union, but on Western European nations for abandoning them.<sup>154</sup>

Kundera ultimately defines Central Europe as “[a]n uncertain zone of small nations between Russia and Germany.”<sup>155</sup> Unlike their neighbors, “the small nation is one whose very existence may be put in question at any moment; a small nation can disappear and it knows it.”<sup>156</sup> At the time of writing, all of Central Europe, except Austria, had been swallowed up by the Soviet Union.<sup>157</sup> Kundera argues that once the majority of Central Europe was subjugated by the Soviet Union, the region was forgotten by Western Europe.<sup>158</sup> Leading up to the true independence of Central European states, the region faced immense political strife: Nazi invasion, a long stretch of failed revolutions, and finally, Soviet domination. Even prior to the Second World War, Hungary’s struggles were often overlooked by the larger European community. Unlike larger Western nations, such as the United Kingdom, France, or Germany, the histories of these small Central European nations have been “turbulent and fragmented.”<sup>159</sup> Their histories, including Hungary’s, have been ones of frequent invasion and in turn, “[t]heir traditions of statehood have been weaker and less continuous than those of the larger European nations.”<sup>160</sup> Hungary’s history includes long occupations by the Mongols, the Ottoman Empire, the Hapsburg Empire, the Third Reich, and the Soviet Union.<sup>161</sup> Unlike their western neighbors, Central European nation-states experience constant existential threat. István Bibó aptly reflects:

“The death of the nation” or “the annihilation of the nation” rings empty in West European ears; Westerners can imagine extermination, subjection, or slowly going native, but political “annihilation” overnight is sheer bombast to them, yet it is a *palpable reality* for the nations of Eastern Europe. Here there is no need to exterminate or expel a nation to

---

<sup>154</sup> *Id.* at 38 (“The real tragedy for Central Europe, then, is not Russia but Europe: this Europe that represented a value so great that the director of the Hungarian News Agency was ready to die for it, and for which he did indeed die. . . . He did not suspect that the sentence he was sending by telex beyond the borders of his flat country would seem outmoded and would not be understood.”).

<sup>155</sup> *Id.* at 35.

<sup>156</sup> *Id.*

<sup>157</sup> *Id.* at 36.

<sup>158</sup> *Id.*

<sup>159</sup> *Id.* at 34.

<sup>160</sup> *Id.*

<sup>161</sup> *Hungary Timeline*, BBC NEWS (Feb. 14, 2012), [http://news.bbc.co.uk/2/hi/europe/country\\_profiles/1054642.stm](http://news.bbc.co.uk/2/hi/europe/country_profiles/1054642.stm).



make it feel endangered; it is enough to *call its existence into doubt* with a sufficiently aggressive rhetoric.<sup>162</sup>

This persistent existential threat to independent nationhood has impacted the democratic development of Central Europe and “has been the decisive factor in making democracy and democratic development waver in these countries.”<sup>163</sup> The embedded identity of the existential anxiety of the nation-state guides political and legal decision-making and may even interfere with Hungary’s democratic prospects.

### B. *Hungarian Political History (Hapsburg—1989)*

Developments in Hungarian history uniquely positioned the nation to be susceptible to the rise of the far-right. Wilkin contends that “the roots of illiberalism in the modern world-system are a reaction, in part, to the threat that liberalism presented to established social hierarchies, secular or religious”<sup>164</sup> in the seventeenth and eighteenth centuries. Elements of illiberalism have been persistent throughout Hungarian political history. Wilkin argues that four major periods of history shaped the development of the Hungarian nation-state and its relationship to illiberalism. First, the restoration of the monarchy and the establishment of the Austro-Hungarian Empire, following the defeat of the democratic Hungarian Revolution of 1848, was reactionary and anti-modern.<sup>165</sup> Monarchical systems are, unsurprisingly, adverse to the classical liberal values that threaten the social hierarchies entrenched in traditional monarchies.<sup>166</sup> Second, the crumbling of the Austro-Hungarian Empire after the First World War bred animosity within the nation-state after “it suffered drastic loss of territory and population through the Treaty of Trianon.”<sup>167</sup> Hungary’s experience during the inter-war period mirrors that of other Central Powers, like Germany, whose reactionary fascist forces rose to power as a response to the aftermath of the war.<sup>168</sup> This massive loss of territory “left a lasting

---

<sup>162</sup> ISTVÁN BIBÓ, *The Miseries of East European Small States*, in *THE ART OF PEACEMAKING* 130, 150 (Iván Zoltán Dénes ed., Péter Pásztor trans., 2015).

<sup>163</sup> *Id.* at 151.

<sup>164</sup> Peter Wilkin, *The Rise of ‘Illiberal’ Democracy: The Orbánization of Hungarian Political Culture*, 24 *J. WORLD-SYS. RES.* 5, 9 (2018).

<sup>165</sup> *Id.* at 13.

<sup>166</sup> *Id.* at 13–14 (“[The Hungarian monarchy was] understandably[] deeply hostile to liberal ideas of universality and equality, preferring instead to entrench social life in traditional social hierarchies shaped through the church and respect for secular authority in the forms of the King and the aristocracy.”).

<sup>167</sup> *Id.* at 14.

<sup>168</sup> ARENDT, *supra* note 53, at 308 (“After the first World War, a deeply anti-democratic, predictatorial wave of semitotalitarian and totalitarian movements swept Europe; Fascist movements spread from Italy to nearly all Central and Eastern European countries . . .”). Also note, Austro-Hungary lost two-thirds of its territory and two-thirds of its population. *Treaty of Trianon*,

legacy of resentment amongst the country's right-wing social forces, which still manifests itself rhetorically today with both Fidesz and Jobbik.<sup>169</sup> Third, Hungary's experience with fascism and Nazism during the Second World War impacted the rise of illiberal democracy in the country today by bringing anti-Semitism, racism, and other prejudices to the fore.<sup>170</sup> The Second World War presented an opportunity for Hungary to reclaim its lost territory; but, by 1944, Hungary had become a puppet state for Nazi Germany after a *coup d'état* by the far-right fascist Arrow Cross Party.<sup>171</sup> Finally, after the defeat of the Axis forces in the Second World War, the Soviet Union army invaded Hungary and instituted another authoritarian régime. Between 1945 and 1989, Hungary was occupied by Soviet forces that quashed democratic revolutions<sup>172</sup> and decimated civil society.<sup>173</sup> These eras of Hungarian history shaped its development as a nation-state in the *longue durée* and influenced the rise of Orbán's illiberal democratic project.

### C. *Fledgling Democracy: The Interim Years (1989–2010)*

Prime Minister Orbán's rise to power directly resulted from the missteps of the governments in power during the democratic transitional period.<sup>174</sup> Just as democracy emerged, neo-fascist and anti-Semitic

---

ENCYCLOPÆDIA BRITANNICA, <https://www.britannica.com/event/Treaty-of-Trianon> (accessed Mar. 19, 2019).

<sup>169</sup> Wilkin, *supra* note 164, at 14.

<sup>170</sup> *Id.* at 15 (noting the role that Hungarians and the Hungarian government played in the Final Solution).

<sup>171</sup> JACK R. FISCHER, HISTORICAL DICTIONARY OF THE HOLOCAUST 122–23 (1999). Ferenc Szálasi, the Prime Minister of Hungary and leader of the Arrow Cross Party, was found guilty for war crimes and was sentenced to death. Militiamen of the Arrow Cross Party conducted a reign of terror that resulted in the violent murder of many Hungarian Jews. *Hungary After the German Occupation*, U.S. HOLOCAUST MEMORIAL MUSEUM, <https://encyclopedia.ushmm.org/content/en/article/hungary-after-the-german-occupation> (accessed Mar. 18, 2019). By the end of the Second World War, an estimated 563,000 Hungarian Jews died at the hands of the Arrow Cross Party and the Nazi régime. *Id.* Wilkin notes that Szálasi's "enthusiasm for the Final Solution was implacable. . . . [A]t times SS officers in Hungary had to restrain the Hungarian forces from their enthusiastic slaughter of the Jewish population." Wilkin, *supra* note 164, at 15.

<sup>172</sup> See Ben Cosgrove, *A Rip in the Iron Curtain: Photos From the Hungarian Revolution, 1956*, TIME (Oct. 22, 2013), <http://time.com/3878232/the-hungarian-revolution-of-1956-photos-from-the-streets-of-budapest/> (collecting photographs from the 1956 Hungarian Revolution).

<sup>173</sup> See Péter Krasztev, *Social Responses to the "Hybridization" of the Political System: The Case of Hungary in the Central and Eastern European Context*, in THE HUNGARIAN PATIENT: SOCIAL OPPOSITION TO AN ILLIBERAL DEMOCRACY 167, 171 (Péter Krasztev et al. eds., 2015) ("The historian Stephen Kotkin has called state socialist societies 'uncivic societies,' and the Hungarian example certainly proves this: the Kádár regime was efficient enough to make opposition seem futile and thus eliminate social solidarity and autonomy, and these attitudes still live on today.").

<sup>174</sup> András Bozóki, *Broken Democracy, Predatory State, and Nationalist Populism*, in THE HUNGARIAN PATIENT: SOCIAL OPPOSITION TO AN ILLIBERAL DEMOCRACY 3, 4 (Péter Krasztev et al. eds., 2015) ("This antiliberal turn did not emerge out of the blue: it was a direct response to the hectic

movements, fueled by nationalism, simultaneously rose to mainstream discourse.<sup>175</sup> This Section will explore how internal failures in Hungarian government and external factors influenced the rise of illiberal democracy under Orbán by using Huntington's factors<sup>176</sup> that contributed to prior reverse waves of democracy. The most relevant factors for this period of time are: (1) the weakness of democratic values; (2) severe economic setbacks; (3) social and political polarization; (4) the exclusion of non-elites by elites; and (5) the breakdown of law and order.<sup>177</sup>

The weakness of democratic values and increasing political polarization in the transitional period were instrumental in prompting the rise of the far-right. The post-Cold War Hungarian democratic process was deeply flawed. Bozóki identifies three institutional factors that contributed to Orbán's success: the qualified majority vote, informal rulemaking, and partocracy.<sup>178</sup> The Hungarian Founding Fathers placed an emphasis on "strengthen[ing] the new democratic order, its stability, and its governability"<sup>179</sup> when drafting the new democratic institutional system. The Founding Fathers attempted to achieve these goals by instituting required qualified majority votes in many arenas of the decision-making process.<sup>180</sup> Bozóki writes:

These measures created a democracy in which, between elections, the ruling government's power became almost cemented. It became nearly impossible to remove an incumbent government from the outside; however, this simultaneously made effective governance more difficult. The government in power, due to the large number of qualified majority rules, had to rely on the *opposition* to make decisions on basic issues. Paradoxically, the constitution thus both greatly increased the power *and* limited the political responsibility of the government.<sup>181</sup>

By overvaluing stability, the constitutional system that existed between 1990 and 2010 created systematic inefficiencies that, in turn, contributed to the devaluing of democracy. Bozóki also remarks that Hungary's history of occupation "produced a political culture characterized by a prevalence of

---

reforms implemented by previous governments between 2006 and 2010, as well as the ensuing corruption and economic crisis.").

<sup>175</sup> Wilkin, *supra* note 164, at 18.

<sup>176</sup> HUNTINGTON, *supra* note 5, at 290–91.

<sup>177</sup> *Id.*

<sup>178</sup> Bozóki, *supra* note 174, at 5–9.

<sup>179</sup> *Id.* at 5.

<sup>180</sup> *Id.* ("[T]he Founding Fathers believed that they could safeguard freedom by increasing the number of decisions that required a qualified majority vote.").

<sup>181</sup> *Id.*

informal practices and a lack of institutional accountability.”<sup>182</sup> In order to cope with occupation, a dual system of contradictory formal and informal rules developed whereby the official rules of the occupier would be followed while finding loopholes and cutting corners so as to undermine official rule.<sup>183</sup> This dual rule system persisted throughout the Kádár era of communism and, Bozóki argues, made the system more bearable than those of its neighbors.<sup>184</sup> Because of Hungary’s long history of occupation, Bozóki contends that “in 1989 Hungarians broke only with the institutional system of dictatorship, not with the customs and informal procedures associated with that system.”<sup>185</sup> While this dual rule system was vital during occupation, the persistent culture of rule-bending delegitimized democratic rule in post-occupation life. Finally, “partocracy,” the form of government by which a single party rules hegemonically,<sup>186</sup> is culturally endemic in Hungarian politics and at odds with democratic norms.<sup>187</sup> The anti-pluralism that current day Hungary is experiencing is not novel to Orbán’s régime and in fact predates it. Political parties, including those on the left, dominate all aspects of the political process: public discourse, civic duties, and candidate nominations were all controlled by the major political parties.<sup>188</sup> Thus, the Hungarian democratic system, in its most free state, was highly politicized and plagued by the excessive control of areas of public life that should have remained free from government interference

---

<sup>182</sup> *Id.* at 6.

<sup>183</sup> *Id.* (“Hungarians learned that they only had to feign obedience to the rules imposed upon them by foreign invaders: below the surface, they established a system of informal rules governing society and culture. . . . Therefore, Hungarians learned to get their way around these rules in a conniving fashion, finding loopholes and cutting corners, and this behavioral pattern remains deeply engrained in Hungarian society.”).

<sup>184</sup> *Id.* (“The reason [the regime] became more livable is that the system often did not take its own rules seriously. . . . Under Kádárism, citizens grew accustomed to those procedures that made the dictatorship bearable. For Hungarians, the old system was not nearly as bad as it had been for the Poles, the Czechs, or the Romanians.”).

<sup>185</sup> *Id.*

<sup>186</sup> Wilkin, *supra* note 164, at 18–19 (“Hungary’s political system [was] dominated by either neoliberal parties such as the reform communists and liberal parties (MSZP and SZDSZ) who governed between 1994–[19]98 and 2002–2010; or conservative-nationalist coalitions led initially by the MDF who governed from 1990–[19]94, with Fidesz in office between 1998–2002, leading a coalition including the Christian Democrats and the Smallholders Party.”).

<sup>187</sup> See Bozóki, *supra* note 174, at 7 (“During the second decade of democracy in Hungary, party politics superseded almost all other aspects.”).

<sup>188</sup> See *id.* at 7–8 (“Public discourse was based on party allegiance. . . . It was the parties that organized movements; it was the parties that established public benefit foundations, professional groups, and civic circles. . . . The particular features of the Hungarian political system—including the collection of candidate nomination slips, the high threshold to enter parliament, the large number of regulatory areas in which a qualified majority is required in order to create new laws, the opacity of political party financing, the privileged position of political party foundations, and so on—facilitated the survival of existing parties and made it difficult for new political forces to enter parliament.”).

but were instead controlled by the dominant political party.<sup>189</sup> Partocracy only served to further polarize the Hungarian public in an already turbulent time, creating an atmosphere of a “cold civil war” between those on the left and right.<sup>190</sup> Unchanged, all three practices made the chances of a lasting, stable democracy dead on arrival.

Throughout the late 1990s and persisting to the current day, Hungary has faced many economic crises. As the formerly communist nation transitioned to democracy, Hungary also transitioned to capitalism. Hungary slowly privatized and the transition resulted in an unstable economy.<sup>191</sup> Hungary experienced rapid deindustrialization, widening regional inequalities.<sup>192</sup> As a result of this rough transition, Hungary became increasingly reliant on external financial investments, primarily from the European Union and Japan.<sup>193</sup> In Hungary’s first decade of democracy, the country “experienced periods of massive contraction,”<sup>194</sup> but began to steady itself in the new millennium.<sup>195</sup> This quasi-stability was quickly quashed by the unpopular austerity measures pushed through by the MSZP-SZDSZ government<sup>196</sup> in 2006 and the global recession in 2008.<sup>197</sup> Tóth and Grajczjár speculate that “the recovery period was too short to solve the internal societal tensions, poverty and underemployment, to bridge the wide gap between the eastern and western parts of the country, and to stop the deterioration of public institutions.”<sup>198</sup> These austerity measures promoted by elite members of the MSZP-SZDSZ coalition ran contrary to what the majority of Hungarians actually desired.<sup>199</sup> The political scene only became more polarized when a confidential speech by then-Prime Minister Ferenc Gyurcsány in

---

<sup>189</sup> *Id.* at 8–9 (“The Hungarian system was characterized by a highly politicized society and by the excessive sway that political parties held in various areas of public life.”).

<sup>190</sup> *Id.* at 8 (describing the depths of political polarization in Hungary as a result of partocracy).

<sup>191</sup> Wilkin, *supra* note 164, at 20.

<sup>192</sup> *Id.* at 18.

<sup>193</sup> *Id.* at 20.

<sup>194</sup> *Id.*

<sup>195</sup> *Id.*

<sup>196</sup> The MSZP-SZDSZ was a coalition government between the Hungarian Socialist Party (Magyar Szocialista Párt) and the Alliance of Free Democrats’ (Szabad Demokraták Szövetsége – a Magyar Liberális Párt) party which held a close majority before Fidesz won a majority. Csaba Nikolenyi, *Strategic Co-Ordination in the 2002 Hungarian Election*, 56 EUR.-ASIA STUD. 1041, 1041 (2004).

<sup>197</sup> Bozóki, *supra* note 174, at 11.

<sup>198</sup> András Tóth & István Grajczjár, *The Rise of the Radical Right in Hungary*, in *THE HUNGARIAN PATIENT: SOCIAL OPPOSITION TO AN ILLIBERAL DEMOCRACY* 133, 158 (Péter Krasztev & Jon Van Til eds., 2015).

<sup>199</sup> Wilkin, *supra* note 164, at 19 (“The problem was that these austerity policies were against what the majority of the Hungarian population actually wanted at the time.”).

Balatonönszöd was leaked.<sup>200</sup> In the speech, Gyurcsány said that he and the MSZP-SZDSZ government “had knowingly lied to the public concerning the economic situation in Hungary,”<sup>201</sup> causing an eruption of “violent street protests”<sup>202</sup> on the anniversary of the 1956 Revolution.<sup>203</sup> This economic crisis was intensified by the global recession and occurred “when the government was rapidly losing its political credibility domestically.”<sup>204</sup> The mishandling of the economic crisis decimated support for MSZP and opened the door for far-right parties to rise to prominence.<sup>205</sup> In addition to Fidesz, the Jobbik party, a radical far-right party that has been described as neo-fascist,<sup>206</sup> emerged during the economic crises.<sup>207</sup> By the 2010 parliamentary elections, support for MSZP dropped to 20 percent<sup>208</sup> while Jobbik captured 16.67 percent of the vote, becoming the third largest party in parliament.<sup>209</sup>

Finally, during the interim period of democratization, far-right parties peddled “law and order” narratives that targeted the Roma.<sup>210</sup> Tóth and Grajczjár argue that anti-Roma “law and order” rhetoric entered public discourse in 2006 after a tragic murder was committed by a group of Roma in Olaszliszka.<sup>211</sup> This event was the catalyst for an outpouring of anti-Roma sentiments, with Jobbik leading this discourse as “protector of

<sup>200</sup> Philipp Karl, *Network Analysis of Right-Wing Extremism in Hungary*, in *MINORITIES UNDER ATTACK: OTHERING AND RIGHT-WING EXTREMISM IN SOUTHEAST EUROPEAN SOCIETIES* 221, 222 (Sebastian Goll et al. eds., 2016).

<sup>201</sup> *Id.*

<sup>202</sup> *Id.*

<sup>203</sup> Tóth & Grajczjár, *supra* note 198, at 137.

<sup>204</sup> Bozóki, *supra* note 174, at 11.

<sup>205</sup> *Id.* at 3, 4.

<sup>206</sup> *Jobbik*, COUNTER EXTREMISM PROJECT, [https://www.counterextremism.com/sites/default/files/threat\\_pdf/Jobbik-12172018.pdf](https://www.counterextremism.com/sites/default/files/threat_pdf/Jobbik-12172018.pdf) (last visited Mar. 23, 2019).

<sup>207</sup> PETER KREKÓ ET AL., BROOKINGS, *ANTI-MUSLIM POPULISM IN HUNGARY: FROM THE MARGINS TO THE MAINSTREAM* (2019), <https://www.brookings.edu/research/anti-muslim-populism-in-hungary-from-the-margins-to-the-mainstream/>.

<sup>208</sup> Tóth & Grajczjár, *supra* note 198, at 137.

<sup>209</sup> *Id.* at 139.

<sup>210</sup> The Roma are a richly diverse, historically nomadic group of people who likely originated in northern India and migrated towards Europe in approximately the tenth or eleventh centuries. Throughout their history, the Roma have been subjected to abuse, enslavement, and extermination. In Romania, the Roma were enslaved between the fourteenth century until the 1850s. In Nazi Germany, between 250,000 and 500,000 Roma perished in the Holocaust. An additional note: while the Roma are also referred to as “Gypsies,” this is a widely recognized slur and will be omitted from quoted materials, except for in the case of direct quotes from far-right speeches. Samira Shackle, *Roma Holocaust: Amid Rising Hate, ‘Forgotten’ Victims Remembered*, ALJAZEERA (Oct. 30, 2019), <https://www.aljazeera.com/indepth/features/roma-holocaust-rising-hate-forgotten-victims-remembered-191029173851099.html>; see Adrian Marsh, *Gypsies, Roma, Travellers: An Animated History*, OPEN SOCIETY FOUND. (June 5, 2013), <https://www.opensocietyfoundations.org/voices/gypsies-roma-travellers-animated-history> (detailing the history of the Roma people).

<sup>211</sup> Tóth & Grajczjár, *supra* note 198, at 138.

‘the honest, hardworking common people’<sup>212</sup> against “Gypsy crime.”<sup>213</sup> The Jobbik propagation of this racist discourse was ultimately successful in widening their voter base and bringing far-right politics to the fore: “[M]any [Hungarians] felt themselves being finally liberated from the left-liberal stranglehold of political correctness and allowed themselves to give vent to long-suppressed resentment, naming the causes of their real or imagined grievances.”<sup>214</sup> In 2007, Jobbik spurred the creation of the Hungarian Guard (Magyar Gárda),<sup>215</sup> a paramilitary group dedicated to the “defence” of Hungary against “Gypsy criminality.”<sup>216</sup> Other anti-Roma extremist groups, such as Véderő and Szebb Jövőért, scheduled marches so hostile and aggressive that they caused Roma populations to evacuate towns, such as Gyöngyöspata, in fear for their safety.<sup>217</sup> This xenophobic mainstream law and order rhetoric is echoed in the contemporary far-right discourse surrounding the refugee crisis.<sup>218</sup>

#### D. Rise of Hungarian Illiberal Democracy (2010–present)

The tumultuous period of democratic transition preceding Fidesz’s régime greatly contributed to the party’s triumph.<sup>219</sup> The nascent illiberal democracy became cemented as Hungary’s new system of governance through sweeping unilateral legal reforms. These reforms enshrined the illiberal principles of majoritarianism, absolute sovereignty, dictatorship of law, nationalism, and anti-intellectualism.<sup>220</sup> The following section will examine how Orbán instrumentalized the law to claim the government for his own party and apply it to the typical characteristics of an illiberal democracy. The cumulative effect of these reforms is demonstrated in *Freedom House*’s decision to downgrade Hungary’s freedom ranking from “free” to “partially free” in 2019.<sup>221</sup> These legal reforms attacked six arenas

---

<sup>212</sup> *Id.*

<sup>213</sup> *Id.*

<sup>214</sup> *Id.*

<sup>215</sup> The Magyar Gárda was disbanded by Metropolitan Court of Budapest in 2009. After unsuccessful appeals in the national judicial system, the Magyar Gárda brought their case to the European Court of Human Rights and alleged that the national decision violated freedom of assembly as guaranteed by the European Convention of Human Rights. The court held that Article 11 was not violated and upheld the judgments of the Hungarian national courts. *Vona v. Hungary*, App. No. 35943/10, Eur. Ct. H.R. (2013) paras. 16–18, 72.

<sup>216</sup> *Id.* para. 10.

<sup>217</sup> Karl, *supra* note 200, at 223.

<sup>218</sup> *See infra* Part II.D.

<sup>219</sup> *See supra* Part II.C.

<sup>220</sup> *Id.*

<sup>221</sup> FREEDOM HOUSE, FREEDOM IN THE WORLD 2019: HUNGARY (2019), <https://freedomhouse.org/country/hungary/freedom-world/2019> (“Hungary’s status declined from Free to Partially Free due to sustained attacks on the country’s democratic institutions by Prime Minister Viktor Orbán’s Fidesz party, which has used its parliamentary supermajority to impose restrictions on or assert control over the opposition, the media, religious groups, academia, NGOs, the courts, asylum

of democracy: the Constitution, the Constitutional Court, the presidency, independent institutions, the media, and civil society.

Before Orbán's Fidesz came into power, Hungary "was a liberal democracy characterized by a multiparty system, free elections, representational government, a strong opposition, free media, strong, and credible institutions that protected the rule of law (*i.e.*, the Constitutional Court and Ombudsman's Office), and independent courts."<sup>222</sup> In stark contrast, the Orbán government approaches law not as an entity to be respected, but as a body to be manipulated in order to fit the needs of Fidesz's political agenda.<sup>223</sup> This is first and foremost exemplified with Fidesz's unilateral rewriting of the Constitution. The framers of the old Constitution, drafted after the fall of communism, had two concerns when structuring the new government: first, "a fractured parliament in which small parties would be unable to form stable majority coalitions" and second, "a deeply entrenched constitution that would be too hard to change once the new democrats figured out how they wanted to design their political institutions."<sup>224</sup> The resulting constitution was one that favored larger parties with a provision allowing parliament to alter any part of the Constitution so long as they secure a two-thirds majority.<sup>225</sup> This fatal flaw in the Constitution revealed itself after Fidesz secured 53 percent of the popular vote, translating into 68 percent of the seats in parliament.<sup>226</sup> This meant that Fidesz was now capable of unilaterally amending the Constitution. In their first year in power, Fidesz amended the Constitution twelve times, altering more than fifty separate provisions and weakening any and all checks and balances.<sup>227</sup> Fidesz used their two-thirds majority power to erase the last measure restraining constitutional amendments: the requirement of "a four-fifths vote of parliament to set the rules for writing a new constitution."<sup>228</sup> The elimination of this rule allowed Fidesz to draft a completely new constitution while barring any opposing voices from sitting at the table.

---

seekers, and the private sector since 2010."). Cf. FREEDOM HOUSE, FREEDOM IN THE WORLD 2018: HUNGARY (2018), <https://freedomhouse.org/country/hungary/freedom-world/2018> (displaying a 72/100 Global Freedom Score and designating Hungary as "free").

<sup>222</sup> András Bozóki, *Occupy the State: The Orbán Regime in Hungary*, 19 DEBATTE: J. CONTEMP. CENT. & E. EUR. 649, 649 (2011).

<sup>223</sup> See *id.* at 651 ("[T]he new government saw the 1989 Constitution as a heap of purely technical rules, which Orbán has since shaped to fit the needs of his current political agenda. If any of his new laws proved to be unconstitutional, it was not the law, but the Constitution that had to be changed.").

<sup>224</sup> Miklós Bánkuti, Gábor Halmi & Kim Lane Scheppelle, *Hungary's Illiberal Turn: Disabling the Constitution*, in THE HUNGARIAN PATIENT: SOCIAL OPPOSITION TO AN ILLIBERAL DEMOCRACY 37, 37 (Péter Krasztev & Jon Van Til eds., 2015).

<sup>225</sup> *Id.* at 37–38.

<sup>226</sup> *Id.* at 38.

<sup>227</sup> *Id.*

<sup>228</sup> *Id.*



Now in control of the executive and legislative branches, Fidesz's next target was the Constitutional Court. Before Fidesz's reign, the Constitutional Court was a powerful check on the government.<sup>229</sup> First, Fidesz utilized its newfound amendment powers to change the judicial nomination process by allowing the party in power to nominate candidates to be elected to the court by a two-thirds majority, completely eliminating pluralism from the process.<sup>230</sup> The Constitution had previously "required a majority of parliamentary parties to agree to a nomination and then a two-thirds vote of parliament's members to elect the nominee to the court."<sup>231</sup> Second, Fidesz attacked the Constitutional Court's jurisdiction after it declared a retroactive tax law that punished members of the former MSZP-SZDSZ government unconstitutional<sup>232</sup> and retaliated by "amending the Constitution and limiting the Constitutional Court's jurisdiction."<sup>233</sup> Parliament barred the Constitutional Court from reviewing any law about fiscal matters unless it affects "rights to life, dignity, data privacy, thought, conscience, religion, and citizenship."<sup>234</sup> Finally, the Fidesz government packed the Constitutional Court and delegated themselves the power to name seven of the fifteen judges as well as the chairperson of the Constitutional Court.<sup>235</sup> As to be expected, all of the nominees are Fidesz-affiliates.<sup>236</sup>

In addition to the Constitutional Court, Fidesz uprooted the appointment procedure for judgeships in every single court in the country. Before Fidesz, lower court judges were independently appointed by a panel of their fellow judges.<sup>237</sup> Under the new system, Fidesz established the National Judicial Office (KIH) to oversee the judiciary and holds "the power to select new judges, to promote or demote any judge, to begin

---

<sup>229</sup> *Id.*

<sup>230</sup> *Id.*

<sup>231</sup> *Id.*

<sup>232</sup> *Id.* ("In order to plug gaping budget holes, the Fidesz government established a 98 percent retroactive tax on the customary departing bonuses of those who had left public employment in the preceding five years. The Constitutional Court, before it could be packed with a working majority of new judges, struck down this tax as unconstitutional." (footnote omitted)).

<sup>233</sup> Bozóki, *Occupy the State*, *supra* note 222, at 651.

<sup>234</sup> Bánkuti et al., *supra* note 224, at 39. Bánkuti et al. note that the Constitutional Court is explicitly barred from reviewing "budget or tax laws if they infringe other rights that are much easier to limit with fiscal measures, such as the right to property, equality under the law, the prohibition against retroactive legislation, and the guarantee of fair judicial procedure." *Id.* Fidesz also sought to nationalize private pensions, directly resulting in the European Court of Human Rights being flooded with 8000 cases on the matter due to the Constitutional Court's inability to act. *Id.*

<sup>235</sup> See Bozóki, *Occupy the State*, *supra* note 222, at 651–52 ("In addition, the number of judges was increased from eleven to fifteen, and the Court was packed with right-wing personalities and former politicians known to be close to Fidesz.").

<sup>236</sup> *Id.*

<sup>237</sup> Bánkuti et al., *supra* note 224, at 42.

disciplinary proceedings, and to select the leaders of each of the courts.<sup>238</sup> The president of the KIH has a nine year term and is selected by a two-thirds majority vote in parliament, again guaranteeing that the head of this “independent” body would be a friend of Fidesz.<sup>239</sup> In this case, current Chairperson Tünde Handó is quite literally “a close friend of Prime Minister Orbán and the wife of József Szájer,” the Vice President of Fidesz and principal architect of the new constitution.<sup>240</sup> In contrast to other legal systems, Chairperson Handó also has the sole authority to reassign cases throughout the country at will.<sup>241</sup> As such, the legitimacy of the entire Hungarian judicial system has been decimated in under a decade.

The checks delegated to the president’s office and purportedly independent accountability institutions have also been delegitimized. First, under the old constitution, the president’s main check was that of the suspensive veto power.<sup>242</sup> In the case of the president’s office, parliament simply elects hardline Fidesz supporters, like Pál Schmitt and János Áder, who refuse to veto Fidesz legislation.<sup>243</sup> Second, Hungary’s old ombudsman system comprised of “four separate ombudsmen with separate staffs and independent jurisdictions”<sup>244</sup> that monitored human rights issues. Now, there is a single parliamentary commissioner for fundamental rights that operates with a severely reduced staff.<sup>245</sup> The Data Protection Supervisor was abolished and a new, non-independent office was established.<sup>246</sup> In 2014, the Court of Justice of the European Union found that in doing so, Hungary failed to fulfill its obligations under the 1995 Data Protection Directive.<sup>247</sup> Third, the State Audit Office, once an independent body with the power to investigate the misuse of public funds, is now led by a former Fidesz MP with no professional auditing experience.<sup>248</sup> Fourth, the National Election Commission (NVB), the

---

<sup>238</sup> *Id.*

<sup>239</sup> *Id.* at 42–43.

<sup>240</sup> *Id.* at 42.

<sup>241</sup> *Id.* at 43.

<sup>242</sup> *Id.* at 40.

<sup>243</sup> *Id.*

<sup>244</sup> *Id.* at 44.

<sup>245</sup> *Id.*

<sup>246</sup> Case C-288/12, *Comm’n v. Hung.*, 2014 E.C.R. I-nyr, <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2012:227:0015:0016:EN:PDF>.

<sup>247</sup> Bánkuti et al., *supra* note 224, at 44.

<sup>248</sup> *Id.* at 39. Independent Hungarian media allege that the State Audit Office is selectively enforcing auditing regulation aimed at harming opposition parties, while dismissing like investigations on Fidesz. Jobbik is facing fines of 662 million forints (\$2.4 million) for illegal in-kind campaign contributions, while three parties are facing fines for “renting office space below market rates”: The Democratic Coalition, 16 million forints (\$58,000); Együtt, 16 million forints (\$58,000); and Politics Can Be Different 8.8 million forints (\$31,900). Christopher Adam, *The Hungarian State Audit Office’s Assault on Democracy*, HUNGARIAN FREE PRESS (Jan. 9, 2018, 3:35 PM), <http://hungarianfreepress.com/2018/01/09/the-hungarian-state-audit-offices-assault-on-democracy/>.

independent body charged with regulating election law, has been filled with a Fidesz majority.<sup>249</sup> In addition to monitoring elections and drawing electoral maps,<sup>250</sup> the NVB also has the power to decide what referenda will be voted on in elections.<sup>251</sup> This is particularly notable as referenda are one of the most substantial areas that civil society can attempt to influence the Fidesz government.<sup>252</sup> Finally, the new constitution created the Budget Council that may “veto any budget produced by parliament that adds even a single forint [(0.004¢)] to the national debt.”<sup>253</sup> The terms for Budget Council officials exceed that of a standard parliamentary election cycle, therefore allowing the Budget Council to “exercise dead-hand control over future elected governments.”<sup>254</sup> Even worse, “if parliament fails to agree on a budget by March 31 of each year, then the president may dissolve parliament and call new elections.”<sup>255</sup> If the Budget Council utilizes its veto power right before the deadline, it could force a new election.<sup>256</sup> Thus, Fidesz’s partocracy extends beyond the three branches of government and invades purportedly independent institutions as well.

Fidesz has usurped power horizontally by controlling the vast majority of mainstream media. The Fidesz government established the National Media and Infocommunications Authority (NMHH), a regulatory agency and an “independent” Media Council, charged with monitoring media outlets and fining outlets that do not have “balanced” news programming.<sup>257</sup> Like other “independent” government agencies, Orbán appointed a former Fidesz MP to lead the NMHH, and parliament elected Fidesz loyalists fill all the seats on the Media Council.<sup>258</sup> Although the Hungarian Constitution guarantees freedom of speech and freedom of the press,<sup>259</sup> the public television broadcaster is biased in favor of Fidesz and

---

<sup>249</sup> Bánkuti et al., *supra* note 224, at 39.

<sup>250</sup> Fidesz’s ability to modify election law without cooperation from outside parties has allowed them to manipulate previously independent institutions to preserve their two-thirds majority. For example, in 2014, Fidesz won only 44 percent of the popular vote, but still maintained their two-thirds majority in parliament. Martino Comelli & Vera Hovárth, *What Orbán Knows and His Enemies Don’t*, JACOBIN MAG. (Mar. 9, 2018), <https://www.jacobinmag.com/2018/03/viktor-orban-hungary-fidesz-authoritarian-opposition>.

<sup>251</sup> Bánkuti et al., *supra* note 224, at 39.

<sup>252</sup> *Id.* at 44.

<sup>253</sup> *Id.*

<sup>254</sup> *Id.*

<sup>255</sup> *Id.*

<sup>256</sup> *Id.*

<sup>257</sup> *Id.* at 40.

<sup>258</sup> *Id.*

<sup>259</sup> MAGYARORSZÁG ALAPTÖRVÉNYE [THE FUNDAMENTAL LAW OF HUNGARY], ALAPTÖRVÉNY art. IX [hereinafter HUNGARIAN CONSTITUTION] (“(1) Everyone shall have the right to freedom of expression. (2) Hungary shall recognize and protect the freedom and diversity of the press, and shall ensure the conditions for the free dissemination of information necessary for the formation of democratic public opinion. . . . (6) The detailed rules relating to the freedom of the press and the organ

actively undermines the opposition; Fidesz selectively awards advertising contracts and radio station frequencies to supporters and pressures critical news outlets into silence or closes them.<sup>260</sup> *Népszabadság*, the largest opposition newspaper, was unexpectedly suspended from operation after it uncovered Fidesz scandals and its parent company was subsequently sold to Optimus Press, a firm owned by Fidesz allies.<sup>261</sup> The firm has no plans to reopen *Népszabadság*.<sup>262</sup> Fidesz affiliates and pro-government media currently dominate the market,<sup>263</sup> and much of the opposition media has been pushed to the internet.<sup>264</sup> Approximately 90 percent of all media in Hungary is owned by either the government or allies of Fidesz and use their publications to push pro-government views.<sup>265</sup> For example, a study by Democracy Reporting International and Mérték Media Monitor studied broadcasts by television stations about the refugee resettlement program referendum in Hungary, supported by the European Union, and found that 91 percent of programming by public television stations took anti-referendum positions.<sup>266</sup> Currently, Freedom House has ranked Hungary's freedom of the press as only "partly free."<sup>267</sup>

---

supervising media services, press products and the communications market shall be laid down in a cardinal Act.”)

<sup>260</sup> FREEDOM HOUSE, FREEDOM OF THE PRESS 2017: HUNGARY (2017), <https://www.refworld.org/docid/59fc67e6a.html>. Additionally, the “donation” of most major media outlets—over four hundred—to a pro-Fidesz foundation run by Orbán’s childhood friend is yet another example that Orbán’s “economic patriotism” is merely a guise for his crony capitalism. Of course, this mass donation to the Central European Press and Media Foundation must be approved by regulatory authorities (led by Orbán appointees). If approved, “the deal will place most leading private Hungarian outlets under the control of a single, state-friendly entity.” Patrick Kingsley, *Orban and His Allies Cement Control of Hungary's News Media*, N.Y. TIMES (Nov. 29, 2018), <https://www.nytimes.com/2018/11/29/world/europe/hungary-orban-media.html>.

<sup>261</sup> FREEDOM HOUSE, FREEDOM OF THE PRESS 2017, *supra* note 260.

<sup>262</sup> *Id.*

<sup>263</sup> *Id.* (“Opimus Press bought the publisher Mediaworks, which had recently merged with Pannon Lapok Társasága, which controlled numerous regional newspapers. The business weekly Figyelő, was acquired by Mária Schmidt, a government ally. *Ripost*, a progovernment print tabloid, was launched in the fall of 2016. The free daily newspaper *Metropol* shut down in June 2016, after it lost a contract that had allowed its distribution at metro stations. A government-affiliated free newspaper, *Lokál*, soon emerged in its place. The license of Hungary’s only national private radio station, Class FM, was not renewed in 2016.”).

<sup>264</sup> See Krasztev, *supra* note 173, at 176 (“Independent media have been relegated to the Internet.”).

<sup>265</sup> Marius Dragomir, *The State of Hungarian Media: Endgame*, MEDIA@LSE (Aug. 29, 2017), <https://blogs.lse.ac.uk/mediapolicyproject/2017/08/29/the-state-of-hungarian-media-endgame/>. A second study demonstrated that in 2015, only 31 Hungarian news outlets took a pro-government stance, but by 2018, the number increased to 500. In addition, many of these news sources were owned by fourteen pro-Fidesz oligarchs, with Lőrinc Mészáros owning 205 media titles. Bátorfy Attila, *Infographic: Explore the Media Empire Friendly to the Hungarian Government*, ATLATSZO (May 30, 2018), <https://english.atlatszo.hu/2018/01/16/infographic-explore-the-media-empire-friendly-to-the-hungarian-government/>.

<sup>266</sup> FREEDOM HOUSE, FREEDOM OF THE PRESS 2017, *supra* note 260.

<sup>267</sup> *Id.*

Hungarian civil society is also under siege. Like Fidesz's strategy with the media, the party prevents civil society from genuinely functioning by smothering oppositional NGOs and replacing them with Fidesz allies or by discrediting them based on their affiliation with George Soros.<sup>268</sup> Scholars have argued that Hungary's "historically based tradition of a strong central state, the restored (or rather surviving) authoritarian hierarchies . . . successfully hindered the emergence of civil independence and autonomy."<sup>269</sup> Prior to the Fidesz régime, civil society traditions in Hungary were beginning to grow, for example, with the success of impartial government watchdog groups.<sup>270</sup> Many of these civil society organizations have a high resource dependency and rely on government grants to operate.<sup>271</sup> Once Orbán took power in 2010, "the system of partiality became legitimized, and grant distribution became overtly biased as a 'necessary restoration' of the national and traditional value system, which strictly excluded a number of values, critical voices, and watchdog views."<sup>272</sup> In effect, the Orbán government was able to "dismiss" opposition organizations by withdrawing funding and "replace" them with new organizations run by Fidesz allies.<sup>273</sup> Government Decree 49/2011 (III.30.) was enacted to achieve similar aims by ordering "direct provision of financial support through some of the ministries to 525 organizations, visibly recognizable from their names as NGOs that highlight national, family, and other traditional values and share these with the government parties."<sup>274</sup>

Fidesz is keen to target "opponent" NGOs that are funded through Soros's charitable contributions and Open Society Foundations. The Hungarian far-right perpetuates anti-Semitic myths claiming that Soros is part of an international conspiracy to force "globalism" on unwilling nations vis-à-vis civil society organizations.<sup>275</sup> In 2018, Fidesz passed a law informally known as the "Stop Soros" law that both imposes "a 25

---

<sup>268</sup> See *infra* Part IV.B for a more in-depth discussion of George Soros and Fidesz.

<sup>269</sup> Ágnes Kövér, *Captured by State and Church: Civil Society in Democratic Hungary*, in *THE HUNGARIAN PATIENT: SOCIAL OPPOSITION TO AN ILLIBERAL DEMOCRACY* 81, 84 (Péter Krasztev & Jon Van Til eds., 2015).

<sup>270</sup> See Zsolt Enyedi, *Democratic Backsliding and Academic Freedom in Hungary*, 16 *PERSP. ON POL.* 1067, 1070 (2018) ("These watchdogs have been critical of government practices for many years, no matter which party was in power.").

<sup>271</sup> Kövér, *supra* note 269, at 82–83.

<sup>272</sup> *Id.* at 83.

<sup>273</sup> *Id.* at 84 ("As a result, hundreds of formerly successful organizations disappeared, a process which can be followed by searching the web, where the virtual corpses of once-flourishing NGOs are scattered.").

<sup>274</sup> *Id.* at 83–84.

<sup>275</sup> See Enyedi, *supra* note 270, at 1070 ("The 'observation' that civil-society organizations backed by Soros are part of a global conspiracy is not unique to Hungary.").

percent tax on foreign donations to NGOs that back migration”<sup>276</sup> and also criminalizes the vague practice of “promoting and supporting illegal migration.”<sup>277</sup> This is another instance in which the illiberal value of absolute sovereignty appears. In a recent speech, Orbán said: “We are up against media outlets maintained by foreign concerns and domestic oligarchs, professional hired activists, troublemaking protest organizers, and a chain of NGOs financed by an international speculator, summed up by and embodied in the name George Soros.”<sup>278</sup> We can again observe the law and order rhetoric, similar to the anti-Roma beliefs discussed earlier,<sup>279</sup> but this time deployed as a weapon against humanitarian aid organizations.

The rhetoric put forth by Fidesz and Prime Minister Orbán is steeped in ethnic nationalism and easily distributed as propaganda vis-à-vis the enormous amount of media either owned by the state or by Fidesz allies. The content of these messages is comprised “of nationalism and Christian and patriarchal family values with demands for law and order.”<sup>280</sup> The Constitution’s preamble has been revised to emphasize themes of “Christian values, national history, and a united nation as a cultural and political community with state interests.”<sup>281</sup> Much of Orbán’s rhetoric invokes notions of “the family,” both with regard to valuing the ethnic Hungarian nuclear family<sup>282</sup> as well as referring to the nation-state as a

---

<sup>276</sup> Krisztina Than, *Civil Organizations in Hungary Brace for Government Crackdown on NGOs*, REUTERS (Apr. 25, 2018, 10:33 AM), <https://www.reuters.com/article/us-hungary-orban-ngos/civil-organizations-in-hungary-brace-for-government-crackdown-on-ngos-idUSKBN1HW1ZN>.

<sup>277</sup> Zack Beauchamp, *Hungary Just Passed a “Stop Soros” Law That Makes It Illegal to Help Undocumented Migrants*, VOX (June 22, 2018, 1:30 PM), <https://www.vox.com/policy-and-politics/2018/6/22/17493070/hungary-stop-soros-orban>.

<sup>278</sup> Prime Minister Viktor Orbán, Orbán Viktor’s Ceremonial Speech on the 170th Anniversary of the Hungarian Revolution of 1848 (Mar. 15, 2018). Orbán then went on to claim:

[T]here are those who want to take our country from us. Not with the stroke of a pen, has happened one hundred years ago at Trianon; now they want us to voluntarily hand our country over to others . . . . They want us to hand it over to foreigners coming from other continents, who do not speak our language, and who do not respect our culture, our laws or our way of life: people who want to replace what is ours with what is theirs. . . . Those who do not halt immigration at their borders are lost: slowly but surely they are consumed. External forces and international powers want to force all this upon us, with the help of their allies here in our country.

*Id.*

<sup>279</sup> See *supra* Part II.C (discussing how “far right parties peddled ‘law and order’ narratives that targeted the Roma”).

<sup>280</sup> András Bozóki, *The Illusion of Inclusion: Configurations of Populism in Hungary*, in THINKING THROUGH TRANSITION: LIBERAL DEMOCRACY, AUTHORITARIAN PASTS, AND INTELLECTUAL HISTORY IN EAST CENTRAL EUROPE AFTER 1989, at 275, 308 (Michal Kopeček & Piotr Wciślik eds., 2015).

<sup>281</sup> *Id.*

<sup>282</sup> See, e.g., Viktor Orbán, Prime Minister, Hung., Speech at the 29th Bálványos Summer Open University and Student Camp (July 29, 2018) (“Our second tenet is that every country has the right to

family to be protected.<sup>283</sup> Frequently, Orbán characterizes Hungary as a homogenous Christian nation under siege by refugees—and that “migration and mass population movements are bad, dangerous things which [Hungary] want[s] no part of. . . . In consequence [of migration] we will not be able to keep Hungary as it has been for the past 1,100 years.”<sup>284</sup> Orbán’s speeches conjure images of the Hungarian fictive ethnicity: that of a traditional white, Christian Magyar that embraces traditional values and rejects the moral decline of the West. This fictive Hungarian rejects the multiculturalism “imposed” on them by the West: “We must state that we do not want to be diverse and do not want to be mixed: we do not want our own colour, traditions and national culture to be mixed with those of others. . . . We do not want to be a diverse country. We want to be how we became 1,100 years ago here in the Carpathian Basin.”<sup>285</sup> This fictive ethnicity can be seen in play when a 2010 law granted citizenship rights, including the right to vote in elections, to ethnic Hungarians living in neighboring countries. Although many of these newly enfranchised ethnic Hungarians have never visited the country, they account for approximately ten percent of the electorate and vote for Fidesz at a rate of 95 percent.<sup>286</sup>

The Orbán administration is also in the process of reconstructing Hungary’s history. The much-hated Treaty of Trianon that caused Hungary to lose two-thirds of its territory after the First World War has become a rallying cry for the far-right. In 2016, the central square of Pomaz, a small town outside of Budapest, was renamed “Trianon Square” and features a monument that is a map of greater Hungary prior to its loss of territory.<sup>287</sup> Since Orbán’s 2010 election, the régime has sought to rewrite the nation’s past vis-à-vis the construction of monuments that glorify what were once

---

defend the traditional family model, and is entitled to assert that every child has the right to a mother and a father.”).

<sup>283</sup> See, e.g., Viktor Orbán, Prime Minister, Hung., “State of the Nation” Address (Feb. 10, 2019), <https://visegradpost.com/en/2019/02/11/prime-minister-viktor-orbans-state-of-the-nation-address-full-speech/> (“Thirty years after the fall of communism, on the eve of a pan-European parliamentary election, Europe finds itself in the position that we must stand up for our Hungarian identity, for our Christian identity, protect our families and communities, and also protect our freedom. . . . We have our own future which is the continuation of the lives of our parents and grandparents, the preservation of the traditions of a thousand years, the protection of our economy, our families and our Christian culture.”).

<sup>284</sup> Viktor Orbán, Prime Minister, Hung., Speech at the Annual General Meeting of the Association of Cities with County Rights (Feb. 8, 2018).

<sup>285</sup> *Id.*

<sup>286</sup> See Lili Bayer, *Viktor Orbán Courts Voters Beyond ‘Fortress Hungary’*, POLITICO (Aug. 22, 2017), <https://www.politico.eu/article/viktor-orban-courts-voters-in-transylvania-romania-hungarian-election-2018/> (“In Hungary’s 2014 election, over 95 percent of votes cast by non-domestic citizens went to Fidesz.”).

<sup>287</sup> Rosa Schwartzburg & Imre Szijarto, *The Ghosts of a Fascist Past*, JACOBIN (Jan. 26, 2019), [https://www.jacobinmag.com/2019/01/hungary-fidesz-viktor-orban-memory-history?fbclid=IwAR0PXU19MpJhYbFGAsoQxodTt-HP\\_PUjIHipRrXnZakPUj0eumF1g7dT4MA](https://www.jacobinmag.com/2019/01/hungary-fidesz-viktor-orban-memory-history?fbclid=IwAR0PXU19MpJhYbFGAsoQxodTt-HP_PUjIHipRrXnZakPUj0eumF1g7dT4MA).

considered dark times in history. These include monuments celebrating Miklós Horthy, a right-wing leader and ally of Hitler, and the 2018 replacement of a statue of Imre Nagy, the martyred leader of the 1956 Revolution, with Gyula Gömbös, a fascist Prime Minister during the interwar period.<sup>288</sup> Other monuments erected under the régime downplay Hungary's culpability in the Holocaust.<sup>289</sup> Orbán and Fidesz are in the midst of constructing a new collective mythology of the nation-state, through both legal instruments and national symbols, in order to support their régime of illiberal democracy. Their new conception of what constitutes a Hungarian and Hungarian values has been deployed against intellectuals and academia in order to delegitimize their opposition and tighten their grip on public discourse.

### E. *Emergency and Temporal Uncertainty in Hungarian Illiberal Democracy*

Orbán, like authoritarian leaders across the globe, received another opportunity to consolidate power when the novel coronavirus arrived in Hungary.<sup>290</sup> On March 30, Parliament approved the “Corona Bill,” allowing Orbán to indefinitely rule by decree with effectively unchecked power.<sup>291</sup> Under rule by decree, Orbán may bypass the national assembly completely.<sup>292</sup> The law granting rule by decree also stalled all elections and created two to five year prison sentences for those that “distort facts” or publish “false information.”<sup>293</sup> Abuse of emergency power is not

<sup>288</sup> *Id.*

<sup>289</sup> *Id.* One such monument erected under the cover of night in Budapest's Szabadság tér (Freedom Square) depicts “innocent Hungary” as Archangel Gabriel being attacked by the German imperial eagle and has been the site of a counterprotest exhibit due to the memorial's downplaying of Hungary's complacency in the Holocaust. Former Prime Minister Ferenc Gyurcsány said “with this gesture, Orbán is falsifying [history] and dishonouring all the Jewish, Roma and gay people who died in the Holocaust.” This monument prompted thirty Jewish U.S. congressmen to pen a letter to Orbán asking him to cancel the construction of the monument. Orbán responded to the criticism by claiming that the memorial is “not a Holocaust memorial but a tribute to all the victims of the German occupation.” Daniel Nolan, *German Occupation Memorial Completed Under Cover of Darkness*, BUDAPEST BEACON (July 21, 2014), <https://budapestbeacon.com/german-occupation-memorial-completed-under-cover-of-darkness/>.

<sup>290</sup> WHO Coronavirus Disease (COVID-19) Dashboard: Hungary, WORLD HEALTH ORG., <https://covid19.who.int/region/euro/country/hu> (last accessed Oct. 30, 2020).

<sup>291</sup> Imre Szijarto & Rosa Schwartzburg, *Viktor Orbán Is Using the Coronavirus Emergency to Crush Minorities*, JACOBIN MAG. (Apr. 8, 2020), <https://jacobinmag.com/2020/04/viktor-orban-coronavirus-pandemic-hungary-authoritarianism>; Guy Verhofstadt, *Is COVID-19 Killing Democracy?*, BALKAN INSIGHT (May 18, 2020), <https://balkaninsight.com/2020/05/18/is-covid-19-killing-democracy/>.

<sup>292</sup> Verhofstadt, *supra* note 291.

<sup>293</sup> Edward Szekeres, *Hungary 'No Longer a Democracy' After Coronavirus Law*, BALKAN INSIGHT (Mar. 31, 2020), <https://balkaninsight.com/2020/03/31/hungary-no-longer-a-democracy-after-coronavirus-law/>.



unfamiliar to Orbán. The government, for example, has renewed the state of emergency declared for a “crisis situation due to mass migration” eight times since the European refugee crisis began in 2015.<sup>294</sup> Subverting the division between emergency and normalcy serves two functions: to expand executive authority beyond legislature oversight and to normalize the régime. Through enduring vacillations between emergency and normalcy, “the authorities may turn to rule by decree as easily as switching on a lightbulb.”<sup>295</sup>

When authoritarian governments fabricate indefinite emergency, diminished speech and protest rights often follow. For Orbán, the Corona Bill serves as an effective mechanism to suppress speech while painting dissenters as active threats to the health of the nation. Opposition MP Tímea Szabó argued that in reality, the Corona Bill lends “a free hand to do away with even what’s left of the press and practically imprison journalists, doctors, and opposition lawmakers if we say things that you don’t like—namely, *the truth*.”<sup>296</sup> Since the “fake news” and assembly components of the Corona Bill passed, police launched about one hundred investigations against individuals, though cases have yet to make their way into court.<sup>297</sup> The Fidesz government has also instrumentalized COVID-19 precautions to limit protest and assembly. For example, protesters in a series of car demonstrations against Orbán’s rule by decree were subjected to extreme fines of up to 750,000 forints (about \$2,500).<sup>298</sup> Protesters were fined under an array of charges ranging from violating traffic laws to COVID-19 assembly restrictions.<sup>299</sup> Under illiberal rule, official sites of

---

<sup>294</sup> Edit Inotai, *Pandemic-Hit Hungary Harps on About ‘Migrant Crisis’*, BALKAN INSIGHT (Mar. 19, 2020), <https://balkaninsight.com/2020/03/19/pandemic-hit-hungary-harps-on-about-migrant-crisis/>; Akos Keller-Alant, *Hungarian Police Accused of Abusing Powers to Arrest Critics*, BALKAN INSIGHT (May 13, 2020), <https://balkaninsight.com/2020/05/13/hungarian-police-accused-of-abusing-powers-to-arrest-critics/>.

<sup>295</sup> Orsolya Lehotai, *Hungary’s Democracy Is Still Under Threat*, FOREIGN POL’Y (July 17, 2020), <https://foreignpolicy.com/2020/07/17/hungary-democracy-still-under-threat-orban-state-public-health-emergency-decree/>.

<sup>296</sup> Szijarto & Schwartzburg, *supra* note 291.

<sup>297</sup> Keller-Alant, *supra* note 294. For example, the police arrested two people in rural Hungary for criticizing Orbán’s COVID-19 response on Facebook. *Id.* One detainee, János Csóka-Szűcs, is disabled and was left without transportation home after his release. Csóka-Szűcs was forced to walk home without his cell phone or money, which were still in police custody. *Id.*

<sup>298</sup> Keller-Alant, *supra* note 294. Szijarto notes that these fines are “several times the average worker’s monthly income, and more than enough to cause serious difficulties even for relatively well-off households.” Imre Szijarto, *The Decline of Democracy in Hungary Is a Troubling Vision of the Future*, JACOBIN (June 14, 2020), <https://jacobinmag.com/6/2020/viktor-orban-hungary-democracy-covid-19>.

<sup>299</sup> Gabriella Horn, *Car Honking Protests Cancelled Due to Astronomical Fines Handed Out By Budapest Police*, ATLATSZO (May 25, 2020), <https://english.atlatszo.hu/2020/05/25/car-honking-protests-cancelled-due-to-astronomical-fines-handed-out-by-budapest-police/>.

public discourse such as media<sup>300</sup> and universities<sup>301</sup> are no longer free. The advent of COVID-19 allows illiberal leaders to restrict access to the abstract and literal town square under the façade of public health.<sup>302</sup>

This state of public health emergency also granted Fidesz the ability to quickly pass expansive measures unrelated to COVID-19. On March 31, just one day after the Corona Bill passed, Deputy Prime Minister Zsolt Semjén, introduced and passed an omnibus bill proposing fifty-seven legislative changes.<sup>303</sup> Although this bill was passed under the guise of COVID-19, in reality, it functions to further consolidate power.<sup>304</sup> First, the bill financially enriches Orbán and his allies by ordering the construction of Orbán's pet projects: "the construction of new museum buildings in one of the capital's biggest public parks"<sup>305</sup> and a new Budapest-Belgrade railway.<sup>306</sup> Second, the bill expands illiberal control over the arts by packing the theater supervisory board with government appointees.<sup>307</sup> Third, the bill attempted to suspend municipal autonomy.<sup>308</sup> Finally, the bill "forces trans people to have the same gender as they were assigned at birth and bans gender reassignment altogether."<sup>309</sup> This aspect of the bill not only functions as a literal attack against transgender Hungarians, but also as an abstract attack against their identities.<sup>310</sup> Denying trans existence as deviant to the traditional heteropatriarchal norms of the Hungarian nation-state serves to Otherize from "the people" of the imagined community.<sup>311</sup> This dual role of the new law serves two purposes: cast aside trans people as non-members of the community and routinely out them in regular aspects of national life.<sup>312</sup>

---

<sup>300</sup> See *supra* notes 258–70 and accompanying text (analyzing media freedom under the Orbán régime).

<sup>301</sup> *Infra* Part III.

<sup>302</sup> Szijarto & Schwartzburg, *supra* note 291.

<sup>303</sup> *Id.*

<sup>304</sup> *Id.*

<sup>305</sup> *Id.*

<sup>306</sup> *Id.*

<sup>307</sup> Szijarto & Schwartzburg, *supra* note 291. Régime control over the arts is directly related to the anti-intellectual aspect of illiberal democracies. By restricting the arts and academic freedom, authoritarians seek to manipulate and regulate national discourse. See *supra* Part I.C.3.

<sup>308</sup> This portion of the omnibus bill was reversed after extensive outcry from the opposition, "who made significant gains in the last municipal contests in fall 2019, including [Gergely] Karácsony's election in Budapest" as mayor. Szijarto & Schwartzburg, *supra* note 291.

<sup>309</sup> *Id.*

<sup>310</sup> See Szijarto, *supra* note 298 ("Since Hungarian society at large is anything but trans accepting, this move is not only an attack on trans people's right to their 'identities' in an abstract sense. It is likely to turn regular interactions with society into recurring rituals of humiliation.").

<sup>311</sup> *Supra* Part I.C.1.

<sup>312</sup> See Szijarto, *supra* note 298 ("This means outing trans people not only to employers and landlords but even to receptionists and cashiers, whenever they use a credit card.").

On June 16, proposed legislation ending Orbán's rule by decree passed unopposed.<sup>313</sup> The Orbán government issued about one hundred decrees since March, many of which are completely unrelated to COVID-19 measures.<sup>314</sup> This brief reprieve would not last for long: Keeping in line with tradition, Parliament again pushed the nation into a state of emergency. This time, Parliament passed legislation allowing Orbán to enter a "state of medical emergency" and revert back to rule by decree absent a mandated end date.<sup>315</sup> Under a state of health emergency, the government may restrict fundamental rights for a maximum of six months, but critics argue that they may be extended indefinitely in practice.<sup>316</sup>

The Hungarian Helsinki Committee wrote that the formal June 20 end to rule by decree "is nothing but an optical illusion: if the Bills are adopted in their present form, that will allow the government to again rule by decree for an indefinite period of time, this time without even the minimal constitutional safeguards."<sup>317</sup> This looming state of emergency destabilizes any normalcy that existed before the régime. Crisis legitimizes authoritarian control. "When the coronavirus arrived in Hungary, Orbán used it to illustrate that he was already fully in control of his system."<sup>318</sup> These states of emergency are cast as an offensive against the perceived invasions of disease or the Other<sup>319</sup> into the imagined community.<sup>320</sup> The aim is that the régime citizens will more readily accept losing civil liberties—the "liberalism"<sup>321</sup> of the old liberal democratic order—and the tightening authoritarian grasp will begin to feel natural. Orbán's illiberal democracy is emblematic of this threat. Suspended in the strings of state

<sup>313</sup> Lehotai, *supra* note 295.

<sup>314</sup> Shaun Walker, *Hungarian Government to End Orbán's Rule-By-Decree Legislation*, GUARDIAN (May 26, 2020), <https://www.theguardian.com/world/2020/may/26/hungarian-government-to-end-orbans-rule-by-decree-legislation-emergency-coronavirus>.

<sup>315</sup> *Never-Ending Story?*, HUNGARIAN HELSINKI COMMITTEE (May 27, 2020), <https://www.helsinki.hu/en/never-ending-story/>. The legislation also amended the Disaster Management Act to authorize "the government [to] order any measures it deems necessary if the measures previously specified by parliament are inadequate." *Id.*

<sup>316</sup> *Id.*

<sup>317</sup> *Id.*

<sup>318</sup> Nick Schiffrin & Layla Quran, *How Authoritarianism Has Spread Since the Coronavirus Pandemic Began*, PBS (Aug. 4, 2020), <https://www.pbs.org/newshour/show/how-authoritarianism-has-spread-since-the-coronavirus-pandemic-began>.

<sup>319</sup> Edward Said's theory of Orientalism explains that European societies Otherize non-European societies in order to legitimize their imperialist projects. EDWARD SAID, *ORIENTALISM* 7 (1994) ("Orientalism is never far from what Denys Hay has called the idea of Europe, a collective notion identifying 'us' Europeans as against all 'those' non-Europeans, and indeed it can be argued that the major component in European culture is precisely what made that culture hegemonic both in and outside Europe: the idea of European identity as a superior one in comparison with all the non-European peoples and cultures.").

<sup>320</sup> See *supra* Part I.C.1 (applying Benedict Anderson's *Imagined Communities* to illiberal democracy).

<sup>321</sup> See *supra* notes 65–73 (explaining liberal democracy and constitutional liberalism).

emergency, the régime order reifies amorphous crises into accepted domination. As such, after the pandemic, we emerge only to find that the system had long collapsed, replaced by hollow illiberal democracies.

### III. THE INTERSECTION OF ANTI-INTELLECTUALISM, ILLIBERAL DEMOCRACY, AND ACADEMIC FREEDOM

#### A. *Lex CEU: A Background*

Central European University (CEU) is a highly-ranked graduate research university in Budapest founded in 1991 by George Soros, Václav Havel, and other intellectual members of the democratic opposition as a direct response to the dissolution of the Soviet Union.<sup>322</sup> Their vision was to establish a university dedicated to promoting the values of open society and democracy as envisioned in Karl Popper's philosophy.<sup>323</sup> Since its founding, Central European University has been regarded as a liberal intellectual bastion of Central Europe and Hungary. Central European University is accredited both in Hungary and in the state of New York and leads all other Hungarian universities in receiving European research grants. As such, it has become the target of Prime Minister Orbán's attacks on freedom of thought, academic freedom, and liberal opposition.

In April of 2017, the Hungarian parliament adopted amendments to the existing Act CCIV of 2011 on National Higher Education ("Lex CEU").<sup>324</sup> The new criteria for foreign universities operating in Hungary directly targeted only Central European University.<sup>325</sup> The most onerous component of Lex CEU is the requirement of an international agreement between Hungary and the university's country of origin (in CEU's case, America).<sup>326</sup> This requirement is particularly problematic as it "practically means that the right to conduct educational activities will no longer depend on professional criteria (e.g., on the decision of accreditation boards), but on the preferences of the government."<sup>327</sup> This law also requires that all foreign-accredited universities provide higher education services in their country of origin and "restricts the possibility for non-European universities to enter into cooperation with Hungarian universities."<sup>328</sup>

---

<sup>322</sup> *Our Mission*, CENT. EUR. U., <https://www.ceu.edu/about/our-mission> (last visited Jan. 12, 2020); *History*, CENT. EUR. U., <https://www.ceu.edu/about/history> (last visited Jan. 12, 2020).

<sup>323</sup> *Our Mission*, *supra* note 322.

<sup>324</sup> This law quickly gained the nickname of "Lex CEU" because Central European University (CEU) is the only higher education institution affected by the law.

<sup>325</sup> Enyedi, *supra* note 270, at 1067.

<sup>326</sup> *Id.*

<sup>327</sup> *Id.*

<sup>328</sup> Griff Witte, *Amid Illiberal Revolution in Hungary, a University with U.S. Roots Fights to Stay*, WASH. POST (Sept. 3, 2018, 3:08 PM), <https://www.washingtonpost.com/world/europe/amid-illiberal->

Finally, Lex CEU “eliminates exemptions for work permits and requires that the name of the university differs clearly from the name of already registered universities even in foreign languages.”<sup>329</sup> If a university fails to comply with any aspect of Lex CEU, it will lose its license to operate.<sup>330</sup>

Although the stipulations of this amendment appear to be neutral, they disparately impact Central European University. CEU is an American-chartered university that is connected to the Hungarian higher education system via a legal entity called Közép-európai Egyetem.<sup>331</sup> The university has a substantial population of non-European professors, primarily from America and Canada, that relied on the recently eliminated work permit exception.<sup>332</sup> In order to fulfil the requirement of an international agreement, the successful negotiation of *two* treaties—one between Hungary and New York state and the other between Hungary and the United States federal government—were now required by January 1, 2018, giving CEU less than a year to comply.<sup>333</sup>

The Orbán government faced immense backlash in response to this absurd law. On April 2, 2017, approximately 10,000 people marched throughout Budapest to parliament in support of CEU.<sup>334</sup> Then, on April 4, parliament voted in favor of Lex CEU in spite of a petition to the government with over 30,000 signatures from 134 different countries.<sup>335</sup> Finally, on April 9, an estimated 80,000 demonstrators took to the streets in peaceful protest, with hundreds of international universities and over twenty Nobel laureates expressing solidarity.<sup>336</sup> Central European University worked with the state of New York and launched an academic exchange program at Bard College in Annandale-on-Hudson, New York. Despite complying with Lex CEU and successfully negotiating a mutually accepted draft agreement with Governor Cuomo, the Orbán government refused to ratify the treaty. This left the university in legal limbo by extending the deadline of compliance to January 1, 2019, prompting international condemnation.<sup>337</sup> Without approval by the Orbán government,

---

revolution-in-hungary-a-university-with-us-roots-fights-to-stay/2018/09/03/7061771c-a547-11e8-a656-943eefab5daf\_story.html?utm\_term=.9a2aa60e3ca5; Enyedi, *supra* note 270, at 1067.

<sup>329</sup> Enyedi, *supra* note 270, at 1067.

<sup>330</sup> *Id.*

<sup>331</sup> *Id.*

<sup>332</sup> *Id.*

<sup>333</sup> *Id.* at 1068.

<sup>334</sup> *CEU Thanks Organizers, Supporters of Today's Event*, CENT. EUR. U. (Apr. 2, 2017), <https://www.ceu.edu/node/17917>.

<sup>335</sup> *Timeline of Events*, CENT. EUR. U., <https://www.ceu.edu/istandwithceu/timeline-events> (accessed Mar. 23, 2019).

<sup>336</sup> Enyedi, *supra* note 270, at 1068.

<sup>337</sup> *Id.* (“Widespread public and international criticism of the law followed. During visits to CEU at the height of the crisis, both former German President Joachim Guack and Nobel Prize-winning writer Mario Vargas Llosa, condemned the law. Many other public figures contributed to the growing

the university has been forced to shutter its United States-accredited operations in Budapest and move its students to a satellite campus in an old state hospital on the outskirts of Vienna, Austria for the fall 2019 academic year.<sup>338</sup>

### B. *Anti-Intellectualism and Illiberal Democracy: A Hungarian Reprise*

Illiberal democracies, our modern authoritarianism, are deeply anti-intellectual. As discussed in Part I, critical intellectual discourse poses a direct threat to authoritarian control.<sup>339</sup> Illiberal leaders like Orbán that espouse populism in their political platform are wont to espouse anti-intellectual rhetoric. Populism can be defined as “[a] thin-centered ideology that considers society to be ultimately separated into two homogenous and antagonistic groups, ‘the pure people’ and ‘the corrupt elite,’ and which argues that politics should be an expression of the *volonté générale* (general will) of the people.”<sup>340</sup> Historically, the intelligentsia have been classified as “the elite” and have become the scapegoat for the ire of populist leaders. Other than the general inaccessibility of expensive university degrees to the masses, academia and its intelligentsia are targeted to suppress anti-régime discourse and preemptively neutralize the opposition. This line of thought makes Central European University an ideal target for Orbán. CEU is an institution that was founded with a

---

worldwide protest. On April 19, nine American senators, including John McCain and Chuck Schumer, sent a letter to Prime Minister Viktor Orbán, arguing that the legislation threatens academic freedom and calling on him to work with CEU to find a solution. The most important political statements came from the European People’s Party (EPP, the center-right bloc in the European Parliament) and the spokesperson for German Chancellor Angela Merkel. The U.S. State Department also expressed its clear support for CEU and for academic freedom, and rejected the threat to an American university abroad. . . . Parallel to American diplomatic efforts, the European Commission condemned the law as a violation of EU regulations and core European values, including academic freedom.”); *see also* Letter from Sen. Dick Durban, Sen. Jeanne Shaheen, Sen. Chris Murphy, Sen. Kirsten Gillibrand, Sen. Amy Klobuchar, Sen. Ron Wyden, Sen. Charles E. Schumer, Sen. Cory A. Booker, Sen. Edward J. Markey, Sen. Bernard Sanders, U.S. Senate, to Ambassador David B. Cornstein (Dec. 20, 2018) (on file with Sen. Dick Durban) (supporting CEU and academic freedom in Hungary); Letter from Dr. Laura Schultz et al., Professor Cognitive Sci., MIT, to H.E. Réka Szemerkényi, Ambassador of Hung. to the U.S. et al. (Mar. 31, 2017) (on file with Central European University) (letter from 1061 cognitive scientists in support of Central European University); *Statements of Support*, CENT. EUR. U., <https://www.ceu.edu/istandwithCEU/support-statements> (accessed Mar. 23, 2019).

<sup>338</sup> *CEU Forced Out of Budapest: To Launch U.S. Degree Programs in Vienna in September 2019*, CENT. EUR. U. (Dec. 3, 2018), <https://www.ceu.edu/article/2018-12-03/ceu-forced-out-budapest-launch-us-degree-programs-vienna-september-2019>; *see also* Rosa Schwartzburg, *School’s Out Forever*, SLATE (Feb. 20, 2019), [https://slate.com/news-and-politics/2019/02/authoritarians-education-universities-hungary-brazil-populism.html?via=rss\\_socialflow\\_facebook&fbclid=IwAR3kYxKbgCTBPcNH04be6fici61661ewsRhOgwXvdcCn2d\\_n3FWetNn5VBk0](https://slate.com/news-and-politics/2019/02/authoritarians-education-universities-hungary-brazil-populism.html?via=rss_socialflow_facebook&fbclid=IwAR3kYxKbgCTBPcNH04be6fici61661ewsRhOgwXvdcCn2d_n3FWetNn5VBk0) (describing CEU’s ousting to Vienna).

<sup>339</sup> ARENDT, *supra* note 53, at 339.

<sup>340</sup> CAS MUDDÉ, *THE OXFORD HANDBOOK OF POPULISM* 29 (Cristóbal Rovira Kaltwasser et al. eds., 2017).

dedication to open society, with programs such as their Nationalism and Gender Studies departments<sup>341</sup> that directly, and loudly, challenge illiberal democracy. Universities are particularly dangerous to authoritarians because their purpose “is [to] produce knowledge that is often critical of the established ways of doing things . . . . And in the social sciences it’s quite dangerous . . . because the knowledge that’s produced is calling into question the habits and ‘ordinary ways’ that we go about doing things.”<sup>342</sup> Likewise, the government-controlled press launched a campaign against the “intellectual elite” that attacked philosophers associated with the Georg Lukács School, like Ágnes Heller, by falsely claiming that they had received overly generous government research grants.<sup>343</sup> Authoritarians rely on the closure of critical discourse to create a one-dimensional arena of thought that is uncritical of the régime in order to successfully quell opposition and maintain societal control.

Anti-intellectualism, especially in Europe, is inherently tied to anti-Semitism. Arendt’s history of anti-Semitism in *Origins of Totalitarianism* notes that European nation-states were hostile to Jewish intellectuals in particular as an attempt to prevent Jewish assimilation in the nineteenth century.<sup>344</sup> But by the early twentieth century, the most notable Central European intellectuals were Jewish: Sigmund Freud, Edmund Husserl, Gustav Mahler, Franz Kafka, and so on.<sup>345</sup> George Soros, the primary founder of Central European University, is a Hungarian-born Jewish financier and billionaire and has been the subject of anti-Semitic conspiracy theories by the Hungarian far-right. Most recently, Fidesz has launched a taxpayer-funded<sup>346</sup> campaign against Soros and the European Union that includes peddling the myth that Soros was “allegedly

---

<sup>341</sup> *Nationalism Studies Program*, CENT. EUR. U., <https://courses.ceu.edu/unit/nationalism> (accessed Jan. 12, 2020); *Department of Gender Studies*, CENT. EUR. U., <https://courses.ceu.edu/unit/gender> (accessed Jan. 12, 2020).

<sup>342</sup> Schwartzburg, *supra* note 338 (quoting interview with Joan Wallach Scott, Professor Emerita at the School of Social Science at the Institute for Advanced Study).

<sup>343</sup> Bozóki, *The Illusion of Inclusion*, *supra* note 280, at 306.

<sup>344</sup> ARENDT, *supra* note 53, at 32 (“From then on, the nation-state, equipped with conservative arguments, drew a distinct line between Jews who were needed and wanted and those who were not. Under the pretext of the essential Christian character of the state—what could have been more alien to the enlightened despots!—the growing Jewish intelligentsia could be openly discriminated against without harming the affairs of bankers or businessmen. This kind of discrimination which tried to close the universities to Jews by excluding them from the civil services had the double advantage of indicating that the nation-state valued special services higher than equality, and of preventing, or at least postponing, the birth of a new group of Jews who were of no apparent use to the state and even likely to be assimilated into society.”). Ultimately, this project was a failure: “What the nation-state had once feared so much, the birth of a Jewish intelligentsia, now proceeded at a fantastic pace.” *Id.* at 52.

<sup>345</sup> Kundera, *The Tragedy of Central Europe*, *supra* note 143, at 35.

<sup>346</sup> Lili Bayer, *Hungary Launches Campaign Targeting Jean-Claude Juncker*, POLITICO (Feb. 18, 2019), <https://www.politico.eu/article/hungary-launches-campaign-targeting-jean-claude-juncker-george-soros/>.

responsible for the wave of migrants to Europe.”<sup>347</sup> Deputy Prime Minister Zsolt Semjén claimed in a recent radio interview that: “[T]he Soros-type extreme liberalism which hates Christian traditions and, if possible even more, nation states,”<sup>348</sup> is dangerous to Hungarians, and he went on to conclude that “leftist and liberal parties needed to import millions of foreigners in order to defeat their nationalist opponents and, in cooperation with immigrant Islamic forces, to rule the nations of Europe.”<sup>349</sup> Jobbik spokesman Ádám Mirkóczki pushed this narrative even further by claiming that high-ranking Fidesz politicians, including Orbán, are implicated in this conspiracy theory due to the Soros funding that they received for their educations.<sup>350</sup> The anti-Soros conspiracy theory directly entered the parliamentary debate on Lex CEU when “the minister responsible for education stated that ‘we are committed to use all legal means at our disposal to stop pseudo-civil society spy groups such as the ones funded by George Soros.’”<sup>351</sup> Although Central European University would still have been a likely target of the Fidesz régime, Soros’s involvement threw fuel into the fire.

### C. *Illiberal Legal Challenges to Academic Freedom*

Since Prime Minister Viktor Orbán was elected in 2010, the Fidesz government has centralized education and enacted a wide array of laws that shrink academic freedom in the country.<sup>352</sup> Fidesz has pushed through a gradual overhaul of the public education system, slowly growing government influence over school curricula.<sup>353</sup> In 2014, legislation was adopted that permits government-appointed chancellors to make significant financial decisions at public universities,<sup>354</sup> and in some cases, the Minister of Education has even “imposed his own candidate for rector of universities and political appointees without any academic record were promoted to professorship at state-controlled universities.”<sup>355</sup> Likewise, a slash in government funding for the Hungarian Academy of Science and

---

<sup>347</sup> Enyedi, *supra* note 270, at 1069.

<sup>348</sup> *Id.*

<sup>349</sup> *Id.* (summarizing the Semjén interview).

<sup>350</sup> *Fidesz Gets a Reminder Who Was Funded By George Soros*, JOBBIK (Jan. 27, 2017), [https://www.jobbik.com/fidesz\\_gets\\_a\\_reminder\\_who\\_was\\_funded\\_by\\_george\\_soros](https://www.jobbik.com/fidesz_gets_a_reminder_who_was_funded_by_george_soros) (“He added that when Szilárd Németh identified Soros’ protégés as organizations and individuals to be ‘cleaned away’, the Fidesz politician forgot to mention that from 1984 through 1994 the Hungarian-born US businessman funded such persons as PM Viktor Orbán and several other individuals with ties to the current government or the ruling party, including Tamás Deutsch, József Szájer, Zsolt Németh, István Stumpf, Zoltán Cséfalvay, József Pálincás, János Áder or László Kövér.”).

<sup>351</sup> Enyedi, *supra* note 270, at 1070.

<sup>352</sup> *Id.* at 1071.

<sup>353</sup> FREEDOM HOUSE, *FREEDOM IN THE WORLD 2018: HUNGARY*, *supra* note 221.

<sup>354</sup> *Id.*

<sup>355</sup> Enyedi, *supra* note 270, at 1071.



general reallocation of significantly less funds for academic research institutions threatens institutional closures in the near future.<sup>356</sup> Fidesz leveraged the Corona Bill to forcibly transfer control over the University of Theater and Film Arts (SZFE) to a private foundation, prompting students to occupy campus buildings, professors to resign, and tens of thousands to protest in Budapest.<sup>357</sup> The battle for academic freedom in Hungary has manifested itself as a culture war between the nationalistic illiberal democracy and the liberal intelligentsia.

At the same time that the Orbán government refused to acknowledge Central European University's compliance with Lex CEU, Prime Minister Orbán signed a decree revoking accreditation and funding for gender studies departments, effectively banning the discipline.<sup>358</sup> This decree only targets programs from two universities in Hungary: Central European University and Eötvös Loránd University.<sup>359</sup> Effective immediately, the accreditation of all gender studies programs has been revoked, although the government is allowing currently enrolled students to finish their programs. Anti-intellectual and misogynistic rhetoric surrounding the gender studies ban date back to Fidesz's seizure of power. In 2015, László Kövér, one of the founders of Fidesz, stated:

We don't want the gender craziness. We don't want to make Hungary a futureless society of man-hating women, and feminine men living in dread of women, and considering families and children only as barriers to self-fulfillment ... And we would like if our daughters would consider, as the

---

<sup>356</sup> Szabó András Péter, *Greetings from the Sinking Boat of Hungarian Academia*, INDEX (trans. Zoltán Kovács, Feb. 8, 2019), [https://index.hu/english/2019/02/08/greetings\\_from\\_the\\_sinking\\_boat\\_of\\_hungarian\\_academia/](https://index.hu/english/2019/02/08/greetings_from_the_sinking_boat_of_hungarian_academia/).

<sup>357</sup> See *supra* Part II.F (discussing the Corona Bill and Hungarian states of emergency). SZFE is a prestigious 155-year-old university that has educated prominent Hungarian artists. In March, the Fidesz government passed the Corona Bill, which included a provision “that transferred ownership of the public university to a private foundation. The government also appointed a new board of trustees—actions that raised fears that the university, long a target of the government, will be forced to hew more closely to Mr. Orbán’s nationalistic and conservative vision for Hungary.” Benjamin Novak, *Student Blockade Protests Viktor Orbán’s Reach at a Top Arts University*, N.Y. TIMES (Sept. 6, 2020), <https://www.nytimes.com/2020/09/06/world/europe/hungary-students-blockade-orban.html>. At the time of writing, the students have refused to end their occupation “until [they] are guaranteed the university’s autonomy.” *Hungarian Students, Artists Protest Government’s Takeover of Famed Film School*, RADIO FREE EUR./RADIO LIBERTY’S HUNGARIAN SERV. (Oct. 19, 2020), <https://www.rferl.org/a/hungarian-students-artists-protest-government-s-takeover-of-famed-film-school/30901261.html> (internal quotations omitted).

<sup>358</sup> Anna Zsubori, *Gender Studies Banned at University—the Hungarian Government’s Latest Attack on Equality*, CONVERSATION (Oct. 9, 2018), <http://theconversation.com/gender-studies-banned-at-university-the-hungarian-governments-latest-attack-on-equality-103150>.

<sup>359</sup> *Id.*

highest quality of self-fulfillment, the possibility of giving birth to our grandchildren.<sup>360</sup>

Kövéér went even further earlier that year to argue that “‘genderism’ is ‘an intellectual founding of such a human experiment that is nothing better than, let’s say, eugenics in Nazi times.’”<sup>361</sup> It is obvious that the government’s decision to ban certain academic fields—and compare them to Nazism—stems not from a genuine desire to improve the academic endeavors of universities but to control freedom of thought based on political ideology. Nationalist populist movements tend to yearn for a return to “traditional” society and reject modernism, and as such, embrace “traditional” gender roles.<sup>362</sup> Hungary is a very patriarchal country invested in traditional family structures,<sup>363</sup> and Fidesz’s politics “signals opposition to the moral-cultural transformation of developed societies.”<sup>364</sup> The gender studies ban is emblematic of the culture war between the Fidesz government and Central European University. The ban primarily affects Central European University, an institution that is one of the most diverse universities in the world<sup>365</sup> and embodies the spirit of open society, multiculturalism, and cosmopolitanism. Accordingly, the ban, and its targeting of CEU, functions as the Fidesz government’s rebuke of the moral decline<sup>366</sup> of the West. Academic institutions in other illiberal democracies have experienced similar treatment, especially with regard to teaching gender studies and other related subjects.<sup>367</sup> This war on academia

---

<sup>360</sup> *Id.* (alteration in original).

<sup>361</sup> *Id.*

<sup>362</sup> See Joane Nagel, *Masculinity and Nationalism: Gender and Sexuality in the Making of Nations*, 21 *ETHNIC & RACIAL STUD.* 242, 251–53 (1998) (discussing the distinct roles that men and women tend to play in nationalist narratives).

<sup>363</sup> See *Religious Belief and National Belonging in Central and Eastern Europe*, PEW RES. CTR. (May 10, 2017), <https://www.pewforum.org/2017/05/10/social-views-and-morality/> (finding that Hungarians tend to hold conservative views on LGBTQ+ and gender issues: 54 percent believe that homosexuality should not be accepted by society, 27 percent favor same sex marriage, and 77 percent believe that women have a social responsibility to bear children).

<sup>364</sup> Enyedi, *supra* note 270, at 1069.

<sup>365</sup> *CEU Named #2 Most International University in the World by THE*, *CENT. EUR. U.* (Apr. 27, 2017), <https://www.ceu.edu/article/2017-04-27/ceu-named-2-most-international-university-world>.

<sup>366</sup> See Orbán, *supra* note 282 (“[I]n liberal Europe being European means nothing at all: it has no direction, and it is simply form devoid of content.”).

<sup>367</sup> In Russia, the European University at St. Petersburg lost its license to teach for two years due to an administrative ruling that began with a complaint by MP Vitaly Milonov, who was the key author of the homophobic 2013 “gay propaganda” bill. Milonov said of the university: “[An alleged student complaint] was the teaching of gender studies at the school. I personally find that disgusting, it’s fake studies, and it may well be illegal.” Fred Weir, *Why Is Someone Trying to Shutter one of Russia’s Top Private Universities?*, *CHRISTIAN SCI. MONITOR* (Mar. 28, 2017), <https://www.csmonitor.com/World/Europe/2017/0328/Why-is-someone-trying-to-shutter-one-of-Russia-s-top-private-universities>. In Turkey, at least 146 academics face individual and separate trial hearings for signing a 2016 petition condemning Turkey’s treatment of Kurds on the southeastern border and calling to resume peace talks with the Kurdistan Workers’ Party. This petition was initially signed by a group of 1128 academics

is not only ideological, but wrapped up in the very legal institutions of Hungary. Further, Lex CEU flies in the face of the Hungarian Constitution, which purports to protect academic freedom from government intrusion, and is a reminder of the hollowness of democratic institutions in illiberal democracies.<sup>368</sup>

Orbán and Fidesz's attacks on Central European University exemplify illiberal democratic principles in action. As has been demonstrated, the majoritarian Fidesz partocracy quickly seized control of all government organs, including purportedly independent ones, in order to bypass all opposition checks in every stage of the democratic process. Fidesz's reach has extended beyond democratic institutions to control virtually all mainstream media and restrict the operation of civil society to further insulate the régime from opposition.<sup>369</sup> In addition to manipulating the law, they are controlling societal norms to produce the Magyar fictive ethnicity and suppress non-conservative discourse.<sup>370</sup> At the same time, Hungarian history is actively being rewritten in a manner that glorifies past fascist leaders and stokes anger over the perceived injustices of the Treaty of Trianon.<sup>371</sup> Orbán invokes principles of absolute sovereignty when he

---

known as Academics for Peace. A number of these signatories are facing between one and five years' imprisonment for "spreading terrorist propaganda." Over 360 academics have been fired from their jobs and barred from teaching. *Turkey: Academics on Trial for Signing Petition*, HUM. RTS. WATCH (Dec. 5, 2017, 12:00 AM), <https://www.hrw.org/news/2017/12/05/turkey-academics-trial-signing-petition>. On January 24, 2019, Professor Yonca Demir of Bilgi University was sentenced to three years imprisonment, the longest prison sentence to date for members of Academics for Peace. John K. Wilson, *Turkey's Ongoing Attack on Academic Freedom*, ACADEME BLOG (Jan. 24, 2019), <https://academeblog.org/2019/01/24/turkeys-ongoing-attack-on-academic-freedom/>. In Brazil, President Jair Bolsonaro has "dismantl[ed] the Education Ministry, threaten[ed] to send the military to take over public schools, and publish[ed] stringent new textbook guidelines. . . . Bolsonaro has made ridding the Brazilian education system of 'Marxist rubbish' one of his administration's top priorities, with students being encouraged to report on teachers who attempt to 'indoctrinate' them." Bolsonaro's new Education Minister, Ricardo Vélez Rodríguez, swore in his inauguration speech to fight the "aggressive promotion of gender ideology." Schwartzburg, *supra* note 338. The Thai junta has detained American academics and maintains a watchlist of academics and researchers. Suluck Lamubol, *Junta Steps Up Harassment of International Academics*, U. WORLD NEWS (Mar. 1, 2019), <https://www.universityworldnews.com/post.php?story=20190301083420506>. In Egypt, Alexandria University has required that thesis proposals for masters and doctoral students adhere to the government's incomplete Vision 2030 development plan. Ashraf Khaled, *Concern Over Freedoms as University Curbs Thesis Topics*, U. WORLD NEWS (Dec. 5, 2018), <https://www.universityworldnews.com/post.php?story=20181205120855294>.

<sup>368</sup> "Article X: (1) Hungary shall ensure the freedom of scientific research and artistic creation, the freedom of learning for the acquisition of the highest possible level of knowledge and, within the framework laid down in an Act, the *freedom of teaching*. . . . (3) Hungary shall protect the scientific and artistic freedom of the Hungarian Academy of Sciences and the Hungarian Academy of Arts. *Higher education institutions shall be autonomous in terms of the content and the methods of research and teaching*; their organization shall be regulated by an Act." HUNGARIAN CONSTITUTION, *supra* note 259 (emphasis added).

<sup>369</sup> See *supra* notes 260–70 and accompanying text.

<sup>370</sup> See *supra* notes 278–82 and accompanying text.

<sup>371</sup> See *supra* notes 290–92 and accompanying text.

treats Central European University as yet another invasion of the West and its “degenerate” ideology and paints Hungary as a nation that is still being constantly invaded, this time by liberals under George Soros’s watch.<sup>372</sup> Orbán has also deployed the “dictatorship of law” when passing Lex CEU as a direct assault on the “elite intelligentsia” that, through critical discourse and academia, threaten the legitimacy of the illiberal democracy.<sup>373</sup> Far-right nationalist and anti-intellectual rhetoric characterizes Central European University and its intellectuals as a danger to “the people” of Hungary, pitting them against each other in a culture war that CEU did not sign up for.

While the situation in Hungary is grim, action can be taken. Individuals can continue to protest Fidesz’s régime and draw international attention to Hungary. Right now, “activists on the ground are setting up underground education lectures and organizing queer theory readings and poetry nights in people’s living rooms and basement bars.”<sup>374</sup> Independent media and the academic community can further support democracy and academic freedom by publicizing the situation in Hungary and supporting communities and individuals under threat. Likewise, the international community can pressure illiberal democracies through diplomatic measures. Political and economic unions such as the European Union can adopt measures condemning the actions of illiberal democratic states and impose sanctions on noncompliant governments. On September 12, 2018, the European Union voted to pursue disciplinary action against Hungary under Article 7 of the European Union Charter. This is the first time that the EU has pursued action invoking Article 7, which lays out the ways that EU bodies can act if a member state violates the core values of the European Union.<sup>375</sup> On March 20, 2019, the European People’s Party<sup>376</sup> suspended Fidesz from the party in response to their campaign attacking Soros and European Commission President Jean-Claude Juncker, as well as its actions against Central European University. EPP President Joseph Daul stated: “We cannot compromise on democracy, rule of law, freedom of press, academic freedom or minorities’ rights. And anti-EU rhetoric is

---

<sup>372</sup> See *supra* notes 278–82, 349–54 and accompanying text.

<sup>373</sup> See *supra* Part III.B.

<sup>374</sup> Rosa Schwartzburg, *I’m a Gender Studies Student in Hungary & Americans Need to Know What’s Happening Here*, BUSTLE (Jan. 7, 2018), <https://www.bustle.com/p/im-a-gender-studies-student-in-hungary-americans-need-to-know-whats-happening-here-15574938>.

<sup>375</sup> Maïa de la Baume & Ryan Heath, *Parliament Denounces Hungary’s Illiberalism*, POLITICO, <https://www.politico.eu/article/european-parliament-approves-hungary-censure-motion/> (last updated Apr. 19, 2019, 1:39 AM).

<sup>376</sup> The European People’s Party is a center-right transnational organization founded by Christian democratic parties in Europe. The EPP is currently the largest party in the European Parliament, the European Council, and the European Commission. *History*, EUR. PEOPLE’S PARTY, <https://www.epp.eu/about-us/history/> (accessed Mar. 23, 2019).

unacceptable. The divergences between EPP and Fidesz must cease.<sup>377</sup> Fidesz must end the anti-Junker and Soros campaign, solve the legal dispute over Central European University, and “recognize the damage it caused and refrain from similar action”<sup>378</sup> in order to regain full membership to the EPP. As a result from pressure by the EPP, on March 20, Orbán and Prime Minister Söder of Bavaria exchanged letters supporting a partnership between CEU and the Technical University of Munich on the condition that CEU be allowed to remain in Budapest.<sup>379</sup> At the time of writing, Orbán has yet to give CEU a legal guarantee to remain in operation and “an international agreement guaranteeing the freedom of CEU to operate in Budapest as a US degree granting institution”<sup>380</sup> is needed before celebration. If Central European University is ultimately successful, it will be a testament to the sway of influence that European institutions still hold over this illiberal democracy. Even if Central European University is allowed to remain in Budapest, Hungary’s democratic institutions are still under siege and many other liberties enshrined under constitutional liberalism are still threatened.

#### CONCLUSION

We are currently in the throes of Huntington’s third reverse wave. After decades of democratic prosperity, global democracy has entered a “decade of decline.”<sup>381</sup> Between 2006 and 2016, Freedom House’s *Freedom in the World* report discovered that 105 countries suffered net declines in their scores, while 61 demonstrated improvement. Given these grim statistics, we have undoubtedly entered Huntington’s third reverse wave. Based on the first and second reverse waves, Huntington concluded that each reverse wave will give rise to a new form of authoritarianism. Illiberal democracy is that form of authoritarianism.

Illiberal democracy differs from other forms of authoritarianism due to the fact that some semblance of a democratic system remains. In many cases, the authoritarian government comes to power through the democratic process. However, once elected, they quickly consolidate control, and greatly weaken democracy and the safeguards of civil liberties

---

<sup>377</sup> Maïa De La Baume & Lili Bayer, *Hungary’s Orbán Clings on to Europe’s Power Center*, POLITICO (Mar. 20, 2019), <https://www.politico.eu/article/orbans-party-suspended-from-european-peoples-party/>.

<sup>378</sup> *Id.*

<sup>379</sup> *CEU Needs Legal Guarantees If the Bavarian Solution Is to Work*, CENT. EUR. U. (Mar. 20, 2019), <https://www.ceu.edu/article/2019-03-20/ceu-needs-legal-guarantees-if-bavarian-solution-work>.

<sup>380</sup> *Id.*

<sup>381</sup> See PUDDINGTON, *supra* note 74, at 5 (“According to *Freedom in the World*, the crucial indicators of democracy experienced setbacks in each of the 10 years in question. In all, 105 countries suffered net declines, while 61 showed some measures of improvement. The decade marked the longest democratic slump of its kind in more than 40 years of Freedom House analysis.”).

as guaranteed by constitutional liberalism. These régimes tend to value majoritarianism, absolute sovereignty, and the dictatorship of law while embracing nationalism and anti-intellectualism. When illiberal democracies operate under these values that run counter to the very nature of democracy, their democratic institutions become hollow and operate in name only. At the same time, illiberal democracies fail to protect the individual with liberties typically ascribed to constitutional liberalism and instead value the perceived needs of the community. This definition of community, however, is guided by nationalism. Illiberal democratic leaders determine which citizens are worthy of belonging to their imagined community—and thus the nation-state—through fictive ethnicity and national mythologies. For populist leaders like Orbán, “elite” intellectuals, non-ethnic Magyars, and the opposition are not welcome. Intellectuals pose a specific threat to authoritarians because critical discourse threatens to destabilize their régimes. Taken together, we can begin to recognize the warning signs of an illiberal democracy.

Huntington’s second generalization is particularly interesting as applied to Hungary’s illiberal democracy and is worth expanding. Huntington contends that transitions from democracy to authoritarianism tend to take place either through a military *coup d’état* or through an executive coup whereby a head of government concentrates power in the executive by declaring a state of emergency or instituting martial law. As opposed to concentrating power on an individual executive, Hungary’s partocracy has concentrated Fidesz’s power. Fidesz controls all three branches of government, independent government institutions, and the media, and while strangling civil society and academia. And, not a single action Fidesz has taken has been illegal under national law. Therefore, Huntington’s second generalization should be expanded to include the centralization of power by a group or party.

Finally, the large majority of Huntington’s factors under the first generalization have manifested themselves in Hungarian illiberal democracy. During the period of democratic transition, political scandals, majoritarianism, and partocracy weakened public regard for democratic values. Instability was intensified during this period due to the numerous economic crises experienced during Hungary’s transition to capitalism and in the 2008 economic recession. Third, the country is politically polarized, in part due to partocracy, the failures of MSZP, and Orbán’s inflammatory rhetoric. Fourth, MSZP’s missteps pushing through austerity measures and other initiatives that were unpopular with the general public fueled populist desires. Fifth, the use of law and order rhetoric, first with regard to the Roma and then the refugee crisis, has sparked ethno-nationalist sentiments and moved racist discourse into the political mainstream. Sixth, Hungary’s long history of occupation by nondemocratic régimes and the dramatic loss of territory resulting from the Treaty of Trianon, strengthened calls for

absolute sovereignty. Now, Fidesz has identified the European Union, “Western values,” and George Soros as the next invaders. Finally, the illiberal democratic régime has snowballed and propagated itself both in neighboring nation-states and internationally.

Hungary’s polarized political condition is exemplified in the clash between Orbán and Central European University. This clash pits Central European University—a symbol of Western liberalism and cosmopolitanism—against Orbán’s ethno-nationalist illiberal democracy. Other attacks on academic freedom, like the gender studies ban, slashed funding for research institutions, and campaigns against the Hungarian intelligentsia, demonstrate the serious threat that is posed to not only academic freedom in Hungary, but also academic freedom globally as illiberal democracies spread. There is some hope, however. Central European University’s ousting caught international attention, prompting criticism and solidarity from individuals, academia, powerful politicians, and government institutions. Orbán experienced severe backlash from the European Union for his actions—facing Article 7 disciplinary proceedings—as well as from the conservative European People’s Party that suspended Fidesz from the party. The European People’s Party’s actions even went so far as to prompt a dialogue between Orbán and EPP leaders that may result in the university being permitted to stay in Budapest and remain U.S. accredited.

For some reason, academic freedom in Hungary has received immense international attention that has been critical in CEU’s fight to exist in Hungary. Perhaps academia is more alluring than amendments to parliamentary procedure or executive power. Perhaps the international community still holds high regard for universities. Or perhaps people have rallied around Central European University for another reason. If CEU is successful, the outcome could provide us with a potential strategy to protecting academic freedom in Hungary and in other illiberal democracies, and even allow us to glean perspective on how illiberal democracies operate.



**Statement of Ownership, Management, and Circulation  
(All Periodicals Publications Except Requester Publications)**

1. Publication Title <b>Connecticut Law Review</b>	2. Publication Number	3. Filing Date
	664-550	9/9/19

4. Issue Frequency <b>November, January, May</b>	5. Number of Issues Published Annually <b>45</b>	6. Annual Subscription Price <b>\$42 Domestic; \$63 Int'l</b>
---	---	--

7. Complete Mailing Address of Known Office of Publication (Not printer) (Street, city, county, state, and ZIP+4®) <b>Connecticut Law Review; 65 Elizabeth Street, Hartford, CT 06105</b>	Contact Person <b>Alexandria Madjeric</b> Telephone (Include area code) <b>860-570-5331</b>
--	--

8. Complete Mailing Address of Headquarters or General Business Office of Publisher (Not printer)  
**Connecticut Law Review; 65 Elizabeth Street, Hartford, CT 06105**

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor (Do not leave blank)  
Publisher (Name and complete mailing address)  
**Western Newspaper Publishing Company, Inc.; 929 W. 16th Street; Indianapolis, IN 46202**  
Editor (Name and complete mailing address)

**Alexandria M. Madjeric; 65 Elizabeth Street, Hartford, CT 06105**  
Managing Editor (Name and complete mailing address)

**Adam J. Kuegler; 65 Elizabeth Street, Hartford, CT 06105**

10. Owner (Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)

Full Name	Complete Mailing Address
Connecticut Law Review Association	Connecticut Law Review; 65 Elizabeth Street; Hartford, CT 06105

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box  None

Full Name	Complete Mailing Address

12. Tax Status (For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)  
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:  
 Has Not Changed During Preceding 12 Months  
 Has Changed During Preceding 12 Months (Publisher must submit explanation of change with this statement)



13. Publication Title		14. Issue Date for Circulation Data Below	
<b>15. Extent and Nature of Circulation</b>		<b>Average No. Copies Each Issue During Preceding 12 Months</b>	<b>No. Copies of Single Issue Published Nearest to Filing Date</b>
a. Total Number of Copies ( <i>Net press run</i> )		419	336
b. Paid Circulation ( <i>By Mail and Outside the Mail</i> )	(1) Mailed Outside-County Paid Subscriptions Stated on PS Form 3541 (Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)		
	(2) Mailed In-County Paid Subscriptions Stated on PS Form 3541 ( <i>Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies</i> )		
	(3) Paid Distribution Outside the Mails Including Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Paid Distribution Outside USPS®		
	(4) Paid Distribution by Other Classes of Mail Through the USPS (e.g., First-Class Mail®)	254	158
c. Total Paid Distribution [ <i>Sum of 15b (1), (2), (3), and (4)</i> ] ▶		254	158
d. Free or Nominal Rate Distribution ( <i>By Mail and Outside the Mail</i> )	(1) Free or Nominal Rate Outside-County Copies included on PS Form 3541		
	(2) Free or Nominal Rate In-County Copies Included on PS Form 3541		
	(3) Free or Nominal Rate Copies Mailed at Other Classes Through the USPS (e.g., First-Class Mail)		
	(4) Free or Nominal Rate Distribution Outside the Mail ( <i>Carriers or other means</i> )	84	78
e. Total Free or Nominal Rate Distribution ( <i>Sum of 15d (1), (2), (3) and (4)</i> )		84	78
f. Total Distribution ( <i>Sum of 15c and 15e</i> ) ▶		338	236
g. Copies not Distributed ( <i>See Instructions to Publishers #4 (page #3)</i> ) ▶		140	100
h. Total ( <i>Sum of 15f and g</i> )		478	336
i. Percent Paid ( <i>15c divided by 15f times 100</i> ) ▶		75%	67%

\* If you are claiming electronic copies, go to line 16 on page 3. If you are not claiming electronic copies, skip to line 17 on page 3.



**Statement of Ownership, Management, and Circulation  
(All Periodicals Publications Except Requester Publications)**

16. Electronic Copy Circulation	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Paid Electronic Copies <span style="float: right;">▶</span>		
b. Total Paid Print Copies (Line 15c) + Paid Electronic Copies (Line 16a) <span style="float: right;">▶</span>		
c. Total Print Distribution (Line 15f) + Paid Electronic Copies (Line 16a) <span style="float: right;">▶</span>		
d. Percent Paid (Both Print & Electronic Copies) (16b divided by 16c × 100) <span style="float: right;">▶</span>		

I certify that 50% of all my distributed copies (electronic and print) are paid above a nominal price.

17. Publication of Statement of Ownership

If the publication is a general publication, publication of this statement is required. Will be printed  
in the November issue of this publication.

Publication not required.

18. Signature and Title of Editor, Publisher, Business Manager, or Owner

Date

9/9/19

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

