EMERGING
THREATS

FACEBOOK V. SULLIVAN

BY KATE KLONICK

KNIGHT
FIRST AMENDMENT
INSTITUTE
at Columbia University

EMERGING
THREATS

ABOUT EMERGING THREATS

The Knight First Amendment Institute's *Emerging Threats* series invites leading thinkers to identify and grapple with newly arising or intensifying structural threats to the system of free expression. These threats may be caused by changes in the forms and applications of technology, in the means and economics of communication, in the norms and practices of politics, or in legal doctrine. The papers in the series explore ways to address these threats and preserve the foundations of democracy essential to healthy open societies, including the United States.

The *Emerging Threats* series is edited by David Pozen, professor at Columbia Law School and inaugural visiting scholar at the Knight Institute.

ABOUT THE KNIGHT INSTITUTE

The Knight First Amendment Institute is a non-partisan, not-for-profit organization established by Columbia University and the John S. and James L. Knight Foundation to defend the freedoms of speech and press in the digital age through strategic litigation, research, and public education.
For more information, please visit www.knightcolumbia.org.

ABOUT THE AUTHOR

Kate Klonick is an assistant professor at St. John's University School of Law, where she teaches property, internet law, and a seminar on privacy. Klonick holds a Ph.D. from Yale Law School and a J.D. from Georgetown University Law Center. Her research focuses on emerging issues in law and technology, including the ways in which private internet platforms govern online speech. Klonick's work has appeared in the *Harvard Law Review*, *Maryland Law Review*, *New York Times*, *Atlantic*, *Slate*, *Guardian*, and numerous other publications.

TABLE OF CONTENTS

Facebook v. Sullivan

I. "Public Figures" and "Matters of Public Interest" in Communications Torts **3**

II. Facebook's Free Speech Doctrine **5**

A. Facebook's Public Figure Exception for Bullying *6*

B. Facebook's General Newsworthiness Exception *8*

C. The Current Public Figure and Newsworthiness Standards *9*

III. Facebook Versus *Sullivan* **10**

A. The Problems with Defining "Public Figure" Algorithmically *11*

B. The Rise of the Involuntary Public Figure *12*

C. Straddling the Role of Governor and Publisher *15*

IV. Conclusion **17**

FACEBOOK V. SULLIVAN

Kate Klonick*

In August 2017, shortly after the Unite the Right rally in Charlottesville, Virginia, a post began circulating on Facebook about Heather Heyer, the woman who was killed while protesting the rally.[1] "Heather Heyer, Woman Killed in Road Rage Incident was a Fat, Childless 32-Year-Old Slut" was shared over 65,000 times.[2] To some Facebook users, the post seemed like obvious hate speech that violated the company's "Community Standards" and therefore ought to be deleted. To other users, it might have seemed like controversial but permissible commentary on a person whose tragic death had turned her into a public figure. Ultimately, Facebook hedged. The company announced that the post would generally be removed because it originated from the neo-Nazi website *Daily Stormer* but that the post could still be shared if accompanied by a condemnatory caption.[3]

As this episode reflects, the United States now has two systems to adjudicate disputes arising from harmful speech about other people. The first is older and more familiar: the tort system in which judges resolve claims brought under state defamation and privacy law. The second is newer and less well understood: the content moderation policies and practices of private platforms such as Facebook. These platforms are not, as a general rule, bound by the First Amendment. Yet as this episode also reflects, they have come to rely on some of the same concepts used by courts to resolve tensions between regulating harmful speech and preserving free expression, including the concepts of "public figures" and "newsworthiness."

This paper analyzes Facebook's use of these concepts and the implications for online speech. It begins with a brief summary of the Supreme Court cases that introduced the concepts, with an eye toward the underlying First Amendment theory. It then looks to Facebook's moderation of user speech, discussing how and why exceptions for public figures and newsworthiness were carved out. In developing and applying these exceptions, Facebook has adopted much of the Court's reasoning for creating First Amendment limits to tort liability in cases involving public figures and matters of public concern.

Drawing on this analysis, I argue that comparing these systems reveals three main points that can help both courts and platforms going forward. First, Facebook's partial reliance on online news sources and news aggregators to make public figure determinations runs into many of the same critiques leveled at judges who defer to the media in determining newsworthiness. In some circumstances, this results in Facebook keeping up harmful speech about users who have compelling reasons for wanting the speech taken down. Moreover, these aggregators cannot adequately take into account localized newsworthiness or public figures in smaller communities, and they therefore threaten to over-censor certain other types of speech.

* The author is grateful to Jack Balkin, Molly Brady, Danielle Citron, Thomas Kadri, Margot Kaminski, David Pozen, and colleagues at the Yale Information Society Project and the Cornell Tech Speed Conference for helpful thoughts and comments on earlier versions of this paper.

[1] Julia Angwin, Ariana Tobin & Madeleine Varner, *Have You Experienced Hate Speech on Facebook? We Want to Hear from You.*, ProPublica (Aug. 29, 2017), https://www.propublica.org/article/have-you-experienced-hate-speech-on-facebook-we-want-to-hear-from-you.

[2] Casey Newton, *Facebook Is Deleting Links to a Viral Attack on a Charlottesville Victim*, Verge (Aug. 14, 2017), https://www.theverge.com/2017/8/14/16147126/facebook-delete-viral-post-charlottesville-daily-stormer.

[3] *Id.*

Second, factual situations arising in the unique environment of online culture reveal—for perhaps the first time since the Court imagined them in 1974—the existence of "truly involuntary public figures."[4] The internet has eroded some of the traditional reasons for specially protecting speech concerning public figures, based on the assumption that people become public figures by choice and that, as public figures, they have greater access to channels of rebuttal. These assumptions are becoming increasingly outdated in the digital age, given the dynamics of online virality and notoriety and given the ubiquity of channels for engaging in counterspeech.

Finally, comparing these systems reveals something significant about Facebook's role in society. Whereas courts apply the concept of newsworthiness to resolve private disputes and newspapers apply the concept to decide what to print, platforms like Facebook rely on it for both tasks. Like a court, Facebook responds to claims involving allegedly defamatory, hateful, or otherwise harmful speech. Like a media company, Facebook curates content and decides which sorts of statements reach a large audience and which don't. From the perspective of its users, Facebook functions as a speech regulator, adjudicator, and publisher all at the same time.

Ultimately, Facebook must determine whose interests it wants to prioritize and what theory of free expression will animate the speech standards it sets. I conclude by suggesting that Facebook's approach ought to vary depending on context. When Facebook acts more like a court in evaluating individual claims of harmful speech, the company should focus on threats to individual users. In contrast, when Facebook acts more like the press in evaluating general newsworthiness exceptions, the company should err on the side of allowing as much content as possible to stay up.

I. "Public Figures" and "Matters of Public Interest" in Communications Torts

On March 29, 1960, L.B. Sullivan, an elected commissioner of Montgomery, Alabama, sued the *New York Times* for defamation after the newspaper published an advertisement criticizing the way in which police in Montgomery had treated civil rights demonstrators. Writing for the Court in *New York Times Co. v. Sullivan*,[5] Justice William Brennan explained that while the ad did contain false statements, criticism of government was at the core of the speech protected by the First Amendment. Public officials alleging defamation, accordingly, must prove that the offending statement was made with "'actual malice'—that is, with knowledge that [the statement] was false or with reckless disregard of whether it was false or not."[6] In deciding that constitutional values of free speech outweighed liability for harmful speech in the absence of such malice, the Court identified two main concerns: the democratic imperative of protecting "debate on public issues" and the practical ability of "public officials" to rebut remarks made against them.[7]

---

[4] *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 345 (1974).

[5] 376 U.S. 254 (1964). The suit also included four African-American clergymen.

[6] *Id.* at 279–80.

[7] *Id.* at 268–83. For an excellent overview of the history of the public figure doctrine, see Catherine Hancock, *Origins of the Public Figure Doctrine in First Amendment Defamation Law*, 50 N.Y.L. Sch. L. Rev. 81 (2005).

Today, *Sullivan* is frequently described as a case about "public officials" or "public figures." This characterization is somewhat misleading. To a significant extent, the public figure doctrine has come to focus on whether speech relates to debate on public issues, reflecting the *Sullivan* Court's overriding concern with what Brennan called a "profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open."[8] The result is that the inquiry into whether a tort plaintiff is a "public figure" is now essentially an element of a larger inquiry into whether the speech in question is sufficiently a matter of public concern.

The *Sullivan* Court itself gave little guidance on the meaning of "public official."[9] Just a few years after *Sullivan*, in the 1967 case *Time, Inc. v. Hill*, the Court applied "the First Amendment principles pronounced in [*Sullivan*]" to a privacy suit brought against *Life Magazine*.[10] Stressing that "[f]reedom of discussion . . . must embrace all issues about which information is needed,"[11] the *Hill* Court declined to rely on "the distinction which has been suggested between the relative opportunities of the public official and the private individual to rebut defamatory charges."[12] In *Hill*, one sees the Court begin to move away from the "public official" concept in justifying its use of the First Amendment to limit libel, privacy, and related tort claims. Instead, the justices were concerned with preserving debate on "matters of public interest,"[13] broadly defined, and to that end they "declared an expansive view of the First Amendment as protection for all newsworthy material."[14] *Hill* signaled that virtually any matter that would be considered of public concern in a defamation action would be considered newsworthy in a privacy tort action, and vice versa.[15]

The Court continued to move away from the "public official" concept in subsequent defamation cases. Later that same year, in *Curtis Publishing Co. v. Butts*,[16] the Court extended First Amendment protection to media reports concerning plaintiffs who were not government officials but who were nonetheless sufficiently prominent in their communities to be considered "public figures." Four years later, in *Rosenbloom v. Metromedia, Inc.*, a plurality of the Court attempted to extend *Sullivan* to all matters of public concern, regardless of whether the plaintiff was a public or private figure.[17] Writing for the plurality, Justice Brennan reasoned that a matter "of public or general interest . . . cannot suddenly become less so merely because a private individual is involved, or because in some

---

[8] *Sullivan*, 376 U.S. at 270.

[9] In a footnote mid-decision, the *Sullivan* Court brushed away the question of defining public officials, stating: "We have no occasion here to determine how far down into the lower ranks of government employees the 'public official' designation would extend for purposes of this rule, or otherwise to specify categories of persons who would or would not be included." *Id.* at 283 n.23.

[10] 385 U.S. 374, 390 (1967); *see also* Hancock, *supra* note 7, at 105–12 (discussing the relationship of *Hill* to *Sullivan*).

[11] *Hill*, 385 U.S. at 388 (internal quotation marks omitted).

[12] *Id.* at 391.

[13] *See id.* at 387–88 ("We hold that the constitutional protections for speech and press preclude the application of the New York statute to redress false reports of matters of public interest in the absence of proof that the defendant published the report with knowledge of its falsity or in reckless disregard of the truth.").

[14] Samantha Barbas, *When Privacy Almost Won:* Time, Inc. v. Hill, 18 U. Pa. J. Const. L. 505, 508 (2015).

[15] On the overlap between inquiries into matters of public concern and newsworthiness, see Richard T. Karcher, *Tort Law and Journalism Ethics*, 40 Loy. U. Chi. L.J. 781, 824–30 (2009); Mary-Rose Papandrea, *Citizen Journalism and the Reporter's Privilege*, 91 Minn. L. Rev. 515, 578–81 (2007). *See also, e.g.*, Fla. Stat. § 90.5015(1)(b) (2017) ("'News' means information of public concern relating to local, statewide, national, or worldwide issues or events.").

[16] 388 U.S. 130 (1967).

[17] 403 U.S 29 (1971).

sense the individual did not 'voluntarily' choose to become involved."[18] To "honor the commitment to robust debate on public issues . . . embodied in the First Amendment," in Brennan's view, the *Sullivan* rule should be applied "to all discussion and communication involving matters of public or general concern, without regard to whether the persons involved are famous or anonymous."[19]

Rosenbloom's doctrinal simplicity—if the jettisoning of the "public figure" notion for the equally ambiguous notion of the "public interest" can be considered simplicity—was short-lived. In the 1974 case *Gertz v. Robert Welch, Inc.*, the majority expressly rejected the extension of *Sullivan* to private defamation plaintiffs.[20] Yet despite disavowing Justice Brennan's approach in *Rosenbloom*, the *Gertz* Court's reconstruction of the public figure concept nonetheless incorporated some of its logic. The Court imagined at least two, and perhaps three,[21] types of public figures: (1) general public figures, (2) limited-purpose public figures, and (3) involuntary public figures. General public figures are those who "occupy positions of such persuasive power and influence that they are deemed public figures for all purposes."[22] Limited-purpose public figures "thrust themselves to the forefront of particular public controversies in order to influence the resolution of the issues involved."[23] Notably, the Court left the third (questionable) category particularly undefined, opining that the "instances of truly involuntary public figures must be exceedingly rare."[24]

Although now glossed with a new taxonomy, the Court's basic rationale for affording constitutional protection to speech concerning public figures remained much the same as the rationale suggested in *Sullivan*. The powerful and notorious—be they so from wealth, fame, or public office—have greater access to counterspeech than private individuals. At the same time, the power, fame, or celebrity of such figures makes their behaviors inherently a matter of public interest, just as the behaviors of someone who thrusts herself into a public controversy are inherently a matter of public interest. And the First Amendment must protect robust public debate on all such matters.

II. Facebook's Free Speech Doctrine

For Facebook, the idea of making exceptions for newsworthy speech or speech concerning public figures did not arise from tort litigation but rather from the company's efforts to deal with situations in which one user alleges that another user's speech has violated "Community Standards." Community Standards are Facebook's public rules about the types of speech that users may post on the platform. Because huge amounts of user content are posted each day, Facebook cannot proactively police all speech violations and must rely to a large extent on users to flag speech that might be in violation. The flagged speech is then reviewed by human content moderators—individuals

[18] *Id.* at 43. "The public's primary interest is in the event," Justice Brennan continued; "the public focus is on the conduct of the participant and the content, effect, and significance of the conduct, not the participant's prior anonymity or notoriety." *Id.*
[19] *Id.* at 43–44.
[20] 418 U.S. 323, 339–48 (1974).
[21] *See* W. Wat Hopkins, *The Involuntary Public Figure: Not So Dead After All*, 21 Cardozo Arts & Ent. L.J. 1, 21 (2003) ("[T]here is disagreement as to whether the Supreme Court identified two or three categories of public figure status.").
[22] *Gertz*, 418 U.S. at 345.
[23] *Id.*
[24] *Id.*

trained to apply Facebook's rules and determine whether the reported speech actually runs afoul of them. Speech that is found to be in violation is removed. The rest stays up.[25]

Somewhat like a common law system, Facebook updates its Community Standards and the internal guidelines used by moderators, both in response to new factual scenarios that present themselves and in response to feedback from outside observers. The first iterations of the Community Standards and content moderation policies were created in 2009 largely by Dave Willner, who was then part of the Site Integrity Operations team. Willner later transferred to a team focused on "organic" content (user-generated content, as opposed to advertising or commercial content) under Jud Hoffman, who joined Facebook in 2010 as global policy manager. Hoffman and Willner were the principal players in a six-person group established to formalize and consolidate the informal rules that Facebook's content moderators had been using, thereby enhancing their consistency and transparency.

A. Facebook's Public Figure Exception for Bullying

Whereas *Sullivan*'s public figure doctrine grew out of claims of defamation, both Hoffman and Willner describe Facebook's concept of public figures as emerging from claims about bullying.[26] In 2009, Facebook was facing heavy pressure from anti-cyberbullying advocacy groups to do more to prevent kids from being bullied online.[27] The problem, however, was that traditional academic definitions of bullying seemed impossible to translate to online content moderation. "How do we write a rule about bullying?" recounts Willner. "What is bullying? What do you mean by that? It's not just things that are upsetting; it's defined as a pattern of abusive or harassing unwanted behavior over time that is occurring between a higher power [and] a lower power. But that's not an answer to the problem that resides in the content—you can't determine a power differential from looking at the content. You often cannot even do it from looking at their profiles."[28]

The apparent impossibility of employing a traditional definition of bullying meant that Facebook had to make a choice. It could err on the side of keeping up potentially harmful content, or it could err on the side of removing all potential threats of bullying, even if some of the removed content turned out to be benign. Faced with intense pressure from advocacy groups and media coverage on cyberbullying, Facebook opted for the latter approach, but with a caveat. The new presumption in favor of taking down speech reported to be "bullying" would apply only to speech directed at private individuals. "What we said was, 'Look, if you tell us this is about you, and you don't like it, and you're a private individual, you're not a public figure, then we'll take it down,'" said Hoffman. "Because we can't know whether all of those other elements [of bullying] are met, we had to just make the call to create a default rule for removal of bullying."[29]

---

[25] The following section borrows from the research and conclusions in my previous work on the history of content moderation, Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 Harv. L. Rev. 1598 (2018).

[26] Telephone Interview with Jud Hoffman, former Global Policy Manager, Facebook (Mar. 6, 2018); Telephone Interview with Dave Willner, former Head of Content Policy, Facebook (Mar. 5, 2018). All interview notes are on file with the author.

[27] Telephone Interview with Dave Willner, *supra* note 26.

[28] *Id.*

[29] Telephone Interview with Jud Hoffman, former Global Policy Manager, Facebook (Jan. 22, 2016).

Although he denies borrowing directly from the First Amendment public figure doctrine, Hoffman's justification for creating this exception tracks the reasoning of *Sullivan* and subsequent cases in treating certain targets of allegedly harmful speech differently on account of their public status and the public interest in their doings. According to Hoffman, this approach reflected Facebook's mission statement, which at that time was "Make the world more open and connected." "Broadly, we interpreted 'open' to mean 'more content.' Yes, that's a bit of a free speech perspective, but then we also had a concern with things like bullying and revenge porn," Hoffman recalls. "But while trying to take down that bad content, we didn't want to make it impossible for people to criticize the president or a person in the news. It's important there's a public discussion around issues that affect people, and this is how we drew the line."[30]

In trying to resolve these dilemmas, Hoffman and his colleagues sought to "focus on the mission" of Facebook rather than adopt "wholesale . . . a kind of U.S. jurisprudence free expression approach."[31] They quickly realized, however, that the mission had to be balanced against competing interests such as users' safety and the company's bottom line. While Hoffman and Willner were at Facebook, the balance was often struck in favor of "leaving content up," but they were always searching for new ways to address concerns about harmful speech. "We felt like Facebook was the most important platform for this kind of communication, and we felt like it was our responsibility to figure out an answer to this," says Hoffman.[32]

The policy required a new way of determining if someone was a public figure. Facebook told its moderators that when reviewing a piece of content flagged for bullying, they should use Google News.[33] If the user who was the subject of the allegedly bullying content came up in a Google News search, she would be considered a public figure—and the content would be kept up.

By tying the public figure determination to the algorithmic calculations of Google News, Facebook sought to maintain vibrant discussion on matters of public concern while creating a temporal constraint on the "limited-purpose public figure" concept, as individuals who had thrust themselves (or been thrust by circumstance) into a particular public controversy would likely remain in Google News search results only as long as the controversy was topical and newsworthy. For this reason, Willner reflects, "calling the exception [an exception for] 'public figures' was probably a mistake. A more accurate way of thinking about it is as a newsworthy person."[34]

Despite the arguable conflation of "public figure" and "newsworthiness" occasioned by relying on Google News, both Hoffman and Willner were opposed to the idea of a general exception to the Community Standards that would prevent any "newsworthy" piece of content from being taken down. Facebook's approach to this issue would develop on a slightly different track.

---

[30] Telephone Interview with Jud Hoffman, *supra* note 26.
[31] *Id.*
[32] *Id.*
[33] Telephone Interview with Dave Willner, *supra* note 26.
[34] *Id.*

B. Facebook's General Newsworthiness Exception

For most of the history of Facebook's content moderation, no exceptions were made for content that violated Community Standards but was newsworthy. Overtly sexual, graphically violent, or "extremist" content would be taken down regardless of whether it had cultural or political significance as news. This was a deliberate choice made by Hoffman and Willner. But this policy came under increasing pressure.

Members of the policy team recall an incident in 2013 concerning a graphic picture from the Boston Marathon bombing as a turning point toward the creation of an exception for newsworthy content. The image in question was of a man in a wheelchair being wheeled away with one leg ripped open below the knee to reveal a long, bloody bone. The picture had three versions. One was cropped so that the leg was not visible. A second was a wide-angle shot in which the leg was visible but less obvious. The third, and most controversial, version clearly showed the man's "insides on the outside"—the content moderation team's shorthand rule for when content was graphically violent. Despite being published in multiple media outlets, Facebook policy dictated that any links to or images of the third version of the picture must be removed.[35] "Philosophically, if we were going to take the position that [insides on the outside] was our definition of gore and we didn't allow gore, then just because it happened in Boston didn't change that," remembers one of the team members on call that day.[36] Policy executives at Facebook disagreed, however, and reinstated all such posts on the grounds of newsworthiness.

For some members of the policy team, who had spent years trying to create administrable rules, the imposition of such an exception seemed a radical departure from the company's commitment to procedural consistency. Some of their reasoning echoes the *Gertz* Court's rationale for reining in *Rosenbloom*.[37] In his opinion for the Court in *Gertz*, Justice Lewis Powell worried openly about allowing "judges to decide on an *ad hoc* basis which publications address issues of 'general or public interest' and which do not."[38] Many at Facebook worried similarly that "newsworthiness as a standard is extremely problematic. The question is really one of 'newsworthy to whom?' and the answer to that is based on ideas of culture and popularity."[39] The result, some feared, would be a mercurial exception that would, moreover, privilege American users' views on newsworthiness to the potential detriment of Facebook's users in other countries.

Although there were other one-off exceptions made for incidents like the Boston Marathon bombing, Facebook's internal content moderation policies continued to have no general exception for newsworthiness until September 2016, when a famous Norwegian author, Tom Egeland, posted a well-known historical picture to his Facebook page. The photograph, "The Terror of War," depicts a nine-year-old Vietnamese girl naked in the street after a napalm attack (for this reason, the photo is often called "Napalm Girl"). In part because of its graphic nature,

[35] Simon Adler, *Post No Evil*, Radiolab (Aug. 17, 2018), https://www.wnycstudios.org/story/post-no-evil.
[36] Telephone Interview with former member of Facebook Policy Team (Aug. 28, 2018).
[37] *See supra* notes 17–24 and accompanying text.
[38] *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 346 (1974).
[39] Telephone Interview with former member of Facebook Policy Team, *supra* note 36.

the photo was a pivotal piece of journalism during the Vietnam War.[40] But it violated Facebook's Community Standards.[41] Accordingly, Facebook removed the photo and suspended Egeland's account. Because of Egeland's stature, the takedown itself received news coverage. Espen Egil Hansen, the editor-in-chief of the Norwegian newspaper *Aftenposten*, published a "letter" to Zuckerberg on *Aftenposten*'s front page calling for Facebook to take a stand against censorship. Hours later, Facebook's chief operating officer Sheryl Sandberg admitted that the company had made a mistake and promised that the rules would be rewritten to allow for posting of the photo.[42] Shortly thereafter, Facebook issued a press release underscoring the company's commitment to "allowing more items that people find newsworthy, significant, or important to the public interest—even if they might violate our standards."[43]

The "Terror of War" incident led Facebook to start looking more broadly at how it evaluated newsworthiness outside the context of bullying. "After the 'Terror of War' controversy, we realized that we had to create new rules for imagery that we'd normally want to disallow, but for context reasons that policy doesn't work," says Peter Stern, head of Product Policy Stakeholder Engagement at Facebook. "And that's led us to think about newsworthiness across the board. When we do these things, we have two considerations: safety of individuals on the one hand and voice on the other."[44] But how exactly should "voice" be taken into consideration? Here, again, Facebook has increasingly aligned itself with the Court's public figure doctrine. "When someone enters the public eye," Stern explains, "we want to allow a broader scope of discussion."[45]

C. The Current Public Figure and Newsworthiness Standards

Over the last two years, Facebook's content moderation policies have continued to evolve and to become somewhat less mechanical and more context-sensitive. For example, in recent months Facebook has modified its rules on bullying and harassment of public figures. "Our new policy does not allow certain high-intensity attacks, like calls for death, directed at a certain public figure," members of the Facebook policy team told me on a recent call.[46] In the past, they explained, a statement such as "Kim Kardashian is a whore" would never be removed for bullying or harassment (whereas a statement calling a private individual a "whore" would be). But now, Facebook allows some speech directed at public figures, when it is posted on their own pages or accounts, to be removed depending on the severity of the language. Under this new regime, public figures are defined as people elected or assigned through a political process to a government position; people with hundreds of thousands of fans or followers on a social media account; people employed by a news or broadcast organization or who speak publicly;

---

[40] *See* Kate Klonick, *Facebook Under Pressure*, Slate (Sept. 12, 2016), http://www.slate.com/articles/technology/future_tense/2016/09/facebook_erred_by_taking_down_the_napalm_girl_photo_what_happens_next.html.

[41] The photo was likely removed because of the nudity, not because it was child pornography. *See* Kjetil Malkenes Hovland & Deepa Seetharaman, *Facebook Backs Down on Censoring 'Napalm Girl' Photo*, Wall St. J. (Sept. 9, 2016), http://www.wsj.com/articles/norway-accuses-facebook-of-censorship-over-deleted-photo-of-napalm-girl-1473428032.

[42] *See* Claire Zillman, *Sheryl Sandberg Apologizes for Facebook's 'Napalm Girl' Incident*, Time (Sept. 13, 2016), http://time.com/4489370/sheryl-sandberg-napalm-girl-apology.

[43] Joel Kaplan & Justin Osofsky, *Input from Community and Partners on Our Community Standards*, Facebook Newsroom (Oct. 21, 2016), https://newsroom.fb.com/news/2016/10/input-from-community-and-partners-on-our-community-standards.

[44] Telephone Interview with Peter Stern, Head of Product Policy Stakeholder Engagement, Facebook (Mar. 7, 2018).

[45] *Id.*

[46] Televideo Interview with Idalia Gabrielow & Peter Stern, Policy Risk Team, Facebook (Aug. 15, 2018).

or people who are mentioned in multiple news articles within a certain recent time period as determined by a news search.

Facebook's current policies on newsworthy content are somewhat harder to pin down. Unlike the term "public figures," which is primarily used by Facebook for purposes of its bullying standards, "newsworthiness" is now a possible exception to all of the company's general guidelines for removing offensive content. And unlike the public figure determinations made in the bullying context, determinations of newsworthiness do not rely on news aggregators. Instead, every suggestion of possible newsworthy content is made by a person and evaluated by Facebook employees on a case-by-case basis.

In deciding whether to keep up otherwise removable content on the basis of its newsworthiness, Facebook officials stress that they weigh the value of "voice" against the risk of harm. Assessments of harm are informed by the nature as well as the substance of the objectionable content. Hateful speech on its own, for instance, might be seen as less harmful than a direct call to violence. Facebook officials maintain, however, that most of the newsworthiness decisions are around nudity. Difficult decisions include what to do about nudity in public protests. "Just a few years ago, we took that down," states David Caragliano, a policy manager at Facebook. "But it's really important to leave this up consistent with our principles of voice. That's led to a policy change that's now at scale for the platform."[47] The non-hateful, nonviolent expressive conduct of public protesters, it seems, will today almost always be considered newsworthy and therefore will not be taken down.

III. Facebook Versus *Sullivan*

As the foregoing discussion reflects, the two systems in the United States for adjudicating claims of harmful speech—through tort lawsuits and through private online speech platforms like Facebook—share a number of similarities. Both have developed rules to weigh individual harms against a baseline commitment to enabling as much speech as possible. In the courts, defamation law gives plaintiffs recourse for untruthful speech against them, except that plaintiffs who are public figures have a substantially higher burden to meet. On Facebook, an anti-bullying policy gives users who have been harassed the ability to have the harassing speech removed, except that users who are public figures can rarely avail themselves of this option. In the courts, privacy law allows plaintiffs to hold defendants liable for certain invasions of privacy, except when the underlying information is deemed to be of sufficient public interest. On Facebook, users can request that disturbing content such as graphically violent or hateful speech be taken down, except when the underlying information is deemed to be of sufficient public interest. Federal judges and Facebook executives justify these exceptions in similar terms, citing the importance of preserving public discourse and the special capacities of people who are powerful, famous, or at the forefront of a particular public controversy to rebut speech against them. Even though Facebook is not generally bound by the First Amendment, its content moderation policies were largely developed by U.S. lawyers trained and acculturated in U.S. free speech norms, and this cultural background has invariably affected their thinking.

[47] Televideo Interview with Ruchika Budhraja, David Caragliano, Idalia Gabrielow & Peter Stern, Policy Risk Team, Facebook (Oct. 4, 2018).

The observation that First Amendment principles like "public figures" and "newsworthiness" have wended their way into Facebook's content moderation policies is interesting in its own right. But it also suggests several broader lessons about the structure of digital discourse. First, Facebook's use of Google News for its public figure determinations underscores the dangers of reducing such judgments to mechanical calculations. Second, Facebook and other digital speech platforms have helped bring into being the elusive "involuntary public figure" imagined in *Gertz*, even as they undermine the access-to-counterspeech justification for keeping up more speech about public figures. Finally, Facebook's struggle to create principled exceptions for newsworthy content underscores how the company straddles the roles of regulator, adjudicator, and publisher in controlling access to speech for both speakers and listeners.

A. The Problems with Defining "Public Figure" Algorithmically

Facebook's use of Google News to determine whether a person is a public figure provides a vivid illustration of the problems that may be raised when such definitions are outsourced to the media marketplace.

Although it has been suggested to me that this policy may be changing, Facebook's method for ascertaining "public figure" status has traditionally turned in part on the presence or absence of an individual's name in news search results, which are effectively an averaging algorithm of media outlets' publication decisions. (Facebook's "newsworthiness" determinations, in contrast, involve multiple layers of human review.) This runs straight into the threat of what Clay Shirky has called "algorithmic authority," insofar as "an unmanaged process of extracting value from diverse, untrustworthy sources" is treated as authoritative without any human second-guessing or vouching for the validity of the outcome.[48]

As commentators have pointed out for over fifty years in a closely related context, if "newsworthiness" is defined solely in terms of news outlets' publication decisions, then granting a special legal privilege for newsworthy content is liable to swallow torts such as invasion of privacy. "The publisher has almost certainly published any given report because he judged it to be of interest to his audience . . . and believed that it would encourage them to purchase his publications in anticipation of more of the same," a student comment observed in 1963. "A plaintiff in a privacy action would thus have lost almost before he started."[49] Partly for this reason, courts making these determinations have considered a range of factors[50] and, especially in recent years, have been unwilling to defer entirely to the media.[51]

---

[48] Clay Shirky, *A Speculative Post on the Idea of Algorithmic Authority*, Clay Shirky (Nov. 15, 2009), http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-idea-of-algorithmic-authority.

[49] Comment, *The Right of Privacy: Normative-Descriptive Confusion in the Defense of Newsworthiness*, 30 U. Chi. L. Rev. 722, 725 (1963).

[50] *See, e.g.*, *Snyder v. Phelps*, 562 U.S. 443, 453 (2011) ("Deciding whether speech is of public or private concern requires us to examine the content, form, and context of that speech, as revealed by the whole record." (internal quotation marks omitted)).

[51] *See, e.g.*, Amy Gadja, *Judging Journalism: The Turn Toward Privacy and Judicial Regulation of the Press*, 97 Calif. L. Rev. 1039, 1041–42 (2009) (explaining that some courts have become less deferential to the media in determining newsworthiness, perhaps on account of "growing anxiety about the loss of personal privacy in contemporary society" or "declining public respect for journalism"); Sydney Ember, *Gawker and Hulk Hogan Reach $31 Million Settlement*, N.Y. Times (Nov. 2, 2016), https://www.nytimes.com/2016/11/03/business/media/gawker-hulk-hogan-settlement.html (describing the groundbreaking jury verdict that awarded former professional wrestler Hulk Hogan $140 million after the gossip news site Gawker.com published a sex tape featuring him). *But cf.* Erin C. Carroll, *Making News: Balancing Newsworthiness and Privacy in the Age of Algorithms*, 106 Geo. L.J. 69, 77–81 (2017) (discussing cases in which the courts have "largely left the role [of determining what is newsworthy or of legitimate public interest] to the press").

Outsourcing public figure determinations to Google News may well result both in too much harmful speech being kept up *and* in too much benign speech being taken down. As for the former, consider the case of an "involuntary public figure." Although a rare phenomenon in the physical world, the involuntary public figure is far from an unusual occurrence in the online realm (more on this below). Countless stories exist of relatively unknown individuals being filmed or photographed and then finding themselves subject to widespread online shaming and related news coverage.[52] Should such an individual report any particularly offensive posts to Facebook for violating the company's anti-bullying rules, the Google News search results would indicate that the individual is a public figure at the center of a newsworthy event and—at least until recently—the posts would stay up. Google News is unequipped to distinguish between situations where people have voluntarily "thrust themselves to the forefront of particular public controversies"[53] and situations where speech platforms have themselves facilitated the creation of "news."

As for the problem of taking down too much benign speech, consider that the vast majority of content that gets flagged for Facebook moderators, according to many different Facebook employees with whom I have spoken over the years, is not speech that amounts to bullying, defamation, or a privacy violation but rather content that certain users simply don't like. Moreover, a great number of virtual communities and groups have formed on Facebook, many of which have their own distinctive cultures and social structures. Provocative content flagged in these communities may not seem to involve any "public figures" when judged against a global Google News search and therefore may be removed even it involves a matter of intense interest within that local community.

In short, relying exclusively on algorithmic news aggregators to determine who is and who is not a public figure is an invitation to over- and under-removal of content. Whereas the Supreme Court has essentially folded the determination of whether a tort plaintiff is a public figure into a larger inquiry into whether the speech in question is sufficiently a matter of public concern, Facebook seems to have done the opposite by using newsworthiness as an element of defining public figures. It may be impossible to define either of these concepts in an entirely satisfying way. At the very least, however, both courts and platforms should start to rethink the treatment of people involuntarily thrust into the spotlight, as the next section describes.

B. The Rise of the Involuntary Public Figure

In attempting to define the concept of public figures, the *Gertz* Court expressly included not only individuals who "have assumed roles of especial prominence in the affairs of society" but also individuals who have "thrust [themselves] into the vortex of [a] public issue."[54] Both of these formulations seemed to assume that public figure status is something voluntarily attained. But could there be an involuntary public figure? The Court gave a vague and dismissive answer: "Hypothetically, it may be possible for someone to become a public figure through no purposeful action of his own, but the instances of truly involuntary public figures must be exceedingly rare."[55]

---

[52] *See generally* Kate Klonick, *Re-Shaming the Debate: Social Norms, Shame, and Regulation in an Internet Age*, 75 Md. L. Rev. 1029 (2015).
[53] *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 345 (1974).
[54] *Gertz*, 418 U.S. at 345, 352.
[55] *Id.* at 345.

Cases since *Gertz* have done little to clear up this vagueness. As a general matter, these cases require that a person who is not in a role of especial prominence "take voluntary, affirmative steps to thrust himself or herself into the limelight," and "they make it difficult for anyone to be found an involuntary public figure."[56] The *Restatement (Second) of Torts* defines involuntary public figures as "individuals who have not sought publicity or consented to it, but through their own conduct or otherwise have become a legitimate subject of public interest. They have, in other words, become 'news.'"[57] The only examples given by the *Restatement* of such figures are "victims of crime" and "those who commit crime or are accused of it."[58]

For the first time since the Court imagined them in *Gertz*, the democratization of publishing platforms on the internet has created a generation of truly involuntary public figures. Some of these public figures are more compellingly "involuntary" than others.[59] One of the clearest examples in recent internet history is that of "Alex from Target." On November 2, 2014, an anonymous Twitter user "tweeted a picture of a Target employee wearing the name tag 'Alex' and bagging items behind the cashier. In the following 24 hours, the tweet gained over 1,000 retweets and 2,000 favorites."[60] Over the next day, "the hashtag #AlexFromTarget was mentioned more than one million times on Twitter while the keyword 'Alex From Target' was searched over 200,000 times on Google."[61] Shortly thereafter, Twitter users started an effort to identify the "Alex" in the photo, which resulted in the publication of his Twitter handle, @acl163, at which time he amassed more than 250,000 followers.[62] Two days later, he appeared on the television talk show *Ellen*. Death threats, denigrating posts, and "fabricate[d] stories" soon followed.[63]

It is hard to argue that Alex from Target, a "global celebrity" with hundreds of thousands of social media followers,[64] is merely a private figure. Similarly, it is hard to argue that Alex from Target is a *voluntary* public figure who thrust himself into the vortex of a public issue by bagging groceries at his part-time job. Moreover, Alex from Target does not fall into the one category of involuntary public figures that has been clearly established in the case law thus far: people who have been victims of crimes or accused of committing crimes. Because of the strong public interest in crime and criminal justice, courts have been very reluctant to allow liability for harmful speech about such individuals, even when their stories are highly sympathetic.[65]

---

[56] Erwin Chemerinsky, *Constitutional Law* 1474 (3d ed. 2009).

[57] *Restatement (Second) of Torts* § 652D cmt. f (Am. Law Inst. 1977).

[58] *Id.*

[59] *Compare, e.g.*, Jon Ronson, *How One Stupid Tweet Blew Up Justine Sacco's Life*, N.Y. Times Mag. (Feb. 12, 2015), https://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html (describing the instant notoriety a thirty-year-old communications executive received after a racist tweet), *with* A.J. Willingham, *Why the Mom of a Child with a Facial Deformity Fought to Take Down Just One Cruel Tweet*, CNN (Feb. 7, 2018), https://www.cnn.com/2018/02/07/health/sophia-weaver-natalie-facial-deformities-advocates-medically-fragile-kids-trnd/index.html (describing online abuse directed at a child with a rare genetic disease that causes facial deformities).

[60] *Alex from Target/#AlexFromTarget*, Know Your Meme, https://knowyourmeme.com/memes/alex-from-target-alexfromtarget (last visited Sept. 24, 2018).

[61] *Id.*

[62] *Id.*

[63] Nick Bilton, *Alex from Target: The Other Side of Fame*, N.Y. Times (Nov. 12, 2014), https://www.nytimes.com/2014/11/13/style/alex-from-target-the-other-side-of-fame.html.

[64] *Id.*

[65] *See, e.g.*, *Florida Star v. B.J.F.*, 491 U.S. 524 (1989) (finding that public access to lawfully obtained truthful information outweighed the privacy interests of a rape victim, even where state law barred the press from reporting the names of sexual assault victims and publication had led to the victim's harassment).

How should First Amendment law treat an involuntary public figure such as Alex from Target in whom the public interest arguably has less relevance to democratic deliberation? One option is to decide that this collection of facts makes such a person more like a private figure than a public one and therefore to give lower constitutional protection to harmful speech about her. This approach recognizes the importance of not allowing the actions of others to extinguish ordinary people's privacy rights. A second option would be to treat Alex from Target like a crime victim. Under this approach, judges would not have to make controversial judgments about the degree to which speech about Alex from Target does or does not contribute to public debate; all the memes, posts, and cultural engagement generated by his celebrity would be enough to trigger First Amendment protection for anything that is said about him. For better or worse, the privacy and reputational interests of "Alex the Person" would give way to the public interest in Alex from Target.

Perhaps most persuasively, however, the Alex from Target episode might be seen as a reason to dispatch with the "voluntary" and "involuntary" concepts altogether. The Court's original rationale for granting special First Amendment protection to speech about public figures, as discussed in Part I, emphasized that such figures had greater access to means of rebuttal and had assumed the risk of unwanted publicity through their own actions. Yet in both systems of speech regulation discussed in this paper—courts and private platforms—these rationales are becoming increasingly outmoded. The ease of publishing speech online means that virtually every person now has access to means of rebuttal. The powerful and the well-connected may be better equipped to amplify their speech, at least in some cases. But for everyone who feels victimized by the expression of others, projecting one's counterspeech has never been easier, even as raising one's voice above the fray has never been harder.[66] Moreover, the enormous volume of hateful and harassing speech that circulates online has raised the stakes for victims of such speech. To require the targets of this speech to surrender their privacy rights whenever they speak out in response risks "blaming the victim"[67] and chilling their expression. The growing number of truly involuntary public figures, meanwhile, means that there will be more and more victims who have not so much thrust themselves into the vortex of a public controversy as been consumed by it.

In recent months, as discussed in Section II.C, Facebook appears to have adopted several new criteria for identifying public figures and to have moved away from its prior position that public figures receive no protection from bullying or harassment. Even if speech concerns public figures, the company's content moderators will now consider whether it is directed "toward" an individual on a personal page or group page, rather than in general

[66] *Cf.* Tim Wu, Knight First Amendment Inst., *Is the First Amendment Obsolete?* 2 (2017), https://knightcolumbia.org/sites/default/files/content/Emerging%20Threats%20Tim%20Wu%20Is%20the%20First%20Amendment%20Obsolete.pdf ("The most important change in the expressive environment can be boiled down to one idea: it is no longer speech itself that is scarce, but the attention of listeners.").

[67] For discussion of this general problem online and in other areas of law, see Danielle Keats Citron, *Hate Crimes in Cyberspace* 77–78 (2014); Heidi M. Hurd, *Blaming the Victim: A Response to the Proposal That Criminal Law Recognize a General Defense of Contributory Responsibility*, 8 Buff. Crim. L. Rev. 503 (2005); Josephine Ross, *Blaming the Victim: 'Consent' Within the Fourth Amendment and Rape Law*, 26 Harv. J. Racial & Ethnic Just. 1 (2010); JoAnne Sweeny, *Gendered Violence and Victim-Blaming: The Law's Troubling Response to Cyber-Harassment and Revenge Pornography* 8 Int'l J. Technoethics 18 (2017).

conversation elsewhere on the site, as well as whether the language rises to the level of an "attack."[68] This new policy continues to permit "involuntary" public figures to be subject to more bullying and harassment than nonpublic figures. Given that platforms like Facebook must make millions of content moderation decisions a day, there is no way for the company to analyze every case of arguable "involuntariness" in a rigorous manner. The new policy does, however, allow for more sensitive judgments about speech concerning the full range of public figures. Going forward, both Kim Kardashian and Alex from Target should have an easier time convincing Facebook to remove the most abusive posts made directly against them.

C. Straddling the Role of Governor and Publisher

In a long line of cases, U.S. courts have developed the public figure and newsworthiness doctrines to help balance promotion of uninhibited speech with protection from harmful speech. Newspapers and, later, other types of media companies were given substantial deference in deciding what to publish. Platforms like Facebook do not map neatly onto this traditional model. They are both the governors, setting speech policies and adjudicating speech disputes, and the publishers, controlling access to speech on behalf of speakers and listeners. They are the *Sullivan* Court, and they are the *New York Times*.

As a "new governor" of the public sphere, Facebook has created intricate rules and systems to weigh individuals' rights to control what is said about them against the values of free expression.[69] These rules reflect the strong pro-speech orientation of modern First Amendment theory, and some of the key concepts they use—including public figures and newsworthiness—originated in First Amendment law. Unlike the courts, however, Facebook does not adjudicate legal claims for monetary damages based on factual scenarios in which it has no direct stake. It adjudicates extralegal claims about whether allegedly harmful speech should be kept up or taken down *on its own platform*. In this sense, Facebook is more like a newspaper, deciding what is or isn't published within its pages.

The courts and the press have sought to preserve robust public debate on matters of public concern. Facebook seems to have this same goal, but it must also satisfy shareholders and take into account the interests, expectations, and concerns of billions of users worldwide. Facebook's mission statement is interestingly ambiguous in this regard. "Giv[ing] people the power to build community and bring the world closer together"[70] could be seen as a mandate either for vigorous suppression of antisocial content or for free speech absolutism.

It is not clear what Facebook itself makes of this ambiguity. A few weeks before the 2016 presidential election, the company's founder and chief executive Mark Zuckerberg held a closed-door "town hall" in Menlo Park, California. The meeting came in response to weeks of agitation by employees charged with creating and enforcing the company's content moderation policies. Some of these employees were upset that even though then-candidate Donald Trump's posts about a Muslim ban appeared to violate Facebook's policies on hate speech, the posts

[68] In Facebook's vernacular, "attacks" include calls for death or serious disease or disability, statements of intent or advocacy to engage in sexual activity, claims about sexually transmitted diseases, and terms pertaining to sexual activities used to describe an individual. Email from Ruchika Budhraja, spokesperson, Facebook, to author (Sept. 10, 2018) (on file with author).
[69] *See generally* Klonick, *supra* note 25, at 1630–58.
[70] *About*, Facebook, https://www.facebook.com/pg/facebook/about (last visited Sept. 24, 2018).

were not taken down. Over the course of the two-hour meeting, Zuckerberg justified this decision on the ground that Trump's status as a "public figure" made the posts "newsworthy or in the public interest."[71] "In the weeks ahead," Facebook announced shortly afterward, "we're going to begin allowing more items that people find newsworthy, significant, or important to the public interest—even if they might otherwise violate our standards."[72]

A little less than two years later, on July 26, 2018, Facebook removed four videos from pages belonging to Alex Jones, the controversial radio host, conspiracy theorist, and founder of the fake news site Infowars. The removed videos were found to have violated Facebook's Community Standards by encouraging physical harm (bullying) and attacking "someone based on their religious affiliation or gender identity" (hate speech).[73] A few weeks later, more content was removed, triggering all of Jones's pages to be unpublished on Facebook. In a statement made after the decision, Facebook explained that "upon review, we have taken [Jones's pages] down for glorifying violence, which violates our graphic violence policy, and using dehumanizing language to describe people who are transgender, Muslims and immigrants, which violates our hate speech policies."[74]

Both the Trump and the Jones decisions concern high-profile figures with massive followings who made statements that violate Facebook's policies against hate speech, yet one was blocked and the other not. Facebook's explanation of its policies suggests that the two decisions can be reconciled through a simple test that balances "the public interest value of the content against the risk of real-world harm."[75] Jones's speech was bullying, hateful, and low-value, on this view, and it was specifically harmful to other users on the site. To many at Facebook, blocking Jones may have appeared a relatively easy call. But it seems equally easy to argue that Trump's anti-Muslim rhetoric, which led more or less directly to a travel ban once he became president, did much more significant "real-world harm" than Jones's conspiracy theories. Similarly, it is unclear whether Jones's Facebook pages would have stayed up had he simply run for political office and thereby *amplified* his offending messages and their potential negative impact on society.

There are no easy answers to the question of how Facebook should strike the balance between protecting users and protecting free expression. Debates like the ones raised by these examples—whether to publish inflammatory content and what factors to weigh in that decision—are usually thought to be the province of newsrooms. Today, alongside the adjudication of individual speech grievances, they are also the province of platforms.

---

[71] Deepa Seetharaman, *Facebook Employees Pushed to Remove Trump's Posts as Hate Speech*, Wall St. J. (Oct. 21, 2016), http://www.wsj.com/articles/facebook-employees-pushed-to-remove-trump-posts-as-hate-speech-1477075392. Although Zuckerberg used the term "public figure," it appears that he did so casually, as a proxy for newsworthiness and not in reference to Facebook's public figure policy, which applies only to bullying and harassment and would not be applicable here.

[72] Kaplan & Osofsky, *supra* note 43.

[73] Casey Newton, *How Conspiracy Sites Keep Outsmarting Big Tech Companies*, Verge (July 28, 2018), https://www.theverge.com/2018/7/28/17623290/facebook-youtube-infowars-ban-discipline-policies.

[74] *Enforcing Our Community Standards*, Facebook Newsroom (Aug. 6, 2018), https://newsroom.fb.com/news/2018/08/enforcing-our-community-standards.

[75] *Introduction*, Facebook Community Standards, https://www.facebook.com/communitystandards/introduction (last visited Sept. 24, 2018).

IV. Conclusion

The realization that Facebook is playing the part of the *Sullivan* Court and the *New York Times* not only is descriptively illuminating but also provides tools for more nuanced normative assessments of how Facebook should develop its speech policies going forward. These two roles, of governor and publisher, can supply guiding values. Facebook should be somewhat less concerned with avoiding the suppression of speech when it finds itself acting like a court and evaluating claims that specific content has harmed a specific individual, and more concerned with avoiding suppression when it finds itself acting like the press in evaluating general newsworthiness.

Under its existing policies, Facebook acts most like a court when making "public figure" determinations for purposes of deciding whether content that violates the site's anti-bullying rules will nevertheless stay up. In this role, Facebook should maintain a robust commitment to free expression but also recognize that features of the internet—including the dynamics of online virality, shaming, and involuntary notoriety—can amplify the harms of abusive remarks directed at particular public figures. In contrast, when Facebook acts like the press in balancing considerations of newsworthiness against more generalized concerns of social harm, the company should be guided by a strong preference for keeping up as much human speech as possible consistent with its users' safety and security—perhaps more so than a traditional media outlet that doesn't also double as a public forum. And across each of these domains, Facebook must continue to develop more consistent and transparent protocols for content removal. It should not take so many hours of academic sleuthing to figure out basic facts about what these protocols look like and how they are being administered.