

WRITING IN RACE: CULTURAL DEMOCRACY IN THE DIGITAL AGE

Natalie T. Crawford[†]

I. INTRODUCTION

Social media platforms change the social conditions of speech, allowing historically oppressed groups to mobilize in new ways. For instance, in 2013, the Black Lives Matter (BLM) Movement was formed online by Patrisse Khan-Cullors, Alicia Garza, and Opal Tometi via hashtags like #BlackLivesMatter following the killing of Trayvon Martin in 2012.¹ The movement pushed back on media's portrayal of young Black men as thugs, thereby "recoding" society's perception of Black death into a continuation of racism and oppression of the Black community. More recently, in 2020, BLM came back into the cultural forefront following the deaths of George Floyd, Ahmaud Arbery, and Breonna Taylor. Protests were ignited not only in the streets but also online in the form of social media blackouts,² brand boycotts,³ and rallied support for Black artists, creators, and businesses.⁴ In doing so, social media has become a vehicle for the cultural participation of the Black community as they voice not only the harms done to the Black community but also its successes. This highlights that as technology evolves, so do methods of meaning making, as well as the means of participation. In other words, not only can the meaning of Black death in America change, but it can be changed due to the increased participation of Black persons in the cultural ecosystem.

DOI: <https://doi.org/10.15779/Z38MW28G08>

© 2021 Natalie T. Crawford.

[†] J.D. Candidate, 2022, University of California, Berkeley, School of Law.

1. *Black Lives Matter*, BRITANNICA (Aug. 13, 2020), <https://www.britannica.com/topic/Black-Lives-Matter>.

2. Sandra Gonzalez, *Music industry leaders vow to pause business for a day in observation of Blackout Tuesday*, CNN (Jun. 2, 2020), <https://www.cnn.com/2020/06/01/entertainment/blackout-tuesday-the-show-must-be-paused/index.html>.

3. Coral Murphy, *'Boycott for Black Lives': People plan to stop spending in companies that don't support BLM*, USA TODAY (Jun. 18, 2020), <https://www.usatoday.com/story/money/2020/06/18/boycotts-people-plan-stop-spending-stores-dont-support-blm/3208170001/>.

4. Alaa Elassar, *TikTokers stand in solidarity with black creators to protest censorship*, CNN (May 19, 2020), <https://www.cnn.com/2020/05/19/us/tiktok-black-lives-matter-trnd/index.html>; Tina Bellon, *Uber Eats features black-owned restaurants on U.S., Canadian delivery platform*, REUTERS (Jun. 4, 2020), <https://www.reuters.com/article/us-uber-eats-protests/uber-eats-features-black-owned-restaurants-on-u-s-canadian-delivery-platform-idUSKBN23B3DL>.

Cultural democracy theory explores this long-established process of cultural production.⁵ While utopic, cultural democracy “describe[s] a world where audiences freely and widely engage in the use of cultural symbols...A [cultural] democracy enables the audience...to ‘resist,’ ‘subvert,’ and ‘recode’ certain cultural symbols to express meanings that are different from the ones intended by their creators.”⁶ However, not all audiences are able to “freely and widely” engage in culture. While social media has amplified Black voices, it has also allowed those who seek to oppress the Black community to continue to do so in a larger and louder way than ever before.

For example, Facebook has become a haven for militia groups akin to modern day KKK chapters.⁷ Following the shooting of Jacob Blake in Kenosha, Wisconsin, the Kenosha Guard Facebook group raised a call to arms against African Americans and their BLM allies.⁸ As a result, social media’s lack of geographical limitations worked against the Black community as many who answered the call, including Kyle Rittenhouse from Illinois, travelled to Kenosha to intimidate protestors. Once there, Rittenhouse killed Anthony M. Huber and Joseph Rosenbaum and wounded Gaiige Grosskreutz.⁹ Despite Facebook’s Community Standards preventing “dangerous individuals and organizations” on its platform and the 455 complaints filed against the group prior to the violence in Kenosha, its presence on Facebook allowed racist rhetoric to fester and manifest into real world violence.¹⁰

5. JOHN FISKE, TELEVISION CULTURE 239 (Routledge 1999) (1987). In Fiske’s theory, he uses the phrase “semiotic democracy,” which is equivalent for cultural democracy in this Note.

6. Sonia K. Katyal, *Semiotic Disobedience*, 84 WASH. UNIV. L. REV. 489, 489-90 (2006).

7. Tyler Sonnemaker, *Facebook ‘did nothing’ about violent militia and hate groups for 5 years despite being warned at least 10 times, advocacy group says*, INSIDER (Sep. 17, 2020), <https://www.businessinsider.com/facebook-ignored-warnings-violent-anti-muslim-militia-hate-groups-2015-2020-9>; see also Ewan Palmer, *Facebook Sued Over Kenosha Shootings Using 19th Century Law to Combat KKK*, NEWSWEEK (Sep. 24, 2020), <https://www.newsweek.com/facebook-lawsuit-kenosha-protest-kkk-1534003>.

8. Ryan Mac and Craig Silverman, *How Facebook Failed Kenosha*, BUZZFEED NEWS (Sep. 3, 2020), <https://www.buzzfeednews.com/article/ryanmac/facebook-failed-kenosha>; see also Russell Brandom, *Facebook takes down ‘call to arms’ event after two shot dead in Kenosha*, VERGE (Aug. 26, 2020), <https://www.theverge.com/2020/8/26/21402571/kenosha-guard-shooting-facebook-deplatforming-militia-violence>.

9. Brad Parks and Artemis Moshtanghian, *Kenosha shooting suspect Kyle Rittenhouse posts \$2 million bail*, CNN (Nov. 20, 2020), <https://www.cnn.com/2020/11/20/us/kyle-rittenhouse-kenosha-shooting-bail-release/index.html>.

10. Mac & Silverman, *supra* note 8’ *Dangerous Individuals and Organizations*, Facebook: Community Standards, https://www.facebook.com/communitystandards/dangerous_individuals_organizations (last visited Sep. 30, 2022).

While racist speech may not always result in physical violence as in Kenosha, the harm is no less *real*. When confronted with racist speech, it affects someone at not only the level of individual dignity, but it is also felt on the level of the collective, as members of a group that has been historically oppressed.¹¹

As a result, the digital age emphasizes two forms of speech: one that seeks to engage in cultural production and another that limits the cultural production of the former, intentionally or otherwise. Therefore, within a *race-conscious* cultural democracy, some control of racist speech is necessary in order to allow everyone to freely and widely engage.

Social media platforms are well-positioned to rethink the way in which racist speech is controlled because while the First Amendment shields some racist speech, it only prevents government encroachment. Platforms, as private actors, are therefore given breadth of discretion as to how they can conceptualize and control racist speech. Facebook in particular has repeatedly found itself at the intersection of race and speech online. In May 2020, the platform established an Oversight Board to help Facebook manage speech on its sites (Facebook and Instagram).¹² Two of the Board's first six cases concerned hate speech content.¹³

This Note uses Facebook and its Oversight Board to examine how a race-conscious cultural democracy might function online. It concludes that Facebook's current definition of hate speech stands open to abuse and that more transparency and methods of information dissemination would enable Facebook to better trust its users, thereby allowing users to more effectively regulate themselves. However, while community control is the most democratic, it stands open to the criticisms of cancel culture and an inability to effectuate substantive change against racism. Therefore, some form of corporate oversight is needed, and the Oversight Board is a promising method of such control.

Part II of this Note lays out the cultural democracy framework as dictated by current legal theorists and proposes how it should be reworked to account for inequalities between persons and ideas, particularly those of historically oppressed groups. It then establishes a theory of race-conscious cultural

11. Richard Delgado, *Words that Wound: A Tort Action For Racial Insults, Epithets, and Name-Calling*, 17 HARV. C.R. C.L. LAW REV. 133, 136 (1982).

12. Nick Clegg, *Welcoming the Oversight Board*, FACEBOOK: NEWSROOM (May 6, 2020), <https://about.fb.com/news/2020/05/welcoming-the-oversight-board/>.

13. *Announcing the Oversight Board's first cases and appointment of trustees*, OVERSIGHT BD. (Dec. 2020), <https://www.oversightboard.com/news/719406882003532-announcing-the-oversight-board-s-first-cases-and-appointment-of-trustees/>.

democracy, which recognizes that in order for all persons to participate in cultural production, racist speech must be controlled. Part III outlines the spectrum that racist speech functions on and establishes that if speech is a message (1) “of racial inferiority” (2) “directed against a historically oppressed group” (3) that is “persecutorial, hateful, and degrading,” then it should be considered actionable racist speech that warrants control.¹⁴ With this as a guide, this Note examines Facebook’s own definition and finds it insufficient. Part IV introduces the hybrid model for speech control in the digital age: community control buttressed by corporate oversight. It examines the way in which community control works both in theory and already in practice on Facebook and then evaluates the benefits of corporate oversight. It then concludes that Facebook should institute more tools to allow it to better trust its users and that Facebook’s Oversight Board is an effective method of corporate oversight that promotes a race-conscious cultural democracy.

II. THE CULTURAL DEMOCRACY FRAMEWORK

Culture is the glue of society and the result of collective meaning making.¹⁵ It is a set of conditions that provide a baseline for conversation in society, and it influences everything from dating practices and Twitter trends to political systems and international relations.¹⁶ It is produced in part through a process of democratization. Information, knowledge, and ideas are under constant production or are otherwise remixed (built upon and manipulated into something new).¹⁷ This process maps onto the entire evolution of technology, and the intertwined histories of the printing press and the Bible provides a fitting example to give practical context to this process of democratization.

The printing press, while a centuries-old precursor to modern day social media platforms, emphasizes a history of culture production (and

14. Mari J. Matsuda, *Public Response to Racist Speech: Considering the Victim’s Story*, 87 MICH. L. REV. 2320, 2357 n. 36 (1989).

15. YOCHAI BENKLER, *THE WEALTH OF NETWORKS* 279 (2006) (“It is difficult to specify how [culture] functions in terms readily amenable to a conception of individuals whose rationality and preferences for their own good are treated as though they preexist and are independent of society. A conception of culture requires some commonly held meaning among these individuals. Even the simplest intuitive conception of what culture might mean would treat this common frame of meaning as the result of social processes that preexist any individual, and partially structure what it is that individuals bring to the table as they negotiate their lives together, in society or in a polity.”).

16. *Id.* at 282 (“[Culture] is a frame of meaning from within which we must inevitably function and speak to each other, and whose terms, constraints, and affordances we always negotiate.”).

17. *See generally* LAWRENCE LESSIG, *REMIX: MAKING ART AND COMMERCE THRIVE IN THE HYBRID ECONOMY* (2008).

reproduction). In comparison to the 30,000 books available in all of Europe in the early 1450s, Gutenberg managed to produce an astounding 180 Latin Vulgate Bibles of consistent quality in a matter of years.¹⁸ Not even a hundred years after the Gutenberg press revolutionized printing, Martin Luther became the first “best-selling author” with his Bible translation seeing more than 40 editions.¹⁹ Needless to say, the production potential of the printing press not only fueled Luther’s individual success but also the Protestant Reformation as a whole. In Luther’s own words, “Printing is the ultimate gift of God and the greatest one.”²⁰

Like Luther, the consumers of the Bible prescribe their own meanings onto the text. Then as producers, these once-consumers reproduce those meanings for others via new translations and versions, reformulating religious culture. With each iteration of this cycle, power is decentralized. While the Church was once relied on as the sole source of salvation, vernacular translations put the power of salvation in the hands of the consumers. As a result, as of 2019, there were around 670 translations of the Bible.²¹

This is the process of the cultural democratization—as means of production improve, the amount of cultural material in the marketplace increases and the reconfigurations of that material proliferate. In other words, ideas spread, information multiplies, and power decentralizes. A simple book burning of a single heretic no longer suffices to silence opposition to the Church. Books print faster than they are burned, and eBooks and PDFs fan the flames of a growing cultural democracy.

A. EXISTING CULTURAL DEMOCRACY THEORY

While cultural democracy, as the above Biblical example shows, is nothing new, the theory that revolves around it is a creature of media studies initially developed by Professor John Fiske in 1987.²² At that time, consumers had access to more television networks and therefore more shows, and it was this televised information buffet from which Fiske excavated his theory of semiotic democracy. Under this theory, audiences were not blank canvasses, incapable of critical consumption; instead, according to Fiske, the identities of individual

18. *Id.*

19. Dave Roos, *7 Ways the Printing Press Changed the World*, History: History Stories (Aug. 28, 2019) <https://www.history.com/news/printing-press-renaissance>.

20. Pearl Goyal, *Gutenberg's Printing Press*, WOODWARD JOURNAL (May 21, 2020), <http://www.woodwardjournal.com/gutenbergs-printing-press>.

21. Daniel Esparza, *How many times has the Bible been translated?*, ALETEIA (Jan. 1, 2019), <https://aleteia.org/2019/01/06/how-many-times-has-the-bible-been-translated/>.

22. Fiske, *supra* note 5 at 239. John Fiske used the phrase “semiotic democracy,” which was then interchanged with “cultural democracy” by other theorists.

consumers allow them to superimpose their own meanings onto cultural texts like television shows—meanings that might differ from those intended by show writers and producers—and then reinsert those new meanings into the marketplace.²³ As a result, meaning making is not necessarily a monopolistic hierarchy, but rather a dialogue between cultural producers and consumers in which the consumers become producers themselves.²⁴

Jack M. Balkin and other legal scholars²⁵ brought Fiske's semiotic democracy into the legal realm to re-envision what rights are and, perhaps more importantly, what rights should be when it comes to speech.²⁶ The encroachment of the digital revolution only further disrupted the hierarchies that Fiske was concerned with a decade prior. In particular, the digital revolution further cut costs of the initial distribution of cultural materials, increased content flows across cultural and geographic borders, cut costs of remixing content, and as a result, further democratized culture.²⁷ For comparison, prior to Gutenberg, a book had to be transcribed by hand over days, if not months, and struggled to meet a growing demand.²⁸ However, in 2022, filming a TikTok, posting a Facebook post, or sending a Tweet takes a matter of seconds, is free, and is distributed worldwide. Therefore, as technology changes, so must the rights that secure and encourage participation; in other words, technology clarifies and alters social relations, requiring rights such as free speech to adapt as well.²⁹

To this end, free speech is not integral to only political democracy but also cultural democracy. While free speech encourages political debate and the democratic political process, free speech is integral to the cultural democracy as well. According to Balkin, the Internet has several important traits related to free speech, which suggest that a conception of free speech as merely related to political democracy is too limited:

Speech ranges over a wide variety of subjects, including not only politics but also popular culture. The speech of ordinary people is full of innovation and creativity. That creativity comes from building

23. See generally Fiske, *supra* note 5.

24. Katyal, *supra* note 6 at 490.

25. See generally 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, 2 (2004); Katyal, *supra* note 6; Benkler, *supra* note 15; Lessig, *supra* note 17.

26. Balkin, *supra* note 25 at 52 (“Rights dynamism is the claim that the nature, scope, and boundaries of rights, and in particular fundamental rights like speech, are continually shifting with historical political, economic, and technological changes in the world.”).

27. *Id.* at 6-9.

28. See Richard W. Clement, *Medieval and Renaissance books*, UTAH ST. U. LIBR. FAC. & STAFF PUBLICATIONS 1, 8 (1997).

29. Balkin, *supra* note 25 at 52.

on what has come before. Speech is participatory and interactive as opposed to mere receipt of information. It merges the activities of reading and writing, of production and consumption. Finally, speech involves cultural participation and self-formation. The Internet reminds us how central and important these features are to speech generally. It reveals to us in a new way what has always been the case.³⁰

Understanding free speech in this way—as central to both political and cultural democracy—allows democracy to be as much about culture as it is about governance. Free speech in the political sphere allows people to debate and create politics, laws, and public policy.³¹ Free speech in the cultural sphere allows people to create and progress the culture via meaning making, which in turn produces people’s identities.³² This extends beyond consumerism and deciding which brand of cereal to buy. In reposting a Facebook post, information, knowledge, and/or ideas are not only shared but can be modified through the inclusion of a caption or primer that (re)frames how a future viewer might understand the content. A civil rights activist can share a racist Facebook post in order to educate and counteract the harm done by the original post by providing context and educational resources.³³ Free speech in the cultural democracy is the freedom to create. It is “an active engagement with the world.”³⁴

In addition, in a robust cultural democracy, speech “is interactive and appropriative.”³⁵ Speech is interactive because it relies on the exchange between persons and is not a one-way street. As the speaker communicates, the receiver internalizes, digests, and responds. In doing so, one either assimilates to or rejects the culture created in that speech.³⁶ Speech is appropriative because to move culture forward, it remixes—builds on and manipulates—existing cultural materials. For this reason, speech is as collective

30. *Id.* at 32.

31. In December, 1791, James Madison wrote in the *National Gazette*, “Public opinion sets bounds to every government, and is the real sovereign in every free one.” James Madison, *For the National Gazette*, FOUNDERS ONLINE (DEC. 19, 1791) <https://founders.archives.gov/documents/Madison/01-14-02-0145>.

32. Balkin, *supra* note 25 at 33.

33. The nuance that arises between racist speech on its own and racist speech used as a tool to inform and advocate creates a grey area of speech that is often difficult for platforms to sift through. Facebook even recognizes this in its Community Standards, and therefore, content that is meant to raise awareness must have its intent clearly indicated. *Hate Speech*, FACEBOOK: COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/hate_speech (last visited Sep. 22, 2022).

34. Balkin, *supra* note 25 at 33.

35. *Id.* at 4.

36. *Id.*

as it is individual; it enacts a cultural system—“a network of people interacting with each other, agreeing and disagreeing, gossiping and shaming, criticizing and parodying, imitating and innovating, supporting and praising.”³⁷

The insurgence of social media has decreased the cost of participating in cultural democracy and increased access to means of cultural production to the benefit of historically oppressed groups. In the 18th century, slave codes criminalized gatherings of three or more slaves and prohibited slaves from riding horses.³⁸ By controlling bodies, slaveholders were able to effectively control the sharing of ideas, information, and knowledge—thereby controlling the cultural collective of Black Americans. However, in the digital age, space between bodies no longer affects the ability to congregate and engage in cultural production. The Black Lives Movement is a concrete example of this—multiple protests are able to erupt in various cities on the same day based solely on Facebook event pages. Black Twitter is an even more fluid example of the growing cultural participation of Black voices online. Through platform engagement via coded language and hashtags, Black Twitter users can find each other, share experiences, and communicate. One prominent focus of Black Twitter every year is the Oscars. In 2015, 86 percent of top films featured white actors and all 20 acting nominations were white, leading to an annual hashtag of #OscarsSoWhite.³⁹ When Black individuals speak as a collective via Black Twitter, they are able to garner more attention to the dearth of Black representation in film and work to reshape culture in a way that accounts for systematic racism. Nonetheless, it would be inaccurate to view pockets of cultural participation as equal participation—in 2022, the numbers have yet to improve with only one Black nominee.⁴⁰

B. RACE-CONSCIOUS CULTURAL DEMOCRACY THEORY

In part due to its utopic vision of cultural production and reproduction, the implications and limitations of current cultural democracy theory are far reaching and conflict with the actual American cultural ecosystem. Not all speech is interactive—in particular, racist speech limits the interaction of historically oppressed persons in culture. Then, when historically oppressed persons do engage in cultural participation, the appropriation of their speech

37. *Id.*

38. Jill Lepore, *The Invention of the Police*, NEW YORKER (Jul. 20, 2020), <https://www.newyorker.com/magazine/2020/07/20/the-invention-of-the-police>.

39. Reggie Ugwu, *The Hashtag That Changed the Oscars: An Oral History*, N.Y. TIMES (Sep. 20, 2020), <https://www.nytimes.com/2020/02/06/movies/oscarssowhite-history.html>.

40. See Katey Rich, *Oscar Nominations 2022: Full List of Nominees Led by The Power of the Dog*, VANITY FAIR (Mar. 27, 2022), <https://www.vanityfair.com/hollywood/2022/02/oscar-nominations-2022-list>.

does not have the positive effect that Balkin espouses. Balkin approaches speech with the false assumption that culture is a singular expression despite its pluralistic nature. Therefore, a race-conscious cultural democracy must recognize and account for the harms of certain speech by instituting methods of control.

Balkin's conception of speech is as interactive disregards the harm that can result from some speech; in particular, within a cultural democracy, speech that is anti-productive fails to be interactive. Some believe that in order to protect the speech of oppressed persons, the speech of the oppressors (or historically dominant group) must also be protected, making counter speech the most effective way to prevent racist speech.⁴¹ In other words, by protecting both sides' right to speech, a dialogue is created between the two.⁴² However, this stance is unpersuasive because, when it comes to racist speech, dialogue is not productive and therefore is not interactive. For example, the speech espoused by the KKK is intended to incite fear and silence the voices of African Americans, and therefore it does not promote dialogue between members of society. Fearmongering impedes healthy dialogue and debate. In addition, the dialogue between the KKK and African Americans is not productive because the former seeks to not only delegitimize but also *dehumanize* the latter.

Furthermore, even unintentional racist speech has a capacity of harm that prevents it from being interactive as intended under Balkin's theory. A general influx in racist speech online has routinely been shown to result in an influx of real world violence, regardless of the intent of any one racist speaker.⁴³ Studies have also shown that repeated exposure to racist speech takes a mental toll on individual members of a historically oppressed group regardless of the intention of the original speaker.⁴⁴ In particular, historically oppressed persons

41. Matsuda, *supra* note 14, at 2326-27 n. 36.

42. *See id.* at 2357.

43. *See e.g.*, Katarzyna Bojarska, *The Dynamics of Hate Speech and Counter Speech in the Social Media: Summary of Scientific Research*, CTR. FOR INTERNET AND HUM. RTS., https://cibr.eu/wp-content/uploads/2018/10/The-dynamics-of-hate-speech-and-counter-speech-in-the-social-media_English-1.pdf (last visited Sep. 30, 2022); *Hate Speech on Twitter Predicts Frequency of Real-life Hate Crimes*, NYU (Jun. 24, 2019), <https://www.nyu.edu/about/news-publications/news/2019/june/hate-speech-on-twitter-predicts-frequency-of-real-life-hate-crim.html>; Rachel Hatzipanagos, *How online hate turns into real-life violence*, Wash. Post (Nov. 29, 2018), <https://www.washingtonpost.com/nation/2018/11/30/how-online-hate-speech-is-fueling-real-life-violence/>; *In Germany, online hate speech has real-world consequences*, ECONOMIST (Jan. 12, 2018), <https://www.economist.com/graphic-detail/2018/01/12/in-germany-online-hate-speech-has-real-world-consequences>; Melissa Fisher Paoni, *Understanding the harm in hate speech*, ILL. TIMES (Nov. 8, 2018), <https://www.illinoistimes.com/springfield/understanding-the-harm-in-hate-speech/Content?oid=11447249>.

44. Bojarska, *supra* note 43.

have to be constantly prepared to encounter racist speech, and because racist speech functions on a spectrum, in a variety of forms, such preparedness requires a heightened level of vigilance.⁴⁵ This translates to increased stress, known as “minority stress,” which can in turn lead to adverse health conditions such as anxiety and depression.⁴⁶ Therefore, repeated engagement with racist speech is not only ineffective at controlling it but also increases the physical and mental harm done to historically oppressed persons. As a result, it fails to be interactive within the scope of Balkin’s definition of speech in a cultural democracy.

Balkin’s conception of speech as appropriative is equally harmful because it ignores the dynamic interplay between different cultures. When there is a power imbalance between the appropriator and those being appropriated and the appropriator profits from their actions, the harm done is incongruous with Balkin’s conception of cultural democracy. For example, in 2014, fashion brand DKNY launched “Urban Fabulous” at New York Fashion Week, featuring white women with laid edges.⁴⁷ Not only was the use of “urban” coded language for Black, but the hairstyles themselves, historically deemed “ghetto,” were reimagined to be a signifier of high fashion when donned by white women.

This is not to say all appropriation is bad; to the contrary, conversation between two cultures can enhance the lived experience of people in both and particularly those that live at their intersection. For example, in 2020, a white TikTok user posted a video asking Black users for methods of controlling her baby hairs in a way that did not appropriate the storied history behind laid edges.⁴⁸ Within the Black and Latina communities, laying edge has gone beyond a way of taming baby hairs and has been transformed into art—a form of speech between Black and Latina women, expressed through complex swirls and swoops.⁴⁹ The response from Black TikTok was entirely supportive,

45. *Id.*

46. See e.g., *What is Minority Stress?*, CNTR. FOR COMMUNITY PRAC., <https://www.urccp.org/article.cfm?ArticleNumber=69> (last visited Mar. 3, 2021); Michael Dentato, *The minority stress perspective*, AM. PSYCHO. ASS’N (APR. 2012), <https://www.apa.org/pi/aids/resources/exchange/2012/04/minority-stress>.

47. Kara Brown, *The problem with Baby Hairs, ‘Urban’ and the Fashion Industry*, JEZEBEL (Sep. 17, 2014), <https://jezebel.com/the-problem-with-baby-hairs-urban-and-the-fashion-indu-1635947700>.

48. Alex Lasker, *People are loving this TikToker who asked about a hairstyle before potentially appropriating it*, IN THE KNOW (Nov. 13, 2020), <https://www.intheknow.com/2020/11/13/tiktok-baby-hairs/>.

49. Ali King, *Four Celebrity Natural Hairstylists Share Their Go-To Baby Hair Tips and Tricks*, VOGUE (Aug. 6, 2019), <https://www.vogue.com/article/celebrity-natural-hairstylists-baby-hairs-tips-yara-shahidi>.

providing her with product suggestions and tips to manage her “shorty hairs.”⁵⁰

The cultural exchange in the DKNY example is invariably different from the TikTok example because (1) there is a greater power imbalance between the Black/Latina communities and the high fashion sphere that has routinely denied them a role in setting beauty standards and (2) DKNY sought to profit off their culture. As a result, while not all speech is appropriative, it can very easily become so when it profits off the cultures of historically oppressed groups.

If the foundation of cultural democracy is speech, then cultural democracy theory must account for the anti-interactive nature and harmful appropriation of certain speech. This is especially true when it comes to racist speech and appropriation across racial lines. A race-conscious approach to cultural democracy realizes this complex interplay between race and cultural production. Therefore, while it may seem anti-democratic and counterintuitive to control some speech to promote cultural democracy overall, this must be the case. By controlling racist speech, historically oppressed persons can more freely engage in their own cultural production without worry of harmful appropriation, and they are empowered to interact across cultures in a productive way.

III. DEFINING RACIST SPEECH

As cultural democracy theory recognizes, speech is the bread and butter of social interaction and thus embodies the complicated nature of social relationships. Race, ethnicity, gender, sexuality, religion, place of origin, etc., all affect social relationships and therefore are fundamental to speech. However, given the various histories that mold the identities of different historically oppressed groups, it is important to consider them with their own unique histories at the forefront. In other words, while a discussion on race can reciprocally inform a discussion on gender, the Venn diagram does not always overlap, and to cast a net wide enough to encompass all historically oppressed groups would necessarily dilute the particular experiences of any one group.⁵¹ Therefore, in order to think about what a more equitable cultural democracy theory would look like, racist speech is taken as the proxy for this Note, but as Part V will discuss, once the framework of a race-conscious

50. Lasker, *supra* note 48.

51. See Kimberlé Crenshaw, *Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color*, 42 STANFORD L. REV. 1241, 1242 (1991) (“The problem with identity politics is not that it fails to transcend difference, as some critics charge, but rather the opposite—that it frequently conflates or ignores intragroup difference.”).

cultural democracy has been established, it can be an effective tool for examining other types historically oppressed groups and hate speech directed toward them.

A. THE SPECTRUM OF RACIST SPEECH

Racist speech functions on a spectrum, making it difficult to define. A conceptualization of the spectrum of racist speech places clear threats of racial violence on one end and covert, sanitized speech on the other. In the middle lies an ambiguous zone with historically laden symbols, slurs, and overt microaggressions.

Table 1: Spectrum of Racist Speech

Direct Calls for Violence	Historically laden Symbols and Slurs	Overt Microaggressions	Covert, Sanitized Speech
---------------------------	--------------------------------------	------------------------	--------------------------

Following the events in Kenosha, a lawsuit was filed against Facebook for the failure to uphold its Community Standards on militia groups on behalf of Hannah Gittings (life partner of Anthony Huber), Christopher McNeal (allegedly harassed by militia members), Carmen Palmer (allegedly threatened by militia members), and Nathan Peet (present at one of the killings).⁵² Calls for violence—such as those used by the militia groups targeting African Americans and advocating for their and their allies’ death—are the easiest to identify and therefore should be the easiest to control.

On the other side of the spectrum, is covert, sanitized speech, and the rhetoric of the CEO of Wells Fargo, Charles Scharf, is a salient example. Scharf stated in a memo that “while it might sound like an excuse, the unfortunate reality is that there is a very limited pool of black talent to recruit from.”⁵³ While this rhetoric plays into and reinforces the stereotype of African Americans as uneducated and poor, it is too deep into nuanced waters for an enforcer to attempt to regulate it. This type of covert, sanitized speech is a harm commonly remedied by public accountability alone. In the case of Wells

52. *Gettings v. Mathewson*, No. 2:20-cv-1483 (E.D. Wis. filed Sept. 22, 2020); see also Jeff Richgels, *4 file federal lawsuit against Facebook, Kyle Rittenhouse, militia groups in wake of Kenosha protest shootings*, Wisc. St. J. (Sep. 24, 2020), https://madison.com/wsj/news/local/crime-and-courts/4-file-federal-lawsuit-against-facebook-kyle-rittenhouse-militia-groups-in-wake-of-kenosha-protest/article_646089ba-dbec-5972-9e8b-acc9ddac2559.html.

53. Imani Moise, Jessica DiNapoli, and Ross Kerber, *Exclusive: Wells Fargo CEO ruffles feathers with comments about diverse talent*, REUTERS (Sep. 22, 2020), <https://www.reuters.com/article/us-global-race-wells-fargo-exclusive/exclusive-wells-fargo-ceo-ruffles-feathers-with-comments-about-diverse-talent-idUSKCN26D2IU>.

Fargo, Scharf's speech resulted in boycotts and people withdrawing their money from the bank in favor of Wells Fargo's competitors, in particular Black-owned banks.⁵⁴

However, there is a middle ground between a racially motivated militia group's call to action and a CEO's inability to understand the complexity of race in hiring pools. This middle ground is composed of historically laden racist symbols, slurs, and microaggressions. An example of a historically laden symbol is a burned cross left in the yard of an African American family's home.⁵⁵ An example of a microaggression is an African American interviewee being informed by her interviewer that her hair would be more professional if it were straightened.⁵⁶

Even within these periphery categories of racist speech, there is often a blurring that reinforces the assertion that racist speech must be viewed as a spectrum rather than distinct buckets of types of speech. In particular, microaggressions often worm their way into covert, sanitized speech just as historically laden racist symbols and slurs meld into direct calls for violence. For example, an ad agency refusing to include African Americans in a campaign because their hair does not convey the idea of luxury would lie somewhere between a microaggression and covert, sanitized speech.

B. CRITICAL RACE THEORY DEFINITION OF RACIST SPEECH

There is no universal definition of hate speech under U.S. law,⁵⁷ nor a definition of racist speech. Commonly, racist speech is defined under an umbrella definition of hate speech, which is "speech, writing, or nonverbal communication that attacks, threatens, or insults a person or group on the basis of national origin, ethnicity, color, religion, gender, gender identity, sexual orientation, or disability."⁵⁸ However, a definition like this provides little

54. Isabella Jibilian, *People are calling for a boycott of Wells Fargo after CEO blames 'very limited pool of Black talent' for the bank's trouble hitting its diverse hiring goals*, Business Insider (Sep. 23, 2020), <https://www.businessinsider.com/wells-fargo-boycott-ceo-charles-scharf-black-talent-hiring-comments-2020-9>.

55. See Dennis Romero, *Virginia man accused of burning cross on Black neighbors' front lawn*, NBC NEWS (Jun. 27, 2020), <https://www.nbcnews.com/news/us-news/virginia-man-accused-burning-cross-black-neighbors-front-lawn-n1232345>.

56. See Chanté Griffin, *How Natural Black Hair at Work Became a Civil Rights Issue*, JSTORE DAILY (Jul. 3, 2019), <https://daily.jstor.org/how-natural-black-hair-at-work-became-a-civil-rights-issue/>.

57. "Hate Speech and Hate Crime," AM. LIBR. ASS'N, <http://www.ala.org/advocacy/intfreedom/hate> (last visited Sep. 30, 2022).

58. "Hate Speech," DICTIONARY.COM, <https://www.dictionary.com/browse/hate-speech> (last visited Sep. 30, 2022).

guidance in differentiating racist speech along the spectrum in a way that controls the most problematic, unwanted racist speech.

Nonetheless, if certain speech is to be systematically controlled as necessary within a race-conscious cultural democracy, then a clearer line must be drawn between racist speech that should be controlled and speech that, while still unacceptable, should not be controlled. In other words, actionable racist speech must be separated from unactionable racist speech. And while communities can still hold accountable speech that is unactionable (as in the case of the Wells Fargo CEO), racist speech rising to the level of actionable speech requires a heightened level of control. Such control can be brought by corporate oversight in addition to public accountability, as discussed later in Part V.⁵⁹

Mari Matsuda provides two workable definitions for racist speech—one qualitative and the other engrained in distinct, value-laden elements.⁶⁰ The first definition Matsuda puts forward is that racist speech presents “an idea 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.”⁶¹ However, this first understanding of racist speech does not easily cut through the contextual ambiguity that often shrouds racist speech. Matsuda’s other definition of racist speech, however, has three well-defined characteristics: it is a message (1) “of racial inferiority” (2) “directed against a historically oppressed group” (3) that is “persecutorial, hateful, and degrading.”⁶² This second definition can be utilized to cut through the ambiguity of racist speech.

While this definitional line in the sand oversimplifies the interwovenness of racist speech outlined in Section III.A, it at least casts a net that lands between direct calls for violence and covert, sanitized speech. Furthermore, Matsuda’s conception of racist speech in this way also recognizes the importance of context, as not all speech verbal or written. In the case of symbols and imagery especially, context is important, and in these cases, “if the historical message, known to both victim and perpetrator, is racist persecution, then the sign is properly treated as actionable racist speech.”⁶³ By using Matsuda’s definition of racist speech, speech that works against a race-conscious cultural democracy because it is anti-interactive and problematically

59. *Infra* Part V.

60. Matsuda, *supra* note 14 at 2357.

61. *Id.*

62. *Id.*

63. *Id.* at 2365-66.

appropriative can be controlled, allowing historically oppressed persons to more freely and widely engage in cultural production.

Furthermore, under Matsuda's conception of actionable racist speech, expressions of hatred directed toward the oppressor—such as when Malcolm X spoke of the white devil—would not be restricted.⁶⁴ This language only arises as a byproduct of racism against African Americans, and while offensive, those targeted are a part of a historically non-oppressed group and are able to more easily retreat and reaffirm their personhood.⁶⁵ Conversely, the rhetoric espoused by neo-Nazi leader David Duke would fall within prohibited speech. In other words, so long as there is an unequal distribution of power along race lines, the harm done by the speech of the oppressed will never rival the harm done by the speech of the oppressor, and by limiting the oppressor's racist speech, the cultural participation of historically oppressed can not only increase but prosper.⁶⁶

C. FACEBOOK'S DEFINITION OF RACIST SPEECH

Facebook has used its Terms of Service (ToS) and Community Standards to delineate the types of speech unacceptable on its platform. The platform defines racist speech under the broader umbrella of hate speech and provides two main elements for objectionable speech.⁶⁷ First, it must be a "direct attack," defined as "violent or dehumanizing speech, harmful stereotypes, statements of inferiority . . . calls for exclusion or segregation."⁶⁸ Second, it must target people with "protected characteristics," with race being one such characteristic.⁶⁹ A direct attack on persons with protected characteristics aligns

64. *Id.* at 2362.

65. *Id.*

66. Furthermore, returning to the DKNY example from earlier, the coded language of "urban fabulous" would not be considered racist speech under Matsuda's definition because of a failure to meet the second prong (directed against a historically oppressed group). "Urban fabulous" teeters too far over the line into cultural appropriation. This is harmful because it effectively erases the cultural origin, history, and meaning behind laid edges, thereby negatively affecting the cultural engagement and production of African Americans. Cultural appropriation is typically driven by ignorance and is often more focused on the benefit for the appropriating culture rather than the harm done to the appropriated culture. *What is cultural appropriation and how can you spot it?*, THE WEEK (Jun. 1, 2020), <https://www.theweek.co.uk/cultural-appropriation>. As a result, this type of cultural speech lies closer toward covert, sanitized speech on the spectrum of racist speech, making it far harder to systematically control.

67. PART III: OBJECTIONABLE CONTENT, FACEBOOK COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/objectionable_content (last visited Mar. 3, 2021).

68. *Id.*

69. *Id.*

with the first and third elements of Matsuda's definition—a message of racial inferiority that is persecutorial, hateful, and disregarding.⁷⁰

However, without that second element (directed against a historically oppressed group), Facebook's definition stands open to abuse and fails to rectify the biggest concern of race-conscious cultural democracy: that racist speech silences historically oppressed persons, thereby limiting their cultural participation. Returning to the example of Malcolm X versus David Duke, if Facebook does not clarify the difference between the speech of the oppressed from the speech of the oppressor, then both Malcolm X and David Duke's speech would be removed from Facebook. This would bring to fruition the fear that attempts to control speech will be disproportionately affect the people they are meant to protect.⁷¹ By specifying the group that a definition of racist speech is intended to protect, this type of coopting can be remedied. Therefore, Facebook should amend its definition to read “we define hate speech as a direct attack on people *of a historically oppressed group* based on what we call protected characteristics.”

IV. A HYBRID MODEL OF SPEECH CONTROL FOR THE DIGITAL AGE

Reexamining how Facebook defines actionable racist speech is only the first step in working toward a race-conscious cultural democracy where all members can freely and widely engage in cultural production. A system must be put in place to ensure racist speech is controlled. To that end, the community, as the most democratic option, must be the first line of defense. However, relying solely on the community to govern speech inevitably stands open to criticism of cancel culture and inability to effectuate substantive change. Therefore, there needs to be a second layer of protection like corporate oversight; and social media platforms like Facebook are well positioned to control racist speech online.

This Part examines how the community functions to control speech—historically and online today—and then looks to how Facebook already allows its community to manage racist speech. It then turns in Section IV.B to how corporate oversight can reinforce the community in order to support a race-conscious cultural democracy online. And finally, it uses two of Facebook's

70. Matsuda, *supra* note 14 at 2357.

71. Cindy Cohn, *When Limiting Online Speech to Curb Violence, We Should be Careful*, WIRED (Aug. 9, 2019), <https://www.wired.com/story/when-limiting-online-speech-to-curb-violence-we-should-be-careful/>. Often, policies aimed at controlling racist speech disproportionately affect the speech of historically oppressed persons.

Oversight Board's first decisions to examine how corporate control might successfully work in practice.

A. ON THE FRONTLINE: THE ROLE OF THE COMMUNITY

The most democratic means of speech control is relying on the community, and when laws are not enough to adequately control racist speech—such as in the United States under the First Amendment—the community must take a larger role in controlling speech online. And given the vast amount of speech posted online, it is a matter of practicality to first rely on the on-the-ground readers and watchers to determine what speech is unwanted. This way, it becomes easier for an overseeing body to institute systematic change based upon the community response. Furthermore, this is something communities are already doing under different methods of public accountability, and it has the power to positively affect the lives of historically oppressed persons.

1. *In Theory: Benefits and Detriments of Community Speech Control*

The benefits of public accountability are often overshadowed by cancel culture rhetoric as a modern tyranny of the majority.⁷² Unsurprisingly though, the term “cancelling” itself was an internet meme on Black Twitter that was coopted by journalists, politicians, celebrities, and the like to misconstrue community accountability into feared censorship as a means of “delegitimiz[ing] the dissension that echoes from society’s margins.”⁷³ As a result of this cancel culture rhetoric, while public accountability is an important mechanism for racist speech control online, its success is limited, and its effectiveness must be amplified by additional measures like that of corporate control as discussed in Section IV.B.

Public accountability is nothing new. Historically, public accountability manifested as public shaming. For example, in the medieval era, there was the pillory, a wooden framework with three holes intended to restrain someone’s head and wrists, and those most often subjected to the pillory include perjurers, forgers, counterfeiters, and blasphemers.⁷⁴ While someone might not be displayed on the pillory for more than a couple hours, it was located in

72. Edwin J. Feulner, *Preventing “The Tyranny of the Majority,”* THE HERITAGE FOUNDATION (May 7, 2018), heritage.org/conservatism/commentary/preventing-the-tyranny-the-majority.

73. Meredith D. Clark, *DRAG THEM: A brief etymology of so-called “cancel culture”*, 5 (3-4) COMM AND THE PUB. (COMMENT. FOR SPECIAL F.) 88, 89 (2020).

74. *Pillory*, ENCYCLOPEDIA.COM, <https://www.encyclopedia.com/history/modern-europe/british-and-irish-history/pillory> (last visited Mar. 23, 2021).

a central, busy part of town to allow for the most visibility.⁷⁵ While often the result of some judicial proceeding, the crux of the pillory's effectiveness was audience engagement. People would often throw whatever they had on hand at those subjected to the pillory, from rotten food to rocks. The effects of being sentenced to the pillory went beyond the couple hours someone spent berated by passersby.⁷⁶ Especially in small towns, where reputation was everything, a pillory sentence could result in job loss and crushed marriage prospects—effectively, a social cancellation.⁷⁷

For modern comparison, the drama within YouTube's beauty influencer, or gurus, community is riddled with prime examples of public accountability. There have been numerous “[enter name] is over parties,” where an influencer's racist comments are exposed—typically on Twitter—followed by a massive unfollowing spree.⁷⁸ In the case of “Dramageddon I” in 2018, numerous beauty gurus were exposed for their previous racist tweets.⁷⁹ For example, Gabriel Zamora, Nikita Dragun, and Laura Lee all had racist tweets resurface from 2012.⁸⁰ Public uproar akin to the medieval pillory proceeded, minus the literal rotten food and rocks. Lee in particular lost half a million of her 4.5 million subscribers within the month following the exposé. She lost an estimated \$25,000–\$65,000 in income from viewership alone, in addition to losses in lucrative partnerships and brand deals.⁸¹

Dramageddon I highlights two main concerns critics have with “cancel culture.” First, it can happen to anyone at any time for anything. In particular, cancel culture revolves around the belief that no material is too old; any previous tweet can become the basis for a current cancellation, and the idiom “the Internet never forgets” stands to be abundantly true. It raises the question if some things are better left forgotten, as the cultural climate of 2012, some argue, was different than that of 2018.⁸² As a result, cancel culture can create a

75. Mathew Green, *A Grim and Gruesome History of Public Shaming in London: Part 1*, LONDONIST (Jan. 19, 2017), <https://londonist.com/2015/12/publicshaming1>.

76. *Pillory*, BRITANNICA (Oct. 1, 2022) <https://www.britannica.com/topic/pillory-penology>.

77. *Id.*

78. Lilly Gelman, *Twitter Explained: Who is Over and Where's the Party?*, MOMENT (Aug. 15, 2015), <https://momentmag.com/twitter-explained-who-is-over-and-where-the-party/>.

79. Alex Abad-Santos, *Laura Lee, Jeffree Star, and the racism scandal upending the YouTube Beauty community, explained*, VOX (Aug. 31, 2018), <https://www.vox.com/2018/8/28/17769996/laura-lee-jeffree-star-racism-subscriber-count>.

80. *Id.*

81. *Id.*

82. It is worth noting in 2012, racism was as unacceptable as it was in 2018. 2012, the publication year of most of the racist tweets concerned with in Dramageddon I, is the same year a national uproar occurred following the death of Trayvon Martin. Therefore, it was not

sense of panic because anything said in the past can be brought up as problematic in the present and because anything said now might become a point of contention in the future.⁸³

Second, in addition to the fear incited by cancel culture, there are concerns that it is not an effective method of changing racist behavior. For example, Nikita Dragun, one of the influencers exposed during *Dramageddon I*, has repeatedly been called out for racially problematic behavior since 2018—blackfishing in particular.⁸⁴ Yet despite flooded comment sections and trending on Twitter numerous times for her behavior, Dragun’s Instagram is still composed of an array of artificial skin tones. In fact, she has even leaned into the criticism by tweeting “which race is nikita gonna be today ?” in October 2020.⁸⁵

For another example of cancel culture’s perceived ineffectiveness, in 2019, Shane Gillis was fired from *Saturday Night Live (SNL)* over previous racist comments.⁸⁶ Then, in a statement Gillis released soon after his removal from *SNL*, Gillis failed to reckon with why the public sought to hold him accountable, equating racism with pushing boundaries and taking risks.⁸⁷ Like Laura Lee’s, Gillis’s cancellation is largely reminiscent of the effects of the

racism that changed but people’s ability to put into action effective counters for that speech. People’s ability to use social media as a tool of collective action to shed light on racism immensely improved over those six years.

83. Austin M. Hooks, *Cancel Culture: Posthuman Hauntologies in Digital Rhetoric and the Latent Values of Virtual Community Networks 16* (Aug. 2020) (M.A. thesis, University of Tennessee at Chattanooga). Furthermore, the panic connected with cancel culture has been problematically associated with a fear of actual harm, “adding a neologic twist on the origin of the practice by associating it with an unfounded fear of censorship and silencing.” Clark, *supra* note 73 at 89.

84. Blackfishing is when non-black women use fake tanners, dark makeup, and sometimes wigs to appear black or at least racial ambiguous. Emma Nolan, *Who is Nikita Dragun? YouTuber Under Fire For Race Remarks*, NEWSWEEK (Oct. 5, 2020), <https://www.newsweek.com/nikita-dragun-youtuber-racist-tweets-instagram-blackfishing-1536416/>. Even though blackfishing, like the DKNY example earlier in this Note, would not meet Matsuda’s definition for actionable racist speech, social media users’ inability to alter Nikita’s behavior is still a salient example of how cancel culture is viewed as ineffective.

85. Sophia Thompson, *Nikita Dragun called out for insensitive tweet mocking her blackfishing past*, PopBuzz (Oct. 6, 2020), <https://www.popbuzz.com/internet/youtubers/nikita-dragun-blackfishing-tweet-racism/>.

86. Jacob Stolworthy, *SNL star Shane Gillis’ firing criticised as ‘cancel culture’ by Bill Burr and Jim Jeffries*, INDEPENDENT (Sep. 17, 2019), <https://www.independent.co.uk/arts-entertainment/tv/news/shane-gillis-snl-saturday-night-live-fired-bill-burr-jim-jeffries-a9108071.html>.

87. Following his removal from *SNL*, Gillis tweeted a statement saying, “I’m a comedian who pushes boundaries. I sometimes miss. If you go through my 10 years of comedy, most of it is bad, you’re going to find a lot of bad misses. I’m happy to apologize to anyone who’s actually offended by anything I’ve said. My intention is never to hurt anyone but I am trying to be the best comedian I can be and sometimes that requires risk.” *Id.*

pillory—he suffered immediate societal attention and even financial loss. However, when Gillis and other members of the comedy profession do not confront the harm done by the racism they perpetuate, is social condemnation and stifled streams of income a solution or no more than a temporary, individualized Band-Aid?

As a result, the end goal of public accountability is sometimes dubious. If it is to remove someone’s platform in order to mitigate the damage done by their racism, there is nothing stopping them from finding new avenues, such as creating new accounts or switching social media platforms entirely, or simply reoffending as in the case of Nikita Dragun. If it is to cut off income streams as a form of retribution, it is unlikely to elicit changed behavior absent conversation on why the language or behavior was problematic in the first place, as in the case of Shane Gillis.

However, public accountability cannot be so easily dismissed as an effective tool of controlling racist speech online. Rather, if the end goal of public accountability is refocused from changing the offender to minimizing the impact felt by those affected, then it can be a positive tool in controlling racist speech online. In particular, the act of holding someone accountable can allow for those in a historically oppressed group to reclaim power, and this is evident in displays of public accountability within Black, Latinx, and LGBT communities.

For instance, “reading,” the open critique of someone’s appearance, within drag culture is a prime example of productive public accountability. Reading dates back to queens of color in the 1950s but was mainstreamed by the film *Paris is Burning*, a documentary that illuminated the New York City drag scene and house culture in the 80s.⁸⁸ It has also become a reoccurring feature on the popular television show *RuPaul’s Drag Race*.⁸⁹ Reading is usually done in a dramatic, rhythmic way; and yet, at the heart of reading is unconditional community love and awareness.⁹⁰ It is less of an insult and more of an act of displaying each other’s flaws before the community and proclaiming love regardless of those flaws for not just the individual being read but the community as a whole. In formal sessions, as in the case of *RuPaul’s Drag Race*, an announcer begins by “opening the library,” making reading a communal

88. PARIS IS BURNING (Jennie Livingston and Barr Swimar 1991).

89. *RuPaul’s Drag Race* (VH1).

90. Ted Mentele, *The history of drag queen slang*, READER’S DIGEST, [https://www.readersdigest.co.uk/inspire/life/the-history-of-drag-queen-sang](https://www.readersdigest.co.uk/inspire/life/the-history-of-drag-queen-slang) (last visited Oct. 1, 2022).

act, where even the person read is in on the activity.⁹¹ Reading highlights the potential for accountability to strengthen a community overall.

Public accountability was also a primary driver for boycotts in the Civil Rights Era.⁹² In December of 1955, the Montgomery Improvement Association with Martin Luther King, Jr. at its helm led the Montgomery bus boycott, which led to *Bowder v. Gayle*, which in turn ended segregation on buses.⁹³ In King's memoir *Strides Toward Freedom*, King highlights the importance of the "growing self-respect" within the Black community that came with the boycotts.⁹⁴ Over 60 years later, communities still rely on boycotts as a means of public accountability as a reflection of self-respect—as was the case for Goya in 2020. When Robert Unanue, Goya's CEO, praised former president Donald Trump, Goya's largely Latino consumer base withdrew support given "everything...going on with [the Trump] administration and the border, the family separations and DACA."⁹⁵

Effectively, the tools that historically oppressed groups have been using to assert their own voice have remained the same—the read, the boycott, the cancel. Public accountability is driven by the collective experience of targeted historically oppressed persons as they discuss, morally evaluate, and prescribe remedies for harm done unto them.⁹⁶ Furthermore, as put by Professor Meredith Clark, "Not every critique can come wrapped up in niceties and polite speech. Nor should it. Sometimes, the urgency and weight of oppression requires us to immediately cry out."⁹⁷

Nonetheless, the potential for cancel culture to run awry has led some to support an alternative form of public accountability: the *call-in*. The call-in is "a call-out done with love."⁹⁸ It recognizes that some conflicts can be resolved privately, rather than in the public pillory or in the Twitter trending page.

91. *Id.*; Gary Hartley, *A Beginner's Guide to Drag Terminology*, CAPE TOWN MAGAZINE, https://www.capetownmagazine.com/arts-culture/a-beginners-guide-to-drag-terminology/104_22_19320#:~:text=A%20subtler%20form%20of%20reading,%2C%E2%80%9D%20explains%20Manila%20von%20Teez (last visited Oct. 1, 2022).

92. *Id.*

93. *Montgomery Bus Boycott*, STANFORD: THE MARTIN LUTHER KING, JR. RESEARCH AND EDUCATION INSTITUTE, <https://kinginstitute.stanford.edu/encyclopedia/montgomery-bus-boycott> (last visited Mar. 25, 2021).

94. *Id.*

95. Raul A. Reyes, *Latinos boycotting Goya say it's not about politics. It's about standing against Trump's 'bete'*, ABC NEWS (Jul. 18, 2020), <https://www.nbcnews.com/news/latino/latinos-boycotting-goya-say-it-s-not-about-politics-it-n1234052>.

96. Clark, *supra* note 73 at 89.

97. *Id.*

98. Loretta Ross, *I'm a Black Feminist. I Think Call-Out Culture Is Toxic.*, N.Y. TIMES (Aug. 17, 2021), <https://www.nytimes.com/2019/08/17/opinion/sunday/cancel-culture-call-out.html>.

Furthermore, it supports the belief that open conversations and debate allow people the opportunity for growth and change.⁹⁹ This argument is familiar—it is the same one used to justify robust free speech rather than control of racist speech.¹⁰⁰ As a result, the counter remains the same: the harm caused by racist speech is such that it dissuades historically oppressed persons from engaging in culture production. The harm done is felt physically, as in the case of Kenosha, but it is also felt psychologically as evident through studies on minority stress.¹⁰¹ Therefore, while open conversation, especially between private individuals, is one effective method of accountability for racist speech, it is not the only one, nor should it be. Public accountability not only tells the individual wrongdoer that their racist speech is unwanted but it tells others who might favor similar speech that that speech is unwelcome within the cultural ecosystem.

2. *In Practice: Facebook's Current Community Speech Control Regime*

Facebook users are the producer-consumers of culture on the platform and are therefore best suited to be the first line of defense when it comes to evaluating racist speech that warrants control. Utilizing its own definition of racist speech outlined in Section III.C, Facebook uses multiple approaches to monitor speech: user reporting, artificial intelligence, and third-party fact-checking. This Section discusses the first approach (user reporting), leaving the latter two for evaluation in Section IV.B.2. User reporting is the bread and butter of community control of racist speech; however, the crux of Facebook's current regime falls on the platform's lack of trust in its own community.

When a user encounters offensive speech on Facebook, there are three ways that the content can be removed. First, a user can directly delete a racist comment on their post. However, when the user deletes the comment, they are not presented with an opportunity to also report it to Facebook. In addition, the user can only report the offender via the offender's profile for "harassment and bullying" or "posting inappropriate things."¹⁰² At no point is the user allowed to flag the offender for racism or expand on their justification for reporting besides these broad categories. After deleting the comment or reporting the offending user via their profile page, the user can either block, unfollow, or unfriend the offending user.

99. *Id.*

100. *Id.*

101. *See* note 52 and accompanying text.

102. Reporting Form, FACEBOOK, <https://www.facebook.com/> (The reporting form can be located following the steps described within the text of this Note.) (last visited Oct. 1, 2022).

The second way racist content can be removed is if a third-party user sees a racist comment on another's post and reports it for review by Facebook. On the report screen, the user is presented with various problem options, including hate speech. The user is only allowed to select one problem, so if the content in question is both harassment and hate speech (as is common), the reporting user must decide which problem to file the content under. If the user selects hate speech, the user is able to clarify that the speech in question is related to race. Then, the reporting user is presented with an abridged version of Facebook's Community Standards, stating:

We only remove content that directly attacks people based on certain protected characteristics. Direct attacks include things like:

- Violent or dehumanizing speech

For example, comparing all people of a certain race to insects or animals

- Statements of inferiority, disgust or contempt

For example, suggesting that all people of a certain gender are disgusting

- Calls for exclusion or segregation

For example, saying that people of a certain religion shouldn't be allowed to vote¹⁰³

Notably, while Facebook includes harmful stereotypes in its Community Standards as a direct attack, it is left out of the above list presented to users upon reporting. After reviewing the definition provided of a direct attack, the user can either select "I don't know. I'd like to see other steps I can take." or "Yes, I'd like to continue filing this report."¹⁰⁴ The first option will lead the user to other steps they can take, including blocking the offending user or hiding all their posts. The second option will file the report with Facebook.

The third way racist content can be removed on Facebook is through group administrators. Within Facebook groups, users can report racist speech directly to the administrators rather than going through the corporation itself. Once the content is reported, an admin can remove it and the offender from the group. A user still has the opportunity to report the content via Facebook's reporting system outlined above. However, reporting to a group administrator will likely receive a quicker response than reporting to Facebook itself. In addition, a reporting user is likely more interested in the speedy removal of the

103. *Id.*

104. *Id.*

particularly offensive content and the offender than in improving the overall environment on the platform. Unless a user or administrator reports directly to Facebook, however, the offender is still capable of joining other groups and committing similar offenses, all while staying below Facebook's radar.

In order for the community to have a stronger systematic effect, Facebook must (1) improve the toolkit it provides users and (2) modify its current definition of racist speech to align with critical race theory, as discussed in Section III.B. Doing so would allow Facebook to have more trust in the decisions made by the community, which in turn would result in a more race-conscious environment on the platform overall.

Facebook should institute better tracking of users that are removed from groups over Community Standards violations like racist speech. This way, an offending user cannot repeatedly jump groups and commit similar violations without Facebook being notified. Then, to provide heightened transparency, the offending user should be notified of how many particular strikes they have against them; as of now, Facebook does not preemptively notify users when their content has been removed, let alone why.¹⁰⁵

However, Facebook administrators cannot be the only vehicle for community control given the repeated issues Facebook has had with groups themselves that violate community guidelines. When 455 users reported the Kenosha Guard to Facebook, it slipped through Facebook's control mechanisms, despite the group blatantly violating Community Standards.¹⁰⁶ Moreover, the Kenosha Guard incident also came a year after Facebook claimed it was clamping down on groups by removing the option to create secret groups and limiting groups to be either public or private.¹⁰⁷ Therefore, more work needs to be done to properly monitor groups. For example, Facebook should institute a screening page when someone seeks to create a group, reminding them of the Community Standards and prohibition on certain kinds of groups like militias.

Facebook should also modify its current definition of racist speech to ensure that only historically oppressed persons are explicitly protected as

105. This is also a policy recommendation that the Board includes in its decision on case 2020-003-FB-UA. The case and the Board's policy recommendation is discussed in depth in the following Section.

106. Mac & Silverman, *supra* note 8; *Dangerous Individuals and Organizations*, FACEBOOK: COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/dangerous_individuals_organizations (last visited Oct. 1, 2022).

107. Brandy Zadronzy, *Facebook has doubled down on groups—now it's looking to clean them up*, NBC NEWS (Aug. 14, 2019), <https://www.nbcnews.com/tech/social-media/facebook-has-doubled-down-groups-now-it-s-looking-clean-n1042336>. This article was published August of 2019 with the Kenosha shooting occurring August of 2020.

discussed in Part III. Without Matsuda's element that actionable racist speech be "directed against a historically oppressed group," Facebook's current definition stands open to abuse, protecting both historically oppressed groups *and* historically oppressive groups.

Furthermore, while the current reporting system is robust, it would benefit from providing its users more information. Facebook should add "harmful stereotypes" to its list of justifications for removal in order to more accurately reflect how its Community Standards define a direct attack.¹⁰⁸ It would also be beneficial to have a link to the full Community Standards upon reporting. Facebook utilizes a three-tier system to further define attacks,¹⁰⁹ and many users might not be familiar Facebook's robust definition unless they are presented with it upon reporting.

Another method of increasing community self-regulation would be by increasing the scenarios in which a user can report racist content. For example, when an administrator goes to remove someone from a group, they should be able to simultaneously report that user to Facebook, or when a user deletes a comment from their post, they should also be able to simultaneously report that comment to Facebook.

By implementing methods of increased transparency and by providing more access to information, Facebook can give its users the tools needed to better hold themselves accountable. Doing so would be the most democratic means of controlling speech online. However, it is inevitable that some incidents will slip through the cracks. Having a secondary level of review can help alleviate the fear that some have that they are being silenced for simply being in opposition. In addition, this secondary level of review can also work to reaffirm the community's belief that what is deemed racist is in fact so. This can help reinforce the benefit of public accountability: not only is the reported user made aware that their speech is unwanted, but so are others who might be of similar mind.

B. REINFORCEMENT: THE ROLE OF CORPORATE OVERSIGHT

Platforms are best suited to be the second level of review for content that slips through the cracks of public accountability because of the flexibility they have under the First Amendment and Section 230; and platforms alone—without the assistance of the community—cannot sift through the vast amount of content uploaded online every day. Therefore, only when platforms work with the community, implementing community decisions systematically and

108. *Hate Speech*, FACEBOOK: COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/hate_speech (last visited Oct. 1, 2022).

109. *Id.*

ensuring that content is moderated in alignment with critical race theory, can race-conscious cultural democracy online prosper.

1. *In Theory: Corporate Oversight of Speech Control*

First Amendment jurisprudence does not apply to social media platforms because platforms are private actors. Unless a private actor is found to be exercising “a function traditionally exclusively reserved to the State,” it is not obligated to observe another’s freedom of speech.¹¹⁰ For instance, in *Manhattan Cmty. Access Corp. v. Halleck*, the Court held that simply because a public access channel on a cable system provides a space for the speech of the others, it does not mean it is a state actor under the state-action doctrine.¹¹¹

The government could reach users’ content on social media platforms by extending the state-action doctrine. In doing so, social media platforms would be regarded as stepping into the shoes of the government, thereby becoming incapable of limiting the speech that occurs on their sites. To some, this would be an easy extension, especially after the strong language Justice Kennedy used in his opinion in *Packingham v. North Carolina*.¹¹² Justice Kennedy stated:

A fundamental First Amendment principle is that all persons have access to places where they can speak and listen, and then, after reflection, speak and listen once more. Today, one of the most important places to exchange views is cyberspace, particularly social media, which offers “relatively unlimited, low-cost capacity for communication of all kinds,”...to users engaged in a wide array of protected First Amendment activity on any number of diverse topics. The Internet’s forces and directions are so new, so protean, and so far reaching that courts must be conscious that what they say today may be obsolete tomorrow.¹¹³

However, Justice Alito in his concurrence chalks up Kennedy’s above language to “loose rhetoric,”¹¹⁴ and this language has yet to sway courts in how they approach speech on the Internet. Furthermore, in *Nyabwa v. Facebook*, a federal

110. *Manhattan Cmty. Access Corp. v. Halleck*, 139 S. Ct. 1921, 1926 (2019) (internal quotation omitted). *But see* *Marsh v. Alabama*, 326 U.S. 501, 507 (1946) (holding that “[private] ownership [of a town] does not always mean absolute dominion” and that “[s]ince [the] facilities are built and operated primarily to benefit the public and since their operation is essentially a public function, it is subject to state regulation.”).

111. 139 S. Ct. at 1926.

112. 137 S. Ct. 1730, 1738 (2017) (holding that North Carolina violated a sex offender’s right to free speech by preventing them from accessing sites like Facebook and Twitter that permit minors to become members).

113. *Id.* at 1732 (internal citations omitted).

114. *Id.* at 1743 (Alito, J. concurring).

district court in Texas directly held that Facebook was not covered by the state-action doctrine.¹¹⁵

In addition to the lack of a state-action doctrine extension, Section 230 of the Communications Decency Act shields social media platforms from liability based on third-party speech. Section 230 states “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”¹¹⁶ As a result, platforms like Facebook cannot be held liable for the content shared on their platform. Without Section 230, platforms would be forced to either not host content at all or to employ heavy censorship in order to shield themselves from liability.¹¹⁷ Therefore, with Section 230 in place, platforms are able to navigate their own approaches to speech without the fear of liability that would result if First Amendment obligations were extended to online platforms. This breadth of freedom is what would allow a platform like Facebook to take innovative steps to rethink controlling racist speech.

2. *In Practice: Facebook’s Approach to Corporate Oversight of Speech Control*

Given the flexibility Facebook has to moderate content, the platform is best fit to adapt to and control racist speech in a way that promotes a race-conscious cultural democracy. Facebook’s Oversight Board has the potential to be an important vehicle for promoting this race-conscious cultural democracy on the platform, so long as Facebook gives the Board the leeway to make meaningful change. Of the first six cases the Board reviewed upon creation, two dealt with Facebook’s Community Standard section on hate speech.¹¹⁸ The first case, Case 2020-002-FB-UA, reached the right decision, but it did so for the wrong reasons and has the potential to hinder the flourishing of a race-conscious cultural democracy online. However, Case 2020-003-FB-UA presents a workable rationale that best supports a race-

115. Nyabwa v. Facebook, No. 2:17-CV-24, 2018 U.S. Dist. LEXIS 13981 *2 (S.D. TX Jan. 26, 2018).

116. 47 U.S.C. § 230.

117. CDA 230: THE MOST IMPORTANT LAW PROTECTING INTERNET SPEECH, EFF, <https://www.eff.org/issues/cda230> (last visited Mar. 3, 2021).

118. Thus far, the Board has not reviewed any cases on racist speech in particular. The two hate speech cases involved hate speech based on religion and ethnicity. However, the race-conscious culture democracy framework can still be applied to ask whether the right kind of speech is being controlled using Matsuda’s three factors of actionable racist speech: a message (1) “of racial inferiority” (2) “directed against a historically oppressed group” (3) that is “persecutorial, hateful, and degrading.” This definition can be utilized to cut through the ambiguity of racist speech. Matsuda, *supra* note 14 at 2357. Matsuda herself uses this definition in other contexts such as Anti-Semitism. *See id.* at 2363. The definition is merely changed by inserting “religion” or “ethnicity” in the first element.

conscious cultural democracy online. As a result, so long as the Board continues to approach questions of hate speech in alignment with Case 2020-003-FB-UA, then the Board is a beneficial method of racist speech control in the digital age; it sets an example for other platforms in how to approach speech.

In Case 2020-002-FB-UA, the Board overturned Facebook's decision to remove a Burmese user's allegedly Islamophobic post.¹¹⁹ In doing so, the Board reached the right conclusion under the race-conscious cultural democracy framework, but it utilized a definition of hate speech different than that advocated in this Note and even different from Facebook's own definition in its Community Standards.

For context, the offensive content included two infamous photos from 2015 of a Syrian toddler who drowned in an effort to reach Europe. Attached to the photo was rhetoric on the psychology of Muslims. The Burmese poster criticized the Muslim response to the treatment of Uyghur Muslims in China in comparison to the attacks that followed the cartoon depictions of the Prophet Muhammad in France. As a result, a driving force of Facebook's decision to originally remove the content was the history of intense Islamophobia in Myanmar. Facebook cited to its Community Standards as still prohibiting hate speech even if it does not advocate or incite violence given "the capacity [of hate speech] to trigger acts of discrimination, violence, or hatred, particularly if distributed widely, virally, or in context with severe human rights risks."¹²⁰

To reach its decision, the Board utilized three "bodies of law": Facebook's Community Standards, Facebook's espoused values, and international law. Under its Community Standards, Facebook acknowledges that hate speech can, *but does not always*, result in real world violence and, as discussed in Part III, defines hate speech as an attack based on protected characteristics.¹²¹ An attack includes harmful stereotypes and "generalizations that state inferiority (in written or visual form) [such as] mental deficiencies[, which] are defined as

119. OVERSIGHT BD., CASE DECISION 2020-002-FB-UA (Jan. 28, 2021), https://web.archive.org/web/20210128145029if_/https://oversightboard.com/decision/FB-I2T6526K/.

120. *Introduction*, FACEBOOK: COMMUNITY STANDARDS, <https://www.facebook.com/communitystandards/introduction> (last visited May 4, 2021). In particular, the Introduction states: "We are committed to making Facebook a safe place. Expression that threatens people has the potential to intimidate, exclude or silence others and isn't allowed on Facebook."

121. *Hate Speech*, FACEBOOK: COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/hate_speech (last visited May 4, 2021). As noted previously, while Facebook does a lot of work to define and explain its approach to hate speech, it would benefit from requiring that hate speech be directed at historically oppressed persons to ensure that the definition is not abused and used against the people it is intended to protect.

those about: intellectual capacity...education...mental health.”¹²² However, the Introduction to Facebook’s Community Standards highlights that “Voice” is an important value of the platform, which must be balanced against Facebook’s value of “Safety.”¹²³ Furthermore, the Board also utilized the UN Guiding Principles on Business and Human Rights, a human rights framework intended for private businesses, and used the international human rights standards for the right to freedom of expression, the right to non-discrimination, and the right to life and security to support its decision.¹²⁴

In reaching its decision, the Board rightfully considered the content in its full context and used the full breadth of its tools as outlined in its bylaws. Using context is imperative when approaching hate speech given the nuances of different cultures. This is especially true for Facebook, which is an American corporation with users all around the world. Such a context-driven approach promotes a race-conscious cultural democracy. In this case, the Board consulted with translators and cultural experts and even took public comments into consideration when reaching its decision.¹²⁵ What the case notably highlighted was the importance of translation—while Facebook translated the text as “[i]t’s indeed something’s wrong with Muslims psychologically,” the Board’s translators suggested that the text actually read “[t]hose male Muslims have something wrong in their mindset.”¹²⁶ Though a slight change in language, the Board believed that the latter translation was not derogatory or violent.

In addition, when consulting cultural experts on Myanmar, they found comments on the mental state of Muslims was not a part of typical Islamophobia in the country. Furthermore, the post was within a group that was intended for intellectual and philosophical discussion and meant to be a “commentary pointing to the apparent inconsistency between Muslims’ reactions to events in France and in China.”¹²⁷ Therefore, given this context,

122. *Id.*

123. *Supra* note 120.

124. *See* International Covenant on Civil and Political Rights art. 19, Dec. 19, 1966, 999 U.N.T.S. 171, 175 (entered into force Mar. 23, 1976) [hereinafter ICCPR]; ICCPR Human Rights Comm., General Comment No. 34 on Article 19, U.N. Doc. CCPR/C/GC/34 (2011); United Nations High Commissioner for Human Rights, *Annual Report of the United Nations High Commissioner for Human Rights*, U.N. Doc A/HRC/22/17/Add.4 (Oct. 05, 2012); David Kaye (Special Rapporteur), Promotion and Protection of the Right to Freedom of Opinion and Expression, U.N. Doc. A/74/486 (2019); ICCPR art. 6 and art. 9.

125. OVERSIGHT BD., CASE DECISION 2020-002-FB-UA (Jan. 28, 2021), https://web.archive.org/web/20210128145029if_/https://oversightboard.com/decision/FB-I2T6526K/.

126. *Id.*

127. *Id.*

the Board decided that the post was an expression of opinion and not hate speech.

When examining the post under Matsuda's conception of racist speech as a message (1) "of racial inferiority" (2) "directed against a historically oppressed group" (3) that is "persecutorial, hateful, and degrading,"¹²⁸ the Board's decision was rightfully made. While the comment was directed against a historically oppressed group (Muslims), it was neither a message of inferiority nor was it persecutorial, hateful, or degrading. Instead, when considered in full context, the post intended to draw attention to how Muslims reacted to different situations: one in which there was a cartoon and one in which there was a threat to Muslim lives. Therefore, by allowing this speech on Facebook, race-conscious cultural democracy is supported as others can engage with the poster's ideas and either support or reject them, thereby continuing the cycle of cultural production.

However, the Board's logic diverges from Matsuda's definition when it weighed human rights standards on freedom of expression, citing Article 20, para. 2 of the International Covenant on Civil and Political Rights (ICCPR).¹²⁹ In particular, the Board decided that while the content was offensive, it did not rise to the level of "advocat[ing] hatred or intentionally incit[ing] any form of imminent harm."¹³⁰ As a result, the Board established a high bar for actionable racist speech that is actually above the one established in Facebook's own Community Standards and also above the one advocated in this Note.

The language "intentionally" is not located in Facebook's Community Standards and does not work to support a race-conscious cultural democracy online. The Community Standards explicitly state that "[i]f intention is unclear, we may remove content."¹³¹ Therefore, intent is not a requirement for removal. In addition, as explored in Part III, intent is not a defining factor of racist speech under Mari Matsuda's definition, which only establishes a bar of

128. Matsuda, *supra* note 14, at 2357.

129. ICCPR, *supra* note 124, art. 20(2).

130. *Oversight Board overturns Facebook decision: Case 2020-002-FB-UA*, OVERSIGHT BD. (Jan. 2021), <https://oversightboard.com/news/773985406543781-oversight-board-overturns-facebook-decision-case-2020-002-fb-ua/>.

131. *Hate Speech*, FACEBOOK: COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/hate_speech (last visited May 4, 2021). This section on intent is in reference to content that is perceived as racist when in fact the poster intends for the racist content to be used for educational purposes—a different situation entirely from the one in this case. In other words, the Community Standards require intent to be made clear when the intention is to *not* be racist, while the intent standard used by the Oversight Board is an intent to be racist.

“persecutorial, hateful, and degrading” speech.¹³² Often, intent is not present at all in racist speech but that does not lessen its harm.¹³³

Furthermore, an imminent harm requirement is also not located in Facebook’s definition of hate speech, which hinges on the speech being a direct attack. An attack is “violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing, and calls for exclusion or segregation.”¹³⁴ This is a far different standard than that used to evaluate the content in Case 2020-002-FB-UA, and it even directly conflicts with the rationale of Case 2020-003-FB-UA, which recognizes that Facebook is allowed to “prohibit some discriminatory expressions...absent any requirement that the expression incite violent or discriminatory acts.”¹³⁵

While citing to international law lends legitimacy to the Board’s decision making, it also sets a dangerous precedent for future cases given that hate speech is not always driven by intent and does not always result in imminent harm. Nonetheless, harm is still associated with hate speech, and this harm hinders the cultural participation of those affected, thereby suppressing a robust cultural democracy online.

In Case 2020-003-FB-UA, the Board upheld Facebook’s removal of a post that included a Russian slur regarding Azerbaijanis.¹³⁶ Similar to Case 2020-002-FB-UA, the Board based its decision on Facebook’s Community Standard, Facebook’s values, and relevant international law. However, unlike the previous case, the Board acknowledged Facebook is not under the same obligations as a government when it comes to controlling racist speech. Therefore, it focused its analysis on whether the limitation on freedom of expression was justified under ICCPR Article 19, para. 3, which includes a three-part test: legality, legitimacy, and necessity and proportionality.¹³⁷ This standard, as opposed by the one in the previous case, allows the Board more flexibility to promote a race-conscious cultural democracy online.

132. Matsuda, *supra* note 14, at 2357.

133. *See supra* note 43–46 and accompanying text.

134. *Hate Speech*, FACEBOOK: COMMUNITY STANDARDS, https://www.facebook.com/communitystandards/hate_speech (last visited May 4, 2021).

135. OVERSIGHT BD., CASE DECISION 2020-003-FB-UA (2021), <https://oversightboard.com/decision/FB-QBJDASCV/>.

136. *Id.*

137. *Id.* (“Facebook’s Hate Speech Community Standard prohibits some discriminatory expression, including slurs, absent any requirement that the expression incite violent or discriminatory acts. While such prohibitions would raise concerns if imposed by a Government at a broader level..., the Special Rapporteur indicates that entities engaged in content moderation like Facebook can regulate such speech....”)

The content under question was the term “taziks” to refer to Azerbaijanis. In Russian, the term refers to “wash bowls,” but it is also a play on the word “aziks,” a derogatory term for Azerbaijanis that was on Facebook’s list of prohibited slurs. As a result, Facebook removed the post as a violation of its Community Standards because it was meant to dehumanize Azerbaijanis.

When examining the post in question, the Board agreed that the slur was a violation of Facebook’s Community Standards on hate speech. It also believed that the post violated Facebook’s values of “Dignity” and “Safety” to such an extent as to outweigh Facebook’s countervailing commitment to “Voice.”¹³⁸

The main question that the Board confronted was whether Facebook’s restriction met three requirements: legality, legitimacy, and necessity and proportionality.¹³⁹ To answer the question of legality, the Board asked if the restriction on speech was clear and accessible to user. Because Facebook has an extensive Community Standard section on hate speech that explicitly lists slurs as prohibited, this requirement was met. In order to meet the legitimacy requirement, the Board asked if the restriction on slurs served a legitimate aim. In this case, the restriction on slurs protects “people’s right to equality and non-discrimination” under Article 2, para. 1 of the ICCPR.¹⁴⁰ Finally, in order to meet the necessity and proportionality requirement, the restriction must be in response to a threat to the rights of others and proportional. The Board acknowledged that while international law does not prohibit “insults, ridicule or slander,” Facebook, being a company and not a government, is allowed to make such restrictions.¹⁴¹

In its decision, the Board highlighted the importance of context. The Board consulted cultural experts in the region as to the meaning and use of the slur, determining that it was hateful and dehumanizing. Furthermore, the post was shared in an area that is in the middle of an armed conflict. Therefore, even if offline harm resulting from online speech might not have been imminent or known, the Board concluded that the hate speech risked such a result. The Board also held that Facebook’s response was proportional; mere warning labels would not have sufficed, and Facebook did not look to take

138. *Id.*; *Introduction*, FACEBOOK: COMMUNITY STANDARDS, <https://www.facebook.com/communitystandards/introduction> (last visited May 4, 2021).

139. *Supra* note 135.

140. ICCPR, *supra* note 121, art. 2(1).

141. OVERSIGHT BD., CASE DECISION 2020-003-FB-UA (2021), <https://oversightboard.com/decision/FB-QBJDASCV/>.

more severe measures such as removing the offender from the platform entirely.¹⁴²

Case 2020-03-FB-UA was also the only case of the first six to come with a policy recommendation. The Board suggested that the platform increase its transparency by providing users with reasons for why their content was restricted. In particular, the Board highlighted that “Facebook’s lack of transparency left its decision susceptible to the mistaken belief that it had removed the post because the user was addressing a controversial subject or expressing a viewpoint Facebook disagreed with.”¹⁴³ Additional transparency by Facebook can only help the platform as it would allow users to instill more trust in the platform, and in turn, allow the platform to instill more trust in its users, an important aspect of race-conscious cultural democracy.

This decision fits squarely in line with race-conscious cultural democracy. Under Matsuda’s definition, the slur in this case was a message of inferiority directed against Azerbaijanis that was persecutorial, hateful, and degrading, making it an easy case of actionable hate speech. The Board relied on similar reasoning under both the legitimacy analysis and the necessity and proportionality analysis. The Board also emphasized the harmful effect that hate speech can have regardless of whether it is experienced offline in the form of violence—a key tenant of why hate speech must be controlled within a race-conscious cultural democracy. It also recognized that removal of hate speech is sometimes the best course of action, leaving no room for the argument that open debate is the best solution to hate speech. Importantly, the Board acknowledged that the offender was still able to engage in cultural production as they were “free to engage in discussions on the same issues within the boundaries of the Community Standards.”¹⁴⁴ Therefore, not only is the offender not exiled from culture in the way that cancel culture rhetoric threatens, but those harmed by the slur are also able to engage in a space free from hate, thereby promoting race-conscious cultural democracy.

142. *Id.* Notably, however, the Board acknowledged that the offender in question had previously posted similar hate speech on the platform, raising the question of how many times someone must get reported for individual incidents before they are removed from the platform entirely. The potential for reoffending is a setback for community accountability as discussed in Part IV.D and is also a driving criticism that Facebook received following the 450 reports placed against the Kenosha Guard. Mac & Silverman, *supra* note 8. The Board’s decision does not specify how many times this user was reported, nor does it suggest how many infractions would be necessary rise to the level of platform removal. This is a question that will inevitably need to be resolved by the Board in future cases.

143. *Id.*

144. Oversight Bd., Case Decision 2020-003-FB-UA (2021), <https://oversightboard.com/decision/FB-QBJDASC/> .

<https://>

As a result, while Case 2020-002-FB-UA stands on shaky grounds in regards to both race-conscious cultural democracy theory and Facebook's own Community Standards, Case 2020-003-FB-UA establishes a rationale that is supported by both. As more cases on similar issues come before the Board, it will have to look at its case precedent and decide which approach to follow when analyzing hate speech. Only the latter approach, however, is tenable under race-conscious cultural democracy theory.

V. CONCLUSION

Finding a balance between too much and not enough control over speech has always been difficult. Racist speech works on a spectrum with direct call for violence on one end and covert, sanitized speech on the other. In between is a zone of ambiguity that is colored with context that is often difficult to disentangle from a perspective outside the community affected by the racist speech. Critical race theory can help explain where to draw the line between actionable and non-actionable racist speech. Further, cultural democracy theory can help elucidate how to put that line into practice by emphasizing cultural participation from all members of society. In particular, speech that is a message (1) “of racial inferiority” (2) “directed against a historically oppressed group” (3) that is “persecutorial, hateful, and degrading” is actionable racist speech that should be controlled.¹⁴⁵

However, current cultural democracy theory neglects the harm that racist speech can have on the cultural participation of historically oppressed persons. Therefore, a race-conscious cultural democracy recognizes the need to control some speech to promote the cultural participation of all persons. To that end, the first line of defense against racist speech should be the community—the most democratic form of control. But with community accountability, there is often a fear of cancel culture and a worry that it does not do enough to remedy racism. Therefore, social media platforms like Facebook must step up to ensure that not only that their users have the proper tools to regulate themselves, but also that the correct speech is being controlled.

There is still much work to be done before a race-conscious cultural democracy can be fully realized, and it will inevitably take trial and error. Communities themselves are still learning the best way to use their collective power, as shown by the constant turmoil in YouTube's beauty guru community. Platforms as well are still learning how to craft ToS and Community Standards that capture the right kind of unwanted speech.

145. Matsuda, *supra* note 14, at 2357.

In the case of Facebook, their Community Standards should be amended to state, “we define hate speech as a direct attack on people *of a historically oppressed group* based on what we call protected characteristics.” This will more explicitly distinguish the speech of the oppressor from the oppressed and therefore prevent the hijacking of the section by those who seek to weaponize it against the people it is meant to protect.

In addition, Facebook must trust its users and remove reported content—especially when numerous users report the same content. This is an issue that went unresolved following Kenosha and that has been acknowledged but unremedied by the Oversight Board.¹⁴⁶ Facebook can increase its trust in its users by enhancing the users’ toolkit with more opportunities to report, providing them with more information on Facebook’s definition of racist speech, and increasing the transparency when content is removed.

Facebook’s Oversight Board is a promising mode of corporate oversight that works to buttress community control and promote race-conscious cultural democracy in the digital age. While the decision in Case 2020-002-FB-UA has potentially detrimental reasoning, it reaches the right conclusion and is balanced by the rationales underpinning Case 2020-003-FB-UA. Therefore, going forward, if the Board seeks to create a space on the platform that allows all persons to freely and widely engage in cultural production, it should follow the latter case’s reasoning. However, as of yet, it is unclear what role precedent cases will hold in the Board’s future decision making as it has only recently selected its next batch.¹⁴⁷

Regardless, the installation of the Oversight Board should spur other platforms to consider how they too can better manage speech on their sites. They should note the benefits that such a Board can have when it comes to promoting the ever-growing cultural democracy online.

146. See Mac & Silverman, *supra* note 8; OVERSIGHT BD., CASE DECISION 2020-002-FB-FBR (2021), https://web.archive.org/web/20210128145029if_/https://oversightboard.com/decision/FB-I2T6526K/.

147. *Announcing the Oversight Board’s next cases*, OVERSIGHT BD. (Jan. 2021), <https://oversightboard.com/news/175638774325447-announcing-the-oversight-board-s-next-cases/>.

