
FACEBOOK'S FACES

Chinmayi Arun*

The case of the suspension of former President Donald Trump's Facebook and Instagram accounts clarified the relationship between Facebook and its Oversight Board. To understand how, we have to appreciate the complexity of Facebook's relationships with states, publics, and its own staff.

Scholars have offered brilliant, nuanced accounts of social media platforms' relationships with states and users. This Essay builds on their work and expands their theorization to account for differences among states, the varying influence of different publics, and the complexity and tensions within companies. Theorizing Facebook's relationships this way includes less influential states and publics that are otherwise obscured, and renders visible the agency and influence of Facebook's staff.

Facebook engages with states and publics through multiple parallel regulatory conversations, further complicated by the fact that Facebook itself is not a monolith. This Essay argues that Facebook has many faces — different teams working towards different goals, and engaging with different ministries, institutions, scholars, and civil society organizations. It is also internally complicated, with staff whose sympathies and powers vary and can be at odds with each other. Content moderation takes place within this ecosystem.

This Essay's account of Facebook's faces and relationships shows that less influential actors can sway the company through strategic alliances with stronger actors. It discusses where the Oversight Board is placed in this ecosystem, and how far it can oversee Facebook. It suggests that Facebook's carelessness with those it perceives to be weak or inconsequential can affect the company when powerful allies care or powerful alliances are formed.

I. INTRODUCTION

On January 6, 2021, a mob attacked and breached the United States Capitol, causing deaths and injuries, and threatening the constitutional order.¹ Donald Trump, then President, shared two posts on Instagram

* Fellow of the Information Society Project at Yale Law School. I am indebted to Jack Balkin for his generous feedback on these arguments and grateful to Adam Posluns, Agustina Del Campo, Alvin Padilla, Amos Toh, Ana Vanzoff Robalinho, Anne Cheung, Artur Pericles Lima Monteiro, Brenda Dvoskin, Daniel Markovits, Daphne Keller, David Kaye, Faiza Patel, Gilad Abiri, Jenny Domino, Jisu Kim, Jonathan Ong, Jonathan Zittrain, Kate Klonick, Michael Karanicolas, Nikolas Guggenberger, Nishant Shah, Patricia Cruz Marín, Rafael Bezerra Nunes, Ryan Goodman, Salomé Viljoen, Sandra Magalang, Susan Benesch, Tarleton Gillespie, Timothy Garton Ash, Tom Tyler, Zoe Darmé, the Berkman Klein Center, the Yale Information Society Project, the Global Network Initiative, Stanford's Global Digital Policy Incubator, ARTICLE 19, the Hoover Institution at Stanford University, the Platform Governance Research Network, the Institute for Rebooting Social Media, and the editors of the *Harvard Law Review*.

¹ *Case Decision 2021-001-FB-FBR*, OVERSIGHT BD. 9–10 (May 5, 2021) [hereinafter *Trump Decision*], <https://www.oversightboard.com/sr/decision/2021/001/pdf-english> [<https://perma.cc/3EJT-CH8T>].

and Facebook² during and immediately after the attack that Facebook removed for violating its Community Standards.³ One was a video calling the 2020 presidential election “stolen” and “fraudulent,”⁴ and the other a statement saying: “These are the things and events that happen when a sacred landslide election victory is so unceremoniously & viciously stripped away from great patriots who have been badly & unfairly treated for so long.”⁵ After removing the second post, Facebook blocked Trump’s accounts from posting for an initial twenty-four hours, and then extended the block indefinitely and for at least two weeks until the transition of power took place.⁶ Facebook referred its decision to indefinitely suspend Trump’s accounts to its Oversight Board, also inviting the Board’s recommendations about the suspension of political leaders’ accounts.⁷

The “Trump Ban” case, as I am calling it, illuminates the complexity of Facebook’s content-moderation system and the role of its Oversight Board in unusual ways. This makes the case a helpful trigger for the perspective I propose in this Essay. The case is also arguably an inflection point in the relationship of the Board with the company. Although I focus on Facebook, I hope that this frame of reference is useful to everyone trying to understand global social media companies with sophisticated content-governance systems.

Facebook is a polarizing company. Those who engage with it often have strong opinions about its willingness to change. To reconcile these points of view, it helps to stop thinking of Facebook as a monolith. This is clear from the company’s varied engagement with external actors and from the different priorities and goals of its internal actors. There are conversations and tensions within the company, among teams with different mandates working towards opposing outcomes and staff members with their own visions of how the company should act. In other words, (1) Facebook engages differently with different actors, and (2) different teams and individuals within the company have different priorities. Facebook has many faces.

² In October 2021, the Facebook company changed its name to Meta. Mark Zuckerberg, *Founder’s Letter, 2021*, FACEBOOK (Oct. 28, 2021), <https://about.fb.com/news/2021/10/founders-letter> [<https://perma.cc/PS2C-4763>]. This Essay continues to refer to the company as Facebook.

³ *Trump Decision*, *supra* note 1, at 10–11.

⁴ *Id.* at 10.

⁵ *Id.* at 10–11.

⁶ *Id.* at 11.

⁷ Nick Clegg, *Referring Former President Trump’s Suspension from Facebook to the Oversight Board*, FACEBOOK (Jan. 21, 2021), <https://about.fb.com/news/2021/01/referring-trump-suspension-to-oversight-board> [<https://perma.cc/8RWB-DH4Q>].

The Essay begins in Part II with an introduction to Facebook's content-governance system⁸ and the Oversight Board. The Board was constructed as an independent entity with a limited sphere of influence and an apparent purpose to legitimize the company's self-regulation,⁹ which presumably included Facebook's decisionmaking about political leaders like Donald Trump. CEO Mark Zuckerberg described the Board as an entity that would make Facebook accountable to its community.¹⁰ Part II also introduces the role the Oversight Board was supposed to play in legitimizing the content-governance system, and how this legitimacy is likely assessed.

Part II provides helpful background to understand Facebook's faces, which I discuss in Part III, in terms of the company's engagement with the outer world and its internal tensions. I explain why this perspective is both closer to reality and more helpful than prevailing understandings. It can be easy to miss the complex relationships between the company and external actors if they are described in ways that flatten them into uniformity. In this Part of the Essay, I discuss the differences in power and influence of states and publics, as well as teams inside Facebook, and why this complexity of engagement between these actors matters.

In Part IV, I discuss the Trump Ban case¹¹ in terms of the complex perspective offered in Part III. Here I draw out what this case shows us about the content-governance system and the relationship between

⁸ See generally ROBYN CAPLAN, DATA & SOC'Y, CONTENT OR CONTEXT MODERATION? (2018), https://datasociety.net/wp-content/uploads/2018/11/DS_Content_or_Context_Moderation.pdf [<https://perma.cc/P4AE-RXAJ>]; TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA (2018); Robert Gorwa, *The Platform Governance Triangle: Conceptualising the Informal Regulation of Online Content*, INTERNET POL'Y REV., June 30, 2019, at 1; Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018); Chinmayi Arun, *The Facebook Oversight Board: An Experiment in Self-Regulation*, JUST SEC. (May 6, 2020), <https://www.justsecurity.org/70021/the-facebook-oversight-board-an-experiment-in-self-regulation> [<https://perma.cc/S7D7-RYE5>]; David Kaye, *The Republic of Facebook*, JUST SEC. (May 6, 2020), <https://www.justsecurity.org/70035/the-republic-of-facebook> [<https://perma.cc/7JZV-44B2>]; Siva Vaidhyanathan, *Facebook and the Folly of Self-Regulation*, WIRED (May 9, 2020, 2:58 PM), <https://www.wired.com/story/facebook-and-the-folly-of-self-regulation> [<https://perma.cc/7KWX-PVLD>].

⁹ See evelyn douek, *Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility*, 21 N.C. J.L. & TECH. 1, 46 (2019); see also Michael Karanicolas, *Squaring the Circle Between Freedom of Expression and Platform Law*, 20 PITT. J. TECH. L. & POL'Y 177, 192–93 (2020).

¹⁰ See Mark Zuckerberg, *A Blueprint for Content Governance and Enforcement*, FACEBOOK (May 5, 2021), <https://www.facebook.com/notes/751449002072082> [<https://perma.cc/2ASC-A9HP>]; Letter from Mark Zuckerberg, CEO, Facebook, Facebook's Commitment to the Oversight Board (Sept. 2019), <https://about.fb.com/wp-content/uploads/2019/09/letter-from-mark-zuckerberg-on-oversight-board-charter.pdf> [<https://perma.cc/YER8-NUJR>].

¹¹ *Trump Decision*, *supra* note 1; see also Clegg, *supra* note 7; *Oversight Board Accepts Case on Former US President Trump's Indefinite Suspension from Facebook and Instagram*, OVERSIGHT BD. (Jan. 2021), <https://www.oversightboard.com/news/236821561313092-oversight-board-accepts-case-on-former-us-president-trump-s-indefinite-suspension-from-facebook-and-instagram> [<https://perma.cc/S23R-YNQT>].

Facebook and its Oversight Board. This discussion also highlights what we might learn from the information that whistleblower Frances Haugen revealed after the Trump Ban case, which resulted in the Board reopening some of the questions surfaced by the case.

Part V of the Essay circles back to the content-governance system and the extent to which the Oversight Board is a body that can hold Facebook accountable. The Board is a valuable institution that adds tremendously to global discourse about online speech, modeling how a shared jurisprudence can be created for a global set of rules for global publics. However, this does not necessarily imply that it can offer oversight without Facebook's cooperation.

It turns out that Facebook's painstaking efforts to prepare faces to meet the faces it meets have created a situation in which it is at odds with itself. This does, however, create space for the Oversight Board to ally with other powerful actors and influence the company.

II. THE CONTENT-GOVERNANCE SYSTEM AND THE OVERSIGHT BOARD

Facebook mostly separates its engagement with states from its engagement with its users in the context of decisions about speech on its platform. It does this by using two different systems to regulate content. States expect the company to follow their laws and are interested in the company's implementation of those laws. The local-law system for determining unlawful content leads to the blocking of content in the state where it is illegal, while the content is available and visible on Facebook elsewhere.¹² Decisions made using this system are not subject to appeals and review,¹³ or the Oversight Board.¹⁴ The other system is of Facebook's global rules concerning users — its "Community Standards"¹⁵ — which are described as privately ordered "platform

¹² See *Content Restrictions Based on Local Law*, FACEBOOK: TRANSPARENCY CTR., <https://transparency.fb.com/data/content-restrictions> [<https://perma.cc/D9FP-VMPF>].

¹³ See *How We Assess Reports of Content Violating Local Law*, FACEBOOK: TRANSPARENCY CTR., <https://transparency.fb.com/data/content-restrictions/content-violating-local-law> [<https://perma.cc/R4K7-HSKE>] (outlining how Facebook reviews and responds to reports of content alleged to violate local law).

¹⁴ OVERSIGHT BD., OVERSIGHT BOARD BYLAWS § 1.2.2 (2022), <https://www.oversightboard.com/sr/governance/bylaws> [<https://perma.cc/ZTU4-Y6A3>] (noting that cases "[w]here the underlying content is unlawful in a jurisdiction with a connection to the content" are "not eligible for the board to review"); see also FACEBOOK, OVERSIGHT BOARD CHARTER art. 7 (2019) [hereinafter OVERSIGHT BOARD CHARTER], https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf [<https://perma.cc/AZH3-WX4L>] ("The board will not purport to enforce local law.").

¹⁵ *Facebook Community Standards*, FACEBOOK: TRANSPARENCY CTR., <https://transparency.fb.com/policies/community-standards> [<https://perma.cc/H7UL-PGCW>].

law”¹⁶ and are subject to an increasingly elaborate apparatus of review, including the Oversight Board.¹⁷ Content violating Facebook’s rules is removed globally.¹⁸ As part of this second system, the Oversight Board was created to bring it more legitimacy.¹⁹

To understand why the Trump Ban case was a defining moment for Facebook’s self-regulatory content-moderation system and for the Oversight Board, it is necessary to understand the contours and goals of this system and the role that the Board plays in this context. This Part of the Essay discusses Facebook’s content moderation and the role of the Oversight Board in legitimizing the company’s system. It also discusses how such legitimacy is assessed, and why this is related to the company’s uneven responsiveness to different actors. In Part III, I offer a perspective that makes the complexities of this system visible.

A. Local Laws and Community Standards

In its enforcement of local laws, Facebook implements the laws of each country in which it operates for users within that country.²⁰ As a result, blasphemous content may be blocked in Pakistan, where it is illegal, while remaining visible on Facebook in other parts of the world. This system is applied when content is blocked in connection with laws such as *lèse-majesté*, sedition, and obscenity that are state-set standards for illegal speech. Facebook uses an opaque system to implement the states’ “local laws,” and decisions made in this system are not subject to appeals and review,²¹ or the Oversight Board.²² However, this is not as simple as it seems, as I explain below.

In contrast to the local laws enforced by Facebook, the Community Standards are Facebook’s rules for the entire platform, and they apply globally.²³ They have been described as privately ordered “platform

¹⁶ See David Kaye (Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression), *Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 1, U.N. Doc. A/HRC/38/35 (Apr. 6, 2018), <https://www.undocs.org/A/HRC/38/35> [<https://perma.cc/M2XF-WC8B>]; Molly K. Land, *The Problem of Platform Law: Pluralistic Legal Ordering on Social Media*, in *THE OXFORD HANDBOOK OF GLOBAL LEGAL PLURALISM* 975, 975 (Paul Schiff Berman ed., 2020); see also Klonick, *supra* note 8, at 1634–35 (describing the rise of platform-specific rules for content moderation).

¹⁷ OVERSIGHT BOARD CHARTER, *supra* note 14, art. 1, § 4.

¹⁸ See *Facebook Community Standards*, *supra* note 15.

¹⁹ See douek, *supra* note 9, at 46.

²⁰ See *Content Restrictions Based on Local Law*, *supra* note 12. This system leads to the blocking of content only in the state where it is illegal; the content is otherwise available and visible on Facebook elsewhere. *Id.*

²¹ See *How We Assess Reports of Content Violating Local Law*, *supra* note 13.

²² See sources cited *supra* note 14.

²³ There is a rich body of scholarship about the Community Standards, including, for example, GILLESPIE, *supra* note 8; DAVID KAYE, *SPEECH POLICE: THE GLOBAL STRUGGLE TO GOVERN THE INTERNET* (2019); and Klonick, *supra* note 8.

law”²⁴ and are subject to an increasingly elaborate system of review, including the Oversight Board.²⁵

Facebook’s system for applying local laws is not subject to the same reviews and opportunities for appeal as the privately ordered content-moderation system that Facebook uses for its Community Standards. Facebook explains that it follows the following process while making a decision about content flagged for violating local laws:

If content does not violate our policies, in line with our commitments as a member of the Global Network Initiative and our Corporate Human Rights Policy, we conduct a careful legal review to confirm whether the report is valid, as well as human rights due diligence. In cases where we believe that reports are not legally valid, are overly broad, or are inconsistent with international human rights standards, we may request clarification or take no action.²⁶

It often falls to social media companies to interpret and implement local laws.²⁷ It is worth noting that Facebook’s interpretive choices may not always align with the state’s or the judiciary’s reading of laws.²⁸ This means that the company’s over- or undercompliance, or willingness to contest the law should be of interest to us. I discuss this in more detail in section III.A of this Essay.

These decisions about whether content violates laws are largely opaque except for the information found in transparency reports, which offer examples of how the company handles requests or orders to remove content for violating local laws.²⁹ There is no independent audit or assessment of this system except to the extent that the Global Network Initiative (GNI), a multi-stakeholder, self-regulatory institution, performs assessments of Facebook’s good faith efforts to implement the GNI Principles.³⁰ The GNI Principles are part of the organization’s

²⁴ See sources cited *supra* note 16.

²⁵ OVERSIGHT BOARD CHARTER, *supra* note 14, art. 1, § 4.

²⁶ *How We Assess Reports of Content Violating Local Law*, *supra* note 13.

²⁷ See Jack M. Balkin, Essay, *Free Speech Is a Triangle*, 118 COLUM. L. REV. 2011, 2015–21 (2018); see also Julia Black, *Regulatory Conversations*, 29 J.L. & SOC’Y 163, 194–95 (2002) (observing that “interpretive control in regulation” represents “control over a central power resource,” *id.* at 194).

²⁸ See Black, *supra* note 27, at 176 (noting that written norms “are open to continual reinterpretation, depending on the actor’s preoccupations and goals, the context of action, and who else is involved in the encounter”).

²⁹ *Content Restrictions Based on Local Law*, *supra* note 12.

³⁰ GNI’s assessment is a confidential process by design, which means that it is difficult for the public to understand the extent and manner of each company’s compliance with the GNI Principles. See GLOB. NETWORK INITIATIVE, GNI ASSESSMENT TOOLKIT 16, 67 (2021), <https://globalnetworkinitiative.org/wp-content/uploads/2021/11/AT2021.pdf> [https://perma.cc/WFR8-gT8K]. In the interest of full disclosure, I am a GNI Board member. Board members are invited to review all assessment reports of GNI member companies.

“evolving framework for responsible company decision making in support of freedom of expression and privacy rights.”³¹ Every two or three years, member companies undergo assessments of their progress in implementing these principles by GNI-accredited assessors,³² who generate confidential reports that are subject to review by the GNI Board.³³

Facebook enforces its own rules for speech in the form of its “values” and “Community Standards” on its platform globally.³⁴ These decisions are subject to its appeals mechanism, and now sometimes subject to the Oversight Board.³⁵ Speech violating these rules, like adult nudity,³⁶ is usually restricted worldwide. These Community Standards were the justification offered for eventually suspending Donald Trump’s accounts.³⁷ The decision was made through the company’s internal process.³⁸

The bifurcated system of rules described above looks simple at first glance — global Community Standards are between Facebook and its users, while local laws are about governments’ relationships with their citizens, which Facebook respects. Facebook follows local laws in the countries in which it operates, as does Twitter. However, as I argue in Part III, Facebook’s complex relationships with states and different publics, and its own competing internal priorities, result in the somewhat puzzling range of outcomes in how it implements laws and its own Community Standards.

As the implementation of the Community Standards evolved over time, the company built a bureaucracy to interpret and enforce the

³¹ *About GNI*, GLOB. NETWORK INITIATIVE, <https://globalnetworkinitiative.org/about-gni> [<https://perma.cc/LA76-YWRR>].

³² GLOB. NETWORK INITIATIVE, *supra* note 30, at 4, 8. GNI companies may select any assessor from GNI’s pool of accredited assessors and enter into an agreement with them, determining and bearing the cost of the assessment. *Id.* at 9. The GNI Board “determine[s] whether a company is making good-faith efforts to implement the GNI Principles,” but it is the role of the independent assessor “to provide the [B]oard with the information it needs to make this determination.” *Id.* at 5.

³³ *Id.* at 20, 67.

³⁴ *Facebook Community Standards*, *supra* note 15.

³⁵ OVERSIGHT BOARD CHARTER, *supra* note 14, art. 1, § 4.

³⁶ *Adult Nudity and Sexual Activity*, FACEBOOK: TRANSPARENCY CTR., <https://transparency.fb.com/en-gb/policies/community-standards/adult-nudity-sexual-activity> [<https://perma.cc/KZ94-ZHAQ>].

³⁷ *See* Guy Rosen & Monika Bickert, *Our Response to the Violence in Washington*, FACEBOOK (Jan. 7, 2021, 8:05 AM), <https://about.fb.com/news/2021/01/responding-to-the-violence-in-washington-dc> [<https://perma.cc/7RMV-6AJ3>] (noting Trump had committed two “policy violations”); *Trump Decision*, *supra* note 1, at 2 (attributing Facebook’s actions to Trump’s violations of the Community Standard on Dangerous Individuals and Organizations).

³⁸ Clegg, *supra* note 7; *see also* Alex Hern & Kari Paul, *Donald Trump Suspended from Facebook Indefinitely, Says Mark Zuckerberg*, THE GUARDIAN (Jan. 7, 2021, 1:59 PM), <https://www.theguardian.com/us-news/2021/jan/07/donald-trump-twitter-ban-comes-to-end-amid-calls-for-tougher-action> [<https://perma.cc/B2WF-ERDC>]; Kate Klonick, *Inside the Making of Facebook’s Supreme Court*, NEW YORKER (Feb. 12, 2021), <https://www.newyorker.com/tech/annals-of-technology/inside-the-making-of-facebooks-supreme-court> [<https://perma.cc/2ERV-JXRK>].

rules.³⁹ Content moderators are employed and trained,⁴⁰ and there are multiple layers of review before decisions are made about content. Currently, the company uses a combination of algorithms and human review to implement its Community Standards.⁴¹ Layered over review and implementation is an appeals mechanism for users who disagree with the company's decisions. This was introduced after an outcry about the opacity and finality of Facebook's decisionmaking.⁴² The Oversight Board is arguably the topmost layer of decisionmaking, but it hears a fraction of the total number of appeals. It is Facebook's experiment with introducing independent experts to weigh in on how it interprets and implements its own rules.

B. Enter the Oversight Board

In 2020, Facebook added the Oversight Board to the content-moderation system consisting of detailed Community Standards, human and algorithmic enforcement of these standards, and the apparatus for appeals and reviews.⁴³ The Board was constructed as an independent entity to legitimize the company's self-regulatory system.⁴⁴ It reviews carefully chosen Facebook decisions about content, including choices to leave it up, remove it, or designate it in particular ways, weighs in on difficult questions, and helps the company think through the interpretation of its Community Standards.⁴⁵ It also offers recommendations about other questions that Facebook refers to it,⁴⁶ such as account suspension in the case of Donald Trump.

The scope of the Board's decisions is limited to interpreting Facebook's values, Community Standards, and other related content policies, and taking action in relation to these policies.⁴⁷ This may include, for example, instructing Facebook to allow or remove content or

³⁹ See Balkin, *supra* note 27, at 2021.

⁴⁰ Klonick, *supra* note 8, at 1642; see also SARAH T. ROBERTS, BEHIND THE SCREEN: CONTENT MODERATION IN THE SHADOWS OF SOCIAL MEDIA 80 (2019) (describing content-moderation training at a pseudonymous tech giant).

⁴¹ See Klonick, *supra* note 8, at 1635–42.

⁴² Casey Newton, *Facebook Makes Its Community Guidelines Public and Introduces an Appeals Process*, THE VERGE (Apr. 24, 2018, 5:00 AM), <https://www.theverge.com/2018/4/24/17270910/facebook-community-guidelines-appeals-process> [<https://perma.cc/W4S4-M6E8>]; see also GILLESPIE, *supra* note 8, at 209–13; Kaye, *supra* note 16, ¶ 38.

⁴³ See Hannah Bloch-Wehba, *Global Platform Governance and Private Power in the Shadow of the State*, 72 SMU L. REV. 27, 71 (2019); douek, *supra* note 9, at 3; Klonick, *supra* note 38.

⁴⁴ Mark Zuckerberg and Facebook staff have confirmed this, as reported by Klonick, *supra* note 38. See also douek, *supra* note 9, at 9–10, 18.

⁴⁵ OVERSIGHT BOARD CHARTER, *supra* note 14, art. 1, § 4.

⁴⁶ *Id.*

⁴⁷ See *id.*

to share information with the Board.⁴⁸ The Board can offer policy advice to Facebook about cases before it, or at Facebook's request.⁴⁹

Although Facebook has committed to carrying out the Board's decisions about specific cases, it has not committed to following the Board's policy guidance, except to the extent of considering such guidance carefully and transparently communicating its own decision.⁵⁰ Legal sanctions for noncompliance would not apply to Facebook if it chose to ignore the Board's decisions. They also do not apply when Facebook refuses to answer the Board's questions or fails to share accurate information, as was the case during the Trump Ban hearing. While it is possible that calls for regulation precipitated the creation of the Board, as they have for other self-regulatory systems, there is no government oversight of any kind here. Self-regulation is therefore the lens through which the Facebook Oversight Board must be viewed.⁵¹

Facebook's content-governance system raises concerns of mandate, accountability, and fairness.⁵² As a result, the Board needed to exhibit its independence in the Trump Ban case,⁵³ and to demonstrate its fairness and willingness to go against Facebook's interests, in order to be seen as a credible and legitimate accountability institution. This is necessary if the Board is to increase the legitimacy of Facebook's content-governance system in the eyes of the public. It is worth noting that since the Board is able to weigh in on only one of the many ways in which Facebook engages with its publics, it is unable to intervene in many controversies involving Facebook. The company's other questionable activities, such as widespread data collection, are not in the Board's purview.

The Oversight Board is an expensive and time-consuming endeavor that affects the reputations of the experts on the Board itself as well as the highly regarded experts in its secretariat.⁵⁴ Facebook invested an irrevocable grant in an independent trust⁵⁵ so that its elaborate system to enforce the Community Standards would gain more legitimacy. This legitimacy might have helped the company by both appeasing its publics and warding off state regulation.

⁴⁸ *Id.*

⁴⁹ *Id.*

⁵⁰ *See id.*

⁵¹ *See* Gorwa, *supra* note 8, at 9; Arun, *supra* note 8; Kaye, *supra* note 8.

⁵² *See* KAYE, *supra* note 23, at 112, Arun, *supra* note 8; Kaye, *supra* note 8.

⁵³ *See* DUNSTAN ALLISON-HOPE ET AL., BSR, HUMAN RIGHTS REVIEW: FACEBOOK OVERSIGHT BOARD 51–52 (2019), https://www.bsr.org/reports/BSR_Facebook_Oversight_Board.pdf [<https://perma.cc/EMJ8-Q2A4>]; Klonick, *supra* note 38. It bears noting in the interest of full disclosure, I was among the external peer reviewers invited by BSR to give feedback on this report before it was published. The BSR review discussed legitimacy in the context of the UN Guiding Principles on Business and Human Rights. ALLISON-HOPE ET AL., *supra*, at 3.

⁵⁴ *See* Klonick, *supra* note 38.

⁵⁵ *See* Kate Klonick, *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*, 129 YALE L.J. 2418, 2467–69 (2020).

C. Legitimizing Platform Law

There is no question of the Oversight Board offering legal legitimacy since it is not established by law.⁵⁶ While some may argue that the Board increases the moral legitimacy of the system,⁵⁷ it is arguably sociological legitimacy⁵⁸ that is among the goals of creating the Oversight Board.⁵⁹ Therefore, the yardstick by which we should assess the Oversight Board's role in legitimizing the content-moderation system is whether the "public regards it as justified, appropriate, or otherwise deserving of support for reasons beyond fear of sanctions or mere hope for personal reward."⁶⁰ This was always going to be a challenge since only a fraction of Facebook's engagement with its publics comes within the Board's purview. The Board is not involved in overseeing Facebook's practices in data collection and sharing, algorithmic manipulation, or even interpretation of local laws in content moderation.⁶¹

In its quest for sociological legitimacy, Facebook has sought to promote procedural justice through the Board and its content-moderation system.⁶² It is important that the Oversight Board and the content-moderation system more broadly are perceived as just and offering a fair procedure.⁶³ People's perceptions of a system are affected by the quality of decisions made, as well as the quality of treatment of people in the process. The Trump Ban brought these issues to center stage.

If sociological legitimacy focuses on how the public sees the system, this begs the question of how Facebook assesses its progress or gains in legitimacy. This assessment is tied to the question of what constitutes

⁵⁶ See Richard H. Fallon, Jr., *Legitimacy and the Constitution*, 118 HARV. L. REV. 1787, 1794–95 (2005) (defining legal legitimacy as “depend[ent] on legal norms,” *id.* at 1794). For a discussion of how Zuckerberg and Facebook staff were thinking of the Oversight Board in terms of legitimacy, see Klonick, *supra* note 38.

⁵⁷ See Fallon, *supra* note 56, 1796–801 (defining moral legitimacy as “moral justifiability or respect-worthiness,” *id.* at 1796).

⁵⁸ See *id.* at 1795–96.

⁵⁹ See Klonick, *supra* note 38; douek, *supra* note 9, at 18.

⁶⁰ Fallon, *supra* note 56, at 1795.

⁶¹ For the range of social media platforms' influence, see RONALD J. DEIBERT, RESET: RECLAIMING THE INTERNET FOR CIVIL SOCIETY 26–30 (2020) (illustrating the various risks of social media); Balkin, *supra* note 27, at 2016–17 (discussing the risk of collateral censorship when private platforms enforce local laws on behalf of states).

⁶² See BEN BRADFORD ET AL., JUST. COLLABORATORY, YALE L. SCH., REPORT OF THE FACEBOOK DATA TRANSPARENCY ADVISORY GROUP 4 (2019), https://law.yale.edu/sites/default/files/area/center/justice/document/dtag_report_5.22.2019.pdf [<https://perma.cc/H2GR-U6FX>]; Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283, 284 (2003).

⁶³ Cf. Tom R. Tyler & Maura A. Belliveau, *Managing Work Force Diversity: Ethical Concerns and Intergroup Relations*, in CODES OF CONDUCT: BEHAVIORAL RESEARCH INTO BUSINESS ETHICS 171, 179 (David M. Messick & Ann E. Tenbrunsel eds., 1996) (noting that authorities seeking to “bridge differences in values and interests” of an increasingly diverse polity must maintain “the perception of . . . fair procedures for decision making”).

the “public” in this context and how the perception of this public is gauged. The company’s responsiveness to its users and its willingness to change its policies are not based on public surveys or consultations accessible to all the people affected by it.⁶⁴ Its users are not able to use exit as a strategy⁶⁵ to voice their dissatisfaction with the company’s policies. It is therefore likely that Facebook uses proxies, reaching out through its staff to consult people and listening to others — such as the media, scholars, nongovernmental organizations (NGOs), and even governments — who make users’ views legible to it.

As I argue in more detail in section III.B, it is not always useful to think of all of Facebook’s users as a single public in whose eyes Facebook is aiming to gain legitimacy. We have to think in terms of publics.⁶⁶ The people affected by Facebook’s content-policy system are not homogenous. They comprise a variety of groups such as the media, political parties, and identity-based groups based, for example, on race, caste, religion, or sexual orientation. Groups affected may have competing priorities and interests. Some individuals and groups may be better than others at communicating their concerns on an ongoing basis with the company. This is likely affected by Facebook staff’s affinities, goals, and sympathies, which I discuss in more detail in section III.C.

It is evident that Facebook’s understanding of what its “community” or its publics want is mediated through powerful actors like its staff, the American press — *The New York Times* or *The Wall Street Journal* for instance — and renowned American scholars. This does not mean that others are entirely lacking in influence, but that their influence tends to be mediated through more powerful actors. Therefore, it is worth paying attention to what persuades stronger actors to see Facebook’s self-regulatory system as more legitimate, and noting that this might encompass a number of factors — including the reaction of weaker actors, to the extent that the stronger actors are connected and aligned with them.

III. FACEBOOK’S FACES

States’, publics’, and even Facebook’s interests and goals are not as consonant or as easily identified as one might think. We need to move from thinking of Facebook’s engagement with states as uniform and consistent, to thinking of how the company presents different faces to and within different states. Similarly, there is no one “community” that Facebook engages with. There are groups and individuals with varying power and influence. Facebook’s own staff and teams work towards different priorities and can vary in their accessibility and interest in

⁶⁴ KAYE, *supra* note 23, at 117–22.

⁶⁵ See generally ALBERT O. HIRSCHMAN, EXIT, VOICE, AND LOYALTY: RESPONSES TO DECLINE IN FIRMS, ORGANIZATIONS, AND STATES (1970).

⁶⁶ See *infra* section III.B, pp. 251–54.

different individuals, groups, and states. Facebook is externally and internally complicated.

States engage with social media platforms in more ways than are easily visible. There are ongoing regulatory conversations in which state actors and companies negotiate, agree, and renegotiate.⁶⁷ In the context of content moderation, Professor Jack Balkin's work shows us that there are two ways in which states influence speech: through laws and by informally pressuring platforms like Facebook to regulate through their own Community Standards.⁶⁸ States also have other carrots and sticks at their disposal, in the form of data-protection regulations, competition regulations, or regulatory approvals for the companies' new products. In section III.A, I argue that the state-platform relationship is not confined to local laws or to political leaders' speech, and that Facebook takes a number of factors into account while navigating its relationship with states. This is further complicated by the fact that some states can influence Facebook more easily than other states.

Facebook's rules or Community Standards can also change in response to pressure from its publics, especially from particular nonstate actors such as select NGOs or a section of the media. Facebook's relationship with its publics, like its relationship with states, is complicated. In section III.B, I offer an account that shows that publics can be strong or weak in their engagement with the company, with states, and with each other.

I tie together all these varied influences on Facebook's creation and enforcement of rules in section III.C. This last section is about why it helps to think of Facebook in terms of its many faces and internal voices, and the variations in its dealings with different states and different publics.

A. Facebook and States

The interpretation and enforcement of local laws, and the framing and enforcement of Community Standards are two major sites of engagement between Facebook and states.⁶⁹ As Professor Balkin has pointed out, nation-states' use of social media platforms to enforce local laws leads to collateral censorship, and states also pressure intermediaries to use their private rules — like Community Standards — to carry out states' agendas of regulating speech.⁷⁰ I expand this narrative by

⁶⁷ See Black, *supra* note 27, at 194.

⁶⁸ Balkin, *supra* note 27, at 2015–16.

⁶⁹ See *id.*

⁷⁰ *Id.* at 2015–25; see also Daphne Keller, *Who Do You Sue? State and Platform Hybrid Power over Online Speech* 3–7 (Hoover Inst. Working Grp. on Nat'l Sec., Tech., & Law, Aegis Series Paper No. 1902, 2019), https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech_o.pdf [<https://perma.cc/8STZ-EBXL>]; Molly K. Land, *Against Privatized Censorship: Proposals for Responsible Delegation*, 60 VA. J. INT'L L. 363, 380–88 (2020).

pointing out that social media platforms like Facebook do not necessarily engage with all states in the same way, and that their negotiations with states take place on multiple fronts.

It is important to remember that Facebook also engages with states in many ways that have little to do with content moderation. These include data protection, regulatory approval for new products, and other matters that affect the company, such as taxation. It may be fair to assume that like most companies Facebook would prefer to retain access to its markets, and to ward off state regulation as far as possible.

Refusing to follow local laws threatens Facebook's access to markets.⁷¹ At worst, the platform can be blocked altogether and its staff threatened with arrest.⁷² However, since content moderation is only one among many ongoing negotiations between Facebook and states, there are other costs to refusing to comply. Facebook risks being regulated and having its activities impeded or questioned if it fails to cooperate with the government or antagonizes powerful political leaders. In short, the interpretation of local laws, and the parallel system for content moderation that it creates, is one of many sites for the platform-government negotiations that implicate other actors who may also have a stake in the outcomes of these negotiations.

Both censorship in compliance with local laws and through the social media platform's privately ordered rules⁷³ are affected by the leverage that different states have with Facebook. This may depend on a range of factors, including Facebook's direct and indirect interests in the market, its reputational concerns, its past engagement with the government, and its staff's appreciation and identification with the state's concerns.⁷⁴ As a result, the company may be willing to shut down its services in some states and may be willing to change its global content policies at the behest of other states.

Facebook, like other social media companies, interprets local laws when it makes decisions about implementing them.⁷⁵ The company may not always interpret these laws in the same way as the state does.⁷⁶ It

⁷¹ See, e.g., Declan Walsh, *Pakistan Lifts Facebook Ban but "Blasphemous" Pages Stay Hidden*, THE GUARDIAN (May 31, 2010, 11:55 AM), <https://www.theguardian.com/world/2010/may/31/pakistan-lifts-facebook-ban> [<https://perma.cc/A5AC-6Q7B>].

⁷² See generally REBECCA MACKINNON, CONSENT OF THE NETWORKED: THE WORLDWIDE STRUGGLE FOR INTERNET FREEDOM (2012).

⁷³ For accounts of platforms' adoptions of rules, see GILLESPIE, *supra* note 8, at 45–73; and Klönick, *supra* note 8, at 1631–35.

⁷⁴ These concerns may be related to a laudable goal like human rights. However, some human rights concerns may be more legible to the company than others depending on the backgrounds and priorities of the staff evaluating the concerns.

⁷⁵ See Balkin, *supra* note 27, at 2015–21; see also Black, *supra* note 27, at 194–95.

⁷⁶ See Black, *supra* note 27, at 176.

might also contest existing and emerging law.⁷⁷ This leads to a back-and-forth with the state until a shared interpretation is reached.⁷⁸ As discussed in section II.A, Facebook's decisions interpreting and implementing local laws are not open to appeal. These opaque decisions, which are discussed sometimes in transparency reports, are not audited or assessed except by GNI, in the manner described above.

The global information companies, including Google and Facebook, also engage closely with the process of framing laws, and offer open resistance to some of these laws on occasion.⁷⁹ From time to time, the negotiations lead to either states or companies acting in ways that make it into the news. For example, Australia and Facebook engaged in negotiation over a proposed Australian law to make Facebook pay the Australian news media for its content, during which Facebook disabled all news pages for Australian users.⁸⁰ Equally, states can persuade the company to "self-regulate" through laws that place the burden on the company to remove certain kinds of content.⁸¹ Informal pressure from states can nudge the company to create new Community Standards. Sometimes the pressure comes from individuals. Facebook's "Cross Check" system provides another layer of review for content posted by high-profile accounts, including certain political leaders.⁸² This is arguably a way for the company to avoid irking those who can bring a state's power to bear on it.

⁷⁷ See, e.g., Hannah Beech, *Facebook Plans Legal Action After Thailand Tells It to Mute Critics*, N.Y. TIMES (Sept. 24, 2020), <https://www.nytimes.com/2020/08/25/world/asia/thailand-facebook-monarchy.html> [<https://perma.cc/SK6L-WGZ5>].

⁷⁸ See Black, *supra* note 27, at 194.

⁷⁹ See, e.g., Julie E. Cohen, *Law for the Platform Economy*, 51 U.C. DAVIS L. REV. 133, 196 (2017).

⁸⁰ See Amanda Meade, *Facebook News Ban Fears Grow as Tech Giant Fails to Sign Deals with Australia's Big Media Players*, THE GUARDIAN (Mar. 10, 2021, 11:30 AM), <https://www.theguardian.com/media/2021/mar/11/facebook-news-ban-fears-grow-as-tech-giant-fails-to-sign-deals-with-australias-big-media-players> [<https://perma.cc/2CSN-SEHB>].

⁸¹ See, e.g., *Netzwerkdurchsetzungsgesetz [NetzDG]* [Network Enforcement Act], Sept. 1, 2017, BUNDESGESETZBLATT, Teil I [BGBl. I] at 3352 (Ger.); Linda Kinstler, *Germany's Attempt to Fix Facebook Is Backfiring*, THE ATLANTIC (May 18, 2018), <https://www.theatlantic.com/international/archive/2018/05/germany-facebook-afd/560435> [<https://perma.cc/RHB6-P3YA>]; see also Balkin, *supra* note 27, at 2028–32 (using the example of Germany's NetzDG law to theorize how private companies have become privatized bureaucracies that implement state laws).

⁸² See Monika Bickert, *Working to Keep Facebook Safe*, FACEBOOK (July 17, 2018), <https://about.fb.com/news/2018/07/working-to-keep-facebook-safe> [<https://perma.cc/MUK9-F8JV>]; Jeff Horwitz, *Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt.*, WALL ST. J. (Sept. 13, 2021, 10:21 AM), <https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353> [<https://perma.cc/H7VE-5KG3>]. Facebook has updated its Cross Check system, expanding the additional layer of review to other eligible accounts in addition to high-profile accounts. See *Reviewing High-Impact Content Accurately via Our Cross-Check System*, FACEBOOK: TRANSPARENCY CTR. (Jan. 19, 2022), <https://transparency.fb.com/enforcement/detecting-violations/reviewing-high-visibility-content-accurately> [<https://perma.cc/5CGV-F5SS>]. This change should not affect the arguments in this Essay.

Some states or coalitions of states are more influential than others.⁸³ For example, the European Commission was able to persuade Facebook, Microsoft, Twitter, and YouTube to follow a “voluntary ‘Code of Conduct,’” which required them to implement new community guidelines for hate speech and expedite review of hate-speech notifications, so that it takes place within twenty-four hours whenever possible.⁸⁴ Professor Ben Wagner has argued that the United States, the United Kingdom, and Germany have been the most influential in the shaping of global content regulation.⁸⁵ There is clearly a variation in cooperation of social media platforms with different states. If Facebook complied with all demands made by all states, it might not have been banned in China and Iran. Zuckerberg obeyed summons from the European Parliament.⁸⁶ But he has refused calls from other parliaments around the world, including those of Canada and India,⁸⁷ and even refused to appear before an “international grand committee” consisting of policy-makers from Argentina, Brazil, Canada, Ireland, Latvia, Singapore, and the United Kingdom.⁸⁸

However, even choices made based on how important a state is to Facebook are more complicated than they seem. The failure to prioritize concern for humanity in a small market like Myanmar may create reputational problems for the company in larger markets, which may also

⁸³ See KAYE, *supra* note 23, at 86.

⁸⁴ Bloch-Wehba, *supra* note 43, at 45.

⁸⁵ BEN WAGNER, GLOBAL FREE EXPRESSION — GOVERNING THE BOUNDARIES OF INTERNET CONTENT 6 (2016).

⁸⁶ See Adi Robertson, *Watch Mark Zuckerberg Testify Before the European Parliament*, THE VERGE (May 22, 2018, 8:15 AM), <https://www.theverge.com/2018/5/22/17377974/mark-zuckerberg-european-parliament-meeting-gdpr-streaming-how-to-watch> [<https://perma.cc/BZS4-X3TS>]; Adam Satariano & Milan Schreuer, *Facebook’s Mark Zuckerberg Gets an Earful from the E.U.*, N.Y. TIMES (May 22, 2018), <https://www.nytimes.com/2018/05/22/technology/facebook-eu-parliament-mark-zuckerberg.html> [<https://perma.cc/DU3R-8GPZ>].

⁸⁷ See, e.g., *Facebook Vice President to Appear Before Parliamentary Panel on March 6*, NDTV (Mar. 4, 2019, 5:00 PM), <https://www.ndtv.com/india-news/facebook-vice-president-joel-kaplan-to-appear-before-parliamentary-panel-on-march-6-2002092> [<https://perma.cc/5WUX-GHVP>]; Alex Hern & Dan Sabbagh, *Zuckerberg’s Refusal to Testify Before UK MPs “Absolutely Astonishing,”* THE GUARDIAN (Mar. 27, 2018, 11:04 AM), <https://www.theguardian.com/technology/2018/mar/27/facebook-mark-zuckerberg-declines-to-appear-before-uk-fake-news-inquiry-mps> [<https://perma.cc/WLQ5-GNXV>]; Christopher Knaus et al., *Australian Politicians Call for Facebook’s Mark Zuckerberg to Appear Before Inquiry*, THE GUARDIAN (Dec. 7, 2019, 2:00 PM), <https://www.theguardian.com/australia-news/2019/dec/08/australian-politicians-call-for-facebooks-mark-zuckerberg-to-appear-before-inquiry> [<https://perma.cc/DQJ2-EJ75>]; Donie O’Sullivan & Paula Newton, *Zuckerberg and Sandberg Ignore Canadian Subpoena, Face Possible Contempt Vote*, CNN (May 28, 2019, 2:12 PM), <https://www.cnn.com/2019/05/27/tech/zuckerberg-contempt-canada/index.html> [<https://perma.cc/694F-ZTAY>].

⁸⁸ Tony Romm, *Facebook CEO Mark Zuckerberg Rejects Request to Testify in Front of Seven Countries’ Lawmakers — But a Lower-Level Official Will Appear*, WASH. POST (Nov. 23, 2018), <https://www.washingtonpost.com/technology/2018/11/23/facebook-ceo-mark-zuckerberg-rejects-request-testify-front-seven-countries-lawmakers-lower-level-official-will-appear> [<https://perma.cc/3RBZ-JWFP>].

impact its profits. In other words, other states or nonstate actors may take an interest in Facebook's activities in a particular state, such that the choices before the company are not necessarily a straightforward question of whether or not to appease a state (since the appeasement of one state may well have a detrimental effect on the appeasement of another state).

B. Facebook and Its Publics

If Facebook does not present itself in exactly the same way in its dealings with different states, it certainly does not present itself the same way in its dealings with its publics. We tend to fall into the trap of thinking about the people who use, discuss, are affected by, work for, work with, and otherwise engage with Facebook as if they all have equal power.⁸⁹ The content-moderation system is evolving based on this implicit assumption, at best reaching for equal access by trying to accommodate language and regional groups. This is not, however, an accurate account of Facebook's publics. Borrowing from Professor Nancy Fraser's critique of Jürgen Habermas and her work on counterpublics,⁹⁰ I argue here that we would be better served to think of Facebook's publics in terms of a variety of groups and communities, sometimes allied and sometimes in opposition to each other.

Facebook may actually be affected more by how some influential publics, which may include groups such as the American media, view its legitimacy. Other groups such as Kashmiri teenagers or refugees (except to the extent that the powerful publics are able to connect with them and reflect their concerns) are likely to have less influence.⁹¹ Facebook's staff, community, and people affected by it include influential groups and individuals as well those who are less able to make themselves heard and understood through the communicative processes made available to them.⁹²

With this perspective, existing accounts of Facebook's decisionmaking yield interesting information. It helps us take a closer look at Professor Kate Klonick's thorough empirical work⁹³ to see that the company's

⁸⁹ This is true not just of Facebook's publics but also of all users of the internet. See, for example, the discussion of the Networked Public Sphere in canonical writing such as YOCHAI BENKLER, *THE WEALTH OF NETWORKS* 181, 271–72 (2006).

⁹⁰ Nancy Fraser, *Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy*, *SOCIAL TEXT*, no. 25/26, 1990, at 56, 57–61.

⁹¹ For a discussion of the formation of different publics and their relationships with the public sphere, see *id.* For a discussion of how some groups' experiences are not acknowledged by content-governance systems, see Sarita Schoenebeck & Lindsay Blackwell, *Reimagining Social Media Governance: Harm, Accountability, and Repair*, 23 *YALE J.L. & TECH. (SPECIAL ISSUE)* 113 (2021).

⁹² See Schoenebeck & Blackwell, *supra* note 91, at 129–32; see also Black, *supra* note 27, at 182–89.

⁹³ Klonick, *supra* note 8, at 1621, 1631–35.

Community Standards were developed by staff who not only were American but also appear to have been predominantly white. In the writing and review of Facebook's Community Standards, the company's staff have moved towards consulting a range of stakeholders. Informally as well as through deliberate engagement, the staff tend to be in touch with a range of influential people, and senior staff will usually have strong relationships within particular communities. The Oversight Board itself, for example, was triggered by a Harvard professor who had a long-standing friendship with Facebook's Chief Operating Officer.⁹⁴

Facebook's efforts at stakeholder engagement, and the voices that are privileged through it, depend on choices made by Facebook's staff. The minutes of the Product Policy Forum indicate that a range of stakeholders across many countries are consulted.⁹⁵ However, these people are chosen by staff and their inputs are filtered into the system by the staff, who end up acting as gatekeepers capable of amplifying or minimizing their inputs if they choose to. It is not just the geographic diversity of stakeholders that is significant, but also the individual diversity of the technical experts within the group of stakeholders.⁹⁶

It is evident that the company's staff are often in a position to translate or transmit public opinions to Facebook,⁹⁷ which makes them a very influential group. There are variations of power between them depending on their roles and influence within the company. When the staff in contact with stakeholders are sympathetic to their concerns, they may find themselves overruled by more powerful colleagues who are likely to never come into contact with these stakeholders. However, sometimes they can also make themselves heard outside their roles. In 2020, for instance, there was a virtual walkout by staff who felt that the company should have acted to restrict Donald Trump's use of the platform.⁹⁸ In 2021, former employees spoke out as whistleblowers, once they felt

⁹⁴ Klonick, *supra* note 38.

⁹⁵ See *Product Policy Forum Minutes*, FACEBOOK (Nov. 15, 2018), <https://about.fb.com/news/2018/11/content-standards-forum-minutes> [<https://perma.cc/AFT9-TRCD>].

⁹⁶ Jenny Domino, *The Facebook Oversight Board: International Law and Expert Rule in Private Global Speech Adjudication* (Jan. 13, 2021) (unpublished manuscript) (on file with the Harvard Law School Library).

⁹⁷ See, e.g., Klonick, *supra* note 8, at 1655 (describing employees' protest of Facebook's handling of content posted by then-presidential candidate Donald Trump in violation of the company's hate-speech policies).

⁹⁸ Sheera Frenkel et al., *Facebook Employees Stage Virtual Walkout to Protest Trump Posts*, N.Y. TIMES (Oct. 10, 2021), <https://www.nytimes.com/2020/06/01/technology/facebook-employee-protest-trump.html> [<https://perma.cc/9Q9U-SMCG>]; Fanny Potkin et al., *Facebook Staffers Walk Out Saying Trump's Posts Should Be Reined In*, REUTERS (June 1, 2020, 7:51 AM), <https://www.reuters.com/article/us-facebook-trump-employee-criticism/facebook-staffers-walk-out-saying-trumps-posts-should-be-reined-in-idUSKBN2382Do> [<https://perma.cc/2XUK-BAPQ>].

that their attempts to persuade the company from within to take an interest in vulnerable publics were doomed to failure.⁹⁹

The American political debate filtered into conversations within Facebook through the staff. It is easy to see how American employees who understand, follow, and care about the situation in their country might advocate for goals that they share with American civil society. This triggers the question of whether staff in other countries are able to perform a similar function for their societies. It is unclear whether there are enough people in roles of sufficient visibility and influence to manage this. Despite the company's outreach efforts to include the voices of other groups, these voices are filtered through the staff in charge and can be affected by their networks and sympathies.

Sometimes, marginalized groups will ally with stronger groups such as the American press to make their voices heard.¹⁰⁰ This is another way to get the company's attention, and to influence it to change its policies. This leveraging of a stronger public is how Facebook's role in the Rohingya genocide came to light, forcing the company to take steps to address its role facilitating incitement to genocide in Myanmar. The people of Myanmar were unable to get the company to take their concerns seriously.¹⁰¹ However, once the role of Facebook in enabling incitement to violence made it to *The New York Times*, and the United Nations Independent International Fact-Finding Mission on Myanmar publicly chastised the company for its role,¹⁰² the Burmese people had the company's attention.

Leveraging influential groups and states is a useful strategy for weak publics in their engagement with the company. Given that it appears to be mostly white Americans who led the initial framing of the platform's Community Standards,¹⁰³ it may be unsurprising that early versions of the rules failed to account for the power relations that are inherent to racist speech and for its region-specific manifestations such as caste.

⁹⁹ See, e.g., Reed Albergotti, *Frances Haugen Took Thousands of Facebook Documents: This Is How She Did It*, WASH. POST (Oct. 26, 2021, 12:00 PM), <https://www.washingtonpost.com/technology/2021/10/26/frances-haugen-facebook-whistleblower-documents> [<https://perma.cc/NL6Q-44U7>].

¹⁰⁰ See, e.g., Billy Perrigo, *Facebook's Ties to India's Ruling Party Complicate Its Fight Against Hate Speech*, TIME (Aug. 27, 2020, 9:00 PM), <https://time.com/5883993/india-facebook-hate-speech-bjp> [<https://perma.cc/UM79-K6C7>]; Newley Purnell & Jeff Horwitz, *Facebook's Hate-Speech Rules Collide with Indian Politics*, WALL ST. J. (Aug. 14, 2020, 12:47 PM), <https://www.wsj.com/articles/facebook-hate-speech-india-politics-muslim-hindu-modi-zuckerberg-11597423346> [<https://perma.cc/Q8W5-G3GS>].

¹⁰¹ See Megan Specia & Paul Mozur, *A War of Words Puts Facebook at the Center of Myanmar's Rohingya Crisis*, N.Y. TIMES (Oct. 27, 2017), <https://www.nytimes.com/2017/10/27/world/asia/myanmar-government-facebook-rohingya.html> [<https://perma.cc/465C-ATZG>].

¹⁰² *Myanmar: UN Blames Facebook for Spreading Hatred of Rohingya*, THE GUARDIAN (Mar. 12, 2018, 9:17 PM), <https://www.theguardian.com/technology/2018/mar/13/myanmar-un-blames-facebook-for-spreading-hatred-of-rohingya> [<https://perma.cc/2GDL-ASHW>].

¹⁰³ See Klonick, *supra* note 8, at 1621, 1631–35.

However, reflection and better policies ensued after the rules were criticized by the American press.

Facebook's content-moderation systems affect, involve, and engage several nonstate entities. Nonstate actors are involved both with the implementation of local laws and with Facebook's privately ordered community-standards system. Facebook receives reports of violations of local laws "from governments and courts, as well from non-government entities such as members of the Facebook community and NGOs."¹⁰⁴ It is evident that state and nonstate actors can in theory engage with the process of policymaking, interpretation, and implementation. It is highly likely that a strong public like *The New York Times* has more influence than a weak state.

This has been confirmed by scholars who have shown that some actors have a greater influence on the standard-setting process than others.¹⁰⁵ Professor Wagner writes in his book of how certain NGOs and networks have developed an outsize influence on these policies.¹⁰⁶ Over time Facebook restructured its process for writing the Community Standards to expand the groups of people who might influence the norms.¹⁰⁷ While the process certainly draws from a larger and more varied pool now, it is important to be aware that it remains the case that some actors have greater influence than others on this process.

Facebook engages with many groups and communities, which do not all have the same power and leverage as far as the company is concerned. Awareness of this enables weaker publics to act strategically to gain influence through alliances. This means that the company should account for these alliances, and resist complacency about policies and actions that endanger groups that do not seem influential. These policies and actions seem to have a way of finding themselves in *The New York Times*, whether through civil society or through Facebook's own staff — who may not have been heard within the company but decided to make themselves heard from outside it.

C. Facebook and Its Staff

It is not just Facebook's engagement with states and publics that takes place on multiple fronts with varying goals. The company is also complex internally, like other large corporations. Evidence of this surfaces from time to time, with employees leaking information when they are unhappy with Facebook's decisions and with staff protests.¹⁰⁸

¹⁰⁴ *Content Restrictions Based on Local Law*, *supra* note 12.

¹⁰⁵ See Professor Kate Klonick's account of how those who created the Community Standards were mostly Americans immersed in American democratic culture in Klonick, *supra* note 8, at 1618–25.

¹⁰⁶ WAGNER, *supra* note 85, at 121–33.

¹⁰⁷ Klonick, *supra* note 8, at 1631–35.

¹⁰⁸ See, e.g., Albergotti, *supra* note 99; Frenkel et al., *supra* note 98; Perrigo, *supra* note 100; Potkin et al., *supra* note 98; Purnell & Horwitz, *supra* note 100.

Although a corporation can be assumed to have a broad profit-making motive, it has to choose from different paths towards that profit, and does have reputational concerns to take into account. While not all rent-seeking behavior is necessarily public, some of the company's choices — like its poor judgment in Myanmar — are very public and affect the company's relationship with other actors. As a result, Facebook — like other corporations — has teams and internal priorities that are directed at ensuring that tragedies like the one in Myanmar do not occur. Apart from joining GNI, the company also hired a Director of Human Rights,¹⁰⁹ who has a global team and commissions reports on the company's effect on human rights.

Facebook has to think of different states, of various publics, and of its own employees while making and communicating decisions. It has complex internal processes and teams with different goals to create the internal debates and friction necessary to balance competing priorities. Researchers who have engaged in painstaking interviews with particular teams — Professor Klonick, and Professors Wolfgang Schulz and Matthias Kettemann — report that these teams are devoted to their goals of creating the best possible Oversight Board, or developing Community Standards that best address harmful speech without affecting voice excessively.¹¹⁰ It is safe to assume that other teams — like those focused on marketing or government relations — have different priorities. Now and then, the internal debates spill over into the public, but they are largely opaque, contained within the corporation and taking place on terms that are not always clear.

These internal competing priorities mean that the company presents the world with different faces. There are teams directed at cultivating government relationships and engaging policymakers. There are also teams with strong relationships with NGOs and the human rights community. People outside the company can better affect the company's decisions if they are able to access the team with both influence and sympathetic priorities. It also means that we can safely assume that there is no agreement within Facebook about a question like the Trump Ban.¹¹¹ In addition to being a polarizing question outside the company,

¹⁰⁹ Joshua Brustein, *Facebook's First Human Rights Chief Confronts Its Past Sins*, BLOOMBERG (Jan. 29, 2020, 1:19 PM), <https://www.bloomberg.com/news/articles/2020-01-28/facebook-s-first-human-rights-chief-seeks-to-tame-digital-hate> [<https://perma.cc/9TH9-H92D>].

¹¹⁰ See Klonick, *supra* note 8, at 1634–35; Matthias C. Kettemann & Wolfgang Schulz, *Setting Rules for 2.7 Billion: A (First) Look into Facebook's Norm-Making System* 13–14, 24–33 (Hans-Bredow-Institut, Works in Progress #1, 2020), https://leibniz-hbi.de/uploads/media/default/cms/media/kogjxdi_AP_WiP001InsideFacebook.pdf [<https://perma.cc/CY7U-SZ6Q>].

¹¹¹ See, e.g., Kevin Roose et al., *Don't Tilt Scales Against Trump, Facebook Executive Warns*, N.Y. TIMES (June 30, 2020), <https://www.nytimes.com/2020/01/07/technology/facebook-trump-2020.html> [<https://perma.cc/C79R-TTBG>].

the deplatforming of political leaders is likely to be a polarizing question internally.¹¹²

Within the company, there are people with varying ideas of what the ideal outcome should be and varying sympathies. While there are staff who are devoted to the bottom line, there are also staff who care about the groups harmed through use of the company's platform. The whistleblowers are evidence that there are employees who tried to advocate for certain points of view from within and then chose to exit to engage in more public advocacy when they did not feel heard.

If the Facebook Oversight Board was meant to legitimize the company's content-moderation system to the outer world, perhaps it was also meant to offer this legitimizing role to the members of the company's own staff who have had concerns about their employer's decisionmaking. If this is so, 2021's whistleblowers suggest that some staff were dissatisfied with the extent to which the Board had been able to address the company's problems on its own.

IV. WHAT THE TRUMP BAN CASE CAN TEACH US

The Trump Ban ended the Oversight Board's period of easing into building a reputation and moved it into more controversial terrain. Before this case, the Board grappled with safer questions that affected Facebook's publics but not states or their leaders. Since this case was about political leaders and heads of state, it was globally polarizing and might have had implications for Facebook's access to global markets. Unsurprisingly, it tested the relationship between the Oversight Board and Facebook in ways that I discuss below in section IV.C.

The Oversight Board upheld Facebook's decision to restrict the former President's accounts, while questioning the indeterminate penalty and indefinite suspension and pointing to the lack of a clearly established procedure for these decisions.¹¹³ The Board chastised Facebook for failing to come up with appropriate rules, and refused to develop the rules for the company.¹¹⁴ In doing so, it sidestepped a perilous decision and confined itself to outlining acceptable parameters for Facebook to craft its own rules. It appears from the comments on the decision by Nick Clegg — then the company's Vice President of Global Affairs and Communications — that Facebook had hoped to get the Oversight Board to take responsibility for its rules for political leaders.¹¹⁵ The

¹¹² See, e.g., Potkin et al., *supra* note 98; Purnell & Horwitz, *supra* note 100; Roose et al., *supra* note 111.

¹¹³ See *Trump Decision*, *supra* note 1, at 3–4.

¹¹⁴ See *id.* at 33, 35–38.

¹¹⁵ See Nicholas Vinocur, *Facebook Top Lobbyist Pushes Back on “Contradictory” Trump Ban Guidance*, POLITICO (May 13, 2021, 6:00 PM), <https://www.politico.eu/article/facebook-vp-nick-clegg-pushes-back-on-contradictory-trump-ban-guidance-oversight-board> [<https://perma.cc/N423-G6WL>].

Board asserted its independence by clarifying that arbitrariness and failure to uphold human rights standards are not acceptable, while letting the company shoulder the responsibility of drafting appropriate rules.

The Trump Ban case attracted comments from around the world about how Facebook should treat political leaders. Opinions were strong and in opposition to each other.¹¹⁶ This made it crucial for the Board's reasoning and the process to be seen as fair, or at least as better than the existing system being used by Facebook. The decision was finely calibrated, and invited Facebook to limit its own power of arbitrary decisionmaking by publishing a clear policy and adhering to it. The decision also recommended that the company limit the extent to which it permits incitement of violence on its platform, specifying norms that Facebook should follow.¹¹⁷ However, the rules for political leaders would be Facebook's rules, not the Oversight Board's rules.

This Part of the Essay discusses the details of the Trump Ban case, beginning with the questions before the Oversight Board and the Facebook rules that applied to them. It then goes over the company's engagement with the Oversight Board's questions in the context of information revealed by Frances Haugen after the case. Bringing together these threads, section IV.C discusses what the Trump Ban case and its aftermath show us about the Oversight Board.

A. Facebook's Rules and Donald Trump

Facebook indefinitely suspended then-President Donald Trump's Facebook and Instagram accounts for violating the Community Standards.¹¹⁸ This decision was made through the company's internal process, which in this particular case meant that it was made by Zuckerberg since it involved a head of state.¹¹⁹ Facebook invited the Oversight Board's views on two of the many questions raised by its decision.¹²⁰ The first question was whether the *indefinite* suspension of Trump's accounts was consistent with the platform's values, and the second sought the Oversight Board's recommendations about suspending the accounts of political leaders.¹²¹

While trying to understand the basis for the suspension of Trump's accounts, the Oversight Board looked into the standards and processes

¹¹⁶ See Caleb Ecarma, "This Is Going to Be a Global Moment": All Eyes Are on Facebook as It Weighs Whether to Ban Donald Trump for Life, VANITY FAIR (Apr. 22, 2021), <https://www.vanityfair.com/news/2021/04/facebook-weighs-lifetime-trump-ban> [<https://perma.cc/Y4EF-RK2P>]; Klonick, *supra* note 38.

¹¹⁷ See *Trump Decision*, *supra* note 1, at 30–32.

¹¹⁸ See Rosen & Bickert, *supra* note 37; see also Clegg, *supra* note 7.

¹¹⁹ Hern & Paul, *supra* note 38; Klonick, *supra* note 38.

¹²⁰ Clegg, *supra* note 7.

¹²¹ *Id.*

Facebook used.¹²² It also raised questions about why, in the past, the former President seemed to be able to share content that other people were not allowed to post on Facebook.¹²³ A Facebook user tested this by copying and pasting all Trump's content verbatim and found a copied post was removed for violating Community Standards while Trump's identical original post stayed online.¹²⁴ The copied post was, however, restored later, and Facebook said its removal was an error.¹²⁵

At first glance the exceptions made for Donald Trump's accounts appeared to have been a result of the newsworthiness exception that is a part of Facebook's published content-moderation standards.¹²⁶ This exception was adopted in 2016¹²⁷ and provides a published standard for leaving up content that otherwise violates Facebook's Community Standards, if the public interest of the content outweighs the harm.¹²⁸ In response to the Oversight Board's questions, Facebook revealed that a different, poorly documented policy for the implementation of Community Standards called "Cross Check" had been used.¹²⁹

Facebook has described the "Cross Check" system¹³⁰ as "a second layer of review" to make sure that its Community Standards have been applied correctly to "high profile" accounts.¹³¹ It is not a standard for what content will be allowed, but a procedural rule for how certain content will be reviewed. Unlike the newsworthiness standard, which applies to all accounts, Cross Check applies to selected high-profile accounts, ensuring that they have an added layer of review if they appear to violate the Community Standards. This means that an account is first designated for Cross Check review as Donald Trump's were. Following

¹²² See *Trump Decision*, *supra* note 1, at 13–16.

¹²³ See *id.* at 12.

¹²⁴ Caitlin O'Kane, *Facebook Page that Copies President Trump's Posts Gets Flagged for Violence — When the President's Didn't*, CBS NEWS (June 12, 2020, 2:09 PM), <https://www.cbsnews.com/news/facebook-donald-trump-copy-account-flagged-inciting-violence> [<https://perma.cc/WXQ9-6JPL>].

¹²⁵ *Id.*

¹²⁶ See Thomas E. Kadri & Kate Klonick, *Facebook v. Sullivan: Public Figures and Newsworthiness in Online Speech*, 93 S. CAL. L. REV. 37, 58–69 (2019); *Product Policy Forum*, FACEBOOK (July 30, 2019) [hereinafter *Product Policy Forum on Newsworthiness*], https://about.fb.com/wp-content/uploads/2019/12/PPF-Final-Deck_07.30.2019.pdf [<https://perma.cc/9A2C-VNXXK>].

¹²⁷ *Product Policy Forum on Newsworthiness*, *supra* note 126.

¹²⁸ This determination was made by Facebook employees on a case-by-case basis. See Kadri & Klonick, *supra* note 126, at 66.

¹²⁹ See Bickert, *supra* note 82; Horwitz, *supra* note 82.

¹³⁰ In the period following the Trump Ban decision, Cross Check was renamed "Early Response Strategic Review," but this Essay uses its legacy name for clarity. A component called "General Secondary Review" was added to this system, which is an effort to extend the added layer of review to other accounts through a dynamic prioritization system that ranks content based on "false positive risk," taking into account "topic sensitivity," "enforcement severity," "predicted reach," and other factors. *Reviewing High-Impact Content Accurately via Our Cross-Check System*, *supra* note 82.

¹³¹ Bickert, *supra* note 82.

this, if the account is found to engage in incitement, it would await an additional layer of review before the incitement is removed. It may turn out that the incitement is seen as newsworthy enough (in terms of the newsworthiness standard) to leave up, but the delay would come from Cross Check's added layer of review while the standard would come from the Community Standards, including the newsworthiness exception.

Cross Check can be applied to celebrities and governments according to Facebook,¹³² but *The Wall Street Journal* has reported that an internal guide for eligibility lists being “‘newsworthy,’ ‘influential or popular[,]’ or ‘PR risky’” as qualifications for Cross Check.¹³³ Political leaders around the world arguably fit this profile. The Cross Check system, previously described as “Shielded Review,” was the reason that U.K. far-right political leaders' content has remained online despite being flagged by content moderators.¹³⁴

In the context of newsworthiness, Facebook's Product Policy Team decided to presume that politicians' speech is newsworthy unless there is a “risk of harm” — which specifically includes the “[p]otential to incite violence” — that outweighs the “public interest value” of the speech.¹³⁵ Although Cross Check is described as “a second layer of review” that does not offer any additional protection,¹³⁶ this review is carried out through a separate system, staffed by full-time employees who are presumably aware that their reviews have to minimize “PR fires.”¹³⁷ *The Wall Street Journal* suggests that users can be “whitelisted” and “rendered immune from enforcement” through this system, and that some are allowed to violate Facebook's rules “pending Facebook employee reviews that often never come.”¹³⁸ A 2019 audit found that at least forty-five teams across the company were “involved in whitelisting” and that most employees were able to add Facebook users to the list.¹³⁹ Donald Trump's accounts were among those on the list.¹⁴⁰ Some of his inflammatory posts that otherwise would have been removed for violating Facebook's rules stayed online as a result.¹⁴¹

The problem of seemingly arbitrary decisionmaking about political leaders arose from the implementation of the Community Standards for

¹³² *Id.*

¹³³ Horwitz, *supra* note 82.

¹³⁴ See *Dispatches Investigation Reveals How Facebook Moderates Content*, CHANNEL 4 (July 17, 2018), <https://www.channel4.com/press/news/dispatches-investigation-reveals-how-facebook-moderates-content> [<https://perma.cc/FN6K-6DJ6>].

¹³⁵ *Product Policy Forum on Newsworthiness*, *supra* note 126.

¹³⁶ Bickert, *supra* note 82.

¹³⁷ Horwitz, *supra* note 82.

¹³⁸ *Id.*

¹³⁹ *Id.*

¹⁴⁰ *Id.*

¹⁴¹ See *id.*

high-profile users, and not necessarily from the standards themselves. The newsworthiness exception is no longer presumed to apply when there is a risk of harm, including the potential to incite violence.¹⁴² If the problem lay in the implementation of the standards, the Oversight Board needed to investigate the Cross Check system which was used on the former President's accounts.

B. Factfinding by the Oversight Board

The Oversight Board asked Facebook questions, of which several were answered, but some received partial or no responses.¹⁴³ For example, twenty pieces of content from Trump's accounts were initially marked for violating Community Standards and then found not to be violating them.¹⁴⁴ The reasons for this are unclear. Also left unanswered were questions about the impact of Facebook's News Feed¹⁴⁵ on the visibility of Trump's content and about the removal of content and suspension of accounts of other political figures.¹⁴⁶ In relation to Cross Check, when the Oversight Board asked for the criteria for adding accounts and pages to the system, Facebook did not provide this information.¹⁴⁷ The company also did not provide information about error rates of enforcement decisions made through Cross Check.¹⁴⁸

It is clear that although Facebook clarified that the former President's accounts were part of the Cross Check system, it did not share other information about this system that was pertinent to its decisions about Donald Trump.¹⁴⁹ Following the hearing and publication of the Board's decision, journalists' reports — based on facts shared by whistleblowers — about the details that Facebook failed to disclose to

¹⁴² *Product Policy Forum on Newsworthiness*, *supra* note 126.

¹⁴³ *Trump Decision*, *supra* note 1, at 21; *see also* Clegg, *supra* note 7; *Oversight Board Accepts Case on Former US President Trump's Indefinite Suspension from Facebook and Instagram*, *supra* note 11.

¹⁴⁴ *Trump Decision*, *supra* note 1, at 12.

¹⁴⁵ In February 2022, Facebook changed the name of its News Feed to "Feed." Facebook App (@facebookapp), TWITTER (Feb. 15, 2022, 12:09 PM), <https://twitter.com/facebookapp/status/1493633545444675589> [<https://perma.cc/CP9R-M9UX>].

¹⁴⁶ *Trump Decision*, *supra* note 1, at 21.

¹⁴⁷ Catalina Botero-Marino et al., *To Treat Users Fairly, Facebook Must Commit to Transparency*, OVERSIGHT BD. (Sept. 2021), <https://www.oversightboard.com/news/3056753157930994-to-treat-users-fairly-facebook-must-commit-to-transparency> [<https://perma.cc/M583-W7NX>].

¹⁴⁸ FACEBOOK, RESPONSES TO THE OVERSIGHT BOARD RECOMMENDATIONS IN THE TRUMP CASE 13–14 (2021) [hereinafter RESPONSES TO THE OVERSIGHT BOARD RECOMMENDATIONS], <https://about.fb.com/wp-content/uploads/2021/06/Facebook-Responses-to-Oversight-Board-Recommendations-in-Trump-Case.pdf> [<https://perma.cc/4XGX-FBPG>] (citing lack of feasibility of tracking this information).

¹⁴⁹ *Trump Decision*, *supra* note 1, at 21, 24; *see also* Botero-Marino et al., *supra* note 147; Chinmayi Arun, *Facebook Oversight Board's Decision on Trump Ban in a Global Context: The Treatment of Political Leaders*, JUST SEC. (May 17, 2021), <https://www.justsecurity.org/76186/facebook-oversight-boards-decision-on-trump-ban-in-a-global-context-the-treatment-of-political-leaders> [<https://perma.cc/DVL2-AYSN>].

the Oversight Board¹⁵⁰ called into question the oversight capacity of the Oversight Board. It did not help matters that Facebook gave the Board wrong information about its application of its newsworthiness policy to Trump's accounts, even if the company corrected itself after the decision.¹⁵¹ This has damaged public perception of the Board as an entity with the potential to hold Facebook accountable.¹⁵²

Following media coverage of the Cross Check system, the company has sought the Oversight Board's policy advice about this system.¹⁵³ It is clear by now that the company is not obligated to comply with this advice. It is also possible that the company may fail to share the information necessary to get meaningful advice from the Oversight Board.

C. *The Trump Ban Case and the Oversight Board*

Facebook's content-moderation system for political leaders, including the Oversight Board, was not ready for a case that surfaced so many of the company's complex relationships. Since the company failed to share statistics and other information pertaining to the Cross Check system, the Oversight Board was unable to access information about the system used to apply Facebook's standards to Donald Trump's accounts — a system that likely applies to other political leaders' accounts.

The case was a pivotal moment for Facebook's self-regulatory content-moderation system and for the Oversight Board because of the fraught question, and especially because Facebook was publicly less than forthcoming with the information necessary for an informed opinion from the Board. It induced the Oversight Board to underline its independence. The Oversight Board asserted itself by refusing to take responsibility for the company's policies and by calling Facebook out when its sharing of information was unsatisfactory.¹⁵⁴ It publicly responded differently to criticism of Facebook than the company did, going so far as to invite a prominent whistleblower for a conversation.¹⁵⁵

¹⁵⁰ See Adam Smith, *Facebook "Repeatedly Lied" to Oversight Board About Secret VIP List That Let Users Break Rules, Whistleblower Says*, THE INDEPENDENT (Oct. 12, 2021, 10:28 AM), <https://www.independent.co.uk/life-style/gadgets-and-tech/facebook-whistleblower-lied-haugen-oversight-board-xcheck-cross-check-b1936673.html> [<https://perma.cc/CS7P-ZAFX>]; Hanna Ziady, *Facebook Kept Its Own Oversight Board in the Dark on Program for VIP Users*, CNN (Oct. 21, 2021, 12:29 PM), <https://www.cnn.com/2021/10/21/tech/facebook-cross-check-oversight-board> [<https://perma.cc/WCA7-5YU4>].

¹⁵¹ RESPONSES TO THE OVERSIGHT BOARD RECOMMENDATIONS, *supra* note 148.

¹⁵² See Nick Clegg, *Requesting Oversight Board Guidance on Our Cross-Check System*, FACEBOOK (Sept. 28, 2021), <https://about.fb.com/news/2021/09/requesting-oversight-board-guidance-cross-check-system> [<https://perma.cc/72JZ-Q7MR>]; Botero-Marino et al., *supra* note 147.

¹⁵³ Clegg, *supra* note 152.

¹⁵⁴ *Trump Decision*, *supra* note 1, at 21.

¹⁵⁵ *Oversight Board to Meet with Frances Haugen*, OVERSIGHT BD. (Oct. 2021), <https://oversightboard.com/news/1232363373906301-oversight-board-to-meet-with-frances-haugen> [<https://perma.cc/NN89-KEE5>].

The questions that Facebook refused to answer, and the questions that it answered incorrectly and partially, raise concerns about how far the Oversight Board can exercise meaningful control over Facebook. The Board has publicly admonished Facebook and taken steps to investigate the implications of the whistleblower's revelations.¹⁵⁶

Facebook has undermined the Board's reputation as an institution to oversee it. Despite its reservations,¹⁵⁷ the Oversight Board could not compel answers from the company about the Cross Check system, leaving this opaque system for political leaders and other high-profile users insufficiently discussed — until the whistleblower and the media changed the company's incentives.

It is evident that Oversight Board cannot control the company. However, this is not to understate the Oversight Board's value or influence. Its expert advice has already improved Facebook's decision-making. But the answer to the more limited question of whether Facebook permits the Oversight Board to hold it accountable is no. Still, the Oversight Board does play a very significant role, which I discuss below, in the last Part of this Essay.

V. OVERSIGHT BOARD OR ADVISORY BOARD?

When I started writing this Essay in January 2021, the Oversight Board appeared to be one of Facebook's faces. It was not that the Board ever lacked independence, but it was offered up by the company as its highly publicized response to its many content-moderation problems.

It is already evident that Facebook made a bad choice in being less than forthcoming about the details of the Cross Check policy during the Oversight Board's Trump Ban hearing. The information that could have been offered up in good faith and discussed with a group of people dedicated to helping the company was revealed to the world through newspaper headlines. This undermined the company's sociological legitimacy, including its commitment to supporting the Board as an independent expert body that holds it accountable. The avalanche of bad press continues, and it is clear that offering up the Oversight Board as a solution will not shield Facebook from criticism.

What then of the Oversight Board? The experts involved have done their best to act with integrity, asking difficult questions and following up on the facts reported by journalists. It is, however, difficult to ignore the fact that disgruntled staff along with *The Wall Street Journal* were able to unearth company policies that the Oversight Board could not access, and that the Board needed allies from other publics and Facebook's staff to reopen the question of the Cross Check policy. Armed with the publicity-generated information about Cross Check by

¹⁵⁶ See, e.g., *id.*; Botero-Marino et al., *supra* note 147.

¹⁵⁷ See *Trump Decision*, *supra* note 1, at 24.

the whistleblower, the Oversight Board was able to persuade Facebook to prioritize the question.

It is becoming clear that it may be more accurate to describe the Board as an entity that exerts great influence over Facebook rather than as one that takes on the responsibility for Facebook's politically charged decisions or oversees the company. The Oversight Board is a highly influential body, consisting of people who are in a strong position to sway Facebook but who do not actually have control over the company. This makes the Board an actor that can be leveraged by weaker publics in theory. It also means that the Oversight Board can be influenced by, receive support from, and influence other powerful actors.

It appears that the Oversight Board is settling into its role as a significant advisory body committed to engaging with Facebook, and showing the company and the rest of the world how Facebook can make better decisions about content. The Board has worked on underlining its independence and commitment to its own goals, such that I, for one, cannot think of it as one of Facebook's faces. Over time perhaps it will increase its capacity to hold the company accountable. If it does, it will not be in the manner the company suggested, because it is clear that it has limited control over Facebook. The Board may come to persuade the company to change through alliances with Facebook staff, publics, and perhaps even states one day, and through the amplified consequences for the company when it is unresponsive.

CONCLUSION

Facebook has a complex web of relationships with states and with its publics. It also has a complex internal structure, which by design puts its teams at odds with each other over some questions. This healthy friction is how government-relations teams can be restrained from facilitating politicians' violations of human rights by other teams whose role it is to anticipate and prevent such tragic outcomes. If Facebook has many faces in its engagement with states and its publics, it arguably tried to present the Oversight Board as a powerful face that is barely tethered to the other faces. However, the Oversight Board has exerted its independence from Facebook firmly and visibly.

The Trump Ban hearing and decision impacted the legitimizing of Facebook's content-moderation system, and the widely communicated idea that the Oversight Board can hold the company accountable. It was followed by revelations from former Facebook employees who lost faith in the capacity of systemic solutions to surface and fix the company's problems. Although the Oversight Board tried its best, the company's failure to cooperate fully with the Board in the Trump Ban hearing suggests that this is not a body that can hold the company accountable. It is an influential group of experts who can offer excellent advice and perspectives when the company shares enough information

with it. There are many instances in which Facebook can benefit from expert advice about its content decisions. But to make the most of the Oversight Board's expertise, Facebook has to share the requisite information with the independent experts and be willing to make uncomfortable changes.¹⁵⁸

¹⁵⁸ This does not mean that the Board is free of inherent structural problems (such as the limited opportunity for users to make their case to the Board, given that Facebook has access to more data about the content than users ever will). However, Facebook's publics may see the Board as acceptable even if they do not see it as perfect.