

Making Sense of Digital Content Moderation from the Margins

Margaret Burke Fernandes

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State  
University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
in  
Rhetoric and Writing

Jennifer Sano-Franchini

Sweta Baniya

Chris Lindgren

Katrina Powell

May 2, 2022

Blacksburg, Virginia

Keywords: digital cultural rhetorics, content moderation, user experience, transparency,  
content creation

Copyright 2022

# Making Sense of Digital Content Moderation from the Margins

Margaret Burke Fernandes

## ABSTRACT

This dissertation, *Making Sense of Digital Content Moderation from the Margins*, examines how content creators who are marginalized by race, sexuality, gender, ethnicity, and disability understand their experiences of content moderation on the social media platform TikTok. Using critical interface and narrative-based inquiry methods with six marginalized content creators on TikTok, I argue that marginalized creators navigate the opaque content moderation infrastructure of TikTok by drawing on their embodied experiences. The key research questions ask how these content creators interpret TikTok's platform policies and processes through their interactions on the app and how these interpretations influence content creation on TikTok and how creators feel about moderation in the absence of platform transparency about how content is moderated. To answer these questions, I conducted narrative-driven interviews with six TikTok creators and analyzed these stories alongside online testimonials in eight Change.org petitions. My analysis revealed that lack of transparency around TikTok's algorithmic curation and moderation contributes to content creators feeling alienated, exploited, frustrated, and unwelcome on the platform and influences content creators to adapt their content to avoid moderation, oftentimes by self-censoring themselves and aspects of their marginalized identities. Over time, the accumulation of content moderation micro-interactions diminishes the ability of marginalized content creators to trust content moderation processes. My analysis also shows how TikTok's user experience design and opaque content moderation practices contribute to an affective platform environment in which creators are compelled to speak out and across creator networks about such gaps in experience and platform policy. I conclude with a discussion of how my findings about content moderation and transparency contribute to conversations in writing-related scholarship, especially as it pertains to writing assessment, technical communication, and algorithmic research methodologies.

# Making Sense of Digital Content Moderation from the Margins

Margaret Burke Fernandes

## GENERAL AUDIENCE ABSTRACT

In recent years, marginalized content creators on TikTok have sounded the alarm about the way that the platform's content moderation and algorithmic recommendation disadvantages marginalized creators. This dissertation, *Making Sense of Digital Content Moderation from the Margins*, examines how content creators who are marginalized by race, sexuality, gender, ethnicity, and disability understand their experiences of content moderation on the social media platform TikTok. The key research questions ask how these content creators interpret TikTok's platform policies and processes through their interactions on the app and how these interpretations influence content creation on TikTok and how creators feel about moderation in the absence of platform transparency about how content is moderated. To answer these questions, I conducted narrative-driven interviews with six TikTok creators and analyzed these stories alongside online testimonials. My analysis revealed that lack of transparency around TikTok's algorithmic curation and moderation contributes to content creators feeling alienated, exploited, and unwelcome on the platform and influences content creators to adapt their content to avoid moderation, oftentimes by self-censoring themselves and aspects of their marginalized identities. Moreover, I found that TikTok isolates user experiences of biased content moderation which compels creators to speak out and across creator networks about discriminatory experiences of platform policy.

## ACKNOWLEDGEMENTS

Throughout this process, I have been fortunate to be supported by so many people. I am especially grateful to my advisor, Dr. Sano-Franchini whose generous mentorship has made this project possible. Thank you for asking so many good hard questions and for sharing so much of your time with me as I stumbled for the answers. You have given me such a powerful example of what it means to be a teacher and scholar, and I'm so grateful for the opportunity to work and write with you during my time at Virginia Tech.

Thank you, too, to my committee members: Sweta Baniya, Chris Lindgren, and Katrina Powell. Your questions and suggestions strengthened my work and methods and encouraged me to take the leap into using interview methods for this dissertation. I also want to thank Dr. Lindgren for showing so much enthusiasm in my work and for showing me new possibilities for research in rhetoric and writing.

I want to thank the wonderful friends I have made here at Virginia Tech. I am so grateful to have learned so much from my fellow graduate students in coursework: Matt Homer, Fai Inthajak, Iris Farrou, Kat Gray, Yemi Awotayo, Amilia Evans, Andre Jones, Chloe Robertson, Luana Shafer, and Marti Williams. To Katie Beth Brooks and Julie Mengert, I am so very thankful to have been supported by such smart, kind scholars. Thank you for being there through the good and the bad, for sharing job materials and laughs, and for making this dissertation experience way less lonely.

A special thanks to all my good friends who have supported me from outside of academia. I am especially grateful to Jordan McNary, Allie Arend, and Torey Green, for being such kind and constant friends. I also want to thank all my friends who have helped me to escape my dissertation work with Dungeons & Dragons: Jeremy, Karlin, Cyanna, Īx, Robb, Taylor, and Sarah, Mirna, Andie, and Anderson. Rolling dice with you all has kept me grounded (although not quite in reality, thankfully).

I am where I am today because of the love and support of my family. To Janet, my brilliant sister, thank you for always being there to chat celebrity gossip and for continuing to help me see the world from new and necessary angles. To my beloved cat Bentley, I could not have done any of this without the constant love, comfort, and joy that you bring to my life. Most of all, I am immeasurably grateful to Vinny, the love of my life and my very best friend. I could not have done this without you, and I'll never be able to thank you enough for supporting this dream. When the work felt impossible, you were always there lifting me up and reminding me that it would all be worth it in the end.

Finally, this project would not be possible without the six content creators who generously shared their stories and their time with me. I hope that this work does justice to all that you have taught me and serves as a reminder that the work of creators like yourself should never be forgotten or taken for granted.

# TABLE OF CONTENTS

<b>LIST OF TABLES .....</b>	<b>viii</b>
<b>LIST OF FIGURES .....</b>	<b>ix</b>
<b>CHAPTER 1 INTRODUCTION .....</b>	<b>1</b>
Introduction .....	1
Literature Review .....	7
Content Moderation and Platform Oppression .....	9
Digital Labor, Platform Capitalism, and the Politics of Transparency .....	12
Affect, Emotion and Algorithmic Imaginaries .....	15
Chapter Overviews .....	18
<b>CHAPTER 2 METHODOLOGIES AND METHODS: CULTURAL DIGITAL RHETORICS, CREATOR STORIES, AND CRITICAL INTERFACE ANALYSIS .....</b>	<b>21</b>
Introduction .....	21
Methodologies .....	24
Creator Stories and Critical Interface Analysis .....	26
Centering Difference, Learning from the Margins .....	28
Methods: Examining Creator Interviews and Testimonials .....	29
Interview Questions .....	31
Participants .....	32
Creator Testimonies in Change.org Petitions .....	36
Analytical Framework .....	37
Critical Interface Analysis of TikTok .....	38
Limitations .....	41
<b>CHAPTER 3 SENSING SUPPRESSION: AN EXAMINATION OF CONTENT CREATOR PERCEPTIONS OF EXCLUSION IN TIKTOK’S INTERFACE .....</b>	<b>45</b>
Introduction .....	45
Mass-Reporting: The Role of Platform Users in Content Moderation .....	48
Sensing Bias: TikTok’s Affective Atmosphere of Inconsistent & Uneven Moderation	52

Sneaky Policing: Content Suppression, Reduced Reach, and Muting .....	60
Discussion: Sneaky Rhetorics of TikTok’s Content Moderation.....	62
<b>CHAPTER 4 DISRUPTED CREATOR TIME(S): UNPACKING CUMULATIVE USER EXPERIENCES OF TIME AND EXCLUSION ON TIKTOK.....</b>	<b>71</b>
Introduction .....	71
User Experience and Creator Expectations of Time on TikTok .....	74
The Evolving Nature of Platform Bans and Restricted Access on TikTok .....	78
Accumulating Content Moderation.....	83
Platform Exploitation .....	90
Disruptions in Creator Time.....	95
Discussion .....	97
<b>CHAPTER 5 COLLECTIVE COMPLAINT FROM THE MARGINS: HOW MARGINALIZED CONTENT CREATORS COUNTER CONTENT SUPPRESSION ON TIKTOK.....</b>	<b>104</b>
Introduction .....	104
Using the Platform to Teach Other Platform Users about Content Moderation .....	106
Speaking Out as a Response to Silencing and Suppression .....	110
Creator Connections for Support and Survival .....	126
Disrupting Sneaky Rhetorics: Understanding How Visibility Moderation Encourages Complaint.....	131
Discussion .....	135
<b>CHAPTER 6 CONCLUSIONS AND IMPLICATIONS FOR RESEARCHING ALGORITHMIC TRANSPARENCY IN WRITING STUDIES.....</b>	<b>138</b>
Summary of Findings .....	138
Limitations .....	141
Implications.....	142
For TikTok and Accountable Content Moderation .....	142
For Writing Assessment .....	145
For Technical Communication .....	147
Conclusion.....	148
<b>BIBLIOGRAPHY.....</b>	<b>151</b>

**APPENDIX A: INFORMED CONSENT FOR PARTICIPANTS..... 165**

**APPENDIX B: RECRUITMENT EMAIL..... 168**

**APPENDIX C: REMINDER EMAIL..... 169**

## **LIST OF TABLES**

Table 1: Types of content moderation on TikTok .....	4
Table 2: Top 8 Change.org petitions about TikTok’s Community Guidelines and Content Moderation .....	112



## LIST OF FIGURES

Figure 1: Visual representation of the temporal dimensions of UX from the Dagstuhl Seminar .....	77
Figure 2: Screenshot of TikTok’s content moderation process, including description of temporary and permanent bans .....	79
Figure 3: Screenshot of TikTok’s Account Safety page: How to appeal an account ban .	80
Figure 4: Screenshot of a TikTok system notification of a temporary post ban .....	80
Figure 5: Screenshot of TikTok’s instructions on appealing content removals from TikTok’s Account Safety .....	88
Figure 6: Screenshot of TikTok’s monthly funds transfer from TikTok’s Creator Fund portal .....	93

# CHAPTER 1

## INTRODUCTION

### Introduction

Since the platform launched in 2017, TikTok has steadily become one of the most popular social media networks, surpassing one billion mobile downloads globally in October 2020 and beating out Google as the most popular social media app in 2021 (Carman, 2020; Rosenblatt, 2021). U.S. popular media have investigated TikTok's algorithm to make sense of the platform's popularity, with news outlets and magazines from the *New Yorker*, *Wired*, and the *Wall Street Journal* offering theories and conducting investigations about how exactly the platform comes to "know" its users better than they know themselves. Pointing to the specific nature of TikTok, Jia Tolentino (2021) wrote in a profile of the platform for the *New Yorker*, "TikTok is a social network that has nothing to do with one's social network." While other social media platforms involve at least some component that emphasize friendship, professional, and other social connections, the cornerstone of the TikTok experience is the "For You" page, an endless scroll of user-recommended videos that are delivered to users based on their interactions with content on the platform.

While TikTok and its algorithm has been the subject of fascination and sometimes horror, the platform has also gained mainstream media attention for its consistent string of algorithmic and human content moderation controversies. A few examples of these problematic patterns of content moderation include the removal of content posted by minoritized creators in which they "stitch" video critiques to the videos of white users

violating the platform's Terms of Service using hate speech and targeted racist bullying and harassment, while the original videos remain visible on the platform. This example represents one of the direct ways platforms are less hospitable to marginalized creators, but in recent years, platform users have begun to identify more indirect forms of content moderation. Spokespersons for the platform have admitted to previously suppressing the content of marginalized creators to limit bullying on the platform (Botella, 2019), a solution that penalizes marginalized creators rather than grapples with the problems of hate speech and harassment on the platform in more meaningful ways. More recently, popular creators have also found evidence of the platform suppressing the visibility of political content, including content tagged and associated with Black Lives Matter and Black History Month (Hale, 2020; McCluskey, 2020), in addition to more targeted instances of content suppression for individual users, even as restricted creators retain other access to the platform in a view-only mode.

This study takes up the concern of content moderation on TikTok from the perspective of content creators from historically marginalized and underrepresented backgrounds. While frustrations with content moderation are not limited to marginalized and underrepresented platform users (Myers West), TikTok's string of suppression controversies evidence the platform's specific problems for creators who are marginalized along the intersections of race, gender, sexuality, disability, ethnicity, and class. Moreover, marginalized creators have been drawing attention to the algorithmic suppression of their content on TikTok since 2018, using their personal experiences to make sense of TikTok's opaque algorithmic environment and the disparity in content moderation experiences for marginalized creators. Drawing on digital cultural rhetorics

and critical race methodologies which center cultural narratives and embodied knowledge of technology, I investigate the user experience of TikTok’s content moderation from the perspective of minoritized TikTok users who are uniquely attuned to the platform’s evolving interface design because of their work as content creators.

Content moderation takes many forms on TikTok ranging from the standard removal of content deemed in violation of the platform’s Community Guidelines to sneakier forms like the removal of sounds and reduced algorithmic visibility. The table below lists the forms of content moderation on TikTok that I discuss in this dissertation: content removal (which can include the removal of user-generated sounds, as I discuss in chapter 3), reduced visibility, temporary ban (which include post, comment, and live-streaming bans), and permanent bans. TikTok justifies each of these types of content moderation as being crucial to the platform’s mission to “prioritize safety, diversity, inclusion, and authenticity” and to ensure that TikTok is “a safe environment” (“Community Guidelines”).

<b>Type of Content Moderation</b>	<b>Description</b>	<b>Example</b>
Content removal	Removal of content—including video, audio, livestream, images, comments, links, or other text—that violates the platform’s Community Guidelines, oftentimes through automatic removal systems.	A creator may have a video taken down for violating the platform’s policies around adult nudity or sexual activity.  The removal of a sound on a creator’s video, while the original video remains on TikTok.
Reduced discoverability / visibility	Suppression of content visibility on the For You page or in TikTok’s search features. There is no system notification from TikTok when content or a creator is rendered less discoverable on the platform.	A creator reports that they noticed drastic drops in content engagement, particularly around content

	This type of content moderation is also referred to on TikTok and other platforms as shadowbanning.	
Temporary ban	Account suspension that prohibits a user's ability to post, comment, or edit their profile for a period of time, usually 24-48 hours. According to TikTok, creators receive system notifications for these bans.	A creator is mass-reported in bad faith by right wing platform users, and then is flagged for a temporary ban.  A creator's account is set to view-only, and they are unable to post, comment, or live-stream for a period of time.
Permanent ban	Disabling of user access to an account. According to TikTok, creators receive system notifications for these bans.	Several of the creators interviewed for this study received permanent bans after experiencing a series of temporary bans.

**Table 1. Types of content moderation on TikTok**

This table is not a comprehensive list, but it explains the main types of content moderation discussed in this dissertation. In the chapters to come, I discuss creator experiences of these different types of moderation on the platform, prioritizing the perspective of marginalized creators to understand how moderation works on the platform. For example, while TikTok's policies indicate that creators are always able to appeals content removals and platform judgments that result in bans, each of the creators I spoke with discussed experiences that made platform appeals impossible. For example, creators shared that they had discovered that TikTok had removed the original sound uploaded with some of their videos without issuing a system notification. According to the creators, these sounds, as I discuss in chapter 3, were user-generated and not something that would be reasonably actionable by the platform on the grounds of intellectual property or copyright infringement, which are clearly outlined in the Terms of

Service and Community Guidelines. Likewise, creators report that they have no available recourse to challenge shadowbanning because there are no system notifications related to reduced discoverability. Indeed, TikTok representatives deny the use of shadowbanning as a moderation practice to enforce the Community Guidelines, while reduced discoverability is outlined in the Terms of Service as within the platform's rights.

In digital environments like TikTok in which public-facing documents like Community Guidelines don't always match up with behind-the-scenes practices, content moderation establishes platform identities for user-bases and corporate partners alike and is the foundation for user experiences of digital platforms. While content moderation is most often discussed, debated, and deliberated in popular discourse in terms of removal and bound up in messy conversations about free speech, this dissertation takes up digital content moderation as a complex infrastructure that mediates platform culture by screening, removing, filtering, and curating content for platform users. Through this definition, I understand content moderation as an assessment technology that amplifies some cultural values and concerns at the expense of others. I have chosen to study TikTok for this study because the centrality of the recommendation system in the user experience design makes it an important site to study how platform users, especially self-identified creators, make sense of content moderation. Indeed, the intrigue surrounding the platform's mysterious, uncanny recommendation algorithms draws attention to how platform moderation mediates content not simply through removal but through curation and sorting. As highlighted in many of the allegations of algorithmic discrimination against TikTok, content creators from historically marginalized and underrepresented backgrounds are not only being subjected to discriminatory content removal; rather, they

are facing various forms of discriminatory content suppression while the platform fails to meaningfully grapple with the ongoing presence of hate speech, harassment, and misinformation.

This dissertation approaches the problem of content moderation from the perspective of content creators from marginalized backgrounds because these are the platform users who have been most impacted by TikTok's pattern of discriminatory content moderation practices. Moreover, content creators have been at the center of these conversations in mainstream media, bringing attention to important examples of algorithmic oppression on one of the most popular social media platforms in the world currently. Specifically, the research questions guiding this dissertation are as follows:

- How do content creators make sense of their moderation experiences in the interface without clear, transparent platform guidelines?
- How does digital content moderation affect content creators from multiply marginalized communities on TikTok?
- How do content creators interpret their experiences of content moderation through their interface interactions?
- Finally, how do content creators respond to their experiences of content moderation on TikTok?

To answer these questions, I interviewed six marginalized content creators on TikTok and analyzed their stories with a focus on how their understandings of content moderation connected to features of TikTok's user experience design. Specifically, my study shows how content creators on TikTok engage in constant interface analysis, affectively reading their experiences with content moderation to make sense of the

platform's hidden, inconsistent, and oftentimes discriminatory patterns of moderation. Affectively, TikTok's interface encourages anxiety, fear, frustration, paranoia, and suspicion for minority content creators, and these feelings related to the platform's content moderation practices impact the content these creators publish to the platform. This study demonstrates that more explicit engagement with affect has various implications for technical communication and the rhetorical study of social media, as it facilitates more nuanced understandings of user experience and rhetorical agency.

### **Literature Review**

This dissertation takes as its focus the role of emotion in understanding digital content moderation on TikTok. One of the key arguments that I present in this dissertation is that TikTok creators affectively make sense of digital content moderation on the platform in the absence of wholly transparent policies and practices. My argument is critically influenced both by two strands of research that argue that digital content moderation fundamentally structures discourse and communication on social media (Gillespie, 2018; Klonick, 2019; Roberts, 2018) and that social media interfaces facilitates affect and emotion (Karatzogianni and Kuntsman, 2013; Kuntsman, 2009, 2021; Papacharissi, 2016; van Dijck, 2013). In content moderation scholarship, the discussion of emotion is often implicit rather than explicit.

Scholarship on content moderation often examines the ways in which digital platforms minimize the impacts of moderation on user experience to promote the values of free speech (Gillespie, Klonick, Lessig) and to obscure the working conditions of the human laborers behind the scenes (Roberts, 2019). Consequently, researchers in the fields of communication and human-computer interaction have examined user responses to



content moderation, often emphasizing the affective impact on users as a consequence of misunderstanding the role of content moderation on these platforms, feeling censored or silenced by platform governance (Myers West, 2018; Eslami, et al, 2019; Jhaver, et al, 2019). In “Censored, Suspended, Shadowbanned: User Interpretations of Content moderation on Social Media Platforms,” Sarah Myers West (2019) discussed survey data that revealed the affective experience of content moderation for everyday platform users on Twitter, Facebook, and Instagram . In addition to examining user perceptions about what or who triggered the moderation of their content, which revealed that most users suspected human intervention rather than algorithmic detection, Myers West identified that the most dominant affective responses to content moderation were frustration and confusion: “by and large, those who interact with content moderation systems report they were unable to learn what triggered the content removal or account suspension. Many of these users expressed a desire to understand why this happened so that they could avoid the experience in the future” (p. 4378). This finding supports other scholarship on the affective impact of obscure content moderation policies, while also importantly highlighting that the experience of content moderation triggers in users concerns about future encounters with moderation, content removal, or account suspensions.

Indeed, platform users engage obscured content moderation policies, and I aim to discuss the affective implication explicitly in this dissertation. In attending to the affective implications of content moderation, I hope to explore the digital cultural rhetorics of platform governance: how do marginalized creators understand platform values through their interactions with TikTok’s unwritten rules? How are inconsistencies in moderation interpreted against the platform’s stated goals for diversity, equity, and

inclusion, especially when marginalized platform users experiences these inconsistencies at greater rates? In what follows, I will begin with a discussion of content moderation as it relates to theories of platform oppression before discussing in greater detail the implications of affect theory on this work.

### *Content Moderation and Platform Oppression*

To date, scholars in digital rhetoric and writing studies have engaged the problems of digital content moderation on social media platforms more implicitly than explicitly by theorizing how digital interfaces enable, promote, and sometimes, prohibit certain types of practices and relations among users, platforms, and interfaces. Such interrogations of platform governance and digital interfaces resonate with scholarly conversations across disciplines about how social media platforms maintain and perpetuate systems of oppression (Amrute, 2019; Chun, 2016; Noble, 2018; Benjamin, 2019; Eubanks, 2018; Wachter-Boettcher, 2017). Pointing to the ways that race is coded online and into digital interfaces and infrastructures, Noble (2018) and Benjamin (2019) argue that manifestations of racism in technology should not be understood as glitches, or errors in code that produce isolated racist outcomes, but as evidence of deeper systemic racism. In pointing to the racial biases of technologies like search engines and platforms, these scholars denaturalize technologies as objective and neutral, revealing instead how racial ideologies are reinforced at all times through digital technologies, not merely during the events of glitches. Glitches in content moderation can include instances of algorithms flagging the use of racial slurs like the n-word used by Black people as hate speech, while similar uses by non-Black people are ignored. Indeed, digital rhetoricians have taken up questions of digital content moderation in numerous ways, including how

platforms regulate copyright and intellectual property (Reyman, 2010; Porter, 2018), how digital interfaces enable hate speech and harassment (Sparby, 2017; Brown and Hennis, 2020; London, et al., 2020; Trice, Potts, and Small, 2020), and how algorithms condition user experience online, with implications for privacy and the formation of cultural identities online (Beck, 2018; Popescu and Baruh, 2018). That said, digital content moderation as the foundational set of practices and policies that structure discourse on social media remains under-researched within digital and cultural rhetorics.

To study digital content moderation on TikTok, I draw from a content moderation scholarship that highlights the profound influence of platform governance on user experiences of social media and discourse online (Gillespie, 2018; Klonick, 2019; Goldman, 2019). Importantly, scholarly attention to content moderation has emphasized its important influence on the Internet as we know it. For example, Tarleton Gillespie defines content moderation as the “essential, constitutional, definitional” function of social media platforms, contrary to the public perception that these online services serve as a mere medium for unfiltered user-generated content (p. 21). This scholarship reveals the influence of dominant American ideals about free expression on the formation of the institutional identities of platforms, as well as their infrastructures. Importantly, scholars have pointed to the entanglement of U.S. law and market forces in the organization and moderation of the Internet. For example, legal scholar Lawrence Lessig (2006) connected the architectural code of the software and hardware of online spaces with the legal codes of civil rights, property laws, obscenity laws, and the intrinsically related economies. Yet content moderation scholars have shown that the governing forces of U.S. law and the increasingly powerful digital platforms are often at odds over concerns of free speech.

Scholars discuss this tension in relation to Section 230 of the U.S. telecommunication law, a small section of legal code which in 1996 dictated that intermediaries cannot be held accountable for the speech of users, is now the subject of intense debate about the future of the internet as we know it. Through examining regulatory laws and platform policies and practices, many scholars argue that platforms moderate principally to protect their corporate interests and to promote a profitable consumer experience, and in paying attention to the *why* of content moderation, or why platforms moderate as they do, Klonick argues that platforms moderate and curate primarily to conform to the “normative expectations of users” (p. 1630). Such a conclusion dialogues with digital rhetoricians who have examined content moderation and Section 230 and found that platforms, empowered by Section 230’s safe harbor, shift moderation responsibilities where they can to the user to maintain their commercial identities as spaces for democratic engagement and free expression, while leaving a great deal of abuse and harassment unchecked (Brown and Hennis, 2020; Trice, Potts, and Small, 2020).

More recently, internet researchers have begun to expand definitions of content moderation to recognize the influence of recommendation in platform governance (Are, 2021; Bucher, 2012; Cotter, 2019, 2021; Gillespie, et al., 2021; Petre, Duffy, and Hund, 2019). In the introduction to a review of content moderation scholarship and the debates around Internet regulation, Tarleton Gillespie and Patricia Aufderheide defined content moderation as “the detection of, assessment of, and interventions taken on content or behaviors deemed unacceptable by platforms or other information intermediaries, including the rules they impose, the human labour and technologies required, and the institutional mechanisms of adjudication, enforcement, and appeal that support it” (p. 2).

This comprehensive definition captures much of the sociotechnical scope of content moderation, although I question the notion that moderation in practice is solely about assessing content as acceptable or unacceptable against platform rules on most platforms that increasingly rely on recommendation algorithms, like TikTok. As Jing Zeng and D. Bondy Valdovinos Kaye theorized in their recent work, TikTok implements a type of content moderation that they term *visibility* moderation, which they define as “the process through which platforms manipulate the reach of user-generated content through algorithmic or regulatory means” (p. 79). Through this new conceptualization of platform moderation, which builds on important scholarship related to the increasing importance of visibility online (Cotter, 2019 and 2021), there is room to trouble the definition of content moderation set forth by Gillespie and Aufderheide. Why might some content be amplified algorithmically while others are suppressed? How might processes of visibility moderation contribute to systemic oppression on social media platforms like TikTok?

With more attention to the increasing influence of recommendation and visibility within the encompassing structure of digital content moderation, it follows that digital content moderation is a digital infrastructure of emotion and engagement, not merely an engine of content removal. In the next section, I review literature on the role of affect and emotion in social media to explore how content moderation has been already discussed as an affective experience.

### *Digital Labor, Platform Capitalism, and the Politics of Transparency*

At the heart of this dissertation and its findings are the experiences and stories of content creators on TikTok. Digital influencers and content creators have received a great deal of critical attention in recent years, with scholars examining the practices and

implications type of “micro-celebrity” on social media platforms like YouTube, Facebook, and Instagram (Senft, 2008; Marwick and Boyd, 2011; Marwick, 2018; Raun, 2018; Usher, 2020). Scholarship on microcelebrity has examined personal branding and the commodification of authenticity, intimacy, and vulnerability. For example, Brooke Erin Duffy (2017) examines the lives and labor of fashion bloggers, beauty vloggers, and designers trying to curate their personal lives for algorithmic environments. Other scholars have discussed how content creators contribute to digital platform governance by influencing broader conversations about content creation, digital labor, and moderation (Cunningham and Craig).

The opacity of content moderation, in other words, serves platform interests as they reduce costs through the exploitation of human laborers around the world employed in the difficult work of commercial content moderation (Roberts) and conceal the code that creates reality as we know it online. To this point, Tressie McMillan Cottom, in her recent work “Where Platform Capitalism and Racial Capitalism Meet: The Sociology of Race and Racism in the Digital Society,” argues that opacity and black boxed technology aid racial capitalism today, explaining that:

Theoretically, obfuscation operates much like willful whiteness that can always claim ignorance of statistical discrimination, for example, because it owns the means of discovery. Obfuscation does not mean that someone or some organization does not know these data. It means that the information is difficult to access and often couched in needlessly complex technical jargon or process. [...] Obfuscation becomes a technique of privatization through two processes. One, it extracts data that would have previously been public, publicly available or legally

discoverable. Two, it expands obfuscation as a logic, even in organizations or institutions that have a public mandate. When full privatization is not possible, obfuscation makes it inaccessible in practice. Information is the vessel for social actions and social facts. If information is inaccessible, the objects of everyday life are too. (p. 443)

In connecting obfuscation to “willful whiteness” and racial capitalism, McMillan Cottom demonstrates the ways in which glitches in content moderation cannot so simply be excused as coding errors. In protecting data and information that could offer explanations and contexts for discriminatory patterns in digital technologies, media corporations are able to control the narrative and employ tactics of what McMillan Cottom identifies as “predatory inclusion,” which she defines as “the logic, organization, and technique of including marginalized consumer-citizens into ostensibly democratizing mobility schemes on extractive terms” (p. 443). As an example, TikTok responded to criticisms of its content moderation that alleged that it suppressed Black Lives Matter content in June 2020 by apologizing for the harm caused by an unfortunate “glitch” and pointing to the success of Black creators and content related to Black Lives Matter as evidence of its commitment to diversity and inclusion. While the apology could be generously read as the platform’s attempt to be accountable to its Black creators, this line of response follows a pattern of social media platforms using the language of glitch, evading meaningful conversations about algorithmic oppression by focusing on individual instances of technological failure (Benjamin, 2019).

At the same time that it is important to recognize the harm that is done in opaque platforms, scholars such as Crawford and Ananny have shown that transparency, too, is

not an unproblematic ideal, particularly when deployed by digital platforms. As I will discuss later in this dissertation, the lack of transparency on TikTok regarding its moderation practices harms content creators, particularly as it impacts creators' monetization efforts; therefore, while this dissertation discusses how creators are often harmed by black-boxed content moderation, I complicate the lack transparency along these lines, as well.

### *Affect, Emotion and Algorithmic Imaginaries*

Finally, this dissertation also draws on affect theory (Ahmed, 2014; Ngai, 2004; Berlant, 2011; Stewart, 2007) to understand the role of emotion in understanding content moderation on social media. Such work has been critical to understanding how feelings move online. In recent years, social media researchers have taken up affect scholarship to make sense of how social media facilitates feelings (Karatzogianni and Kuntsman, 2013; Kuntsman, 2009, 2021; Papacharissi, 2016; van Dijck, 2013). Researchers studying content creators on various social media platforms have also explored the affective dimensions of digital labor (Lehto, 2022; Kopf, 2020; Raun, 2018).

Attention to emotion has also yielded important insights into how platform users rely on emotion to navigate social media platforms and algorithmic environments. Recent scholarship by feminist scholars and digital cultural rhetoricians illustrates how platform users navigate social media platforms like TikTok affectively, as a result of content moderation systems which permit gendered and racialized hate and harassment to go unchecked (Gruwell, 2017; Hutchinson, 2018; Sparby, 2020; Reyman and Sparby, 2020). Work in the collection *Digital Ethics: Rhetoric and Responsibility in Online Aggression* illustrated how content moderation infrastructures enables a culture of digital aggression,



characterized by the affective dispositions of widespread online abuse, conflict, and harassment, and the affective impacts on users minoritized by race, gender, sexuality, ethnicity, disability, and other identity markers (Dieterle, Edwards, and Martin, 2020; Gelms 2017; Gruwell, 2020). In the introduction to *Digital Ethics*, Sparby and Reyman (2020) highlight the infrastructural importance of moderation and outline an ecology of reactive and proactive rhetorical approaches to countering digital aggression via the following strategies:

1. Platform design that offers moderation tools and clear policies;
2. community leaders articulating and following rules and norms;
3. Moderators enforcing rules and norms while modeling behavior; and finally,
4. Community members reinforcing norms and rules and also teaching new members how to behave (Reyman and Sparby, “Introduction,” 2020, 7–8).

These strategies center moderation in terms of design, community practice, and transparent policies, highlighting how content moderation operates as a constellation of platform policies, practices, and user behaviors. Relatedly, in “Reading Mean Comments to Subvert Gendered Hate on YouTube: Toward a Spectrum of Digital Aggression Response,” Sparby (2021) examined how five popular women on YouTube navigate and resist gender-, race-, sexuality-, and disability-based aggression from commenters on a platform. Sparby’s work demonstrated how these women’s identities informed their responses to the hate they received on YouTube and how the failure of platforms such as YouTube to moderate digital aggression impacts the realities of those spaces, as the deluge of mean comments often contribute to the exclusion and silencing of marginalized

creators. The affective and material impacts of digital aggression on marginalized content creators evidences the ways in which content moderation upholds systems of oppression and reinscribes privilege along the lines of ethnicity, race, gender, sexuality, and disability.

This work builds on a theory of sneaky rhetorics developed by Jennifer Sano-Franchini, Matt Homer, and me. Sneaky rhetorics, or the sensibility of mistrust that is characteristic to social media environments currently, contributes to scholarly conversations about how social media affectively facilitates the circulation of disinformation (Bakir and McStay, 2021; Boler and Davis, 2021; Kuntsman, 2009, 2021; Sano-Franchini, 2018; Stark, 2021). In this piece, we brought together Ahmed's (2014) concept of stickiness and Chun's (2016) work on habitual new media to argue that Facebook primes its users to be susceptible to mis- and disinformation. Through a study of our personal engagement with Facebook, we found that sneaky rhetorics shaped our social media habits and contributed to our increased consumption of online media. In this dissertation, the sneaky rhetorics of TikTok are explored through creator interpretations of platform policies and processes, as well as in creator understandings of other platform users. In this way, I bring together sneaky rhetorics with Taina Bucher's concept of the algorithmic imaginary. Bucher draws on Kathleen Stewart's work on ordinary affects, which explores the affective dimensions of everyday life, to theorize her concept of the algorithmic imaginary, or a sense of what the algorithm does and how it works. Bucher explains in her work on Facebook that platform users rely on their affective interactions with algorithmic recommendation systems to make sense of how they work. In this dissertation, I find that the algorithmic imaginary of TikTok as put forth by creators is

animated by sneaky rhetorics of TikTok as a platform that exploits and deceives its creators. In the chapters that follow, I discuss the implications of creator understandings of TikTok as sneaky affectively influence creator habits of content production, networking, and engagement with TikTok.

## **Chapter Overviews**

The following chapters take up the issue of content moderation from the perspective of marginalized content creators in several ways. Specifically, this dissertation unfolds to show how marginalized content creators sense inconsistent uneven moderation practices and the affective effects as those experiences unfold over time, as well as how creators respond to these content moderation conditions in their creative labor on the platform, in private networks of content creators, and through public challenges on TikTok and other digital platforms to the TikTok's practices and policies.

In **Chapter 2, "Methods and Methodologies,"** I discuss how the methodologies and methods for this study respond to these non-transparent policies, paying specific attention to emotion and embodied experience as important means of understanding algorithmic environments.

**Chapter 3, "Sensing Censorship: Examining Content Creator Perceptions of Platform Exclusion in TikTok's Interface,"** examines content creator stories to argue that the lack of transparency around TikTok's algorithmic curation and moderation 1) contributes to content creators feeling alienated, exploited, frustrated, and unwelcome on the platform; 2) influences content creators to adapt their content to avoid moderation, oftentimes by self-censoring themselves and aspects of their marginalized identities; and

3) that the accumulation of content moderation micro-interactions diminishes the ability of marginalized content creators to trust content moderation processes.

**Chapter 4, “Cumulative UX and Creator Time,”** draws on research on the temporality of user experience design to argue that disruptions in platform experience affectively impact marginalized and underrepresented content creators in the following ways: 1) creator perceptions that content moderation interactions increase in frequency and intensity over time; 2) creator perceptions that content creators and their creative labor continue to be exploited even when they are prohibited from interacting with the platform; 3) disruptions in platform access limit the ability of content creators to monetize their labor on a platform that requires constant engagement to be successful; and 4) the lack of open communication and transparency in the appeals process worsened creator experiences of disruptions, even when creators had their access restored. This chapter also builds on the findings in chapter 3 to demonstrate that affective responses to content moderation are intrinsically related to the way that time is cultivated and experienced on the platform.

**Chapter 5, “Collective Complaint and Connection as Responses to Platform Silencing and Suppression,”** explores how creators from historically marginalized and underrepresented communities speak back to and against TikTok’s discriminatory platform moderation practices and policies. Through an analysis of creator stories, I argue that TikTok’s opaque content moderation practices contribute to an affective platform environment in which creators are compelled to speak out and across creator networks about such gaps in experience and platform policy. Drawing on Sara Ahmed’s (2021) work on complaint as a feminist method of institutional intervention, this chapter

demonstrates how creators as digital laborers connect and come together to challenge, navigate, and survive the conditions described in chapters 3 and 4.

Finally, **Chapter 6, “Conclusions And Implications For Researching Algorithmic Transparency In Writing Studies,”** discusses how my findings about the digital cultural rhetorics of digital content moderation contributes to conversations in writing studies research. Additionally, I discuss what media companies like TikTok and user experience designers can learn from content creators.

## **CHAPTER 2**

# **METHODOLOGIES AND METHODS: CULTURAL DIGITAL RHETORICS, CREATOR STORIES, & CRITICAL INTERFACE ANALYSIS**

### **Introduction**

This chapter overviews the methodologies and methods central to this dissertation's investigation of how marginalized content creators on TikTok make sense of content moderation on the platform. The purpose this research is to consider three research questions:

- How did content creators understand their experiences with content moderation on TikTok in the absence of transparent platform communication?
- How do creators justify their rationalizations of these experiences?
- How do creator understandings of content moderation impact their user experiences and uses of TikTok?

These questions emerged from my personal use of TikTok through which I encountered videos related to creators amplifying what they considered to be unfair experiences of content moderation on the platform. In these videos, creators explained how they understood their experiences as unfair and discriminatory, using screen-share recordings to explain how platform disparity manifested in their interface experiences and sharing videos from other platform users of similar types of moderation problems. I became interested in understanding how creators make sense of content moderation on TikTok. To do so, I draw on work in digital and cultural rhetorics to examine how content

moderation and digital interfaces mediate culture on TikTok and how creators navigate and negotiate TikTok's content moderation by drawing on their embodied experiences.

First, I will situate the study with a discussion of my researcher positionality as a casual TikTok user who has watched several content moderation controversies unfold and worsen on the platform since late 2019. From there, I expand on this dissertation's digital cultural rhetorics methodologies before discussing the methods. Specifically, I discuss how I accessed creator stories through interviews, detailing the IRB approval and participant recruitment processes (Appendix A and B). I also detail the decision to supplement creator stories gathered through interviews with creator testimonials shared to Change.org in the form of petitions. Then, I discuss the coding schema I used to analyze these creator stories and online testimonials, as well as the importance of critical interface analysis to this study. Finally, I reflect on the limitations of this research and future directions for research into content moderation on digital platforms.

In discussing the methods and methodologies used in this dissertation, it is important to take a moment to discuss my experience with the platform and my interest in the topic of content moderation more generally. I downloaded Musical.ly early in the summer of 2018 upon a friend's recommendation. A few weeks later, I opened up the app and began scrolling through the For You page, only realizing a few minutes later that the platform had fully rebranded as TikTok (Kundu, 2018). I do not recall many of the changes made and had very little awareness that the platform had merged as the result of a buy-out from another company. Over the next few years, however, I became familiar with the platform as a consumer, keeping up with viral trends and platform conversations related to cultural appropriation, content moderation, and creator labor. In very subtle

ways, I noticed that everyday creators and users develop platform habits to avoid accidental and targeted detection, particularly around language related to marginalized identities.

By the beginning of the COVID-19 lockdown in March 2020, many of the creators I personally followed were speaking about frequent, repetitive negative content moderation interactions on TikTok on a near daily basis for speaking about police brutality and the nature of hate speech and harassment on TikTok. These experiences as a platform user informed my interest in looking at content moderation experiences on TikTok, a topic I have been interested in as a social media user and researcher for a number of years. I see the context of the pandemic as being important to this study because the restrictions to life offline heightened the need for connection in online spaces, such as those that are cultivated by creators on TikTok. While I do not see myself as a content creator, I developed a better awareness of the role that content creators play in my everyday entertainment and education as a result of the pandemic and by witnessing marginalized content creators, particularly Black creators, discuss their erasure and suppression during the early days of the pandemic. Moreover, a year earlier during the summer of 2019, many of the Black creators I followed or encountered in my For You page described their experiences of what TikTok called a “glitch” that suppressed Black Lives Matter content during the summer of 2020 when BLM protests took place across the U.S and globally. These recent events, alongside the platform’s other relevant history of content suppression, inform the stories shared with me, as content creators connected their experiences to these broadly discussed events as they unfolded on the platform. For example, during the summer of 2021, TikTok and its



content moderation were the subject of national conversations, including the one initiated by a creator named Ziggi Tyler who shared a video that went viral on Twitter which showed TikTok's Creator Portal platform prohibiting pro-Black language while allowing explicitly white supremacist language. Only a few months earlier in April, platform users including activist Wagatwe Wanjuki noticed that the hashtag #AbolishThePolice was no longer searchable or usable in video captions for a little over a week and a half (Cohen 2021). While the platform apologized for the glitch, many content creators found the timing of this error in code suspicious as it occurred at the same time when conversations about police brutality in the United States were once again reignited by the ongoing trial of Derek Chauvin, the police officer who killed George Floyd, and the tragic sequence of police killings of Duante Wright (20), Adam Toledo (13), and Ma'Khia Bryant (16). Months later, Black content creators launched the Black TikTok Strike, once more reigniting conversations about cultural appropriation and the exploitation and erasure of Black creators on the platform (Pruitt-Young, 2021).

I discuss some of these controversies in the chapters that follow both as they represent examples of the platform's spotty reputation around equitable content moderation and as content creators reference them in relation to their personal experiences on the platform. In what follows, I discuss the methods of narrative-based inquiry, which focuses on experience, affect, and emotion, and critical interface analysis.

## **Methodologies**

Methodologically, this dissertation is grounded in a digital cultural rhetorics perspective that recognizes the relationship between digital technologies and infrastructures, institutions, and bodies. Angela M. Haas writes that "digital cultural

rhetorics highlight the relationships between the rhetorical situation and the actors (animal, technological, non-animal environmental, and hybridized), institutions, subjectivity, power networks, agency, and action therein” (p. 413). She goes on to say that digital cultural rhetorics also challenge perceptions of technologies as neutral, objective, democratizing, [and] emancipatory. I see digital cultural rhetorics and critical race theory to be complementary in their disciplinary investments to challenging dominant narratives of neutrality, including in the articulation of their research methods that emphasize the need for experiential knowledge (Martinez).

This dissertation follows recent calls for more critical race scholarship that attends to digital media and computer-based technology (Benjamin, 2019; Chun, 2013; Noble, 2018). Aja Y. Martinez defines a critical race methodology as one that challenges dominant narratives of white supremacy and white privilege and “rejects notions of ‘neutral’ research or ‘objective’ research,” emphasizing instead the importance of the epistemologies of people of color (p. 3). Safiya Noble has articulated the importance of critical race theory for challenging beliefs about neutrality and objectivity in technology and, ultimately, “moving toward undoing racist classification and knowledge-management practices” (p. 138). A critical race study of content moderation on social media platforms like TikTok, thus, focuses on the perspectives and understandings of people of color and other marginalized communities.

This study primarily focuses on the stories and experiences shared with me by six content creators from marginalized and underrepresented communities. In addition to highlighting the experiences of content creators from marginalized and underrepresented backgrounds, the decision to focus on content creators as a specific category of platform

user emerged from a desire to better understand the material implications of content moderation as it impacts digital laborers. Moreover, this dissertation works from André Brock Jr.'s assertion that "the internet should be understood as an enactment of whiteness through the interpretive flexibility of whiteness as information" (p. 6).

### *Creator Stories and Critical Interface Analysis*

Recent attention to content moderation oftentimes highlights the lack of transparency and the difficulty of studying opaque platforms. To circumvent the problems of transparency, Internet communication scholars and technical and professional communication scholars forward methods such as human-centered inquiry (Jones, 2016; Green, 2021) and auto/ethnography (Haas, 2020; Sano-Franchini, 2020) to study how digital platforms mediate knowledge, identity, and culture and facilitate power, privilege, and oppression. Likewise, the attention to the embodied interface experiences of creators is in line with the ethnographic methods used by social media researchers to study platform user perspectives of algorithmic environments and shadowbanning, also known as content suppression (Are, 2021; Cotter, 2019 and 2021). While platform transparency would perhaps lend some clarity to policies in practice, the heavy emphasis on transparency in the conversation around content moderation often gives the impression that without platform transparency there is no way to know more about content moderation in practice. For example, Taina Bucher (2018) discusses how the transparency fallacy has contributed to scholars treating algorithms as black boxes, or a system whose technical innerworkings are unknowable. One approach that Bucher offers to know algorithms is to engage them phenomenologically from the user perspective. Bucher's belief in a phenomenological approach to algorithms dialogues

well with digital cultural rhetorics frameworks that prioritize the importance of cultural narratives and stories of technology (Haas 2018). Indeed, bringing into focus the lived experiences of content moderation, which includes algorithmic processes, “is meant to highlight the productive force of personal forms of knowledge, knowledge that is gained through experience and practical engagement with one’s lived-in environments” (Bucher). Throughout her work on the programmed sociality on Facebook, Bucher employs autoethnography and interviews to *know* how algorithms condition social and political life.

Similarly, Jennifer Sano-Franchini employs a critical interface analysis of her personal experiences with the platform to theorize the relationship between UX/interface design and political engagement. Through her findings and discussion, Sano-Franchini comes to know more about her algorithmically-curated Facebook News Feed without knowledge of its inner workings by paying attention to the types of relations and emotional responses fostered by the platform’s interface. Ultimately, Sano-Franchini argues that “it is not just the content posted to the site but also Facebook’s interface design—in shaping that content—that contributes to a deepened polarization around political issues.” While Sano-Franchini’s critical interface analysis deals primarily with user experiences that usually go undetected by people (even as they radically reconfigure experiences of information and the world), this approach highlights how much can be learned by analyzing the interface itself.

In this dissertation, my use of critical interface analysis may most often be characterized as an interpretive lens to understand creator stories and testimonies about how they see content moderation manifesting in the interface. However, I do discuss

TikTok's UI in other cases to present my own interpretation of how the platform affectively facilitates creator understandings of content moderation. Given my positionality and my study's aims to center creator perspectives of content moderation, this use of critical interface analysis feels more appropriate.

*Centering Difference, Learning from the Margins*

This dissertation's focus on the experiences of marginalized content creators follows calls to center difference in technical and professional communication (Moore, 2017; Shelton, 2019; Walton, Jones, and Moore, 2019). Scholars in technical communication and user experience design have highlighted the need to prioritize the needs of specific people or users when designing technologies to be more socially just (Green, 2021; Rose, et al., 2018; Rose, Björling, Kim, and Alvarez, 2018; Walton, Jones, and Moore, 2019). To study such perceptions and experiences, I turned to narrative-based methods, following the work of digital cultural rhetorics and technical and professional communication (TPC) scholars who have pointed to the use of story and narrative to study power, privilege, and oppression in digital design contexts. For example, I followed the example set out by Jones (2016) in her important work on narrative inquiry in human-centered design as a means to promote social justice in human-centered design. The use of narrative inquiry, too, emerges out of Black Feminism, in addition to having important roots in critical race theory in the form of counterstory as a method to disrupt hegemony and oppression through the sharing of stories that center the lived experiences and knowledge of marginalized people (Martinez, 2020).

Narrative inquiry also supports work which aims to better understand the role of affect and emotion in interaction design. As Green (2021) noted in his work, citing the

work of Pernice (2018) on interviews in user experience research: “interviews tend to measure people’s perceptions, beliefs, and values about technologies, rather than measuring how people actually interact with a technology” (p. 337). On the other hand, narrative-based inquiry shifts attention to how users interact with technologies. Drawing on the work of Clandinin and Connolly (2000), Jones’ work forwards an approach to narrative-driven interview questions to the five dimensions (Clandinin and Connolly, 2000, p. 50-52): inward, outward, backward and forward, and place. Each of these dimensions helps researchers to understand the various ways participants interact with technology, ranging from internal and external factors to factors related to time (backward and forward) and place. As an example, Jones describes how attention to the inward dimension could involve questions that ask a participant to describe “their feelings, emotions, and wishes when using a particular technology” (p. 487). Similarly, an attention to the outward dimension can be addressed with questions concerning “external issues like physical responses and environmental concerns” (p. 487). Through myriad attention to the various factors impacting participant engagement with technology, this methodology aids in accessing participant stories that yield important information for promoting more socially-just design.

### **Methods: Examining Creator Interviews and Testimonials**

In addition to highlighting the experiences of marginalized and underrepresented content creators, this study focuses specifically on the experience of content creation as a professional endeavor. While creating videos and content on TikTok might be a hobby for many of the platform’s users, especially because of the COVID-19 pandemic, content creators seeking to monetize their use of TikTok demonstrate a platform savviness and an

investment in the platform's shifting policies and practices that make their insights extremely useful and important to this study. Additionally, given how creator-focused TikTok's experience is, I wanted to center the perspectives and insights of individuals whose creativity and labor are instrumental to the platform's appeal.

As mentioned earlier, I use TikTok for my personal entertainment, so I wanted to respect the intellectual labor of creators who I first learned about these issues from by reaching out to them for interviews. After reaching out to creators I personally followed, I recruited creators who had been quoted in media coverage about TikTok and content suppression of marginalized and underrepresented creators. At the onset of this project, I initially wanted to prioritize the recruitment of Black content creators on TikTok, especially Black women and queer individuals because these creators in particular appeared to be spearheading the conversation on TikTok about racialized content moderation, at least on my For You page. I reached out to participants via email and by private message on TikTok, and while I will discuss the limitations with my recruitment approach later in this chapter, I ultimately had to expand my recruitment criteria to be more inclusive of a wider range of backgrounds. Each of these individuals elaborated in greater detail on the other aspects of their identity in relation to their platforms and their experiences with content moderation and suppression on TikTok. Creators who agreed to participate in this study also shared recruitment materials with friends they believed would be interested in sharing their stories with me.

### *Interview Questions*

In line with the goals of narrative-based inquiry methods, I designed semi-structured interview questions to guide my conversations with content creators, which I have included below:

- How would you describe TikTok's content moderation practices?
- Have you experienced content moderation on TikTok? What forms of content moderation have you experienced?
- How do you identify when TikTok is suppressing your content or content created by people that you follow? When you do discover that TikTok is suppressing content, how do you respond?
- How does the way TikTok moderates and curates content impact your content creation?
- How does the way TikTok moderates and curates content impact the ways you interact with the content of other creators?
- Have you spoken out against TikTok's moderation practices before? What was your goal in speaking out?
- How does the content moderation you've experienced and witnessed inform your understanding of your place on the platform, as well as the TikTok's relationship to creators on its platform?

These questions were designed to tap into creator feelings and perspectives about content moderation and how they made sense of their experiences. Above all, I was interested in hearing participant stories of content moderation and suppression on the story to hear



contextualized accounts of how creator interactions in the interface and technical approaches to making sense of and navigating content moderation on TikTok.

### *Participants*

In this section, I discuss the interview recruitment process for this study, before explaining the rationale for the interview protocol, interview transcriptions and analysis, and my follow-up analysis of eight Change.org petitions created by marginalized TikTok content creators. Then, I briefly introduce the six interview participants.

For my interviews, I first recruited popular creators that I personally followed who had shared their stories of suppression and content removals. I recruited creators who identified as marginalized by race, gender, sexuality, ethnicity, or class; who had spoken about their experiences with content moderation on TikTok or other social media platforms, had over ten thousand followers on the platform, made some or most of their income through content creation on TikTok. I used Zoom to conduct six interviews with TikTok content creators who identified as marginalized, from June to September of 2021. I reached out to forty content creators before ultimately interviewing six content creators from various marginalized backgrounds, each of whom met the following recruitment criteria:

- The creator had used their TikTok account to speak about their experiences or the experiences of other creators with content moderation on the platform.
- The creator had 5000 followers or more and regularly created content with the intent to monetize.
- The creator posted an email to be contacted for collaboration or press.
- The creator was eighteen years old or older.

- The creator had to be based in the United States.

To briefly explain the rationale for these criteria, I wanted to speak with creators who were active on the platform, had wide platform reach, and likely identified with content creation as a profession or side hustle. The requirements that the creators were older than eighteen years old and based in the United States serve both a practical ethical purpose to protect minors but also identify participants who meet most of the eligibility requirements for the Creator Fund, TikTok's monetization program. Finally, I wanted to identify creators who demonstrated an interest in speaking about their experiences on TikTok not only in videos shared to their account but through more professional means. I will discuss the limitations of recruiting creators via email later in this chapter.

All in all, I met with six creators on Zoom for this study between the months of June and August of 2021. I recorded all of these interviews on Zoom with the permission of my participants and used the transcription service Otter.ai for the initial transcription. Within Otter.ai, I cleaned up the transcriptions before I coded my data. All but one of the participants are anonymized in this study. This named creator, Zev Burton, wished to be identified because he has been very public in his campaign to hold TikTok accountable to the marginalized individuals who use its services. I use pseudonyms for the rest of the creators, and I include a brief introduction to them below in order of when I interviewed them:

- **Natalie** is a Black woman based in New York City with a decade's worth of experience creating content online, first on YouTube. Natalie first started using the platform to participate in viral dance challenges at the beginning of the COVID-19 pandemic in March of 2020. Natalie spoke about her

experiences on TikTok and on YouTube, and highlighted problems of the cultural appropriation of Black creativity and labor on TikTok during our interview.

- **Mae** is a Black, queer woman who downloaded TikTok in March of 2020 to pass the time while on bedrest, creating videos about her interests in historical costuming, parenting, and witchcraft. Although Mae didn't originally see herself as a content creator, her first account started steadily growing until it was ultimately deleted for having too many content violations. At the time of our interview, Mae had a little over 150,000 followers.
- **Zev Burton (@zevulous)** is a college student at Georgetown University studying international politics and data science and analytics. Zev first joined TikTok to promote his comedy book at the recommendation of his younger sister and soon after became interested in the problem of content moderation on TikTok after having videos removed which discussed the January 6<sup>th</sup> insurrection at the United States Capitol. At the time of our interview, Zev had around 100,000 followers.
- **Eric** is a Black man in his early twenties who joined TikTok in February of 2020. He first started creating videos giving dating advice for straight men and women after his friends recommended he give the platform a try. Eric's account quickly started to grow after he first started posting videos, and his first account was deleted after a string of content violations following a post he made about the January 6<sup>th</sup> insurrection at the United States Capitol. When

we met for our interview, Eric had around 300,000 followers on his second account and was a participant in TikTok's Creator Fund.

- **Stella** is a white queer, nonbinary, and disabled comedian who uses they/she pronouns. Drawing on their background in acting and musical theater, they started using TikTok in November of 2019 to create sketch comedy videos. After their first video went viral in January of 2020, Stella became more invested in content creation to support themselves financially when at the time their epilepsy made other career options less viable. During the pandemic, content creation became their sole source of income. At the time of our interview, Stella had over two million followers on TikTok.
- **David** is a man of color and comedian / content creator who has made Internet videos since he was twelve for YouTube, Vine, and now TikTok, Instagram Reels, and YouTube Shorts. He joined TikTok during the summer of 2019 and quickly reached around 200,000 followers before stagnating until around December of 2020 when his account started steadily growing again. At the time of our interview, David had over three million followers on TikTok and was experiencing cyclical platform bans that were causing him a great deal of stress. David also created a TikTok group chat, known as the TikTok Ban Victims by its members including Zev and Stella, that I discuss later in this dissertation.

While these six creators have a great deal in common related to their experiences of TikTok's content moderation, I want to recognize the specificity of each of their positionalities and embodied experiences. As a result, I am not extrapolating from their

discussion of their experiences broader claims about the specific marginalized and underrepresented communities to which they belong on TikTok. Rather, in my discussion of the collective stories of these individual creators, I recognize important differences in their embodiments and experiences, as I consider how content moderation mediates identity on TikTok in important ways.

### *Creator Testimonies in Change.org Petitions*

During my interviews, I found that several of my participants either published or were associated with petitions calling for changes to TikTok's content moderation published to Change.org. Through analyzing interview data, I learned that speaking out about TikTok content moderation on other digital platforms, like Change.org, represented an important way that creators responded to their experiences on TikTok. As such, I turned to Change.org to supplement interview data. Through my analysis of these petitions, I was able to learn more about content moderation patterns that creators shared with me during our interviews.

All in all, I analyzed eight petitions published to Change.org to supplement the data collected through interviews with creators. These petitions each focused on the ways in which TikTok's content moderation fails its marginalized users. I selected petitions on Change.org that connected the problems of content moderation to the suppression or silencing of marginalized voices on TikTok. I searched the platform to identify the most popular petitions, determining popularity by the number of signatures and only choosing petitions with 15 thousand signatures or more.

## **Analytical Framework**

In my analysis of interviews and testimonials, I coded specifically for three things: (1) interpretation of content moderation, (2) evidence for interpretation, and (3) experiential factors for interpretation, described further below.

1. **interpretation of content moderation:** In other words, I highlighted how content creators interpreted their content moderation experiences, in terms of the platform's motivations for flagging or removal against user understandings of the platform's policies.
2. **evidence for interpretation:** That is, I identified the types of evidence that content creators used to justify their interpretations.
3. **experiential factors for interpretation:** How past interface experiences informed the ways in which content creators engaged and navigated TikTok.

This coding schema was developed during the preliminary stages of coding, as I reviewed participant responses to my questions about their content moderation experiences. As I read through their stories of moderation and suppression on TikTok, I was especially interested and moved by how difficult these experiences had been for each of the creators interviewed for this study. In particular, I was interested in how these emotions seemed tied to TikTok's lack of transparency around the decisions made related to these creators' content. These emotional responses seemed to translate into the creators trying to make sense of their experiences by interpreting the platform's motivations, trying to identify the sources of moderation in the larger context of TikTok and in their content. As a result, I became very interested in how creators read their interface experiences as indicative of the platform's values and unwritten Community Guidelines, which disadvantaged

marginalized creators, and so I coded for interpretations, evidence used to justify these interpretations, and the impacts of experience on how these creators continued to engage and navigate the platform.

I coded the data from the position of a white woman and casual TikTok user who is familiar with TikTok's interface and the conversations about content suppression on the platform. I analyzed the interviews first, developing my coding schema through recursive readings of the stories shared with me by these six content creators. Then, I use these same coding schema on the testimonials shared on Change.org. My interface analysis was also informed by my findings from the creator interviews.

Overall, I found that these content creators become sensitive to platform patterns through both their lived experiences in the platform and careful analysis of TikTok's interaction design. In each chapter of my dissertation, I discuss the different ways that marginalized and underrepresented content creators detect moderation patterns and adapt strategic platform habits to mitigate uneven, inconsistent content moderation practices, including code-switching, networking with other content creators to amplify suppressed content and learn new strategies, and calling out the platform for discriminatory practices.

### **Critical Interface Analysis of TikTok**

Alongside my analysis of creator interviews and Change.org petitions, my analysis is informed by my critical reading of TikTok's user interface. To perform a critical analysis of the discourse surrounding content moderation and platform oppression on TikTok, my critical interface analysis also pays attention to all of the elements that compose a TikTok video (visual, aural, textual). In bringing together narrative-based inquiry and critical interface analysis as methods for this project, I am drawing

inspiration from the methodological approaches of Jennifer Sano-Franchini (2015) and André Brock Jr. (2020). As necessary, I also examine these videos in the context of the interface by paying attention to features like comments, hashtags, view counts, likes, sharing functions (known on this platform as stitches), among others. This interface analysis is especially critical in this project because the creators who are calling out shadowbanning practices and content moderation are conducting this type of analysis to make sense of their place and position within the context of the platform.

Before I continue, I want to take a moment to describe TikTok's interface. The most prominent feature of the home page is the video, which begins to play as soon as the application is launched. At the top of the screen, the user can choose between two experiences of the infinite scroll feature to toggle from the For You page to the Following page, which allows users to only view content uploaded by creators that the user follows. Along the right side of the screen, there is a vertical bar from which users have the ability to interact with the video and its creator. From top to bottom, this vertical bar includes the following white buttons:

- a button in the form of a circular profile picture that leads to the profile of the video creator
- a heart button to like the video,
- a comment button that looks like a speech bubble that raises the comment section of the video,
- an arrow button that offers the users multiple options to share the content across platforms,
- and a button that directs users to the sound used in the video.



Below each of the buttons, there is information about how many times the video has been liked, commented on, and shared in white text. Then, the username of the video's creator and caption of the video, including hashtags, is in the bottom left corner of the video.

Finally, at the bottom of the screen, the horizontal black navigation bar divides the rest of the platform into the following pages from left to right:

- **Home**, where users start the app to view the For You page and their Following page
- **Discover**, where users can search the platform or view content by exploring the trending hashtags and sounds
- **Create**, where users can create or upload short videos and edit videos to include sounds or effects from TikTok's library
- **Inbox**, where users can view activity notifications in order of newest to oldest related to creators they follow, profile views on their personal account, and receive and send messages to other users through the platform; and
- **Profile**, which takes the user to their account page to view their video feed and liked videos, as well as edit their profile information.

This navigation bar appearing at the bottom of the screen is consistent with other mobile social media apps, including TikTok, Instagram, and Facebook.

Like other social media platforms, TikTok delivers content algorithmically and so videos do not appear in chronological sequence in the infinite scroll. Interestingly, while viewing videos from the For You page, there is no indication of when the video was published. More information about when an individual video was published can be found

by viewing the link directly on the creator's profile. This information is presented next to the username and sometimes these times are represented as dates while the publishing date might also be represented as "Days Past" depending on the recency of the video when the user looks at it. In the comments section, comments are dated but not presented in chronological order. On the following page, videos still appear in an algorithmic ordering but they are dated to reflect how long ago something was posted by the creator.

The Discover page, like the For You page, is an infinite scroll page with a white background. The search bar is at the top of the screen next to a QR scan button, but the majority of the page is a list of trending sounds, hashtags, and video effects and previews top videos for each of those categories. These trends are not listed in order of most popular or most most used, although the number of times each of these trends has been used in videos on the platform is indicated next to each of the trends.

## **Limitations**

This dissertation is grounded primarily in the experiences of six content creators on TikTok. I am grateful for the stories and knowledge each shared with me, and I feel that there is more yet to learn from my interviews with each of these individuals and their experiences with content moderation on TikTok. In that way, I see this dissertation as a beginning to future work on the embodied experiences of content moderation on digital platforms. As each of these creators shared with me and has been documented time and time again, digital platforms such as TikTok are evolving everyday, not just in terms of viral content but in how they are governed. For example, as I discussed in chapters 3 and 4, creator theories about platform bans and suppression are evolving on a daily basis and in relation to different user subjectivities and experiences. Indeed, the experiences

represented in this dissertation have important implications for how we can understand the function and effect of content moderation on creators from multiply marginalized and underrepresented groups, particularly as patterns emerge across the experiences of marginalized creators broadly.

At the beginning of this project, I was taking a moment to scroll Tiktok for my own enjoyment and came across a video that has critically informed how I think about this study's methodologies and methods. In the video, the creator, @blickyrice, speaks directly to the camera, me and the other thousands of people who viewed and liked the video, and asked us to pause and consider the uniqueness of our individual feeds:

If you have more than one social media account, stop scrolling for like a minute, I wanna see something. My social media feeds usually consist of a lot of nerd shit, like video games, artwork, etcetera. Nowadays they include politics, a lot of gay people, music, art, and a lot of unique forms of editing. I imagine if you're seeing this, you can probably say the same as well. There's nothing wrong with that...How often do you remind yourselves that we really aren't the majority having this be our customized content?" (@blickyrice)

This call of personal reflection on the singularity of one's user recommended experience is important for individual consumption of online entertainment and for researching algorithmic environments. While I came to this study aware that my algorithmic experiences are not universal, the frame that my content is unlikely to be the majority resonated with me, particularly as I began to consider that conversations about content suppression and shadowbans on TikTok are not only happening in the way that I have

experienced them on the app. Indeed, when I search the top hashtags related to these topics, none of the top videos discuss content suppression as it relates to the experiences of marginalized creators on the app. A few methodological implications can be noted about the absence of these conversations in hashtags related to content moderation (such as shadowbanning, content suppression, Community Guidelines):

1. Content suppression is difficult to study because there is no way to know for sure what has been suppressed.
2. Hashtags on TikTok are used liberally to boost engagement, whether or not hashtags reflect the content of videos, as is evidenced in videos tagged with phrases related to content moderation.
3. The hugeness of these hashtags reflects the importance of this conversation on TikTok across a range of experiences and understandings of TikTok and its algorithm.

Other ways to study conversations on TikTok include tracing the use of sounds and filters, which are linked in the interface so that creators can participate in trends on the platform. While these interface features, perhaps including hashtags, could provide insights into broad conversations on TikTok, these features cannot help us to recover what is lost in totalizing platforms, whether these conversations about content moderation are erased through content removal or buried beneath thousands of videos that aren't relevant to the conversation as a result of suppressed engagement.

These are the problems that brought me to conducting interviews, which proved to be very fruitful and granted me important insights into the experiences of content

creators. Much to my surprise, my random selection of creators in this way yielded a pool of creators in which two of the six creators knew each other prior to our study and spoke about their experiences in relation to one another. While this surprise ended up yielding important insights about how content creators collaborate with one another, the coincidence made me pause. How much of my algorithmic interests are coming through in this study? How has my use of TikTok conditioned this study's findings from the start? While I'm not sure of the answers to these questions, my reflections on the influence of TikTok's algorithmic recommendations have deepened my belief that internet research cannot take place without considering the influence of content moderation.

Although this dissertation also draws on other sources outside of my interviews, it feels important to highlight the extent to which this study is algorithmically tailored. Moreover, I believe that my experience here represents an opportunity to explore more situated study of digital platforms like TikTok. The conversation about TikTok's content moderation is not only taking place on my For You page, but I believe exploring these interviews as partially the product of my For You. Additionally, while there are important patterns that emerge across the stories shared by these creators and creator testimonials shared on Change.org, this study is limited by both the number of participants interviewed for this study and largely reflects the subjective experiences of these participants. Finally, TikTok's content moderation policies and practices are ever-evolving, and so this study reflects content creator experiences in a particular moment on the platform, between the years of 2019 and when these interviews took place in 2021.

## CHAPTER 3

### SENSING SUPPRESSION: AN EXAMINATION OF CONTENT

### CREATOR PERCEPTIONS OF EXCLUSION IN TIKTOK'S

### INTERFACE

#### Introduction

In a March 2020 report for *The Intercept* titled “Invisible Censorship,” Sam Biddle, Paulo Victor Ribeiro, and Tatiana Dias published two internal TikTok moderation documents: one which outlines political and ideological views to be suppressed in live streams and one which lists physical and environmental features that should be suppressed and/or excluded from the For You page. With the criteria for algorithmic exclusion based on “abnormal body shape,” “ugly facial looks,” “too many wrinkles,” or shot in “slum, rural fields” and “dilapidated housing,” the leaked moderation protocol reveals the justification too that “if the character’s appearance or the shooting environment is not good, the video will be much less attractive, not worthing [sic] to be recommended to new users” (Biddle, Ribeiro, and Dias). Unlike internal TikTok moderation guidelines that had been leaked in December of 2019 by the German publication *Netzpolitik* which showed that the platform limited the reach of queer, disabled, and fat users as an anti-bullying measure (Köver and Reuter), these documents connect the moderation criteria to new user retention and platform growth.

At the time that these documents leaked, spokespersons for the platform stated that the moderation guidelines in questions were outdated and had already been revised, had never been implemented, or “represented an early blunt attempt at preventing

bullying” (Biddle, Ribeiro, and Dias). Since then, content creators and platform users have called for increased transparency around moderation guidelines, justifications, and outcomes. As discussed in the overview of TikTok’s transparency policies in the introduction, the platform has yet to disclose internal documents that show exactly how the platform has moved beyond what was revealed in the leaked guidelines. Building on this specific historical context surrounding TikTok’s documented pattern of algorithmically excluding marginalized platform users and the problem of moderation transparency, discussed in more detail in the previous chapter, this chapter examines how content creators sensed algorithmic exclusion through their micro-interactions in the platform’s interface. Indeed, content creators do not only understand the platform’s content moderation through their experiences receiving content violations but also through identifying and interpreting experiences of algorithmic curation, reduced engagement, and suspicious glitches in their content like sound removals. Sound removals, as I learned from my participants, are when the platform takes down the original audio of a video, oftentimes for copyright reasons but not always, as indicated by my participants.

These content moderation experiences are informed by their individual embodied positionalities on the platform and in the world, but across the range of experiences of my interview participants I consistently found that TikTok’s content moderation values, the secret rules that organize the platform, are felt throughout the interface. Moreover, these feelings about the platform’s secret rules inform how content is created and thus what kinds of content are accessible on the platform and how content creators feel about TikTok as a platform for user-generated content.

In this chapter, I specifically discuss how content moderation values are communicated through the following micro-interactions in the interface:

- through platform moderation tools that are weaponized by other platform users;
- through inconsistent and uneven Community Guidelines violations;
- and through sneakier types of content policing like low views, lengthy review processes, and sound removals.

Through my analysis, I argue that a lack of transparency around TikTok's algorithmic curation and moderation

1. contributes to content creators feeling alienated, exploited, frustrated, and unwelcome on the platform;
2. influences content creators to adapt their content to avoid content moderation, oftentimes by self-censoring themselves and aspects of their marginalized identities;
3. that the accumulation of content moderation micro-interactions, regardless of the root cause, diminishes both the degree to which marginalized content creators trust content moderation processes and feel encouraged to continue creating content on the platform

Moreover, these findings are grounded in the theoretical position offered by André Brock Jr. that "the internet should be understood as an enactment of whiteness through the interpretive flexibility of whiteness as information" (p. 6). Taking as given that the Internet is not neutral but is structured through whiteness, I find it important to examine content moderation through an intersectional paradigm as it is felt and experienced and contextualized by the cultural identities of content creators. Through this discussion of



content creator experiences, I ask: in the absence of platform transparency, what can we learn about the cultural implications of content moderation from the available evidence of user experience, and what are the implications of these experiences for anti-racist, inclusive user experience design?

In the subsections below, I analyze the features of TikTok's design that facilitated content creators' feelings about the purpose of content moderation on the platform and by extension who belongs on the platform.

### **Mass-Reporting: The Role of Platform Users in Content Moderation**

In this section, I want to take a moment to discuss the role that platform users play in content moderation through a focus on mass reporting. While the use of platform reporting tools often helps users speed up the removal of harmful content (such as targeted harassment or hate speech), mass reporting is the use of such reporting tools by bad faith actors to censor and suppress the content of creators they don't like. While content creators attribute many of their experiences with censored content to the platform's initial screening of uploaded content, each of my participants mentioned that a significant portion of their content moderation judgments are the result of platform users reporting content that they disagree with or find offensive through the tools made available in the interface. Notably, the reporting tools are not well-signaled in the interface. For example, the button to report a video is accessible through the share button, as shown in the figure below. These insights demonstrate how a comprehensive picture of content moderation practices and policies on TikTok must attend to the way the platform enlists or enables its users in the effort to uphold its Community Guidelines.

Indeed, while it is important for platforms to provide users with tools to flag violative content and abusive behavior, recent scholarship on content moderation and digital technologies have shown how reporting systems and other tools for collecting user feedback across digital platforms are often weaponized by users in bad-faith, often impacting marginalized platform users the most (Fiore-Silvast, 2012; Crawford and Gillespie, 2016; Massanari 2017). In their work which unpacks the political implications of flagging tools in social media, Crawford and Gillespie argue that flags have two key functions: 1) to enlist platform users in the endless labor of content moderation and 2) to evidence community input in the decision-making processes of platforms. They write that “flagging offers a powerful rhetorical legitimation for sites when they decide either to remove or to retain contentious content, as they can claim to be curating on behalf of their user community and its expressed wishes” (p. 412). Reporting systems allow platforms, which operate with very little oversight because of Section 230, to practice and perform “a willingness to listen to users” for both governing bodies and everyday users, while platforms conceal the extent to which these reporting systems materially impact the regulation of content.

Across my interviews, participants shared several stories in which they were penalized by the platform for identifying and exposing discriminatory content on the platform, oftentimes as the original content violations continued to remain unmoderated on the platform. Each of the creators I spoke with mentioned how they often had videos removed when they stitched— shared a video response to someone else’s video or comment—content that they identified as being in violation of TikTok’s Community

Guidelines by being racist or bigoted or by spreading misinformation. For instance, Mae, a Black queer mother who creates a range of content about her personal interests and political beliefs, explained that she generally receives content violations for what TikTok identifies as bullying and harassment, even when she is actually calling attention to bullying, harassment, and other behaviors that are inconsistent with the Community Guidelines. Except for when amplifying creators she likes and supports, Mae shared that she only stitches videos in response to someone attacking her family or uploading something that she identifies as being in violation of TikTok's Community Guidelines:

I tend to get a lot of violations when I'm talking about issues, when I'm talking about racism. When I stitch a video from a person saying something very racist, TikTok hits me with bullying. I'm not the one who says the word like, this person said the word. I'm stitching it, I'm talking about it, but I get hit with the bullying. Maybe if they had people to actually review it. Because what I think they're doing now is that, like, they have certain keywords, you know, whenever they hear that, it just automatically flagged the account and you get hit with a violation.

In this comment, Mae points to algorithmic detection as the source of her violations and shares that just before her first account reached 200,000 followers, Mae was permanently banned after receiving a number of content moderation violations on response videos of this type. Although Mae believes that there are more problems with TikTok's algorithmic moderation that need addressing, she suggests that more human moderators are needed to verify whether violations are genuine or if users are weaponizing mass-reporting against a creator they disagree with. Mae's experiences illustrate how content creators receive vague content violations within an environment

in which other platform users can weaponize TikTok's reporting systems to punish creators with whom they disagree. For this reason, it is key to recognize the role that other platform users play in surveilling marginalized content creators through digital content moderation.

Similarly, Eric, a Black man who creates content about relationships between men and women, spoke to how he adapts his content to prevent other users policing him. For example, Eric noted that he experiences more platform moderation when he dresses certain ways or when he speaks about race. After receiving several violations for adult nudity and sexual activity for videos in which he appeared wearing a tank top, Eric stopped wearing certain types of clothing and accessories and concealed his tattoos in his videos to avoid platform moderation or targeted harassment and reporting from platform users. In the excerpt from our interview below, Eric explains how the wide audience he reaches with his content influences how he carries himself in his videos to avoid mass-reporting: "if I say something that somebody doesn't like, if I say something that might bother somebody that it's like, it's real easy to just report a Black content creator for wearing a wife beater because of the stereotypes that come with that, because I have chains." This example points to the ways in which everyday users on the platform participate in the surveillance and content moderation of marginalized content creators. In our interview, Eric explained that his habits of "code-switching" to be more "palatable" for his potentially hostile audience are not entirely conscious; rather, they come from his "everyday" experiences of code switching to be perceived as less threatening as a tall, athletic Black man: "It's just one of those things that's just in the back of your head." Moreover, Eric discussed his frustration that the most popular

Black creators on the platform have been successful because of their palatability, despite the influence of Black culture on TikTok's most popular trends.

Mass reporting can often represent a coordinated attack on content creators. Sometimes, as a few of my interview participants noted, it is a useful strategy that they hope to mobilize when they share videos of problematic creators on the platform. Yet, it is also important to recognize how this strategy contributes to ongoing oppressive enforcement of content moderation on TikTok and other digital platforms. Moreover, without meaningful transparency about moderation emerging from the platform, content creators—particularly those with large platforms—are subjected to content policing from multiple audiences with fewer tools to protect themselves and their content from being suppressed, silenced, or forced off the platform. As an added layer of moderation, content policing via mass reporting contributes to the discrepancies felt by creators, largely because they are scrutinized not only by oppressive algorithms but also bigoted users on the platform.

### **Sensing Bias: TikTok's Affective Atmosphere of Inconsistent & Uneven Moderation**

Building on the previous section, I discuss the ways in which content creators identified how inconsistent moderation practices, whether originating from the platform's moderation systems or other platform users, contributed to feelings of frustration, anxiety, and stress on the platform, as well as facilitated feelings that TikTok does not support its marginalized creators. This latter feeling emerges in opposition to TikTok's statements about its platform functioning as a space for diversity, where content creators can share their creativity by finding their audiences.

Before I begin to fully describe how inconsistent and uneven moderation manifests in TikTok, I want to define what I mean by these two terms. Specifically, I use *inconsistent* to refer to how content moderation is implemented, whether by algorithmic or human moderators, with a great deal of room for variation, not simply when comparing the experiences of different users but when examining the different experiences of one individual over time. Variation emerges in various forms: communication regarding the violation, interpretation of the Community Guidelines guiding the moderation judgment, penalty to the user, among other factors. While I chose this word very carefully for the purposes of this chapter's analysis, I want to be clear that this is not the primary word that the content creators I spoke with use when describing their experiences with moderation. However, when these content creators shared with me a story of moderation, they described their experiences much more specifically, identifying which biases and values of the platform and its broad user base that they believed contributed to the moderation of their content and the suppression of their voice on the platform. For example, Eric spoke specifically of his experiences as a Black man as limiting what he is allowed to say and do on the platform, especially when wanting to talk about race. On the other hand, Zev shared that he feels that as a white Jewish man that he is afforded a privilege on the platform that other marginalized creators do not have, which he uses to speak about how content moderation impacts marginalized creators most prominently. To be clear, the experiences of these content creators and others who have shared their experiences online indicate the exclusionary impacts of content moderation on TikTok and other algorithmic platforms.

While the language of inconsistency does not communicate the way that content creators do not feel moderation as random errors but as targeted policy enforcement and harassment, inconsistent as a term highlights the temporal experience of content moderation on TikTok. To supplement this key understanding, I use uneven to point to how content moderation on TikTok privileges certain types of people and content within the frame of Western technoculture. Across my interviews, the creators I spoke with often acknowledged the possibility that some of their experiences could be attributed to a simple mistake made by the platform; however, the frequency of mistakes contributed to feelings of being targeted in specific ways. Interestingly, most of my interview participants reported receiving content violations for sexual activity and adult nudity for videos when these videos did not actually include content which violated these guidelines. While each creator receives notifications about these violations, the initial communications and the appeals processes are too limited for creators to feel that they truly understand the platform judgments. For instance, Stella, a white queer nonbinary comedian with over 3 million followers, noticed that their content would often be flagged for adult nudity and sexual activity, or her account would be banned if they said words like “porn” or “dom,” always in the context of their popular joke series. During our conversation, Stella discussed their experiences with a series of videos about sexy anime men that was popular for several months with no problem. However, after gaining more attention, so much so that it inspired copycat videos, they found videos in this series the subject of increased moderation with videos sitting in review for days or being taken down soon after posting. Stella noted that when these videos first started getting taken down that they didn’t understand why previous videos had been acceptable but decided

to self-censor subsequent videos in the popular series to try to make the series more appropriate for the platform.

I started censoring all my words in those videos. When I was writing out captions, I was censoring out the words in the captions, which I also felt bad that I had to do that, because I also have followers who are hard of hearing. I would censor out the words in the audio, with a beep. Those videos were still getting taken down for sexual activity and nudity, and that was kind of crazy to me because I have seen a lot of videos on TikTok, where people are openly saying certain words that I'm saying that don't get taken down. I just wondered, why is this happening to me?

Stella's explanation indicates how, in the absence of clear communication from the platform, that they turned to comparing their videos to other content they saw on the platform to make sense of the rules. Moreover, the failure of these attempts to make the content acceptable made the moderation experiences more frustrating and more indicative that they were the subject of the moderation and not the videos themselves. They also noted that as a fat femme-presenting creator that they were incredibly careful to dress modestly as an additional precaution against violations of this nature. In many cases, Stella noted as well, they would appeal these videos successfully, although they have never received any formal explanation about the initial errors.

Importantly, my findings dialogue with recent studies which show that marginalized social media users are more likely to experience discriminatory forms of content moderation, oftentimes because of algorithmic detection systems struggling to differentiate between counterspeech, reappropriation, and hate speech (Haimson, et



al., 2021). Bianca, a Black woman, discussed frequent difficulties posting content about her Black business unless she censored or coded language with codes like “Bl@ck,” an experience she noted is also common on YouTube and Instagram.

Similarly, Eric, a Black man who creates content about relationships for a following of mostly Gen Z women, discovered soon after he began creating content that he was unable to use the n word casually in his videos. For Eric, this content moderation experience was frustrating to him because while his videos are taken down if he uses the n-word, popular white content creators on TikTok often go viral using songs that feature the n-word. Stella, a white nonbinary queer person, shared similar stories about using language specific to their gender and sexuality, for example when talking about lesbian romances or making jokes about consensual BDSM.

Similarly, Zev, a white Jewish man, shared comparable stories about having his videos about antisemitism and Jewishness policed by the platform. As a Jewish American student attending Georgetown University, Zev began posting about these topics on TikTok after the insurrection at the Capitol on January 6, 2021, after witnessing much of the chaos of that day happening just outside his door. Soon after, Zev began making videos to educate his followers not just about the events of January 6<sup>th</sup> but about white nationalism and antisemitism. Prior to posting about these topics, Zev had never experienced content moderation, but after he began talking about these topics daily, he found that his videos were moderated almost immediately upon being uploaded: “These videos would get taken down, and I thought to myself, that's weird. I would re-upload, they would get taken down, and I'd play the little dance. There was one

time that I was talking about antisemitism on TikTok specifically, and they took down my video within 32 minutes.” Across the interviews, content creators identified uploading difficulties as a sign of content moderation triggered by key phrases and topics, which hasn’t been confirmed by the platform but does correspond with the platform’s most recent quarterly transparency report, which report data on the platform’s responses to legal requests regarding user content in addition to data on how the platform has enforced the Community Guidelines. These reports reveal that platform removes 94% of content before users have the chance to report content to the platform (Community Guidelines Enforcement Report).

Importantly, content creators mentioned throughout our interviews that they felt that criticizing TikTok led to increased content violations, reduced views, and other odd content moderation experiences. For instance, Zev shared that speaking about the antisemitism that Jewish creators such as himself face *on* TikTok resulted in more frequent content violations than talking more generally about content moderation issues on the app: “In these videos, I say directly, look at the comments. I get comments like ‘Hitler missed one,’ comments that use the k-word, that kind of stuff.” In highlighting the harassment that he and other Jewish creators face on the platform, both in comments on videos and in private messages, Zev emphasizes that part of the problem is about what TikTok allows other users to do to creators, particularly marginalized ones who are often attacked when they speak about political and personal issues. As a result of his frequent experiences with antisemitism on the app, both in terms of algorithmic censorship and harassment from platform users, Zev shared that his videos about Jewishness and antisemitism are the only videos that he scripts out before recording. In

particular, he highlighted that these videos require specific language and tone choices to avoid appearing as though he is attacking the platform, which he believes heighten the chance of content removal: “I don’t say, ‘This is what you TikTok are doing to us,’ but it’s more focused on what we experience and trying to center Jewish creators. Even then the posts still get taken down, but I have found that works better.” Zev shared that this is oftentimes his advice to content creators who reach out to him when they experience exclusionary content moderation on the platform.

Overall, inconsistent and uneven interactions with content moderation fostered in each of my interview participants the view that the platform policed representations of difference along axes of race, gender, sexuality, disability, class, ethnicity, and other identity markers. Accordingly, each participant discussed that presenting themselves and their content in ways to appeal to the imagined audience of TikTok was the most successful way to avoid inconsistencies in content moderation. For instance, Zev advocated for more polite tones in videos that call out the platform for its systemic exclusion of marginalized creators. Indeed, Eric’s description of the need to be “palatable” resonated across my interviews, indicating how content moderation experiences communicated that there is a certain type of content desired by the platform—and a certain type of content creator.

The choice to be “palatable” for the platform, to censor oneself as a creator, is in most cases an economic choice. Each of the creators I spoke with reported the massive impacts that content moderation had on their ability to monetize their platforms. While most users might recreationally engage TikTok, for the creators I spoke to, these platforms comprised a significant portion if not all their income, making content

moderation a very real threat to their ability to sustain themselves. When the platform holds a video under review for several hours or removes a video altogether, the creator is no longer able to make money from their work. If a content creator appeals a content removal or violation and wins, and the video is reinstated to the platform, these creators noted that reinstated videos rarely perform as well as other videos, making less money due to the time originally lost since first posting. However, while all these creators suggested that altering their content was a mitigating strategy, none of my participants see it as a long-term solution or a reliable way to avoid frustrating experiences on TikTok.

Regardless of the root cause of the content moderation interactions that have led these content creators to believe that the platform discriminates against its marginalized creators, my findings indicate that repeatedly receiving content moderation violations when speaking back to hate on the platform or about one's experiences as a marginalized person both on and off the platform made content creators feel as though the platform discouraged expressions of difference. This is all the truer in a non-transparent context where policies are communicated through content violations themselves. All my participants felt that they would feel less targeted if policies they disagreed with were communicated transparently and enforced consistently for all users. In that case, they expressed that they would be able to understand their experiences as the result of policy issues as opposed to evidence of exclusion. As it stands, marginalized creators are subjected to layers of biased content moderation as a result of biased users who mass report, biased human moderators, and biased algorithmic detection designed by biased programmers.

## **Sneaky Policing: Content Suppression, Reduced Reach, and Muting**

Thus far, I have discussed direct content moderation interactions in which creators received violations from the platform for videos they posted. In this section, I describe interface interactions that my interview participants identified as sneakier types of platform moderation. As I have discussed in the previous sections, each of the content creators I spoke to have identified certain words or phrases that trigger content violations which lead to videos being removed. Similarly, content creators sense that certain words or phrases contribute to content suppression and reduced algorithmic reach. That the platform has a pattern of screening certain phrases to different effects is common wisdom on TikTok; for example, from my own experience, it is extremely common to scroll through the app and see creators code their captions by integrating numbers and symbols where letters to avoid the algorithmic detection, for example by using *yt* instead of white and Bl@ck instead of Black. My interview participants used the same logic for topics related to social justice. For example, Eric and Mae noticed that when they talked about certain topics their average video view counts would drop drastically and remain low for several days. Eric noted that talking about Black Lives Matter and Black issues would result in reduced engagement, while talking about topics that are not explicitly racialized performs well for him.

Creators shared that they often realized that their content might be shadowbanned or rendered less visible through follower comments on their videos that said they hadn't seen content from my participants in a few days. In other words, TikTok followers sometimes share suspicions about reduced algorithmic reach and suppression. While content creators received alerts about content violations, my

interview participants each reported instances when their followers would alert them to the possibility of content suppression and shiftier methods of moderation. For example, Eric, Mae, Stella, and David mentioned numerous examples of their followers informing them that they had not appeared on their Following page for days. These insights from followers helped these creators to investigate what may have triggered their reduced reach, whether that be mass-reporting or recent videos about controversial topics. Additionally, my interview participants shared that their followers oftentimes notify them of shadier types of moderation, such as muted content. Muted videos, or videos that have had the sound removed, do not affect the visibility of a video on a creator's profile, but the original sound uploaded to the video is listed as no longer available.

While the platform has strict intellectual property guidelines that can lead to videos being removed, creators reported that videos that had the sound removed were not associated with protected property. There is no formal alert that a video has been muted, so several creators shared how disorienting it is to discover that an older video no longer has its original sound. Zev shared a story of muting in which a follower notified him that the sound on one of his earliest videos no longer had a sound. When he reviewed the video, which was the fifth in his series calling on the CEO of TikTok, Vanessa Pappas, to release the content moderation guidelines, he found that it had received more attention since he had originally uploaded it, leading him to believe that the muting was more recent and connected to the video "gaining traction." Given that there is no formal violation or notification associated with muted content, Zev explained, there is also no way to formally appeal what appears to be a platform decision. Without formal options for recourse, Zev explains that when he notices this type of problem soon enough that he

will re-upload the video or “make a video saying weird how they took down the sound two to three months later” when it seemed to be resonating with people on the platform.

### **Discussion: Sneaky Rhetorics of TikTok’s Content Moderation**

In this chapter, I have shared insights from content creators about what aspects of their interface experiences on TikTok contribute to feelings of exclusion, exploitation, suspicion, and unwelcomeness that content creators feel toward the platform and its content moderation processes. Without platform transparency and communication for each of these experiences shared with me, each of these creators’ stories could be reduced to simple errors in code, or platform oversights to the nuances of cultural communication. I argue that the problem of content moderation transparency, as discussed in the introduction, has critical affective implications on content creators on TikTok in terms of the content they create and their attitude toward the platform. However, I argue that platform transparency should not be elevated in importance above the insights that we can learn from platform users, especially those from marginalized backgrounds who by virtue of their embodied experiences are especially attuned to systemic resistance.

Throughout my conversations with creators, I heard each of my interview participants express their feelings that content was removed, suppressed, or otherwise policed for reasons not explicitly stated by the platform. I connect this chapter’s findings to TikTok’s sneaky rhetorics, or, “a sense that there is something going on that is not completely transparent and that thus elicits feelings of fear, risk, and mistrust as a result of unknown audiences...and unknown consequences” (Fernandes, Homer, and Sano-Franchini).

Across my interviews, I hear echoes of sneaky rhetorics animating the frustrations that these six content creators shared with me, particularly in creator suspicions that the platform favors an elite 1% of content creators at the same time that marginalized users are uniquely targeted. For example, when creators shared that they felt their content was being removed at increasing rates as they became more popular on the platform, creators recognized that they were likely experiencing more problems due to their increased reach on the platform; however, they also felt that their experiences possibly signaled TikTok's preference for certain types of content and certain types of creators. The overlapping threats of moderation by platform error, platform exclusion, or mass reporting from other users contributes to content creator beliefs that they are being policed not only by the platform but by fellow users. These beliefs, in turn, factor into content production as content creators attempt to circumnavigate moderation barriers; performing palatability in their content helps to limit some of the resistance they face, but not all, leading many content creators I spoke with to adjust their content less than less.

The insights shared with me by these six content creators have important implications for user experience design of digital platforms. In line with recommendations from content moderation scholars and practitioners (Gillespie, 2018; Roberts, 2019), each of the participants reported that the following changes would reduce their feelings of frustration, anxiety, stress, depression, and confusion when confronted with content violations:

- improved platform transparency about how content violations and their corresponding consequences;



- improved communication about moderation decisions, including throughout the appeals process;
- increased access to human representatives of TikTok, particularly for content creators and especially related to matters of monetization;
- increased transparency of moderation guidelines used by human moderators to remove content;
- or barring improvements that reduce errors, more uniform enforcement of platform policies.

Altogether, these recommendations highlight the need for increased communication and transparency for content creators so that they can more successfully navigate the platform, or at the very least, the need for a consistent environment. As I discuss in the next chapter, the lack of transparency around the possible consequences of content violations, such as platform bans, interface restrictions, and reduce reach, represent a constant anxiety for content creators whose uneven experiences destabilizes access to income generated on the platform. In this way, the labor element of content creation is significant in mobilizing content creator discontent as marginalized creators see white creators benefiting from cultural appropriation while experiencing fewer run-ins with misapplied Community Guidelines.

Furthermore, the lack of transparency on TikTok also facilitates feelings of suspicion in content creators that have important implications for how they interpret the platform's statements supporting diversity and inclusion, particularly as these problems continue to persist on the app. In a report for the *MIT Technology Review*, Abby Olheiser (2021) summarized TikTok's pattern of content moderation controversy in the

following way after the platform came under fire once again for censoring language related to Black Lives Matter:

First, a creator notices a bizarre and potentially harmful issue with the platform's moderation or algorithm, one that often disproportionately impacts marginalized groups. Then they make a video about what's going on, and that video gets a lot of views. Eventually, perhaps, a journalist gets interested in what's going on, asks the company to explain, and the issue is fixed. TikTok then releases a statement saying that the problem was the result of an error, and emphasizes the work they do to support the creators and causes affected.

In describing TikTok's "endless cycle of censorship and mistakes," Olheiser highlights that the larger problem on TikTok is not simply the string of mistakes but that bad press oftentimes is the only way to encourage platform reform. This pattern of platform response feeds into sneaky rhetorics as evidenced by my interview participants describing such patchwork fixes not as evidence of real effort but as damage control. They feel, given the platform's history of ignoring problems as long as they can, that TikTok is not sorry about the impacts of these so-called errors but about getting caught. Despite this, many of the creators I spoke with personally shared their hope that eventually enough media attention would lead to more meaningful investigations into the problems of transparency and content moderation on the platform.

Similarly, sneaky rhetorics, heightened by the lack of transparency, fosters content creators' beliefs that there are secret rules that designed to punish marginalized creators principally. To this point, the sneaky rhetorics of content moderation affect some content creators' ability to trust their perceptions of their content moderation interactions.

While the term gaslighting is typically used to describe abusive relationships, two content creators I spoke with used this language to describe the platform's behavior toward their marginalized creators. Bianca, a Black woman, described the platform's frequent statements for diversity and inclusion as performative when so many Black content creators have their work stolen and monetized by white creators on the platform. Speaking to the frequent theft of Black dances on the platform, Bianca questioned why the platform puts out such statements while refusing to implement changes, like a dance credit button; she explained, "I feel like it's a form of gaslighting, because if I apologize to somebody, but then I go right back to doing the same thing, then what was really the point of the apology?" This critique contributes to current conversations about the meaningfulness of institutional diversity statements, while also pointing to how experiences in the interface oftentimes matter more to how content creators understand the values of platforms like TikTok. Habitual experiences of receiving content violations, Stella noted, led them to feel "gaslit" by the platform as they began to consider maybe they *were* a "problematic creator," especially when they saw other creators—notably, white and straight creators—get away with uploading videos that they couldn't.

At the time that I met with Stella, they had just received a post-ban, a type of ban that only limits publishing rights but not other types of platform access, after receiving a community guidelines violation for saying the word "dom" in a video. After one of their friends shared a raunchy, joke video by two white men creators pointing out how queer creators could not get away with this, Stella reported the video to test the hypothesis:

I know this is a shitty thing to do, but I needed to collect evidence. I needed to see what would happen, so I reported the video for sexual activity. Pretty much an hour later, I got a response back saying that there were no violations. That was a moment where I felt like there was a target on my back and the backs of my friends. A lot of these creators that are having these issues are marginalized in some way. They're either a person of color or they're disabled or they're trans, or they're queer. There are creators who can say whatever they want on the app, and then I must censor out everything.

Stella's decision to report the video emerged from their need to confirm that their feelings of "frustration" that many if not most of their content moderation experiences represented algorithmic bias against them as the creator and not the content. By collecting evidence like this, Stella was able to gain more certainty that there are different rules for different creators. Importantly, the speed at which Stella received the notification about their report is contextualized by their experiences receiving very little direct correspondence from the platform regarding their own content violations.

Relatedly, creators compare their content moderation interactions to the persistence of hate speech, harassment, and mis/disinformation on TikTok. For example, Mae described the daily experience of witnessing other creators of color and queer creators experiencing similar types of discriminatory content moderation as "demeaning and defeating," especially when TikTok fails to moderate hate speech, harassment, and vaccine misinformation. Sneaky rhetorics stick to these encounters with hate speech, harassment, and vaccine misinformation, further eroding platform user trust in TikTok's governance.

Stella was not alone in their desire to “prove” that the problem of content moderation on TikTok impacts marginalized creators the most. Each of the content creators located their gut feelings about discrimination in both their interface experiences and within the context of the platform’s history of algorithmic exclusion of marginalized users, the continued lack of transparency around moderation practices and policies, and their perception of and engagement with other creators experiencing similar types of discriminatory moderation on TikTok. For instance, Zev cited the internal moderation guidelines leaked by The Intercept discussed at the beginning of this chapter as evidence of the platform’s poor history of supporting content creators, not only in our conversations but in the Change.org petition he created that has accumulated over eighteen thousand likes. Similarly, other participants connected their experience to viral examples of content moderation controversies on TikTok, including the documented widespread suppression of Black Lives Matter content at various moments in the past few years (Cohen, 2020; Ghaffary, 2021; McCluskey, 2020). In my view, this triangulation of personal experiences in TikTok’s larger context evidences an important approach to sense making that is useful for interrogating algorithmic bias and discrimination (Benjamin, 2019, Noble, 2018); at the same time, I was left with the impression throughout my interviews that the need to prove that one’s experiences are part of a larger pattern is a consequence of the lack of transparency and the inability to point to evidence of the scale of the problem.

In sharing their experiences of exclusionary and discriminatory content moderation, content creators highlight the unevenness of policy enforcement and the contours of power and privilege on the platform. Moreover, while improvements to

transparency on TikTok are necessary, these stories illustrate the complexity of transparency as a solution. As Mike Ananny and Kate Crawford argue in their explanation of the limitations of transparency as an ideal for algorithmic accountability, “Not only is transparency a limited way of knowing systems, but it cannot be used to explain—much less govern—a distributed set of human and non-human actors whose significance lies not internally but relationally” (p. 12). In other words, while transparency can provide platform constituents the tools to push back against platform decisions and soothe anxieties about control, it is important to recognize that part of the problem content creators face is not limited to the platform’s lack of transparency but is also a consequence how platform users engage content culturally, which can disadvantage marginalized creators.

By highlighting the limitations of transparency as a technological solution to a wicked technological and cultural problem, I believe that we can understand the desire for transparency, as Ananny and Crawford put it, as an emotional response, a desire for control, and as represented in the stories of my participants, perhaps the final confirmation of what my participants hypothesize through their experiences and evidence-gathering. In sum, I argue that efforts to improve transparency in the ways that my participants and others have asked for are important and scholars and practitioners studying these problems must rely on experiential, embodied data of platform users and content creators to address these environments comprehensively. User experiences of content moderation, such as the cases I have discussed here, offer much needed context, and more importantly, understanding of the impacts of emotions and culture on how we engage digital and algorithmic environments.

TikTok's transparency problem, in sum, contributes to a pervasive sense that marginalized content creators do not belong on the platform, communicated affectively not only through the discriminatory patterns of moderation and the gaps in moderation which allow white supremacy and misinformation to thrive on the platform. However, the stories shared in this chapter demonstrate the complexity of resolving transparency problems on digital platforms and complex institutional contexts, which I will discuss in greater detail in the conclusion to this dissertation. Sneaky rhetorics of content moderation reverberate through daily interactions with the platform as content creators encounter, in various forms, evidence that the platform does not value, support, or protect either its marginalized users or creators. Looking forward, in the next chapter, I will build on these findings about user experiences of content moderation in the interface to further discuss the affective influence of platform bans on content creators.

## **CHAPTER 4**

# **DISRUPTED CREATOR TIME(S): UNPACKING CUMULATIVE USER EXPERIENCES OF TIME AND EXCLUSION ON TIKTOK**

### **Introduction**

In the previous chapter, I examined content creator experiences of content moderation within the interface and how those interactions informed creator understandings of TikTok's opaque systems of governance and the platform's commitment to diversity, equity, and inclusion. Here, I move from that discussion of everyday types of content removal to explore creator experiences and perceptions of platform exclusion in the form of temporary platform restrictions and suspensions and permanent bans. Throughout my conversations with creators and my reading of online testimonials, I found that creators were most upset by their experiences with and the ongoing threat of platform bans. In what follows, I unpack these specific anxieties about platform bans and restrictions and what these disruptions reveal about the nature of time on TikTok.

To examine the affective impact of these types of platform exclusion, I draw on my interviews with content creators and an analysis of Change.org petitions about content suppression of marginalized creators on the platform. In this chapter, I found that the everyday inconsistencies and unevenness in content moderation on TikTok I described in Chapter 3 informed how content creators interpret more extreme platform judgements such as temporary and permanent bans. What, then, is the affective impact of frequently receiving inconsistent and uneven platform restrictions? What are the



consequences of the affective disposition fostered in this algorithmic environment on content creators and their relationship to the platform? And, most importantly, what are the material consequences on creators whose digital labor hinges on platform access—and thus, on whom the platform relies to create profits. To answer these questions, I turn to stories from my interview participants, specifically focusing on their out-of-platform experiences, or experiences on TikTok in which their access to posting, commenting, and scrolling were restricted. Importantly, as discussed in the previous chapter, these restrictions were often associated with unclear platform judgments, and the lack of communication about these restrictions worsened creator experiences, as I will discuss later in this chapter. Whether receiving temporary or permanent restrictions to platform privileges like posting or logging on altogether, such disruptions caused creators great emotional distress and frustration.

Therefore, this chapter focuses on the affective responses that emerged in response to recurring interactions with TikTok's banning and suspension practices. I found that the experience of frequently receiving platform restrictions, suspensions, and bans on TikTok affected content creators by intensifying existing feelings of frustration and anxiety. Based on my conversations with six marginalized content creators, I observed that these frustrations and anxieties oftentimes seemed to stem from several. Creator perceptions regarding the relationship between content moderation and creator success on TikTok: escalating moderation, creator exploitation, disrupted engagement, and an opaque appeals process described further below.

- **moderation inevitability:** creator perceptions that content moderation interactions increase in frequency and intensity over time, oftentimes alongside increased creator success on the platform.
- **out-of-time exploitation:** creator perceptions that content creators and their creative labor continue to be exploited even when they are prohibited from interacting with the platform;
- **engagement disruption:** disruptions in platform access limit the ability of content creators to monetize their labor on a platform that requires constant engagement to be successful; and
- **appeals process time:** the lack of open communication and transparency in the appeals process worsened creator experiences of disruptions, even when creators had their access restored.

Importantly, I found that all of these affective responses to platform bans and restrictions are intrinsically related to time. As discussed in the previous chapter, experiences of TikTok over time informed how participants felt and reflected on their personal interactions with the platform. As Kujala et al., explained, “a user can still experience the product and its meaning in *retrospect* through memories,” and memories of experiences can inform future decisions and interactions with UX (p. 561). While this attention to the retrospective experience of user interfaces is important, participants emphasized to me quite directly that time stands out as a concern with respect to bans and restricted platform access as it concerns creation as a means of making money on TikTok and other social media platforms.

Therefore, in this chapter I argue for more attention to time in the rhetorical study of user experience design and content moderation more broadly. I begin with a discussion of the user experiences of time facilitated by TikTok's interface. Then, I briefly discuss the main types of formal bans that a user can receive on TikTok in the section that follows. Finally, I conclude with a discussion of this chapter's findings and their implications for UX designers and scholars.

### *User Experience and Creator Expectations of Time on TikTok*

Through my analysis of creator interviews and online testimonials, it became apparent that user experiences of content moderation on TikTok are deeply related to time. Indeed, time on TikTok is usually experienced as an infinite scroll of neverending content tailored for the individual user. For content creators, participating on TikTok is even more contextualized by time in the sense that their daily participation on the platform is directly related to monetization and maintaining platform relevance. For that reason, platform bans and restrictions disrupt creator expectations around time on TikTok. I highlight two specific creator time expectations disrupted by platform bans and restrictions: 1) the expectation for daily engagement on the platform and 2) the expectation for linear organic growth on TikTok. The first expectation is necessary for the professional work of content creators on the platform who hope to monetize their efforts; communicating with followers and publishing new content on a daily basis is critical for creators who hope to build their platforms on TikTok and other social media platforms. The second expectation is born out of the first expectation in that creators often demonstrate a belief that continued engagement with best content creation practices (such as daily interactions with current and potential followers) will sustain exponential

platform growth. The former expectation emerges out of creators' responsiveness to the flow of time and content on TikTok and the need to be constantly engaged so as to remain relevant (Arriagada and Ibáñez, 2020; Raun, 2020), the latter reveals creator investment in meritocratic thinking that downplays platform forces and complex user behaviors that complicate how much creators can succeed on sheer will alone (Petre, Duffy, and Hund, 2019).

Indeed, several creators spoke about the need to be in constant communication with their followers, a testament to the expectations of micro-celebrity creators as discussed by Raun (2018) in a piece about transgender vlogging. Raun writes:

Micro-celebrities are expected to perform various kinds of labour, many of which are time and energy consuming but not necessarily economically profitable.

Micro-celebrities must signal *accessibility, availability, presence, connectedness* and maybe most importantly authenticity – all of which presuppose and rely on some form of intimacy. I propose that intimacy as genre and as capital is deeply ingrained in the strategies, dynamics and affective labour of micro-celebrities.

Intimacy is an important and necessary signifier in relation to both the form and content of the videos and the relation between the creators and their audience.

Furthermore, intimacy works as an important currency within social media; thus, intimacy can be capitalized in manifold and intersecting ways, for example, for monetary purposes, social recognition and as a tool in advocacy work. (p. 101, emphasis added)

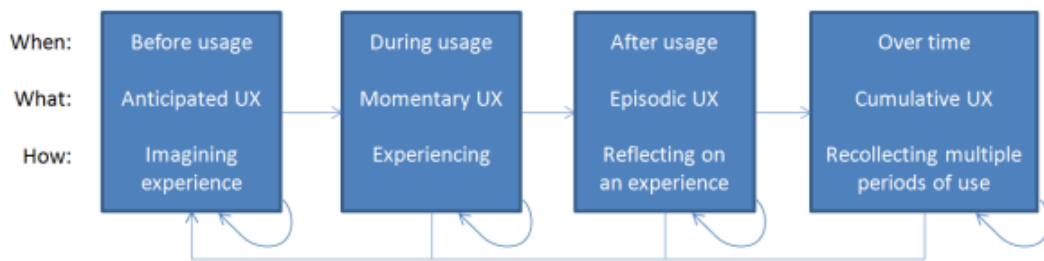
Raun defines intimacy as a crucial currency in social media in terms that indicate the temporal expectations of creators. Accessibility, availability, presence, and connectedness

all demonstrate the importance of immediacy. While Raun's work focuses on YouTube, these implications are perhaps more so relevant for TikTok creators whose videos are contextualized in a never-ending stream of added daily content that is delivered to platform users. Unlike other platforms where the landing page and dominant user experience is centered on creators that users follow and subscribe to, TikTok focuses on new content daily, and the For You page does not always highlight creators that a user follows. For that type of content experience, users must actively engage their Following page. Raun describes the type of work required of creators as a type of affective labor. I see some of the anxieties, frustrations, and stresses associated with platform disruptions on TikTok as an extension of that affective labor. If creators are not able to show up for their followers, how can they guarantee that they will have followers when they return?

In part, I come to these conclusions about the relationship between TikTok creators' feelings about temporal disruptions and emotion through an analysis of how the platform facilitates time. By engaging with creator testimonials and stories, my analysis of TikTok's interface, as I discuss in chapter 2, blends my subjective reading of the platform's design with insights from content creators. Overall, I argue that content moderation facilitates how time is experienced on TikTok in the following ways:

1. Content moderation facilitates time for platform users via an endless scroll of recommended content
2. Content moderation facilitates time for creators by requiring daily engagement to ensure that their work continues to be pushed to the For You page.
3. Platform users infer that moderation is to blame when creators they follow do not appear on their For You pages.

Indeed, in the current study, participants’ past experiences and looming threats of platform bans and restrictions reveal how important it is to consider the influence of time and memories in user experience design (Kapranos et al., 2009; Kujala et al., 2013; Normal, 2007; Pohlmeier et al., 2010; Roto et al., 2011). In the edited white paper on user experience produced by collective efforts at the 2010 Dagstuhl Seminar on Demarcating User Experience, the time span of user experience is defined as including not only “the actual experience of usage” but also indirect experiences that occur before and after usage (p. 8). This white paper outlines several time spans of user experience—such as anticipated UX, momentary UX, episodic UX, and cumulative UX—which they represent visually in the chart I’ve included below.



**Figure 1. Visual representation of the temporal dimensions of UX from the Dagstuhl Seminar**

I find that this visual representation of the temporal dimensions of user experience from the 2010 Dagstuhl Seminar usefully highlights how time structures user experience. Indeed, across my findings, creators spoke about anticipated, episodic, and cumulative user experiences in their stories and reflections, offering important insights about the temporal boundaries of TikTok’s UX design.

## **The Evolving Nature of Platform Bans and Restricted Access on TikTok**

Through my conversations with creators, I came to understand that TikTok bans are more complicated than I originally thought. In addition to the more standard account ban that suspends user profile access, TikTok also employs distinct types of segmented bans including live-streaming bans, comment bans, and post bans. While temporary and permanent bans are widespread practice across social media platforms, TikTok has begun to partially restrict user experiences as a content moderation practice. As shown in the image below, temporary and permanent bans are associated with an account being determined to have committed multiple violations of the Community Guidelines. From my interviews with creators, I learned that such notifications are common but the stated length of sentences are often inaccurate (with creators sometimes regaining suspended privileges earlier than predicted or enduring longer sentences). I also learned that the guidelines violated are not made clear in these system notifications, leaving creators to have to figure out what triggered the punishment for themselves.

### **First violation**

- We'll send a warning in the app the first time your content violates our Community Guidelines.
- If the violation is a zero-tolerance policy, it will result in an automatic ban. We may also block a device to help prevent future accounts from being created.

### **After the first violation**

After your first violation, we can take one or more of the following actions.

#### **Temporary ban:**

- We may suspend your account's ability to upload a video, comment, or edit your profile for a period of time (typically between 24 or 48 hours), depending on the severity of the violation and previous violations.
- We may restrict your account to a view-only experience (typically between 72 hours or up to one week). This means that your account can't post or engage with content during that time.

#### **Permanent ban:**

- After several violations, we'll notify you that your account may be permanently banned. This means that if the behavior persists, the account will be permanently banned.

**Figure 2. Screenshot of TikTok’s content moderation process, including description of temporary and permanent bans**

According to TikTok’s Account Safety policies (represented in Figure 3), platform users will receive a warning in the app when they first violate the Community Guidelines. After the first violation, the policy states that the platform *can* either penalize user accounts with temporary bans (which could include limitations to the account’s ability to upload, comment, or otherwise engage with content on the platform, limiting the account to a “view only experience”) or permanent bans. Importantly, the policy states that there are several violations that are considered “zero-tolerance” and result in automatic bans, although the only example offered of such material is child abuse.

At the time of my interviews in summer of 2021, creators reported that the view-only platform restrictions were a new development to TikTok’s content moderation practices, and that they had discovered them through receiving the in-app system notification that they were no longer able to post, live-stream, or comment. The Account Safety policy also provides instructions on how to appeal temporary or permanent bans, but as I will discuss later in this chapter, it does not describe how successful appeals will be addressed. Given that permanent bans are indicated to be the result of several violations, more information on how the platform measures suspensions would be useful for creators who feel that their platform suspensions are the result of platform errors counting against their account’s record. For example, many creators felt that the platform “counted” content violations and temporary bans “against” creators, even when the appeals process ruled that creators were penalized despite not actually breaking the platform’s rules. In other words, creators felt that the platform penalized them for



mistakes made by algorithmic moderation processes and faulty human moderators, mistakes that are not accounted for in the instructions provided to users who wish to appeal account bans, as shown in Figure 3 below.

### **My account was banned**

Accounts that consistently violate community guidelines will be banned from TikTok.

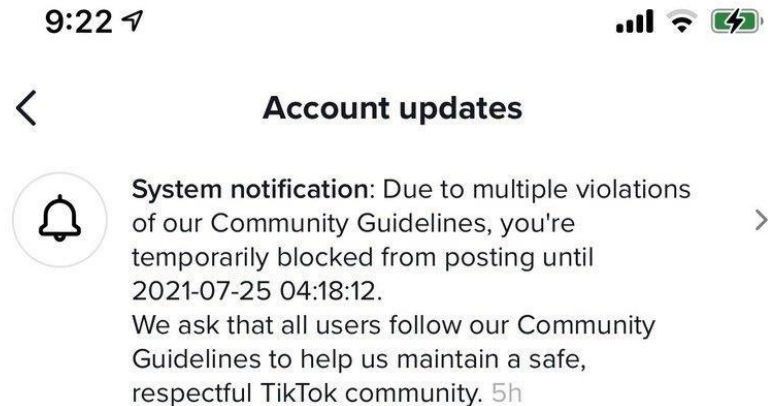
If your account has been banned, you will receive a banner notification when you next open the app, informing you of this account change.

If you believe your account was banned incorrectly, let us know by submitting an appeal.

To submit an appeal:

1. Open the notification.
2. Tap **Appeal**.
3. Follow the instructions provided.

**Figure 3. Screenshot of TikTok’s Account Safety page: How to appeal an account ban**



**Figure 4. Screenshot of a TikTok system notification of a temporary post ban**

While the platform states its ability to limit a range of platform services and features for users who violate platform policy in the Terms of Service, I find it important to note that the platform more frequently gestures to the Community Guidelines as the source of information about content moderation practices, as shown in Figure 4 which

depicts a screenshot of a TikTok system notification of a temporary post ban. The slippage in language related to types of platform restrictions between the Community Guidelines (CG) and the Terms of Service mirrors the discrepancy between TikTok's public statements which deny the use of shadowbanning on the platform and the language in the CG which states that the platform's algorithms "may reduce discoverability, including by redirecting search results or limiting distribution in the For You feed" for content that is under review, considered upsetting or shocking, or classified as spam per the Community Guidelines.

I mention shadowbanning here, which I define as a form of reduce visibility in the introduction to this dissertation (Are, 2021; Cotter, 2019 and 2021), as a corresponding example of how TikTok does not provide its users with a clear picture of its practices and policies; however, in this chapter, I will not be discussing shadowbanning as an experience of platform restriction that is like more traditional bans and restrictions. As I discussed in the previous chapter, perceptions of shadowbanning have important affective implications for content creators' trust in the platform. The sneaky rhetorics of content moderation on TikTok often contribute to creator feelings that shadowbanning is directly associated with content related to the creator's historically marginalized identity. At the same time, this chapter focuses on the more concrete experience of content creators having their usual interactions with the platform interrupted and the frustrating experience of having such interruptions become so frequent that they feel like an everyday part of their TikTok usage. Throughout the findings section, I discuss some of the consequences of these bans, including the effects of these bans on reduced engagement, restricted or lost access to money earned through the Creator Fund, and lost

access to content created on the platform. In addition to these material impacts, creators also reported the emotional toll of these frequent experiences with restricted or suspended access, as repetitive restrictions and bans led creators to worry and stress about future disruptions to their TikTok use.

Importantly, throughout my interviews with content creators, each participant discussed their anxieties that TikTok would implement new forms of bans in the future. In this way, sneaky rhetorics facilitate creator feelings about their present experiences of TikTok, as well as the future of moderation on the platform. Concerns about emergent and future moderation consequences emerged in relation to their discussion of post bans, a temporary restriction of one's ability to upload videos and comment on other people's content, which each creator mentioned was a recent development in their experiences on the platform. As I discussed in the previous chapter, TikTok is known among content creators for its ever-evolving consequences that are usually discovered in the process of receiving them, not through formal announcements or written policy. The emergence of post bans as reported by my participants was no different, meaning that these content creators first heard of post bans through their networks of other content creators or at the time of receiving an alert in the interface indicating a permanent or temporary suspension and the duration of the sentence. Such spontaneous and jarring interactions with platform policy have important implications for content creators and their relationship to the platform as I discuss later in this chapter. While each of these creators has had notable success on the platform in terms of growing their number of followers, all but one of the creators I spoke with experienced regular platform restrictions and suspensions. Often, these creators had their appeal requests approved, but the results of these appeals, which

oftentimes are not communicated by TikTok except through brief in-app system notifications that report the result of the appeal. Because these system notifications, in the event of successful appeals, do not address the cause of the mistake or offer , such notices do little to mitigate the lasting impact of the habitual interactions, as will be shown in the findings which follow.

Having established this context surrounding the discrepancies between TikTok's written policies in the Terms of Service and Community Guidelines on one hand and the lived experiences of marginalized and underrepresented content creators on the platform on the other, I will briefly discuss the relationship between user experience design and time before moving on to my discussion of creator stories about their experiences of platform restrictions and bans and what these stories teach us about how UX design facilitates time on TikTok. Focusing on stories shared in my interviews, I highlight how the accumulation of negative interactions with TikTok's content moderation shapes creator responses to being pulled out of the flow of time on TikTok.

### **Accumulating Content Moderation**

In this section, I discuss the impact of accumulating content moderation interactions on creators' experience of time on TikTok and the implications of this experience on platform trust. Overall, I find that because TikTok is experienced as a fast-paced content machine in which present interactions matter very much, accumulating content moderation interactions which take creators out of that present TikTok fosters anxiety, stress, and frustration related to both present disruptions and future experiences. In other words, creators' trust that they will have stable futures as platform users and creators on TikTok is eroded by accumulated content moderation experiences.

Across my interviews and testimonials shared online, I observed a pattern in the trajectory of a creator's experiences of content moderation and suppression. It goes like this: a creator starts out strong on TikTok, either through sustained efforts to gain followers or by accidentally striking content gold, gaining hundreds of thousands of followers over the course of a few months. Then, after achieving a significant amount of success on the platform, the creator begins to steadily receive increased content violations, many of which do not make sense, possibly due to an increase in user reports as a result of having a wider reach on the platform. At first, the creator finds that the violations are successfully appealed quite quickly through the platform's appeals process. However, over time, the creator is increasingly inundated with content removals and violations, and then eventually, the creator begins to receive additional account penalties, such as temporary suspensions and restrictions. Sometimes, these temporary suspensions appear to add up over time, resulting in a permanent suspension of a creator's account.

While this moderation trajectory may not be the experience of all creators who have struggled with TikTok's content moderation, four of my six participants shared stories that indicated an escalation of moderation interactions over time, as each creator gained more and more followers. In our interviews, Mae, Eric, Stella, and David each identified this pattern and shared that they felt things escalating after reaching a certain amount of visibility on the platform, an assessment that aligns with the previous chapter's discussion of the impact of mass-reporting on content creators. For example, Stella, a white queer nonbinary comedian, shared an experience of having their profile "permanently" banned that after receiving several content violations for sexual activity and pornography, one of the more common violations that they received despite the

categorical inaccuracy. Their description of the events of this experience illustrates part of the pattern I describe earlier quite well:

My account got permanently banned after I had gotten a post-banned from one of my joke videos that they said was porn. They post-banned me for four days. Once I was able to post again, I went on the app, and then it logged me out and said that I was permanently banned. I went to Instagram immediately told people to email TikTok, I had like a list of emails for [my followers] to send an automatic email to, I started emailing them myself.

The next day my account was restored, but it was interesting to me that a lot of my videos would go viral and then I would get a post ban, and then a video taken down, and it just kept going, that was kind of the trend.

As a representative moment in this trend, Stella's experience highlights the experience of receiving a temporary ban from posting after having content flagged inaccurately. While Stella was able to clearly connect their post-ban to an earlier content violation that they received, the subsequent punishment of a permanent ban was less clear and seemed to be an extension of prior consequences. While Stella did not directly explain why the connection between post-bans and viral videos was interesting, I interpret their reflections on this connection as an understanding that going viral opens creators up to more content violations because their content is being viewed by a much larger audience beyond users who actively and purposefully follow them, increasing the likelihood of mass-reporting from bad faith platform users. That type of exposure results in patterns of content being removed and restored. Despite having many videos successfully restored, Stella sensed that the platform held its inaccuracies against their record. I came to this interpretation

from the sense I got from my larger conversation with Stella and from content creators who articulated similar theories about their experiences with the platform.

At the same time, there is sparse information about how TikTok measures content violations for individual creators or whether it marks successfully appealed violations against an account or not. The creators who spoke about this TikTok's policy indicated that multiple violations could result in such actions being taken against user accounts, as shown in the figure below. When discussing their frequent interactions with account bans and posting bans, Stella explained that they felt the platform punishes creators like them for its own faulty system:

It's like this. My videos get taken down, I get a post ban, and then I appeal the video and then they restore it. Like today, it took an hour for them to restore my video, but then I still got another message telling me that I'm still post banned. But that wasn't on me. That was [TikTok's] fault, so why are they punishing creators for things they didn't do?

The post ban thing is new. That didn't always happen. That started a few months ago, and it just blows my mind how you can get punished for something you didn't commit. All of my community guidelines violations are their fault, so I'm not eligible to be verified because this problematic app keeps taking down my videos. Make it make sense.

The general opacity of the appeals process affects how creators like Stella interact with the platform. As in this instance, a successful appeal does not always translate into a better creator experience, especially when lags in restored platform access can be read as continued "punishment." Moreover, Stella interprets their mixed experiences with the

appeals process as a contributing factor to the platform's refusal to grant their account a verification badge. There is extraordinarily little information about the specifications required for verification on TikTok; however, TikTok does specify that accounts must adhere to the platform's Terms of Service and Community Guidelines in order to be eligible for verification. For creators like Stella, this eligibility condition feels unfair because they believe they're being held accountable for TikTok's frequent moderation errors. For verification purposes, creators often feel that there is a higher bar of eligibility for marginalized creators, which speaks to the impact of sneaky feelings about TikTok uneven support and amplification of certain creators over others. The concern that the unevenness of TikTok's verification is heightened by the platform's failure to track its own errors adds to the frustrations that creators like Stella feel. In terms of time and TikTok, the perception that TikTok holds creators responsible for its errors contributes to feelings that marginalized creators are predetermined to fail on the platform as a result of the platform's apathy and lack of accountability to its marginalized creators. Indeed, TikTok specifies that multiple violations result in greater likelihood of future account bans (see Figure 5), but again, there is no policy language that clarifies how platform errors factor into a user account's record.



## My content was removed

If you believe your content was incorrectly removed, let us know by submitting an appeal. Your feedback helps us improve the ways that we keep our community safe.

To submit an appeal:

1. Locate the notification in your TikTok inbox.
2. Tap on the notification.
3. Tap **Submit an appeal**.
4. Follow the instructions provided.

Or

1. Go to the video.
2. Tap **Community Guidelines violation: See details**.
3. Tap **Submit an appeal**.
4. Follow the instructions provided.

Note: If you delete appealed content that is under review, we may not be able to remove the violation from your record if your appeal is successful. Your content will also not be restored if you delete it after submitting an appeal.

Please be aware that multiple violations could result in account penalties.

### **Figure 5. Screenshot of TikTok’s instructions on appealing content removals from TikTok’s Account Safety**

Stella’s perception that the appeals process did not seem to help creators resonated in other creator testimonials, particularly those shared on Change.org. Rosalynne Montoya, a trans activist and content creator who uses TikTok to share educational videos about the trans community, shared her experiences receiving a permanent ban that was lifted after her appeals in a petition she created for Change.org. She traces her problems on TikTok to “transphobic trolls” mass-reporting her videos, sharing that “the vast majority of every TikTok video that was reported and removed was placed back on the app after [she] appealed them, because no rules were broken.” Summarizing her experience and the experiences of creators like David, Stella, Mae, and Eric, “the violation reports added up enough to get my TikTok account with a third of a million followers deleted. This is a MAJOR FLAW.” Rosalynne’s account was restored and even verified later. However, even as a verified creator who has more than doubled

the number of followers she had when her account was first permanently deleted, Rosalynne still receives live-stream bans and temporary suspensions routinely.

Mae shared that she is frustrated that it is obvious that a certain number of bans and violations results in a permanent ban, which she experienced herself on her first account, but that it is unclear how many violations and bans it takes to trigger a deleted account. She explained that knowing specifically how many Community Guidelines violations trigger bans would help her to understand the situation better. A week before our interview, Mae had been banned after receiving a content violation on a video that was ultimately appealed. When we met, she was waiting out a temporary ban after receiving another content violation for a sex joke that was similarly appealed and reinstated.

If TikTok would explain like, “Okay, if you will get banned, if you get reported X number of times, then this will happen.” It always says “too many community guidelines violations,” but I want to know if they’re counting their mistakes. I just came off of a three-day ban for a video that was a joke about Cersei and Jamie Lannister, and you know, I just want to know. Are you counting this one or are you counting the one from before? I just want to know if they’re cumulative.

Taken together, TikTok’s written policy which states that repetitive content violations might result in account bans and the experiences of creators like Mae, Stella, and Rosalynne among others seem to indicate that such violations are tracked cumulatively. In terms of cumulative user experience, each of these creators has intuited ways that the platform works based on their experience in the absence of more explicit policy or communication with human representatives of the platform. At the same time, each of

these creators shared that while they believe that the content violations are counted against them cumulatively that they do not feel that this impacts non-marginalized creators who are less likely to be the victims of mass-reporting or discriminatory moderation. Such feelings led creators to consider their futures on TikTok and as content creators more broadly and to take preventative measures like beginning to develop platform presences on other platforms like Instagram and YouTube shorts and to create backup accounts. These measures demonstrate the temporal and material implications of platform disruptions for content creators.

### **Platform Exploitation**

These creator experiences highlight the common suspicion creators have that content moderation is connected to the success they have on the platform, either as a result of increased exposure to more hostile audiences or as part of a larger problem of TikTok's exploitative relationship with content creators, particularly those who are monetized through the Creator Fund. Two creators, Mae and Eric, who are both Black creators, reported experiencing this type of escalation in content moderation experiences that ultimately resulted in having their accounts permanently deleted. Eric shared that right before his account was deleted, he had been attempting to email TikTok about issues he was having moving money from the Creator Fund that he had earned. At the time that their accounts were deleted, each creator lost access to the content they published on TikTok in addition to money they had earned in the Creator Fund.

When discussing his experience of losing money in the Creator Fund when his account was deleted, Eric shared that he attempted to reverse the permanent ban he received in a few different ways. While he reached out to followers and mutual creators

to advocate on his behalf to the platform, Eric also reached out to the platform, from whom he usually received an automated response. At one point, however, Eric got in touch with an actual human representative by email who responded that he would look into Eric's case, and according to Eric, he acknowledged pieces of evidence supporting his claim that his account didn't deserve to be permanently banned. After sending a few emails back and forth, the representative ceased responding, and Eric's account was not restored. Eric decided to reshoot old videos and begin a new account. After a few months of steady growth until people who followed his old account rediscovered him on the platform. From there, his new account grew until it outgrew his old one.

Despite his ability to recover from having his first account permanently banned, Eric still feels that the platform is rigged against creators, especially Black creators. Likewise, Mae shared that she immediately started to submit appeals to the platform. When the formal appeals process did not help, she began sending emails in the hopes of receiving human assistance. "It sucked," Mae shared, "because I had just gotten into the Creator Program [when my account was banned] and I had like \$400 dollars in my account. At the time, the only time you could withdraw money was 30 days after the end of the month." The policy, shown in Figure 6, that Mae described is TikTok's current funds withdrawal policy and still does not address what funds creators are entitled to should their account be banned before withdrawing earnings.

Because of the frequency of platform bans and suspensions, content creators often have backup accounts as a safety net. Creators use backup accounts, or secondary accounts that they create to use during temporary bans or to prepare permanent bans, to maintain engagement with followers when they are unable to post or comment on their

primary accounts, to challenge wrongful platform bans and suspensions, and to set themselves up if they have to rebuild their platforms. Stella shared that they used their backup account to communicate with followers asking for more videos during stretches of time when the platform suspended their ability to post new content or even respond to followers in the comment sections of their videos. After her first account was permanently suspended, Mae tried to use her backup account to get her account restored. She shared that “it took a month for them to respond to [her] at all” and that all she was told was that “certain behaviors could cause an account ban.” and that her account was flagged for nudity and pornography. After this brief interaction by email, Mae did not hear from anyone else at TikTok again and pivoted to making her backup account a viable platform for her to use. Since restarting her profile and encountering more platform resistance like she did in her first account, Mae has withdrawn from the Creator Fund to see if that improved her situation, encouraging followers to pay her directly when they can through Venmo and Paypal. In a sentiment shared by Eric, another Black creator, Mae explained that she felt that TikTok was more “strict” with Black content creators once they started to have more success associated with the fund.

# When will I get paid?

Funds are available to transfer to your linked payment method no later than 30 days after the end of month in which video views were accrued.

-Funds listed under August 2020 in your Creator Fund Dashboard will be available to withdraw on September 30, 2020.

-Funds listed under September 2020 will be available to withdraw on October 30, 2020.

-Funds listed under October 2020 will be available to withdraw on November 30, 2020.

-Funds listed under November 2020 will be available to withdraw on December 30, 2020.

-Funds listed under December 2020 will be available to withdraw on January 30, 2021.

There is a total withdrawal limit of \$3,000.00 USD per day. You can transfer money into your payment method a maximum of five times each day.

The minimum withdrawal amount is \$10.00 USD. If you have less than \$10.00 in your account, you can transfer the entire balance once from January 1 to June 30 and once July 1 to December 31.

Normally withdrawal requests will be processed within 72 hours. Please note that the financial institutions may withhold the withdrawal amount for an additional period of time at their sole discretion for security review purpose. It is not guaranteed that each withdrawal request will occur within 72 hours.

**Figure 6. Screenshot of TikTok's monthly funds transfer from TikTok's Creator Fund portal**

As described earlier in this chapter, TikTok often restricts user access to the platform following repetitive content violations with a view-only mode. This mode allows content creators to access the platform, but they are restricted from posting content or otherwise interacting with other platform users. Although it would be possible to interpret this view-only mode as a lesser penalty, creators reported in interviews and in testimonials intense feelings of frustration when they received restrictions that limited their ability to comment, post, or live-stream even when they were still able to otherwise use the platform. Such restrictions limit creator abilities to engage their followers, many of whom come to expect certain levels of daily engagement and interaction. Content

creators felt even more frustrated when they found that they were restricted from using the platform while the videos they created and uploaded to the platform were still accessible to be viewed by other platform users.

For example, David, a man of color who does not appear in his videos but instead speaks through a puppet, shared this with me: “When I was banned, people in the comments said I was still showing up on their For You pages. My content was still getting viewed, so TikTok was still benefiting from me existing, but I did not get paid for those days. I got zero dollars in the creator fund for those days.” Other creators have shared similar stories about their content continuing to appear on the For You page while they themselves were banned from the platform, reporting that such occurrences heightened feelings of frustration with the platform. David emphasized that he found this exploitation of his work especially frustrating and disappointing considering how much growth TikTok has had as a social media platform because of the work of its content creators.

Likewise, Stella shared this story with me to demonstrate how this exploitation impacted creators financially but oftentimes hurt as the content shared to the platform reflect their personal stories:

One time I was banned from commenting on posts for a day or two. Anytime someone had a question, I had to go on my secondary account, and then answer questions there. The worst part was that it was Epilepsy Awareness Day, and TikTok was promoting my videos about [my experiences with] epilepsy when they had also comment-banned me. I could do anything, which is infuriating. I

was like, you're using my content to promote this day, seem supportive, and you also just silenced me.

Being silenced while TikTok circulated Stella's videos frustrated them, and while they were still able to post videos, they felt that doing so in response to this problem would invite another suspension that they couldn't afford at the time. In this way, the threat of future bans and restrictions prevented Stella from speaking up about what felt like an exploitation of their lived experiences and the content that they worked hard to create.

### **Disruptions in Creator Time**

Across interviews and online testimonials, content creators reported that disruptions in platform access limited their ability to monetize their labor on a platform that requires constant engagement to be successful. While the previous finding focused more on the feelings of being forced to sit on the sidelines while the platform continued to benefit from their labor while their access to the platform was disabled, this finding related more to perceptions held by creators that time away from the platform was time lost when they could be growing as creators, developing their brand as content creators and increasing their follower counts. As Zev, a white Jewish creator who has dedicated much of his platform to challenging TikTok's management of hate speech on the platform, put it: "when you're a creator, it's post or perish." Even a single day of inactivity can be detrimental for content creators seeking to monetize their platforms, in addition to the long-term effects of these patterns that create anxiety for content creators.

Likewise, Zev shared that although people in general are often emotional about their experiences of exploitation with TikTok, creator concerns are oftentimes more



material. Indeed, while creators often receive the advice to shift to other platforms, like Instagram Reels or YouTube Shorts, Zev explains that it isn't that simple:

No platform has the amount of organic reach that TikTok has. Me, personally, I could never be an Instagram influencer. Like, look at me, I could never be an Instagram influencer. Like, look at me, I'm a scrawny Jewish kid. I can be a fun, scrawny Jewish kid who cares about social issues, and that'll be put onto the For You page. The point I'm trying to get here is that this isn't just about like, wanting to create content and being mad that our content is being used. It is being mad that they're taking away our ability to make money.

Across my interviews, content creators reported what David described as the constant anxiety and dread that the "Big One," or a permanent suspension, was just around the corner: "With TikTok being the way that it is, mental health for creators is at an all-time low. I mean, you have no idea if tomorrow is going to be the day, the platform you spend your whole life to build that your dream, the dream job that you already have is going to disappear for good." The cumulative experience of being repetitively banned unfairly fostered in David and other creators a type of precarity that sticks to their daily interactions with the platform. Such precarity impacts the type of content these creators make as they try to avoid triggering "the big one" and necessitates that creators make back-up accounts on TikTok and other platforms. Interacting with the platform amidst ongoing threats of an inevitable permanent ban

Despite the affordances of TikTok described by Zev, many creators, including David and Stella, who have experienced repetitive account bans are creating platform presences on Instagram and YouTube to prepare for the permanent ban that feels

inevitable. After several months of receiving unpredictable account bans, David shared that he came to the realization that “TikTok does not give a fuck about its creators,” comparing his experiences on TikTok to his better time on YouTube shorts. David shared a wide range of ban experiences and difficulty understanding the appeals process:

I've gone on to completely have my entire account deleted once and then I pressed the appeal button and 24 hours later it was back and they said we've accepted your appeal. And they and then even then they restored the video that they took down that was against me violent guideline violations. And they still left me with a one week posting ban, with this absurd text that reads, “we are so sorry, your videos restored, you are still banned from doing this.” Like it's the same text that would if you got a violation. And then one week later, I got my account back, and five hours after I got my account, I was banned again, except this time, they denied my appeal. And then magically three days after that my account was back. It has been a very frustrating journey.

## **Discussion**

As mentioned at the beginning of this chapter, my findings indicate that the stresses associated with platform restriction are felt in relation to the way time works on TikTok. TikTok’s algorithmic design which centers the For You page as a source of constant innovation and new trends necessitates that creators invest in the platform with daily engagement to stay on top of the trends and connect with followers whose engagement amplifies creators’ content in the FYP. In terms of time, there is a sense that everyday brings new content; the present is exceedingly important for content creators, as being able to consistently participate is a key part of the job. As a result, the design decision

around segmented bans facilitate creator feelings of anxiety, stress, and frustration, feelings that are emphasized by the helplessness that comes from not understanding platform judgments and/or by not being able to reach platform representatives. Put another way, the algorithmic culture on TikTok creates a specific culture of time on the platform in which creators especially feel obligated to be constantly at work on the platform. To revise the common phrase, time is engagement, and engagement is money, or as Zev riffed during our interview: “You know how it’s like publish or perish in academia? It’s post or perish for us.” The feeling of being stuck at a standstill is particularly frustrating for creators, especially when they’re able to see the platform continuing to benefit from their labor via their content that continues to circulate and gain views while they’re banned from publishing. Therefore, time is critical to the work of content creators on TikTok at the base level of engagement. Additionally, these feelings of frustration are heightened when the cyclical bans seem to largely be experienced by creators from historically marginalized and underrepresented communities, leading many creators to believe that the platform doesn’t want marginalized creators to be as successful as white, cis, and straight creators on the platform.

Content moderation is most often sensed when something goes wrong for content creators. In this chapter, these content creator stories demonstrate the lasting effects of habitually wrong content moderation on creators. Platform errors became more egregious to participants who felt that their experiences were part of larger patterns of individual and collective exploitation and exclusion of content creators from historically marginalized communities. Creators have a decreased sense of trust in TikTok’s platform governance in no small part due to the impact of content moderation errors on creator

monetization, and this lack of trust has an immense impact on how creators perceive and respond to glitchy content violations. From the various petitions related to TikTok's censorship of marginalized creators and the collective grievances shared in each of the individual interviews I held, there is a sensibility that creators make sense of their experiences in relation to the stories shared by other creators. Moreover, this chapter highlights several aspects of content moderation experience that creators do expect. Creators expect more accurate algorithmic detection of both violative content and patterns of reporting that indicate the abuse of user-end platform tools to censor and suppress creators from historically marginalized backgrounds. In the event of content moderation inaccuracy, content creators expect timely communication and resolution of content removals and platform suspensions with the confirmation that errors made by the platform will not be held against creators going forward. That is, creators expect that the appeals process will not just result in the restoration of content but also clear communication from the platform that overturned content violations won't be counted against the creator's record. Although TikTok does not explicitly state that creators have a record, the platform does explain

In addition to these expectations creators have for the user experience of moderation on TikTok, I believe that their frustrations with platform bans and restrictions reveal additional tensions in creators perspectives about platform engagement that heighten and complicate their reactions to TikTok's faulty moderation. The denial and disruption of platform access is often felt as the loss of opportunity for creators, both in terms of the material loss of income generated in the Creator Fund or through sponsored content and in terms of engagement and growth. While engagement and growth as

metrics for creator success cannot be separated from the compensation of digital labor, the creators I have spoken to do not discuss their desire for engagement and growth in terms of monetization. Engagement and growth also serve as evidence of creator success on the platform. I suggest that the opacity of TikTok's platform governance affectively contributes to the frustration that content creators feel that their content would otherwise be more successful in the alleged algorithmic meritocracy of TikTok if they were not targeted by the platform, intentionally or otherwise. To describe TikTok as a meritocracy would be to ignore the platform's influence in creator success, as well as the impacts of race, gender, sexuality, class, and ability on content success; however, the creators I spoke with seemed to struggle with the impact that platform moderation started to have on them after they had achieved a certain level of success on the platform despite the obstacles they faced as marginalized creators. In other words, the sneaky rhetorics of content moderation contribute to feelings that the platform's interventions are unfairly preventing these creators from being successful. To be clear, I believe each of the creators who have spoken out about their experiences on TikTok; however, I believe it is worth considering the impact of previous virality on creator expectations of their experiences of TikTok. In other words, it is possible that creators develop expectations for how well their content should perform based on past success.

In this way, the conversation of platform bans and restrictions cannot be separated from creator perceptions of shadowbans, which are often felt as sneakier types of top-down platform governance. Stella shared that while they don't believe the platform strategically shadowbans content creators, they do believe the platform "love-bombs" creators by making engagement more achievable early on so that they become invested in

creating for the platform. Love-bombing was coined originally to describe an early stage of an abusive relationship in which an abuser showers a partner with attention and affection to manipulate them; while I would perhaps be careful with the use of such a term outside of its original context, I found Stella's assessment compelling as it usefully contextualizes both the role of emotion in content creation within algorithmic environments and the affective experience of reduced engagement after reaching the height of virality on a platform like TikTok. That is, Stella's description of love-bombing aptly captures both the highs and the lows that content creators feel at the opposite ends of the virality spectrum:

Starting out on TikTok, you'll make a few videos, and then all of a sudden, like, boom! You go viral, and then—*Oh, my God, like, I just had my first viral video. I am going to continue making videos because like, I want that I want that validation again, and I want to feel like that.* And then they take it away, and your views plummet. You're like, *Oh, no, but like, I want that again,* so you keep making more content until you get it again, and then they give it to you a little bit. Maybe your views are not quite where they were, but like they give you a taste. Have it again. And it's like they sort of just like keep you, just sort of like dangling like views over your head.

In this description of engagement-bombing, Stella provides an explanation for the inconsistent experiences of success that even the largest creators deal. Specifically, their suggestion is that TikTok deliberately manipulates the visibility of content in the For You feed so as to encourage creators to continue innovating and producing to capture that specific feeling of success and the attention of a fickle audience. While Stella has had a

great many viral videos on the platform at this point, they note that they have never had two viral videos back-to-back, sharing that their engagement is always “up and down.” While I do not doubt that creators like Stella experience reduced visibility in the algorithm by the lasting impacts of targeted reporting and faulty moderation, it feels worth examining how prior experiences of virality affect periods of reduced visibility and engagement, particularly for content creators and platform users who have lived experience of marginalization and erasure on digital platforms like TikTok.

These findings contribute to conversations in technical and professional communication and digital rhetoric that highlight how systemic oppression is coded into and maintained by digital technologies (Noble, Benjamin, Chun). At the same time, this study also adds to emergent conversations about the need for more attention to the relationship of emotion and affect to social media, highlighting how time is key to how social media users make sense of and respond to digital interfaces and their cultural values. Anticipatory and cumulative UX offer scholars in technical communication a useful framework for interrogating power, privilege, and oppression in design that spotlights how user feelings and perceptions of technologies are informed by a range of subjective and situational factors before, during, and after immediate UX interactions.

Several creators shared that the frequent interactions with content moderation on TikTok impacted their mental health. Importantly, each person I spoke with viewed content moderation as a stepping stone to a more stable career, citing the instability of content creation on TikTok as the prime reason for wanting to make enough money and connections on the platform that they could afford to leave. This finding is importantly contextualized by the fact that these interviews took place during the ongoing COVID-19

pandemic, and that all but one of these creators started creating content on the platform for both enjoyment and to supplement their incomes when the pandemic first started.



## **CHAPTER 5**

# **COLLECTIVE COMPLAINT FROM THE MARGINS: HOW MARGINALIZED CONTENT CREATORS COUNTER CONTENT SUPPRESSION ON TIKTOK**

### **Introduction**

In the previous chapters, I examined creator experiences with inconsistent and uneven moderation and platform bans and restrictions, highlighting how such experiences reveal the affective impact of time disruptions on the platform on creators as everyday platform laborers. While creators grapple with and navigate shifting platform expectations, policies, and procedures through the platform's interface, they also interact with and develop relationships with other creators and users. These interactions facilitate platform growth, as creators build support networks of creators to collaborate with and learn from other platform users about how to sustain engagement and interest on TikTok. Likewise, these interactions are foundational to the subject of this chapter: the public and private responses of creators on TikTok to the platform's content moderation infrastructure and policies. Thus, in this chapter, I explore the following questions:

- How do TikTok creators speak back to discriminatory experiences with TikTok's content moderation?
- Why do content creators speak out about their experiences with content moderation on TikTok?

- What strategies do content creators use to render their experiences visible and legible?
- How do TikTok creators work together and with platform users to navigate and challenge TikTok's content moderation problems?

To answer these questions, I begin with a discussion of Change.org petitions published and circulated regarding TikTok's content moderation and its impacts on marginalized content creators, before analyzing data collected from reports about TikTok's content moderation controversies and my interviews with content creators. As discussed in chapter 3, creators often make adjustments to their content and platform behaviors to avoid content moderation: censoring themselves by making changes to their personal appearances, censoring their speech in captions and audio, and avoiding certain topics altogether that seem to trigger moderation. This chapter delves into creator practices beyond these personal habits of modifying content, such as educating platform users how to navigate content moderation and support marginalized creators on TikTok, networking with other creators for support and solidarity, creating back-up accounts to prepare for account bans and restrictions, and speaking out against discriminatory content moderation on TikTok and other platforms such as Change.org.

This chapter also relies on Sara Ahmed's critical work on complaint as a feminist methodology to examine oppression within institutional life. As Ahmed notes, "To make a complaint is often to find a gap, a gap between what is supposed to happen, in accordance with policy and procedure, and what does happen. That gap we learn is densely populated" (p. 30). In using Ahmed's work on complaint as central in this

analysis, this chapter examines what happens when content creators find *gaps* in TikTok's platform policies.

### **Using the Platform to Teach Other Platform Users about Content Moderation**

One way that creators respond to their experiences of uneven content moderation is by creating TikTok videos that teach other platform users about content moderation. That is, I have observed a larger pattern of content shared on TikTok in which content creators use the platform to encourage and instruct other platform users how to interact with the platform's interface and algorithmic design to disrupt platform oppression for marginalized and underrepresented content creators. I discuss two examples of this genre of TikTok video below.

As I discussed in the methods chapter of this dissertation, I first was made aware of these gaps between what is supposed to happen based on policy and what does happen in my personal use of TikTok as entertainment. While scrolling through my For You page and through content uploaded by creators I intentionally followed, I encountered countless instances of content creators using the platform to speak out about their experiences with content suppression and moderation and to use their videos to instruct other users on the platform about how to interact with content created by marginalized and underrepresented creators to help counter discriminatory moderation practices. For example, I learned that it was important to watch videos on TikTok all the way through to support creator success in the algorithm through a video posted by someone I followed on my personal account. This theory has been confirmed by an extensive study conducted by the *Wall Street Journal* which identified attention as measured by length of time watched

or repetitive viewing of the same video as the single most important factor for video success on the platform.

As another example, a Black queer content creator shared a video explaining how to submit a help ticket in the mobile app interface so as to mobilize their followers on behalf of another Black creator who had their account banned. Similarly, creators often make instructive videos about using the platform's reporting tools to get content that is dangerous or harmful, like hate speech or misinformation, removed, using their platforms to mobilize fellow users to encourage the platform to carry out its moderation policies. Such videos oftentimes speak both to platform users to mobilize action while also speaking directly to TikTok's failure to catch harmful content without being pushed to do so by the same marginalized platform users who are subjected to routine inconsistent and uneven moderation. In their work on platform governance and creator culture, Cunningham and Craig have identified such practices as "creator practices of sustainability and risk maintenance [that] operate as a form of bottom-up governance" (p. 275). The impact of such creator practices on TikTok are undeniable as strategies designed to avoid algorithmic detection popularized by creators--such as coding language believed to trigger algorithmic detection or commenting single words such as 'boost' on videos to increase the chances of seeing the creator or topic in one's For You page--have become mainstream practices for creators and users alike on TikTok. It is also clear that on TikTok, as on other platforms, that creators with large followings often are able to successfully campaign for the removal of harmful content and creators that are otherwise missed by algorithmic and human moderators. At the same time, none of the creators I spoke with expressed any confidence that educating platform users about how to use the

platform in such ways are enough to negate the impacts of discriminatory platform practices on marginalized and underrepresented creators or that such efforts at amplification and awareness-raising are enough to motivate the platform to address the material concerns of creators on the platform.

### *Overview*

In this chapter, my central claim is as follows: the lack of transparency in TikTok’s content moderation user experience fosters an affective environment in which creators are compelled to speak out and across creator networks about such gaps in experience and platform policy. As such, this chapter examines creator practices of sustainability and risk maintenance not solely as a form of bottom-up governance or as strategies designed to hold the platform accountable to its biggest creators. Rather, I discuss these strategies as affective responses to the felt experiences of inequity on the platform that is oftentimes rendered invisible because of the platform’s algorithmic recommendation interface. In other words, the experiences of suppressed or removed creators and their content is difficult to notice amidst the stream of endless recommended content available for users on the platform to scroll through on a daily basis.

Bringing together Zeng and Kaye’s conceptualization of visibility moderation into conversation with Ahmed helps to make sense of the affective implications of governance that operates on the level of visibility and voice. Indeed, as Ahmed writes, “To make a complaint you have to *keep* making the complaint, to give it *voice*, to give it expression. A complaint can be experienced as the requirement to *become* expressive” (p. 34, emphasis added). Thus, in this chapter, I describe how such experiences described in

chapters 2 and 3 affectively compelled creators to become expressive by speaking publicly against their discriminatory platform experiences and by networking with other creators to make sense of and react to their interactions on the platform. Such expressiveness takes several forms, as I discuss later in this chapter. Importantly, many of the types of expressions that I discuss in this chapter do not happen through formal channels sponsored by TikTok because, as I discuss in chapters 2 and 3, such channels are especially limited.

Therefore, by analyzing creator stories shared online and in interviews, I find that creator experiences of content moderation, influences creators to speak publicly about their experiences and to build networks with other content creators so as to better understand their own experiences at first and then over time to feel less alone while grappling with a hostile platform and the daily tasks of content creation as labor (such as creating content on a daily basis, interacting with followers, networking with brands, etc.). Visibility moderation, as a form of content moderation, renders their interactions with content moderation and suppression invisible to other users on the platform. Put

In doing so, these content creators illuminate trends in platform experiences for marginalized and underrepresented creators on TikTok, as well as share insights into the evolving mechanisms of platform governance with fellow creators and users. By exposing such patterns through personal stories of suppression, these creators both disrupt and critique the sneaky rhetorics of content moderation on TikTok that is highlighted by the platform and its lack of transparency in governance.

In the section that follows, I begin with a discussion of the strategies used by creators to speak out publicly against platform discrimination, bringing together findings

from my analysis of Change.org petitions published by creators and data from my interviews with creators. From there, I then shift to a discussion of the more private ways that creators organize with one another in response to TikTok's treatment of marginalized and underrepresented creators. In my analysis, I emphasize the importance of the emotions motivating creator responses to discriminatory experiences of content moderation and suppression, arguing that their responses have a lot to teach us about the role of emotion in how everyday users and creators affectively interpret and navigate digital platforms like TikTok, particularly when such platforms facilitate interactions that leave creators feeling "harassed," as multiple creators described it, by nameless, faceless platform forces. As I will discuss in the remainder of this chapter, the frustrations of feeling harassed, while experienced individually, are compounded by the perception that this is commonly experienced by their friends and people like them.

### **Speaking Out as a Response to Silencing and Suppression**

In this section, I shift my discussion to focus primarily on findings from my interviews that focus on the impact of opaque content moderation on TikTok on content creators networking with one another on the platform to do the following mitigation strategies

- share strategies to avoid and counter suppression on TikTok
- make sense of emergent content moderation policies and practices
- consider professional options outside of TikTok (such as YouTube Shorts and Instagram Reels),

- and commune with creators who understand the particular experiences of targeted platform moderation and know that these experiences are real, important, and reflect larger platform problems.

I also show how creators turn to other platforms such as Change.org to speak out about their experiences as a way to try to amplify possibly suppressed content by garnering enough visibility to platform users and the platform itself, and to have unfair platform judgments reversed, including? and? to regain control of banned accounts.

Additionally, amplifying personal experiences of TikTok's content suppression on other platforms like Twitter or Change.org has proven to be an important strategy for content creators who believe their content is being unfairly moderated or suppressed. The willingness of creators to speak out about these experiences have contributed to the subject of TikTok's problematic algorithmic curation and moderation gaining mainstream media attention (Botella, McCluskey). Representatives for TikTok have often been pressured to respond after mainstream stories of suppression or censorship on the platform have gained wide enough attention. During our conversations, creators shared with me that the move to discuss their experiences both on TikTok and other platforms was a final resort; they spoke out in TikTok videos and social media posts on other platforms like Change.org after formal processes in the platform either failed to reverse discriminatory judgments or were met with platform silence. Within a few weeks of first researching content moderation on TikTok, I encountered several Change.org petitions shared in creator profiles on TikTok, including one created by Zev, who participated in this study. Change.org promotes itself as a platform for people to use to share stories to



advance their causes, mobilize action and support, and create change at the local, national, and global level every day.

In what follows, I turn to discuss data from my analysis of Change.org petitions. In the table below, I have included some of the more popular petitions related to TikTok’s content moderation, which I analyzed for this dissertation as discussed in chapter 2. I highlight these petitions not as evidence of overwhelming support for these causes or to demonstrate how many people believe that content moderation on the platform is fundamentally discriminatory but to highlight that there is a significant contingency of platform users who believe that content moderation is inherently flawed on TikTok at the expense of marginalized creators.

<b>Date Created</b>	<b>Title</b>	<b>Signatures</b>
4/25/2020	TIKTOK STOP PUNISHING BLACK CREATORS FOR SPEAKING UP!	51,703
5/12/2020	Tell TikTok to abide by their own standards and protect those with disabilities	150,750
7/4/2020	Stop Removing Mental Health Awareness Content on TikTok	34,296
12/10/2020	Stop Antisemitism on TikTok	24,812
12/15/2020	Change TikTok’s Community Guidelines Algorithms	20,865
4/9/2021	Call on TikTok to Release the Content Moderation Guidelines	18,639
8/23/2021	Petition to Unban popular creators on TikTok and reform the Banning & Appeals process	16,522

**Table 2. Top 8 Change.org petitions about TikTok’s Community Guidelines and Content**

**Moderation**

Although Change.org markets itself as “the world’s platform for change,” petitions published to Change.org may best be described as awareness-raising tools that can spotlight social campaigns for news media and decision makers in various positions of power (Change.org). One of the key features of Change.org’s design is that it allows petition creators to tag national organizations and individuals. These petitions regarding TikTok’s content moderation certainly fit this assessment with each being shared with TikTok and large national organizations and powerful individuals, such as the U.S. Department of Health and TikTok CEO Vanessa Pappas among other stakeholders the authors of the petitions thought were relevant in raising awareness or effecting change on the platform. Despite this, my assumption based on the circulation of these petitions on TikTok by creators is that the primary audiences for each of these petitions are TikTok creators and users who support marginalized creators on the platform, as well as TikTok itself and journalists seeking to cover these issues. During our interview, Zev shared that he hoped that his petition would gain enough traction that he could point to it as a sign of the issues importance when speaking with journalists about content moderation on TikTok.

My analysis does not examine these petitions in terms of their effectiveness to generate conversation about the problems facing creators on TikTok, although each of the petitions I discuss has had significant reach online as evidenced by the tens of thousands of signatures each of these petitions has garnered. Rather, I discuss these petitions as supplemental to the testimonials gained through interviews, and as I’ve discussed in previous chapters, these petitions provide great insight into how platform users understand the range of content moderation experiences that are not explained formally in

policy or through communications with the platform. Moreover, I argue that these petitions represent both a strategic rhetorical means of countering the problems of content moderation on the platform and an affective response to the isolating and frustrating experiences of inconsistent and uneven content moderation which are otherwise invisible to other users on the platform.

While each of these petitions approaches the problem of content moderation on TikTok with unique concerns, each petition discusses the disproportionate impact of content moderation on TikTok on marginalized and underrepresented creators and communities on the platform and connects this problem to specific features of the user experience of TikTok. While analyzing these petitions, three key trends emerged. The first trend is that each of the petitions, except for the one specifically invested in the removal of mental health awareness content, pointed to the prevalence of hate speech as a reason for the petition to be created and as a sign that change is needed on TikTok. For example, the petition, “Tell TikTok to abide by their own standards and protect those with disabilities,” was created in response to a harmful trend popular on TikTok that mocked a disability. This petition, which was amplified to TikTok as a decision maker, references the platform’s Community Guidelines definition of hate speech to make the point that the platform is failing users on promises it has made. All in all, four of these petitions gestured to the Community Guidelines to argue that TikTok’s moderation does not align with its own rules as written for users to follow, while one of the other petitions gestured to the community guidelines and policies of other platforms as a way to argue that TikTok should have written policy that disallows Holocaust denialism on the platform. This point illustrates the way that creators do not examine the policies and

practices of different media companies and their approaches to content moderation in isolation.

These petitions also made it clear that the prevalence of hate speech and violative content on TikTok is significantly more problematic when marginalized and underrepresented creators are wrongfully targeted by the platform, particularly when speaking about marginalized people's identities and experiences or speaking against oppressive conditions on TikTok or elsewhere[?]. For example, the petition specific to antisemitism on the platform focuses specifically on the allowance of antisemitic hate speech on the app that frustrates Jewish creators who oftentimes find their videos about being Jewish and antisemitism held in review for long stretches of time before being published or removed after being flagged incorrectly as hate speech. The author of the petition included a list of changes TikTok could make to better train that human moderators and algorithmic software needs to identify antisemitism on the platform, such as the coded antisemitic language used by white nationalists and mass-reporting techniques used to silence Jewish creators and other marginalized creators.

Interestingly, another significant trend that emerged in these petitions was the rhetorical move to share personal stories of suppression alongside demands for changes to the platform's algorithm, standards, and protocols. Of these seven petitions, three began with a detailed personal story of the author's experiences as a creator on TikTok ("Stop Removing Mental Health Awareness Content," "Change TikTok's Community Guidelines Algorithm," and "Petition to Unban popular creators on TikTok"). For example, Rosalynne Montoya began her petition ("Change TikTok's Community Guidelines Algorithm") with a discussion of her experiences. To highlight how she

connects her experiences to the Community Guidelines and the larger problems on TikTok, I will share this extended excerpt from her petition below:

The TikTok algorithm for their community guidelines targets and censors marginalized creators ( ie. that are BIPOC, LGBTQPIA+, and/or living with a disability, etc). My account @RosalynneMontoya with 300K+ followers was just deleted along with many many others. I use my social media to spread education and love. I've never violated TikTok's community guidelines.

I had transphobic trolls report my educational videos about the trans community and my own transition until they were removed from TikTok. Trolls should not have more power than creators. The vast majority of every TikTok video that was reported and removed was placed back on the app after I appealed them, because no rules were broken. These instances shouldn't count against me. After a dozen or so of my videos and live streams were reported and removed by transphobic trolls (even if appealed and deemed to not violate CG), the violation reports added up enough to get my TikTok account with a third of a million followers deleted. This is a MAJOR FLAW. (“Change TikTok’s Community Guidelines Algorithm”)

In this petition, Montoya uses her experiences as the first piece of evidence that she has been algorithmically censored in ways that do not align with the platform’s Community Guidelines. She uses links to provide samples of both the type of content that has made her account popular on TikTok, such as educational videos about the erasure of trans history and videos that affirm trans people on the platform. She also links to videos that have been reported and removed on TikTok, such as a video about the loss of knowledge

about trans medical care as a result of Nazi book burnings in the 1930s (Montoya), and a video discussing her experiences of having the right to live stream on TikTok removed as a result of “transphobic trolls” reporting her content. Through these linked videos that illustrate Montoya’s experience of being a marginalized creator on the platform, she shows how both the Community Guidelines, and the moderation algorithm were at odds, resulting in the deletion of her account. In the remainder of Montoya’s initial petition post, she calls for TikTok to update the algorithm that ensures the guidelines are upheld, to become more transparent about how TikTok enforces its guidelines in a way that does not disproportionately target marginalized creators, and to “create accountability by ... incorporating an advisory board with creators that are members of marginalized communities.” Such a recommendation speaks to a belief echoed by other creators that TikTok needs to demonstrate its commitment to creators by including them in decision making processes (while also compensating them for such efforts, expertise, and time).

Interestingly, several of the petitions not only highlighted personal stories of content suppression on TikTok of the petition authors but brought together the personal stories of several creators, with one petition inviting creators to share their experiences to be shared in updates to the Change.org petitions “Stop Antisemitism on TikTok,” “Change TikTok’s Community Guidelines Algorithm,” and “Petition to Unban popular creators on TikTok”. For example, the petition, “Stop Antisemitism on TikTok,” begins with a collective statement of frustration and a petition for “the app to crack down more on the casual and blatant antisemitism running rampant.” After outlining a list of demands to improve the safety of the app for Jewish users and creators, the petition provides a series of personal stories and user handles of Jewish creators about various

experiences with hate speech in the interface, as well as suppression of Jewish content. Similarly, the petition, “Unban popular creators on TikTok,” begins with personal stories from the petition’s author, Dierdra, a Black woman content creator on TikTok, and another Black content creator named Anthony, both of whom shared difficulties with platform bans, the weaponization of mass-reporting and the appeals process for reversing judgments related to wrongful content removals, and platform suspensions. They share their experiences as representative of the types of treatment that Black content creators face on TikTok, writing that “[s]o many other Black Creators are being targeted, having their accounts permanently banned when they speak about social injustices, blatant racism and pedophiles on TikTok.” As supplemental evidence, the petition shares the handles of several Black content creators who have faced these struggles.

At the core of their frustrations, Dierdra discusses how frustrating and exhausting these challenges are for content creators who are trying to monetize their work on the platform. Below, I include an excerpt from this petition that captures how intensely these problems are felt by some Black creators, particularly when such experiences of suppression and suspension are contrasted to the prevalence of hate speech and harassment on the platform:

**WE ARE TIRED OF TIKTOK ALLOWING BLATANT RACISTS TO  
THREATEN THE LIVES OF BLACK PEOPLE WITHOUT ANY  
REPERCUSSIONS!**

**WE ARE TIRED OF TIKTOK ALLOWING REGISTERED SEX OFFENDERS  
AND PEDOPHILES TO REMAIN ON THE APP!**

WE ARE TIRED OF TIKTOK ALLOWING PEOPLE TO BULLY OTHER  
CREATORS THAT HAVE DISABILITIES WITHOUT RECOURSE!

WE ARE TIRED OF HAVING OUR CONTENT REMOVED! WE ARE TIRED  
OF OUR ACCOUNTS WRONGFULLY SHUT DOWN FOR BEING A VOICE  
FOR THOSE WHO DON'T FEEL LIKE THEY HAVE A VOICE BECAUSE  
TIKTOK IS IGNORING THEM!

WE ARE TIRED OF CONTENT MODERATORS, HIRED BY TIKTOK,  
OVERSTEPPING THEIR AUTHORITY. THE APPEALS PROCEDURE IS  
FLAWED. THERE IS NOWHERE TO TYPE AN EXPLANATION OR  
SUPPORTIVE FACTS AND NEEDS TO BE REFORMED IMMEDIATELY.

(“Petition to Unban popular creators”)

This emphasis on a collective exhaustion points to the impacts of these experiences accumulating over time. The affective condition of being exhausted by discriminatory conditions on TikTok emerges as an important theme in this section and in the other petitions--a feeling that is compounded by the energy, labor, and time that it takes to create content for the platform, to navigate the frustrating appeals process, and to testify to their experiences on Change.org to try to lobby for platform change, as Dierdra notes in her personal story here. Many of these complaints echo the frustrations that I discussed in chapters 3 and 4, regarding the affective impact of inconsistent and uneven moderation and platform suspensions on content creators, but the move to create a petition speaks directly to the affective impact of the appeals process as well. While the petition targets human content moderation as flawed in ways that I cannot verify with absolute certainty in this dissertation, the complaint provides an insight into the frustrating design in the



appeals process that doesn't allow creators to provide personal testimony or evidence on their own behalf. Moreover, the problems of moderation and appeals on TikTok are directly connected to the demonetization and exploitation of creators on the platform, namely creators who have been targeted by mass-reporting and who consequently lost access to money in the Creator Fund at the same time that "videos from these creators are still up on TikTok, generating revenue that TikTok collects money from." The petition expresses a commonly shared frustration that the platform silences creators at the same time that their content is visible on the platform.

On the other hand, other cited Change.org petitions and updates to these petitions take the opposite approach by sharing not just personal stories of harassment, abuse, and wrongful platform judgments, but also by sharing stories of creators who have seemingly evaded the rules, as a way of demonstrating the way the platform privileges certain types of creators (read: creators who are white, cis, straight, wealthy, and able-bodied). For example, one testimony posted to Montoya's petition included a list of "white, cis, heterosexual" creators that the author felt violated Community Guidelines undeterred, while others pointed to notorious creator scandals on TikTok as evidence that some creators are held more accountable than others.

These types of personal stories and testimonials shared to Change.org represent one way creators try to highlight the injustice in their experiences by appealing to TikTok's own guidelines. In my conversations with creators and in my reading of online testimonials, I found that creators expressed that such public efforts are necessary, especially as they have been blocked from sharing their experiences on TikTok itself. In the first update to her first petition, Montoya shared that TikTok rejected her appeal to

restore her TikTok account, including a form message in the mobile interface which read that “it was determined that [her] account violates community guidelines and cannot be restored” after a review of her case (“My appeal was denied”). Montoya followed this update with another which shared that her story had been spotlighted in local Los Angeles media before posting another update that shared that she was able to regain access to her TikTok account, despite never hearing back from TikTok personally. In this post announcing that her account had been restored, Montoya shared that her next goal was to “Gather stories and share them on the petition” and “to sit down one on one with TikTok executive and discuss how they can help make their app truly ‘prioritize safety, diversity, inclusion, and authenticity,’” citing the platform’s community guidelines.

In addition to calling for specific changes to the platform’s moderation and governance to make the platform more equitable for marginalized creators, Montoya invited creators to share their stories with her via TikTok so that she could publish their stories as updates on Change.org. In the TikTok video linked in the petition, Rose faces the camera in front of screenshots of the notifications she received when her account was permanently banned and deleted. She has captioned the video and speaks deliberately to the camera as she shares a little about her experiences on TikTok, the exigence for the petition:

They consistently reported my videos until my account was deleted even though I never violated community guidelines. I would love to share other people’s stories. If you or someone you know has been wrongfully banned from going live, wrongfully deleted, or continued to be targeted, or if you have seen hateful content that breaks community guidelines that is not being removed from the app,

I want your screenshots, I want your stories! Send me a message, send me an email. I'd love to hear from you. I'd love to share your story on the petition. Because this is not about me. This is about changing TikTok's community guidelines to make the app a safer place for everybody.

In this call for stories, Rosalynne uses her experiences as an opportunity to connect with other creators who have similar platform stories. The call for screenshots and stories signals a need for evidence to support the claims in the petition. In the comment to this video, Rosalynne also emphasizes the importance of circulating her story so as to collect as much evidence as possible: "Petition link in bio. Please share and boost this video."

While calling for creator stories, Montoya describes content suppression experiences that mirror patterns described in chapters 3 and 4, including live-streaming bans, wrongful suspensions, and constant targeting from the platform and hostile users. While Montoya does not explicitly describe the reason that she wants to collect stories as a move to generate evidence, I find the move to collect screenshots and stories an interesting one that speaks to how creators navigate the lack of transparency in the platform. Likewise, in calling for stories to be shared in this petition, Montoya uses her personal interactions with content moderation on TikTok as an exigence for others to do the same--and to be more visible to others and the platform, as evidenced by the stated purpose of communicating with TikTok for change.

In addition to garnering just under 21,000 signatures, Montoya gathered and published personal testimonies written by almost forty marginalized content creators as updates to the original petition. In each of these 40 testimonies, the creators introduce themselves and how they identify, describe the content they make for the platform and

their followers, and then share their stories of suppression and moderation on TikTok, including personal stories and screenshots from the app interface. Across these testimonials, creators shared the following reasons for sharing their stories to Montoya's petition included:

- hope for change on the platform
- frustration with constant problems on TikTok
- desperation fueled by platform bans and compromised monetization,
- among other nuanced emotional responses to exclusion and inconsistency.

For example, many of the testimonies shared perspectives on how the platform could “do better,” including recommendations for better communication about content removal and platform ban justifications to help creators avoid similar errors in the future.

In addition to creating videos addressing TikTok's leadership, Zev has also turned to other forms of awareness-raising. During our interview, Zev discussed his decision to create a petition on Change.org in April of 2021, which at the time of writing this dissertation has garnered nearly 19,000 signatures. Like other petitions about TikTok's content moderation problem, Zev's petition points to documented instances of TikTok targeting or silencing marginalized creators and argues for total transparency to improve the experiences of platform users from historically marginalized backgrounds. When I asked Zev about his original goals for the petition, Zev shared that the petition outperformed his original expectations, gaining four thousand signatures within the first three days. He attributed the petition's early success to his efforts boosting the link on TikTok and people choosing to pay money to promote the petition on Change.org, increasing its visibility on that platform. Although Zev spoke about the petition as a

success, he was careful to qualify that he didn't believe that a Change.org would be responsible for improving transparency on TikTok:

I'm not naïve though. I know TikTok isn't going to be like, "There's this scrawny Jewish kid from Indiana creating a petition. Let's listen to him now." I know that. It's just about having those numbers there.

Indeed, taking a slightly different approach than that of Montoya's petition, which emphasized personal stories of creators like herself, Zev explained that his original goal was for the petition to be a tangible representation of how many people believe that the lack of transparency on TikTok is worth addressing. For Zev, the number of signatures was a move to garner "legitimacy" for the cause that would help Zev and others speaking out on the issue of transparency and moderation on TikTok to generate media awareness and attention that could lead to the necessary public pressure to facilitate material change on the platform. Since surpassing the initial goal of ten thousand signatures, Zev has leveraged the petition when speaking with media outlets about these issues.

During our interview, Zev—a current data science student at Georgetown—connected his interest in data collection about this problem to his academic interests in machine learning and neural networks: "What I'm focusing on is machine learning and neural networks, basically taking a lot of data, millions upon millions of rows and columns, and making sense of it all, saying: what do these numbers tell us? What is the story that these numbers tell?" Such a perspective comes through in how he talks about the petition's popularity as key to gaining support for the cause. Likewise, Zev has continued to tell the story of content moderation on TikTok through the creation of a Google Form to collect data about inaccurate content moderation experiences. He

explained that he wants this form to be “a centralized place” to examine and analyze trends: “It is meant to be a bare bones form. I wanted it to be short, and I wanted it to be very much self-described, to get the perception of what people thought was going on [with TikTok’s policies.]” The form, which was published about a week prior to our first interview, was designed by Zev to collect the following information:

- a description of the video that received a CG violation;
- the explanation of CM provided by TikTok;
- whether or not the user has appealed the content moderation judgment;
- if the user’s video was removed;
- if the user believes their video has been suppressed; if the user
- has been reported and banned;
- if the account ban was temporary or permanent;
- how long the temporary ban was; and whether there was an explanation provided for the ban.

For Zev, these pieces of information offer insights into the varied experiences of creators on the platform. In my view of this survey, an interest in these insights and in a central dataset represents an interesting response that highlights the common desire held by creators to generate evidence of their experiences within an opaque platform. While such stories of content moderation and suppression are numerous on the platform, a central dataset about platform judgments signals an effort to make sense of TikTok’s content moderation through creator experiences and perceptions and to render those experiences as both quantitative and qualitative data. The desire for quantitative data might also be

understood as a way to better represent the scale of the problem within a black-boxed algorithmic environment.

Importantly, I see that the move to discuss content moderation problems and experiences on Change.org petitions shifts the conversation from the experiences of an individual toward a collective that may include tens of thousands. While TikTok's user interaction design silos creators' experiences of content suppression, they actively subvert the platform and turn to other platforms to build coalitions or communities and to create public forms of evidence that these problems are not isolated to individuals. This chapter thus looks at the affects of both content moderation and visibility moderation as it is experienced by creators who have limited ability to dialogue with the platform to correct judgments that feel incorrect or to share the behind-the-scenes experiences of content creation known with other creators and platform users.

### **Creator Connections for Support and Survival**

In this section, I discuss how the lack of transparency on TikTok creates the affective conditions in which creators forge connections with one another to counter moderation and suppression, drawing largely on findings from my interviews. I learned that the experiences of content suppression and moderation often served as a point of connection for marginalized and underrepresented creators to network with one another for both professional and personal purposes. Therefore, in addition to connecting with creators who make similar content to collaborate creatively and to share professional knowledge about developing and sustaining platform growth in terms of followers and engagement to support monetization efforts, creators also reached out to one another to gain perspectives on frustrating experiences with content moderation. Each of the

creators I spoke with shared with me that they thought their experiences on the platform were unique until they started seeing people they knew or people with similarly large accounts talking about similar experiences in videos, live-streams, petitions, or private messages.

Zev reported that as his petition for transparency on TikTok took off and his profile grew, more and more creators with large followings started reaching out to him, either to share how their communities were being impacted on the platform or to ask Zev for his take on their problem. Zev shared that he oftentimes tried to help creators figure out the source of the content removal, suppression, or ban and sometimes highlighted creators on his account who were being treated unfairly by the platform, noting that he must be careful not to get himself banned. More than anything, Zev shared that while creators sometimes come to him due to how vocal he has been on the subject, he thinks that most don't really expect him to be able to fix the problem. Rather, he explained that what most people want is some "reassurance" that they did not actually do anything wrong.

Zev shared that because of his videos and his petition on TikTok's content moderation problems, he developed a reputation for being a person to talk to about platform problems with individual creators reaching out to him for advice and emotional support and in the hopes that Zev would amplify their story to his followers. He explained that he thinks that people reach out to him about their stories to generate awareness about the problems facing creators on TikTok: "I don't think creators come to me individually, asking how can I personally help. It's more like, 'I know that you fight for this, this happened to me too.'" Some creators reach out to Zev, who has positioned



himself as an advocate for creators on the platform, for advice about how to reverse or navigate platform problems, but Zev's assessment that many simply want their stories to be seen speaks to the affective impacts of moderation on creators. While creators would ideally prefer platform change, the affective impulse to share stories is a response to the experience of suppression in isolation. Sharing stories and being heard are ways for creators to be validated that they are not causing the problems they're experiencing and are being subjected to an unfairly enforced set of moderation practices and processes.

Other times, creators share and work with one another to strategize ways of navigating the platform for better outcomes despite the disadvantages posed to marginalized creators. One of the interview participants for this study, Natalie, for example, spoke about how she and other Black creators are urging other creators to become more informed about copyright so as to copyright their work and prevent exploitation of their creative labor. In the United States, it is very difficult to copyright creative labor like dance, as Natalie told me in our interview, but cites the idea of dancer Tee Noir on TikTok that the platform should implement a dance credit button to help Black creatives monetize and receive credit for their work. Other creators spoke about how they have learned ways of reaching out to the platform to try to get content or accounts restored after wrongful removals or bans. Much to my surprise, while I recruited creators for this study on an individual basis, I discovered through my interviews that three of my participants were friends with each other and belonged to a group chat, known among the creators as the TikTok Ban Victims. Creators shared that this is one of several creator group chats that was started to talk about the work of content creation on the platform, including its problems related to moderation and suppression.

When discussing his involvement in founding the TikTok Ban Victims group chat, David, one of my interview participants, shared that he first got the idea to reach out to creators when he saw a popular creator talking about her experiences with content suppression on a TikTok live stream. Seeing that other creators with large platforms were affected like he was, David connected individually with many creators before adding them to a group chat on Instagram. He explained that “over time, it grew and grew and grew,” as he reached out to many creators who were struggling with platform suppression. “Originally, we were going to find a solution together,” David shared, “Everyone was discussing their ideas on why this was happening and what was going on.” Over time, the group chat, which involved 32 creators at the time of our interview, became a “useful tool” for creators to discuss brand deals, management, networking strategies and opportunities, in addition to creator feelings about the platform as a space for creators, inclusive of moderation and suppression. With more and more creators facing increased rates of removals and bans, David likened the group chat at the time of our interview to a support group:

But it really is more of a support group than anything else. Whenever somebody gets banned, everyone bands together. They all post each other's videos. When I was taken down, an incredible amount of them posted my [video] onto their [Instagram] story, including people who have left the group chat, because they maybe felt like they weren't being listened to as often because the group chat moves really fast. There are a couple of younger folks who derail the conversation to like random tangential things, but there are also times when you know, we get serious. It's all about showing support.

For creators, support translates to methods of amplifying the content of fellow creators both on TikTok and on other platforms like Instagram. Practices of posting the content of banned creators often contribute to creator efforts to get their accounts restored and to help creators rebuild their followings on backup or new accounts on TikTok or other platforms. I found David's note that he has even felt support from creators that he is less close with, which speaks to the types of solidarity that creators often are able to build with one another through this shared negative relationship to TikTok.

David made it clear later in his interview that relationships with other content creators was important for creators like himself who were not based in entertainment industry cities like New York City or Los Angeles. Connected creators often have the resources and representation to navigate opaque content moderation problems behind the scenes with more direct customer service interventions. He explained:

I think it's important that creators have community that are not just in Los Angeles, because Los Angeles has the protection of everyone you hear works in that industry. They have connected friends and brand deals and the rest of us live out throughout the United States I'm in the middle of bumfuck, nowhere Alabama and I need the same amount of support and community. That is like in every other job. So that's that's a lot of what that group said is, I wish Instagram let us max out the members so I could put more people in there and continue to build this. In between us. At one point between the 32 people in there. We had like 90 million followers between us, you know, and thing at 90 million followers, none of us can make a change, like none of us can help the system, which was one of the more upsetting things.

At the core of this point about region, David highlights the unique circumstances of content creation as a profession that can be very isolating, particularly for creators who haven't had the resources or algorithmic good fortune to have a management team. Without the extra support of professionals who are able to help creators thrive in digital labor, creators have to figure out things like monetization on their own, which can be much trickier when they are also facing inconsistent and uneven content moderation problems. Therefore, forums like this group chat help creators like David share their stories not only to vent but to strategize next steps when problems arise. At the time of our interview, David's account was freshly restored after a permanent ban that had been reversed. Since then, David has faced several temporary suspensions and so-called permanent account bans that have been reversed, and he has shifted creating most of his content for Instagram Reels and YouTube Shorts where he faces far fewer issues. Now, David recommends that creators take similar steps to go to platforms where creators are more supported, like YouTube Shorts, which has a much better customer service system.

### **Disrupting Sneaky Rhetorics: Understanding How Visibility Moderation**

#### **Encourages Complaint**

The sneaky rhetorics of content moderation contribute to content creator feelings that the platform is not to be trusted. Regardless of whatever statements of support the platform might release to quell anxieties that marginalized creators have about the platform's patchy history of suppressing the visibility and discoverability of content created by queer, disabled, Black, and fat creators. Across my analysis of Change.org petitions, creator testimonials [on where?], and my interview data, it became clear that the frustrations and anxieties creators have with TikTok's faulty moderation are

exacerbated by the “disconnect between what is supposed to happen and what does happen, between paper and practice” and the perception that the rules do not apply to everyone (Ahmed, 2021, p. 55). More sharply, the feeling that marginalized creators are more likely to be penalized on TikTok than a chosen elite heightened feelings that TikTok exploits marginalized creators whose creative labor and empowers other users to exploit that labor, too. All of this is compounded by the lack of access in the customer service processes for appeals and complaints about user experiences on the platform. Supporting the findings discussed in chapters 2 and 3, the lack of platform transparency surrounding content moderation affectively encourages creators to compare their experiences with what they can see on the platform. For marginalized and underrepresented creators, the visibility of white creators avoiding platform resistance and achieving great success on TikTok for work that marginalized and underrepresented creators either are penalized for or have innovated on the platform frustrates creators who feel targeted by the platform’s moderation system. The desire to share one’s experiences emerges in this context of zero transparency as a way to combat the messaging of moderation in isolation: that one’s content is wrong, harmful, violative and that the creator under moderation, despite prior success, does not belong on the platform.

Visibility moderation, which governs platforms by increasing or reducing the visibility of content, contributes to the sneaky rhetorics of TikTok, which affect marginalized and underrepresented creators sharply. Importantly, it is not just the visibility of who is being amplified on TikTok (read: straight, cis, white content creators) but also the visibility of the struggles facing marginalized and underrepresented content creators that contributes to the pervasive feelings of mistrust, anxiety, and fear on the

platform for content creators who have faced oppressive content moderation outcomes. As shown in this chapter, the move to speak out publicly or within closed networks of content creators represents an effort to disrupt the sneaky rhetorics of moderation, as well as the rhetorics of meritocracy that dominate content creator culture. The rhetorics of meritocracy refer to what scholars such as Petre, Duffy, and Hund have discussed in their work platform paternalism in which they argue that representations of platforms as disinterested mediators contributes to perceptions of content meritocracy on digital media platforms. They define a content meritocracy as a content media culture “in which the highest-quality content inevitably rises to the top” (p. 2). I see sneaky platform rhetorics and the rhetorics of meritocracy working together on TikTok to de-emphasize the role that platform governance plays in the visibility of content and the success of individual creators to instead affectively convince platform users that visibility and engagement can be controlled through technical savvy and hard work on TikTok. As I will discuss in the conclusion, content creation on TikTok is technical communication, and throughout this dissertation, I have highlighted the strategies and expertise that creators use to be successful and navigate platform oppression. At the same time, their experiences on TikTok speak to the influence of platform governance on how creators understand and feel about the platform’s values.

Such rhetorics of meritocracy on TikTok fosters both frustration and perhaps some defensiveness. Indeed, while many creators believe that the platform targets marginalized creators in particular, others are less certain and feel that there is a certain type of engagement currency that comes with talking about shadowbans. For example, when I asked Zev to speak to his perceptions of shadowbans on the platform, he shared

that he defines a shadowban as when videos are not being pushed to the For You Page and are only being shared with a creator's followers. He also shared that he doesn't believe the practice exists:

It's an interesting balance. because deep down like, oh, if I say I'm settled, then people will comment more. But then I'm actively lying, and supporting something I don't believe in. I think at the end of the day, I really, really appreciate it when followers say, "Oh, I think your shadowbanned or I can't believe how shadow banned you are." Because it means that they care. It means that they care about them, and they are supporting me a whole lot. Well, thank you. But at the end of the day, I don't really believe the shadow bans. I just think either, I'm not making the right content, not posting at the right time, or just luck of the algorithm. Law of large numbers eventually asked to come down. And after a million videos, there's got to be a 2000 views video. I gotta get down to that average somehow.

Zev is not alone in thinking that shadowbans are not a real practice that TikTok deploys against its creators. Rather, he believes that creators find success through some combination of algorithmic luck and savviness. It is striking, too, that Zev believes that discussing shadowbans is an appeal for engagement that many creators use to boost their content in the algorithm. Such an appeal works in the context of an environment steeped in sneaky rhetorics because creators want to support creators who might possibly be suppressed by the algorithm. As a creator, Zev's suspicion speaks to the way that the feeling that shadowy, sneaky things happening in online platforms stick also to other users, not just platform processes. Scholars writing about content creators and algorithmic visibility (Cotter 2019) have discussed what it means to "play the game" of

the algorithm in such ways. Regardless of whether TikTok intentionally shadowbans individual creators, it is interesting that the idea is so pervasive on TikTok that users often ask creators if they're being suppressed directly. Such a dynamic between creators and users speaks to the larger problems of platform trust on TikTok.

## **Discussion**

In examining how content creators use personal stories and other types of evidence grounded in their experiences of TikTok's user experience, I find that these stories demonstrate the importance of UX design being attuned to embodied, experiential knowledge through stories. It also feels important to note that personal stories, when told together, provide a view of platforms like TikTok that are otherwise inaccessible due to the digital isolation of individual user recommendation feeds. As discussed in the second chapter of this dissertation, it would be difficult to capture the scope of these problems on TikTok without following a trail of creator stories. This methodological problem is an area for future research, but in this dissertation, I have found that research methods grounded in story have proven to be vital for creators, including my participants, as part of their work as content creators. Indeed, creators both share and learn from stories to understand their experiences on TikTok through the stories of other creators on the platform. While frequent, repetitive instances of content moderation might be frustrating and hurtful to content creators in isolation, it is through public and private expressions of complaint that creators are able to see their experiences as part of a larger pattern, and in doing so, find affirmation, community, and sometimes relief that they are not alone if not relief that the problems can be resolved on TikTok.



It feels important to note that creators put forth several recommendations for platform change as a result of these conversations with one another and practical study of patterns on TikTok. Creators suggested the following recommendations to TikTok's design to improve creator experiences:

- improved customer service for users involved with the Creator Fund to better support their work on the platform when faced with uneven content moderation
- improved communication about how content violations (including successfully appealed violations) are tracked by account
- more thorough appeals forms to support richer descriptions of cultural contexts and individual circumstances, and more concrete communication about the time span of penalties.

Many of the recommendations put forth by creators are related to better policies and experiences that creators have experienced or witnessed on other platforms, like YouTube shorts.

Creators are not examining their platform experiences on TikTok in isolation -- they're looking at their experiences across platforms and in community with creators with shared lived experiences both online and offline. As such, personal stories shared alongside other types of evidence grounded in the experiences of creators (such as mainstream media reports of TikTok's shadier treatment of creators) both disrupts and heightens the sneaky rhetorics of TikTok as creators are better able to make sense of their experiences. Sharing stories with other creators and with me helps creators who just want their experiences to be acknowledged/seen, particularly as they don't expect change on the platform any time soon. By sharing these stories and talking about their experiences

with other creators who have experienced the same things within the black-boxed system, creators are able to not only imagine possible changes that would make the platform better for everyone; indeed, creators are also able to feel that they are not the only ones experiencing what feels like unfair, discriminatory moderation practices within an isolating digital platform. This matters for creators so that they are able to rationalize their experiences as part of a larger systemic problem, rather than the result of their practices as individuals on the platform.

Digital practitioners can use such methods of data collection to take up the concerns and feedback of creators to improve platform conditions. As Natasha Jones explained, narrative-based inquiries “ask users to discuss the use of technology in the past and ask users about how they would like to see the technology changed for the future” (p. 486). And this dynamic is precisely what I have observed in the example of marginalized content creators on TikTok and described in this chapter. The frustration experienced individually as a disruption in platform usage and content production is compounded by the perception that this is commonly experienced by their friends and people like them while hate speech and harassment is seemingly permitted on TikTok. Perhaps repetitive content moderation glitches would not be so frustrating if they were more uniformly experienced by all platform users and if content moderation did filter out racist, sexist speech and harassment. In the next and final chapter of this dissertation, I discuss the limitations and advantages of such a method of reading across platforms, but here I want to highlight that creators are already demonstrating the potential power of such networking and organizing

## **CHAPTER 6**

# **CONCLUSIONS AND IMPLICATIONS FOR RESEARCHING ALGORITHMIC TRANSPARENCY IN WRITING STUDIES**

### **Summary of Findings**

In this dissertation, I aimed to learn about how content creators from marginalized and underrepresented backgrounds made sense of content moderation on TikTok.

Through a discussion of content creator stories shared with me in interviews and in online petitions circulated through Change.org, I have explored this subject across the preceding chapters. In focusing on these stories, I was able to identify the key feelings and strategies that marginalized creators use to make sense of their experiences and perspectives of content suppression on TikTok and the platform's approach to moderating content. For example, every creator I spoke to compared their experiences of receiving discriminatory or illogical content violations for videos they felt to be allowable per the Community Guidelines to the proliferation of hate speech and harassment on the platform. Creators also spoke to the importance of speaking with other creators and learning from one another on the platform to adapt habits of algorithmic code-switching and self-censoring to avoid algorithmic detection and to strategize with one another to challenge platform suppression in their communities.

The work of creating content on TikTok necessitates that content creators develop what Taina Bucher calls an “algorithmic imaginary” of TikTok, or a sense of what the algorithm does and how it works. Through a careful analysis of creator stories and online testimonials, findings suggest that creators from historically marginalized and

underrepresented backgrounds are sensitive to algorithmic oppression as it manifests in their content moderation interactions. While creators from all backgrounds can become sensitive to algorithmic recommendation systems, the critical lens of marginality (Shelton) enables creators to sense inconsistencies and unevenness in platform moderation, particularly when these inconsistencies and uneven tendencies align with perceptions of platform power. For example, content creators from marginalized backgrounds cite numerous examples of popular white TikTokers evading moderation even when they break written policy as evidence of what several creators called TikTok's 1%. That members of TikTok's 1%, or the perception of this elite category, tend to be white, cis, straight, and wealthy contributes to the perceptions these creators hold that they are held to different, stricter standards, or that TikTok actively targets marginalized creators.

By reading their experiences against TikTok's written policies, such as the Community Guidelines and the Terms of Service, creators determine platform motivations surrounding content suppression, ranging from an indifference for improving algorithmic detection to be less harmful to marginalized creators to an hostility toward marginalized creators thriving on the platform, especially when they choose to use the platform to highlight their experiences as marginalized creators on the platform. The accumulation of negative content moderation micro-interactions over time diminishes the ability of marginalized and underrepresented creators to trust TikTok as a platform that purports to support diversity. Indeed, in the absence of platform transparency, creators rely on their experiences in the interface to make sense of the platform's values and to create content that escapes unfair moderation based on those values. Through a

discussion of content creator stories shared with me in interviews and shared publicly in online petitions circulated through Change.org, I have explored this subject across the preceding chapters.

Indeed, my findings suggest that repetitive negative micro-interactions with content moderation influence how content creators make sense of platform values. Moreover, the frustrations felt by creators when their experiences do not align with expectations of how content moderation should work on the platform reveals a great deal about how time is experienced on TikTok. Everyday platform users expect to have a streamlined experience of endless content, and content creators have similar expectations on top of professional requirements for constant engagement and production. Experiencing content moderation that makes such seamless interactions in the interface contributes to the frustrations felt by creators who already feel penalized, suppressed, or silenced by TikTok as a result of their marginalized identities. Throughout this dissertation, I discussed a range of creator responses to the affective conditions facilitated by TikTok's content moderation, including how anxieties and fears about discriminatory moderation impact the content they create, the relationships they forge with other creators on the platform, and the methods they take to challenge the platform to better support its marginalized and underrepresented content creators.

Overall, the tendency to share stories about content moderation on the platform speak to the importance of experiential and embodied knowledge of black-boxed technologies like digital content moderation on social media platforms. In what follows, I discuss this study's implications for writing assessment and technical communication. Then, I discuss this study's limitations briefly before offering some methodological

insights for the rhetorical study of user experiences in algorithmic environments like TikTok.

### **Limitations**

This dissertation is a beginning to my inquiry into the digital cultural rhetorics of content moderation and what content creators can teach us about the ever-changing nature of platform governance on TikTok and other social media platforms. By working with the stories shared with me by these six creators, I have begun to understand with greater clarity the role of emotion and embodied experience in unpacking the hidden policies that govern platforms like TikTok. I aimed to represent these stories in context so as not to over-interpret the findings and patterns but instead to show how these creators made sense of their experiences from the lens of their specific positionalities. That said, I want to briefly discuss this study's limitations, which include my positionality, the number of interviews grounding this study, and the creep of meritocracy thinking into some of the stories shared with me by creators.

First, this study is limited by my positionality as a white woman and non-creator. Working with creators from marginalized backgrounds has been key to this study, and while I hope this study amplifies the insights shared with me, I do not want to underestimate how my identity has limited what was shared with me and my ability to work with this information. Second, due to time constraints, this study is limited by the number of interview participants. As I discussed in the introduction to this dissertation, this study is limited largely by the subjective experiences of my interview participants, each of whom relates to the platform in different ways as a result of their varied embodied experiences. I believe that the inclusion of online testimonials from

Change.org mitigates this limitation to a small degree, but as a researcher, I found that speaking with creators helped me to better understand not just how they felt about their experiences but also how their interface interactions figured into their interpretations of content moderation on TikTok. Recognizing the limited scope of this study as a result of the small number of creators I spoke with, I intend to speak to more content creators to continue to develop the findings presented in this dissertation.

Finally, I have approached this dissertation from the perspective of believing creators about their experiences with content moderation, amplifying their ways of knowing around platform mechanics and the mismanagement of policies on TikTok. By honoring their lived experiences and perspectives of how their content and their identities have been mediated by content moderation on the platform, I have learned a great deal about how TikTok mediates emotion, identity, and ideologies. That said, as I briefly discuss in Chapter 4, I have found that creator perspectives about content moderation are sometimes made murky by the rhetorics of meritocracy. Such rhetorics about being able to work hard to achieve viral success showed up alongside their stories about moderation and suppression on TikTok.

## **Implications**

### *For TikTok and Accountable Content Moderation*

This dissertation contributes to ongoing conversations about the impacts of visibility moderation. I hope that I have shed light on how the role of emotion in navigating online environments. While users may not notice a favorite creator missing from their feed for a few days, creators feel the invisibility alongside other frustrations: a seemingly endless string of inconsistent and uneven content violations, exposure to hate

speech and misinformation that escapes moderation, and stories from their friends experiencing similar challenges on the platform. Such experiences evidence the material realities and affective experiences of what Zeng and Kaye, building on the work of Cotter, Gillespie, and others, termed “visibility moderation,” or platform governance that amplifies or limits visibility as the primary means of moderation. By using popular content creator experiences as a way to explore this increasingly ubiquitous moderation, I hope I have highlighted the significance of these feelings as they relate to the material realities of digital labor and the content economy in online spaces. I also hope that these stories contribute to important work which interrogates how visibility moderation is aligned with online cultures of surveillance and suppression.

By speaking with content creators about their experiences with content moderation, I have learned a great deal about how creators understand platform values surrounding diversity, equity, and inclusion through their experiences of TikTok’s interface. Although not all of their feelings about TikTok’s user experience can be isolated to features in the interface, each of the creators detailed how encounters that they had in the app impacted their usage and trust in the platform’s approach to content moderation. Throughout these chapters, I have shared recommendations from content creators that they believe would improve the platform experience for marginalized creators. Briefly, here is a list of some of the recommendations discussed throughout this dissertation:

- More specific messaging in system notifications about how content violations were determined;



- An improved appeals process which enables creators to describe the circumstances surrounding their flagged content better, which would help protect creators from the negative consequences of targeted mass reporting;
- An improved appeals process which better communicated to creators how decisions were overturned and the effects of overturned platform judgments on
- A dance credit button to deter the cultural appropriation of work by Black dancers on the platform especially.

Again, these are just some of the recommendations that creators have made, but they represent important insights into how TikTok's moderation structure mediates creators' feelings of trust in the platform and its processes. Overall, creators called for more transparency on TikTok. While full transparency would not address the problems which arise given the complex ways that TikTok's AI works in tandem with the behaviors of the platform's users (Grandinetti, 2021), such changes would go a long way to demonstrate to these creators that the platform does want to support diversity and equity in moderation. At the same time, this dissertation's findings demonstrates that whatever changes TikTok makes to its policies or user experience design would not necessarily result in immediately improved experiences for marginalized creators who have been repetitively hurt by content moderation on TikTok and other social media platforms.

One key way for TikTok to engage creator concerns more meaningfully would be to take the recommendation put forth by Rosalynne Montoya of including creators on advisory boards. The expertise that this dissertation highlights is needed behind the scenes to help inform policy-making and user experience design. Critically, this expertise and the labor creators whose labor makes the platform so successful need to be

compensated more equitably in accordance with how much these creators bring to the platform.

*For Writing Assessment*

This dissertation has probed questions of accountability, transparency, and trust in the assessment of content on TikTok. While speaking with content creators, I was struck by how many of their complaints about their experiences resonated with critiques of assessment practices in writing classrooms, as they spoke about how frustrated they were with vague, subjective standards and the requirement that they perform in specific ways to achieve the desired outcome. Recent work on writing assessment have interrogated how standard grading systems maintain and perpetuate oppression and pointed to alternative assessment systems to increase classroom equity and counter white language supremacy (Condon and Young, 2016; Inoue, 2019; Palomino, 2021).

While I believe in the utility of labor-based assessment and use a variation of this approach in my own classrooms, I believe this study has important implications for studying the affective implications of assessment in writing classrooms, specifically by tracing how students' previous assessment experiences inform their current experiences, even when they are being assessed under new systems. I found in my conversations with content creators that prior experiences of assessment had a lasting impact on how they related to similar assessment contexts. In her article, "Your Grading Contract Ain't It," Craig makes a similar point and troubles the disciplinary move to identify ungrading practices as anti-racist and pro-black. Criag emphasizes the challenges that an unfamiliar assessment system poses to Black students in particular who have no reason to trust that white instructors will treat them fairly, simply because they said they would. Craig

writes, “Contract grading might make my colleagues *feel good* and ease the guilty burden of using practices with deep white supremacist origins in other areas of their courses, but it felt like a trap to me” (146, emphasis added). In this assessment of the conversation around contract grading as social justice, Craig identifies the affective implications of contract grading for students and teachers, highlighting the differences afforded in these different positionalities.

Many writing instructors turned contract grading as a result of feeling that they could better support students, but as Craig notes, feelings about the decision to implement contract grading often reflect not just aspirations for classroom equity but other instructor emotions, including: white saviorism and guilt, curiosity, hopefulness, stress, exhaustion, as well a range of multifaceted anxieties about one’s perception as a teacher to students, peers, and supervisors alike. Moreover, in line with other critiques of contract grading about issues of accessibility and fairness (Carillo; Kryer and Zimmerman), more attention needs to be given to student experiences of contract grading and how those experiences relate to their experiences of assessment in educational contexts more broadly. Bearing these critiques in mind, I believe that writing teachers seeking to implement systems of assessment, such labor-based assessment, need to be more critical of the affective histories students have with assessment in general and to be more attentive to how those feelings and experiences are mediated by new systems. How do new systems of assessment that are less stream-lined in learning management systems contribute to assessment feelings of anxiety, frustration, fear, etc.? Moreover, in our individual classrooms, how can we design user-friendly and accessible assessment tools that challenge white linguistic supremacy and ableism? Through my conversations with

creators, I was once again reminded about the importance of design in assessment, particularly around tools that supposedly give creators the agency to report content from other users or appeal platform judgments related to their content.

Looking forward, I believe writing teachers need to seek more student input about the design of assessment tools and to think more critically about the limitations of assessment reform when enacted in individual classrooms. Without more attention to shaping assessment at the institutional and programmatic levels, alternative assessment paradigms are likely to be mistrusted by students who engage grades and course designs through the lens of their full assessment experiences and how these experiences have been shaped by their intersectional identities.

#### *For Technical Communication*

This work contributes to ongoing conversations in technical and professional communication (TPC) that call for more attention to issues of social justice (Agboka, 2014; Jones, 2016; Shelton, 2020; Williams, 2013). Along these lines, TPC scholars have expanded the field by highlighting how technical communicators mediate information in their communities and challenging narratives based in white supremacy about what constitutes technical communication (Del Hierro, 2018; Mckoy, et al., 2022). TPC scholars have examined how beauty content creators do technical communication through user-tutorials on YouTube (Chong, 2018; Edenfield and Ledbetter, 2018; Ledbetter, 2018). This dissertation follows these important research threads about creators as technical communicators by exploring how creators teach platform users how to navigate unjust and ever-changing moderation environments.

If content creation involves skills in instructing platform users how to engage one's content, then how do creators teach their followers and would-be followers about how to support them in algorithmic environments? In this dissertation, I have discussed how creators model behaviors of engaging content and self-censoring themselves to avoid detection. Crucially, my findings point to the importance of storytelling to incur support for creators on TikTok, to raise awareness about platform problems, and to instruct platform users about how to navigate the platform. Each of the creators I spoke with detailed their experiences of sharing personal stories to mobilize their followers on behalf of themselves or their friends. Storytelling was also a prominent strategy represented within the Change.org petitions that I analyzed. The attention to personal experience is necessary in light of algorithmic opacity and TikTok's user interaction design which silos individual experiences. Storytelling, too, represents an important way that creators share their professional expertise and experience of TikTok with more casual platform users.

## **Conclusion**

Finally, I believe that there is a great deal of room to expand on the work that I have started in this dissertation. In Chapter 3, I discuss how content creators sense inconsistent and uneven moderation through their personal interface interactions with content removals and reduced engagement. In each interview, creators discussed their experiences on TikTok in contrast to their experiences of moderation and engagement on Instagram and YouTube. This study has sought to examine content moderation through a digital cultural rhetorics perspective as the mediation of culture and knowledge. Future work could compare creator perspectives and experiences on these digital platforms to

examine how culture is inscribed visually through moderation on these platforms. As each creator told me, despite TikTok's problems, it still presents the best opportunities for engagement for marginalized content creators, alongside the worst examples of algorithmic suppression and platform punishment. More research is needed to understand this paradox. How does TikTok's specific algorithmic landscape have unique affordances for marginalized content creators at the same time that visibility moderation penalizes these same creators?

At the onset of this project, I was interested in pushing back against content moderation conversations that focused on freedom of expression because, in my view, such a focus fails to make space for the ways in which the legal aspects of this conversation ignores the ways in which free speech is already not free, particularly for marginalized platform users. In this dissertation, I have shared stories of creators who have developed deft affective and rhetorical sensibilities to navigate opaque systems. For these creators, transparency is an ideal feature of digital platforms, but at the same time, there is an awareness that technological transparency a) will never be achieved and b) would only reveal what they already know to be true, that these digital platforms have not invested in the infrastructure to support the free expression of all platform users.

As a scholar, I am interested in continuing to explore the ways in which digital platforms moderate and mediate our conversations about our very use of digital platforms. Along with these insights for inclusive user experience design, this study expands on a theory of sneaky rhetorics in online environments, highlighting the need for greater attention to the role of emotion in discourse online. Indeed, this study shows how a lack of transparency is felt by platform users and how these feelings motivate users to

take certain types of action, such as adapting their content to avoid moderation, and animate online cultures of suspicion. Importantly, this study shows how affect and emotion are embodied and related to subjective experiences of power and oppression. I hope that this dissertation contributes to conversations in digital and cultural rhetorics about the digital mediation of culture, identity, and information and prompt further exploration into how these everyday media platforms are shaping the work we do as digital and cultural rhetoricians.

## BIBLIOGRAPHY

- Ahmed, Sara. (2014). *The Cultural Politics of Emotion*. 2nd Ed., Routledge.
- . *Complaint!*. (2021). Duke University Press.
- Agboka, Godwin Y. (2014). "Decolonial Methodologies: Social Justice Perspectives In Intercultural Technical Communication Research." *Journal Of Technical Writing And Communication*, 44(3), 297-327.
- Amrute, Sareeta. (2019). "Tech Colonialism Today." Epic2019 Keynote Address. Retrieved from <https://points.datasociety.net/Tech-Colonialism-Today-9633a9cb00ad>.
- Ananny, Mike, And Kate Crawford. (2018). "Seeing Without Knowing: Limitations of The Transparency Ideal and its Application To Algorithmic Accountability." *New Media & Society* 20(3), 973–989.
- Are, Carolina. (2021). "The Shadowban Cycle: An Autoethnography of Pole Dancing, Nudity and Censorship On Instagram." *Feminist Media Studies*, 1–18.
- Arriagada, Arturo, and Francisco Ibáñez. (2020). "'You Need At Least One Picture Daily, If Not, You're Dead': Content Creators and Platform Evolution in The Social Media Ecology." *Social Media + Society* 6(3), 1-12.
- Bakir, Vian, and Andrew McStay. (2021). "Empathic Media, Emotional AI, and the Optimization of Disinformation." *Affective Politics of Digital Media*, edited by Megan Boler and Elizabeth Davis, Routledge, 263–279.



- Beck, Estee. (2018). "Implications of Persuasive Computer Algorithms." *The Routledge Handbook Of Digital Writing And Rhetoric*, edited by Jonathan Alexander and Jacqueline Rhodes, Routledge, 291–302.
- Benjamin, Ruha. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity.
- Berlant, Lauren. (2011). *Cruel Optimism*. Duke University Press.
- Biddle, Sam, Paulo Victor Ribeiro, and Tatiana Dias. (2020, March 16). "Invisible Censorship: TikTok Told Moderators to Suppress Posts By 'Ugly' People and the Poor to Attract New Users." *The Intercept*.  
<https://theintercept.com/2020/03/16/tiktok-app-moderators-users-discrimination/>.
- Boler, Megan, and Elizabeth Davis. (2020). *Affective Politics of Digital Media*. Routledge.
- Botello, Elena. (2019, December 4). "TikTok Admits It Suppressed Videos by Disabled, Queer, and Fat Creators." *Slate*. <https://Slate.Com/Technology/2019/12/Tiktok-Disabled-Users-Videos-Suppressed.Html>.
- Brock Jr., André. (2020). *Distributed Blackness: African American Cybercultures*. NYU Press.
- Bucher, Taina. (2018). *If... Then...Algorithmic Power and Politics*. Oxford University Press, 2018.
- (2017). "The Algorithmic Imaginary: Exploring the Ordinary Affects Of Facebook Algorithms." *Information, Communication & Society*, 20.1, 30–44.

- Burton, Zev. (2021, April 9). "Call On TikTok to Release THE Content Moderation Guidelines." *Change.Org*, <https://www.change.org/p/tiktok-call-on-tiktok-to-release-the-content-moderation-guidelines>
- Crawford, Kate, & Gillespie, Tarleton. (2016). "What is a Flag For? Social Media Reporting Tools And The Vocabulary Of Complaint." *New Media & Society*, 18(3), 410–428.
- Cohen, Libby. (2021, April 20). "Why Did #AbolishthePolice Disappear from TikTok?" *The Daily Dot*. <https://www.dailydot.com/debug/tiktok-abolish-the-police-hashtag>.
- Condon, F., & Young, V. A. (Eds.). (2016). *Performing Antiracist Pedagogy in Rhetoric, Writing, and Communication*. WAC Clearinghouse.
- Cotter, Kelley. (2019). "Playing the Visibility Game: How Digital Influencers and Algorithms Negotiate Influence on Instagram." *New Media & Society* 21(4), 895–913.
- . (2021). "'Shadowbanning is not a Thing': Black Box Gaslighting and the Power to Independently Know and Credibly Critique Algorithms." *Information, Communication & Society*, 1–18.
- Chong, Felicia. (2018). "YouTube Beauty Tutorials as Technical Communication." *Technical Communication* 65(3), 293-308.
- Chun, Wendy Hui Kyong. (2016). *Updating to Remain the Same: Habitual New Media*. MIT Press.
- . (2012). Race And/As Technology, Or How to do Things to Race. *Race After the Internet*, 38–60.

- Cunningham, Stuart, and David Craig, eds. (2021). "Creator Rights and Governance." *Creator Culture: An Introduction to Global Social Media Entertainment*. NYU Press.
- Duffy, Brooke Erin. (2017). *(Not) Getting Paid to Do What You Love: Gender, Social Media, And Aspirational Work*. New Haven, CT: Yale University Press.
- Edenfield, Avery. C., and Lehua Ledbetter (2019, October). "Tactical Technical Communication Incommunities: Legitimizing Community-Created User-Generated Instructions." In *Proceedings of the 37th ACM International Conference On The Design Of Communication*, 1-9.
- Edwards, Dustin. (2018). "Circulation Gatekeepers: Unbundling the Platform Politics of YouTube's Content ID." *Computers And Composition*, 47, 61–74.
- Edwards, Dustin, and Bridget Gelms. (2018). "The Rhetorics of Platforms: Definitions, Approaches, Futures." *Present Tense* 6(3).
- Eslami, Motahhare, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert, and Karrie Karahalios. (2019). "User attitudes towards algorithmic opacity and transparency in online reviewing platforms." *Proceedings of the 2019 Chi Conference on Human Factors in Computing Systems*, 1-14.
- Eubanks, Virginia. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- Fernandes, Maggie, Matthew Homer, and Jennifer Sano-Franchini. (2022). "Sneaky Rhetorics, Habitual Media, and the Affective Politics of Lurking on Facebook." *Enculturation: A Journal of Rhetoric, Writing, and Culture*, 33.

- Fiore-Silfvast, Brittany. (2012). "User-Generated Warfare: A Case of Converging Wartime Information Networks and Coproductive Regulation on YouTube." *International Journal Of Communication*, 24(6).
- Gelms, Bridget. (2019). "Volatile Visibility: How Online Harassment Makes Women Disappear." *Digital Ethics: Rhetoric and Responsibility In Online Aggression*, edited by Jessica Reyman and Erika M. Sparby, *Routledge*, 179–194.
- Gillespie, Tarleton. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press.
- Gillespie, Tarleton, Patricia Aufderheide, Elinor Carmi, Ysabel Gerrard, Robert Gorwa, Ariadna Matamoros-Fernández, Aram Sinnreich, and Sarah Myers West. (2020). "Expanding the debate about content moderation: Scholarly research agendas for the coming policy debates." *Internet Policy Review*, 9(4), 1–29.
- Goldman, Eric. (2019). "An Overview Of The United States' Section 230 Internet Immunity." In *The Oxford Handbook of Online Intermediary Liability*, edited by G. Frosio. Oxford University Press.
- Grandinetti, Justin. (2021). "Examining Embedded Apparatuses of AI in Facebook and TikTok." *AI & Society*, 1–14.
- Gruwell, Leigh. (2020). "Feminist Research On The Toxic Web: The Ethics Of Access, Affective Labor, And Harassment." In *Digital Ethics: Rhetoric and Responsibility in Online Aggression*, edited by Jessica Reyman and Erika M. Sparby, *Routledge*, 87–103.

- Haas, Angela. (2018). "Toward A Digital Cultural Rhetoric." In *The Routledge Handbook Of Digital Writing And Rhetoric*, edited by J. Alexander and J. Rhodes. Routledge, 412-422.
- Hale, James. (2020, March 16). "TikTok Moderation Guidelines Show Users Were Punished For Political Content, Being 'ugly' Or 'slummy.'" *Tubefilter*.  
<https://www.tubefilter.com/2020/03/16/tiktok-moderation-guidelines-china-political-content-ugly-slummy/>
- Haimson, O. L., Delmonaco, D., Nie, P., & Wegner, A. (2021). Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, And Black Social Media Users: Marginalization And Moderation Gray Areas. *Proceedings of the ACM On Human-Computer Interaction*, 5(Cscw2), 1–35.
- Hierro, Victor Del. (2019). "DJs, Playlists, And Community: Imagining Communication Design Through Hip Hop." *Communication Design Quarterly Review*, 7(2), 28-39.
- Hutchinson, Les. (2018). "Wielding Power and Doxing Data: How Personal Information Regulates and Controls Our Online Selves." In *The Routledge Handbook Of Digital Writing And Rhetoric*, edited by J. Alexander and J. Rhodes. Routledge, 303-316.
- Inoue, Asao B. (2019). *Labor-Based Grading Contracts: Building Equity and Inclusion in the Compassionate Writing Classroom*. Fort Collins, Co: WAC Clearinghouse.

- Jhaver, Shagun, et al. (2018). "Online Harassment and Content Moderation: The Case Of Blocklists." *ACM Transactions on Computer-Human Interaction (Tochi)* 25(2),1–33.
- Jhaver, Shagun, et al. (2019) "'Did You Suspect the Post Would Be Removed?' Understanding User Reactions to Content Removals on Reddit." *Proceedings of the ACM on Human-Computer Interaction*, 1–33.
- Jones, Natasha N. (2016). "Narrative Inquiry in Human-Centered Design: Examining Silence and Voice to Promote Social Justice in Design Scenarios." *Journal Of Technical Writing and Communication*, 46(4), 471–492.
- . (2016). "The Technical Communicator as Advocate: Integrating a Social Justice Approach in Technical Communication." *Journal Of Technical Writing and Communication*, 46(3), 342-361.
- Jones, Natasha. N., And Williams, Miriam F. (2018). "Technologies Of Disenfranchisement: Literacy Tests and Black Voters in the U.S. from 1890 to 1965." *Technical Communication* 65(4), 371–386.
- Klonick, Kate. (2018). "The New Governors: The People, Rules, and Processes Governing Online Speech." *Harvard Law Review*, 131(6), 2598–2670.
- Kopf, Susanne. (2020). "'Rewarding Good Creators': Corporate Social Media Discourse on Monetization Schemes for Content Creators." *Social Media + Society* 6(4), 2056305120969877.
- Köver, Chris And Markus Reuter. (2019, December 2). *Netzpolitik*.  
<https://netzpolitik.org/2019/discrimination-tiktok-curbed-reach-for-people-with-disabilities/>.

- Ledbetter, Lehua. (2018). "The Rhetorical Work Of YouTube's Beauty Community: Relationship And Identity-Building In User-Created Procedural Discourse." *Technical Communication Quarterly*, 27(4), 287-299.
- Lehto, Mari. (2022). "Ambivalent Influencers: Feeling Rules and the Affective Practice Of Anxiety In Social Media Influencer Work." *European Journal of Cultural Studies* 25(1), 201–216.
- Locatelli, Elisabetta. (2020). "Influencers as socio-technical actors: mapping the paths of influence and the traces of the digital influencers' system in Italy." *Journal of Sociocybernetics* 17(1).
- London, Tabitha M., Joey Crundwell, Marcy Bock Eastley, Natalie Santiago, and Jennifer Jenkins. (2019). "Finding Effective Moderation Practices on Twitch." In *The Routledge Handbook Of Digital Writing And Rhetoric*, edited by J. Alexander and J. Rhodes. Routledge, 51-68.
- Luvaas, Brent. (2021). "Affect and Autoethnography in Social Media Research." *Creator Culture: An Introduction to Global Social Media Entertainment*. NYU Press, 57-74.
- Marwick, Alice E. (2018). The Algorithmic Celebrity: The Future of Internet Fame and Microcelebrity Studies. *Microcelebrity Around the Globe*, 161-169.
- Marwick, Alice. E., & danah boyd (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, 13(1), 114-133.

- Massanari, Adrienne. (2017). # Gamergate And The Fapping: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures. *New Media & Society*, 19(3), 329–346.
- McCluskey, Megan. (2020, July 22). "These TikTok Creators Say They're Still Being Suppressed for Posting Black Lives Matter Content." *Time*.  
<https://time.com/5863350/tiktok-black-creators/>.
- McKee, Heidi A. (2011). "Policy Matters Now and in The Future: Net Neutrality, Corporate Data Mining, and Government Surveillance." *Computers and Composition* 28(4), 276-291.
- Montoya, Rosalynne. (2020, December 15). "Change TikTok's Community Guidelines Algorithms." *Change.Org*, <https://www.change.org/p/tiktok-change-tiktok-s-community-guidelines-1a543d9c-8219-4820-9aa8-8ee809ac059c>
- Myers West, Sarah. (2018). "Censored, Suspended, Shadowbanned: User Interpretations Of Content Moderation On Social Media Platforms." *New Media & Society* 20(11), 4366–4383.
- Nakamura, Lisa. (2002). *Cybertypes: Race, Ethnicity, and Identity on the Internet*. Routledge.
- Nakamura, Lisa., & Chow-White, Peter. (Eds.). (2013). *Race After the Internet*. Routledge.
- Ngai, Sianne. (2004). *Ugly Feelings*. Harvard University Press.
- Noble, Safiya Umoja. (2018). *Algorithms Of Oppression: How Search Engines Reinforce Racism*. New York University Press.



- Olheiser, Abby. (2021, July 13). “Welcome To TikTok Endless Cycle of Censorship And Mistakes.” *The MIT Technology Review*.  
<https://www.technologyreview.com/2021/07/13/1028401/Tiktok-Censorship-Mistakes-Glitches-Apologies-Endless-Cycle>.
- O’Neil, Cathy. (2016). *Weapons Of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- Palomino, Norma. (2021). Labor-Based Grading Contracts: Cultivating a Writing Classroom Ecology of Equity, Inclusion, and Compassion. *Writing Program Administration*, 44(2), 110-117.
- Papacharissi, Zizi. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.
- Payne, Eileen. (2020, May 12). “Tell TikTok to Abide By Their Own Standards And Protect Those With Disabilities.” *Change.Org*, <https://www.change.org/p/a-mom-of-a-wonderful-and-sweet-little-boy-needs-your-help-tell-tiktok-to-abide-by-their-own-standards-and-protect-those-with-disabilities>
- Petre, Caitlin, Brooke E. Duffy, & Emily Hund. (2019). “Gaming The System”: Platform Paternalism and The Politics of Algorithmic Visibility. *Social Media + Society*, 5(4).
- Popescu, Mihaela, and Lemi Baruh. (2019). “Privacy As Cultural Choice And Resistance In The Age of Recommender Systems.” In *The Routledge Handbook Of Digital Writing And Rhetoric*, edited by Jonathan Alexander and Jacqueline Rhodes. Routledge, 280–290.

- R., Sudani. (2020, April 25). "TikTok Stop Punishing Black Creators For Speaking Up!" *Change.org*, <https://www.change.org/p/tiktok-tiktok-stop-punishing-black-creators-for-speaking-up>
- Raine, Izzi. (2020, July 4). "Stop Removing Mental Health Awareness Content on TikTok." *Change.Org*, <https://www.change.org/p/tiktok-stop-removing-mental-health-awareness-content-on-tiktok>
- Raun, Tobias. (2018). Capitalizing intimacy: New subcultural forms of micro-celebrity strategies and affective labour on YouTube. *Convergence*, 24(1), 99-113.
- Reyman, Jessica. (2009). *The Rhetoric of Intellectual Property: Copyright Law And The Regulation Of Digital Culture*. Routledge.
- Reyman, Jessica, and Erika M. Sparby, Eds. (2019). *Digital Ethics: Rhetoric and Responsibility in Online Aggression*. Routledge.
- Roberts, Sarah T. (2019). *Behind The Screen: Content Moderation in the Shadows Of Social Media*. Yale University Press.
- Rose, Emma J., Avery Edenfield, Rebecca Walton, Laura Gonzales, Ann Shivers McNair, Tetyana Zhvotovska, Natasha Jones, Genevieve Garcia de Mueller, and Kristen Moore.. (2018, August). Social Justice in UX: Centering marginalized users. In *Proceedings of the 36th ACM International Conference on the Design of Communication*, 1-2.
- Rose, Emma J., Elin A. Björling, Ada Kim, and Nicole Young Alvarez. (2018, August). Usability testing with teens: Adapting human-centered design and UX methods. In *Proceedings of the 36th ACM International Conference on the Design of Communication*, 1-6.

- Sano-Franchini, Jennifer. (2014). "Cultural Rhetorics and the Digital Humanities: Toward Cultural Reflexivity in Digital Making." In *Rhetoric and the Digital Humanities*, edited by Jim Ridolfo And Bill Hart-Davidson, The University Of Chicago Press, 49–6
- . (2018). "Designing Outrage, Programming Discord: A Critical Interface Analysis of Facebook as a Campaign Technology. *Technical Communication*, 65.4, 387–410.
- Selfe, Cindy L., and Selfe, Richard. (1994). "The Politics of the Interface: Power and its Exercise in Electronic Contact Zones." *College Composition and Communication*, 45.4, 480–504.
- Senft, Theresa M. (2008). *Camgirls: Celebrity and Community in the Age Of Social Networks*. New York: Peter Lang Publishing.
- Shelton, Cecilia D. (2019). *A Techne of Marginality*. East Carolina University, PhD Dissertation.
- Silverstone, Danielle. (2020, December 10). "Stop Antisemitism on TikTok." *Change.Org*, <https://www.change.org/p/tiktok-stop-antisemitism-on-tiktok>
- Sparby, Erika. M. (2017). "Digital Social Media and Aggression: Memetic Rhetoric In 4chan's Collective Identity." *Computers And Composition*, 45, 85–97.
- . (2021). "Reading Mean Comments to Subvert Gendered Hate on YouTube: Toward a Spectrum of Digital Aggression Response." *Enculturation: A Journal of Rhetoric, Writing, and Culture*, 33, <http://enculturation.net/readingmeancomments>.
- Stark, Luke. (2021). "Empires Of Feeling: Social Media and Emotive Politics." In *Affective Politics Of Digital Media*, edited by Megan Boler and Elizabeth Davis, Routledge, 298–313.

- Stewart, Kathleen. (2007). *Ordinary Affects*. Duke University Press.
- TikTok. (2021). Community Guidelines Enforcement Report. Retrieved January 12, 2022. <https://www.tiktok.com/transparency/en/community-guidelines-enforcement-2021-2/>
- . (2021). Account Safety. Retrieved March 15, 2022. <https://support.tiktok.com/en/safety-hc/account-and-user-safety/account-safety>.
- Trice, Michael, Potts, Liza, & Small, Rebekah. (2019). Values Versus Rules: How Platforms Generate Amorality on Facebook and Reddit. In *Digital Ethics: Rhetoric and Responsibility in Online Aggression*, edited by Erika Sparby and Jessica Reyman, Routledge, 33-50.
- Usher, Bethany. (2020). Rethinking microcelebrity: key points in practice, performance and purpose. *Celebrity studies*, 11(2), 171-188.
- Van Dijck, José. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.
- Wachter-Boettcher, Sara. (2017). *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. W. W. Norton & Company.
- Walton, Rebecca, Moore, Kristen R., & Jones, Natasha N. (2019). *Technical communication after the social justice turn: Building coalitions for action*. Routledge.
- Zeng, Jing, and D. Bondy Valdovinos Kaye. (2021). "From Content Moderation to Visibility Moderation: A Case Study of Platform Governance on TikTok." *Policy & Internet*.

Zenon, Dierdra. (2021, August 23). "Petition To Unban Popular Creators on TikTok and Reform The Banning & Appeals Process." *Change.org*,  
<https://www.change.org/p/stop-the-unfair-banning-of-creators-on-app>

## **APPENDIX A: INFORMED CONSENT FOR PARTICIPANTS**

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Informed Consent for Participants

In Research Projects Involving Human Subjects

Title of Project:

Rhetorics of Race and Content Moderation on TikTok

IRB Protocol Number: 21-127

Investigator(s):

Jennifer Sano-Franchini      email: [sanojenn@vt.edu](mailto:sanojenn@vt.edu)

Maggie Fernandes              email: [margaretf@vt.edu](mailto:margaretf@vt.edu)

### I. Purpose of this Research Project

The purpose of this project is to better understand the relationship of content moderation and race on the social media platform TikTok, particularly as this relationship and the material impacts of content moderation are experienced and understood by content creators of marginalized identities.

### II. Procedures

Participants will take part in one 30-minute to an hour long, loosely structured interviews about their experiences as content creators on TikTok. The interviews will be conducted using Zoom, a platform that supports video conferencing and recording. The interviews will be audio and visual.

Participation in this program is based on your willingness to participate.

Information will be collected from your participation in the interview. Interviews will be recorded and transcribed by Maggie Fernandes. By signing this consent form you allow the directors of this program to use any information we obtain via the above methods in research.

### III. Risks

The researchers perceive minimal risk in participating in this project.

### IV. Benefits

The benefits of this study include gaining further insights into how content creators can navigate and counter platform suppression via critical interface analysis to ensure the visibility and circulation of their content on TikTok. Additionally, I am hopeful that beginning to collect stories in this way will be a start to pushing back against TikTok's marginalizing content moderation policies.

No promise or guarantee of benefits has been made to encourage you to participate.

#### V. Extent of Anonymity and Confidentiality

Unless you ask to remain anonymous, all data collected from this program will include your information. All data will be used and stored in a confidential manner: i.e. Data coded into computers.

All video interviews will be deleted as soon as the interview is transcribed.

The parties who will have access to the data are the researchers, collaborating universities, and the study sponsor. We will not release any information without your explicit consent.

The Virginia Tech (VT) Institutional Review Board (IRB) may view the study's data for auditing purposes. The IRB is responsible for the oversight of the protection of human subjects involved in research.

Note: in some situations, it may be necessary for an investigator to break confidentiality. If a researcher has reason to suspect that a child is abused or neglected, or that a person poses a threat of harm to others or him/herself, the researcher is required by Virginia State law to notify the appropriate authorities. If applicable to this study, the conditions under which the investigator must break confidentiality must be described.

#### VI. Compensation

There is no compensation offered for participation in this study.

#### VII. Freedom to Withdraw

It is important for you to know that you are free to withdraw from this study at any time for any reason without penalty. You are free not to answer any questions that you choose or respond to what is being asked of you without penalty.

Please note that there may be circumstances under which the investigator may determine

that a subject should not continue as a subject.

Should you withdraw or otherwise discontinue participation, you will be compensated for the portion of the project completed in accordance with the Compensation section of this document.

### VIII. Questions or Concerns

Should you have any questions about this study, you may contact one of the research investigators whose contact information is included at the beginning of this document.

### IX. Subject's Consent

I have read the Consent Form and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent via electronic signature:

---

Subject signature Date

---

Subject printed name Date

-----  
-----

(Note: each subject must be provided a copy of this form. In addition, the IRB office may stamp its approval on the consent document(s) you submit and return the stamped version to you for use in consenting subjects; therefore, ensure each consent document you submit is ready to be read and signed by subjects.)



## **APPENDIX B: RECRUITMENT EMAIL**

Email Subject: Participation Requested for Content Moderation Study  
IRB Protocol Number: 21-127

Hello, my name is Maggie Fernandes. I am a PhD candidate at Virginia Tech in the Rhetoric and Writing program. I am conducting research on content moderation on TikTok and how content moderation impacts marginalized creators and am seeking interview participants. Overall, I'm really interested in hearing stories and narratives about experiencing, identifying, and navigating content moderation / suppression on this app.

Participation in this research includes a 30-minute to one-hour interview to take place via the video conferencing platform Zoom. The interview will focus on your experiences as a content creator on TikTok and your experiences with content moderation and content suppression on the platform. The video interview will be recorded, but recordings will be deleted once the interview is completed and transcribed. If you participate in this interview, your total time commitment will be between one 30– 60 minutes interviews.

Participants in this study must be 18 years of age or older and based in the United States. If you have any questions or would like to participate in the research, I can be reached at (479)200-9313 or by email at [margaretf@vt.edu](mailto:margaretf@vt.edu).

---

## **APPENDIX C: REMINDER EMAIL**

Reminder email

Email subject: Interview to occur at [specific time / date] via Zoom.

Hello! This is a reminder regarding our interview in two days about the material impacts of content moderation on marginalized creators on TikTok. The interview will be conducted via Zoom at \_\_\_\_ time and will take between 30 minutes and one hour of your time. If you have any questions, please contact me at (xxx)xxx-xxxx or by email at [margaretf@vt.edu](mailto:margaretf@vt.edu).