NORTHWESTERN UNIVERSITY

Caring Together: A Digital Ethnography of How People with Mental Illness Participate Online

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Technology and Social Behavior

By

Jessica L. Feuston

EVANSTON, ILLINOIS

December 2020

ABSTRACT

Caring Together: A Digital Ethnography of How People with Mental Illness Participate Online

Jessica L. Feuston

Social media and online forums provide spaces where people can gather beyond restrictions of geographic proximity. For some individuals with mental illness, these spaces are vital; providing outlets and communities where a multitude of experiences are accepted and understood, rather than judged against normative, often ableist standards. For nearly three years I've studied how people with mental illness use and navigate online spaces, including Instagram, Reddit, and Tumblr.

The work I've conducted through digital ethnography represents hundreds of hours of online observation, online participation, and interviews. Through this ethnography I draw deeply from the perspectives of people with mental illness; aiming to center their experiences, needs, and concerns. In contrast to prior work, which frequently extracts posts explicitly related to mental illness from online accounts, I illustrate the diverse and complex representations of mental illness online, including how experiences may be implied and better understood when contextualized within everyday life and online practices. Through my empirical work I describe the ways in which people with mental illness participate and respond to various forms of online sociotechnical control, including those involving human and algorithmic actors, such as commenting and content moderation. I demonstrate how online processes of control have consequences, sometimes severe. These processes, through largely operating within online spaces, are not removed from scholarship examining mental illness online and generating design recommendations for social media platforms and other spaces.

Following my empirical work, I end by discussing how interventions and recommendations for approaching content related to mental illness online, though motivated from a place of care, can inadvertently result in harmful and oppressive practices and technical features. Rather than caring *for* people with mental illness, as these approaches often do, I introduce an alternative way forward that involves caring together: anti-oppressive care. This approach to research and design involves working with and alongside people with mental illness to develop online spaces that support a multiplicity of experiences with mental illness and address the problems and concerns of the people with mental illness who use them.

Dissertation Committee:

Anne Marie Piper, Chair
Michael S. Horn
Ozge Samanci
Jed R. Brubaker
Alex S. Taylor

Acknowledgements

Though my dissertation research spans several years, I started writing this manuscript in earnest a few months ago. It was painful. Let's chalk it up to a global pandemic, but everything took longer than expected and was seriously exhausting. This sentiment isn't uncommon. As the pandemic continues – bolstered by the wild mismanagement, stupidity, and cruelty of our national government – we all grow just a little bit more weary. It's within these times that I've finished a dissertation, supported and held up by an incredible community of family, friends, and colleagues. Sincerely, thank you. My research, and the compilation of this manuscript, was not an individual endeavor. It's shaped by every conversation, every question, every person I have the opportunity to meet. Know that the gratitude I have for you all is infinitely larger than these few pages.

Let's kick off the specifics with my parents. Mom and Dad, I'm so thrilled to be your daughter and even more so to be your friend. You have been with me every step of the way – from wannabe daddy long-legs doctor to actual PhD. You've so patiently and generously talked with me about my work over the past few years (and always). I'm glad to have shared it all with you and I can't wait to celebrate. Thank you, also, to many of my family members; especially, Aunt Jenny and Uncle Ron Kuczuk, Aunt Donna and Uncle Bobby Murphy, and Erin Murphy.

To my advisor, Anne Marie Piper; I'm so grateful to have learned from you and worked with you for the past five years. This research, and associated program, have never been an end goal for me; rather, a stepping stone. A process of continual learning. You've facilitated that learning, this journey, and shaped my thinking so profoundly. My wonderful committee members – Mike Horn, Ozge Samanci, Jed Brubaker, and Alex Taylor – I'm so grateful to all of you. You've been nothing but generous with your time and kind and thoughtful with your critique. Thank you for joining me during this process.

I've been exceptionally fortunate over the past several years to have such enthusiastic, caring, and dedicated colleagues in the Inclusive Technology Lab; Emily Wang, Katya Borgos-Rodriguez, Mmachi Obiorah, Maitraye Das, and Abir Saha. You have all read countless drafts (some very rough!) and listened to a number of presentations (some also very rough). I appreciate talking to all of you about research approaches and the ways we, more often than we'd like, fumble through the work we're trying to conduct. You have been integral to my experience at Northwestern – thank you! This gratitude also extends to former members of the lab, including Robin Brewer, Amanda Lazar, Jordan Davison, Christina Harrington, and Elham Beheshti. You are wonderful friends, colleagues, and, at many times, mentors.

To Madeleine Agaton, Debbie Fair, Joy Stoyanova, Rebekah Sigman, and many others at Northwestern that make the TSB – amongst others – program possible: Thank you for all of the work that you do on behalf of students, myself included!

There are a number of researchers in the HCI and CSCW communities, within and outside of Northwestern University, that I've been fortunate to get to know throughout my time as a student; Darren Gergle, Jeremy Birnholtz, Madhu Reddy, Stephen Schueller, Courtney Scherr, Maria Wolters, Sean Rintel, Siân Lindley, Anja Thieme, Munmun De Choudhury, and Hitomi Hayashi-Branson. Thank you for your kind words, questions, and guidance. Additionally, to Kate Ringland, I'm so glad to have overlapped at Northwestern! Getting to work in proximity with you is something I've enjoyed. And to one of my most recent collaborators, Stevie Chancellor, I'm thrilled that we've found an opportunity to work together. I'll miss chatting over coffee, but Zoom will have to do.

My time at Northwestern would have been hollow without the many, wonderful friends I've been able to share this experience with. Sarah D'Angelo, since you've moved to Seattle, I've missed teaching together and grading papers (or not) at the Growling Rabbit. But, if we hadn't had all these miles between us, we might never have tried to screenshare Grey's Anatomy. Jim Maddock, our pizza nights at Whole Foods were some of my absolute favorites. Ari Schlesinger, I miss being your desk mate at Microsoft Research, but I'm so grateful for those months of getting to sit beside you. Ceara Byrne, your kindness and humor are entirely delightful; I miss our Atlanta adventures!

Additionally, to Hanlin Li, Leesha Maliakal, Alice Zhang, Ellie Burgess, Scott Cambo, Julia Crane, Ada Ng, Mark Diaz, Sarah Pila, Rey Maktoufi, Will Marler, Michael Ann DeVito, Kyosuke Tanaka, Eureka Foong, and Ashley Walker; thank you all for the various (often hilarious and very cathartic) conversations and gatherings. To those of you within my MTSB cohort, in particular, thank you for sharing support and motivation when it was sorely needed. Also, to my lovely friends Erin Moye and Bridget Gorta, as well as those who've watched me grow up, Ed Sheppard, Dan West, Chris Culberson, and Mary Kay Culberson.

My dissertation work would not have been possible without several other groups of individuals: participants and informants. I understand how difficult and isolating mental illness can be; I appreciate the candor, openness, and vulnerability you all have shared with me while inviting me into this part of your lives. I appreciate the time and generosity of people who've responded to my recruitment calls and participated in my interview studies. Additionally, I'm very grateful for having the opportunity to connect with people online over the past several months. Our conversations and casual interactions, such as through online posts, have been informative, insightful, and, frankly, quite fun.

Lastly, to Dan Bruzzese: You've patiently listened to my indecision, frustration, and anxiety over the past few months (and years). You've generously shared your home, cheered me up and on, dragged me out on runs, and done an infinite number of other things that have, genuinely, made dissertation writing more bearable (and possible!) than it may otherwise have been.

Dedication

For my parents, Dr. Brad Feuston and Dr. Maureen Feuston. Cheers!

Table of Contents

List of Tables

CHAPTER 1

Introduction

> *"When it was deleted, I immediately knew. It felt really crushing and really unfair and very much like we were just being misunderstood."* – Andrea

On November 14th, 2018, Reddit, a community-based content aggregation and discussion social media platform, banned r/ProED, an eating disorder support community or subreddit. This decision was grounded in a particular – and oversimplified – interpretation of pro-ED (i.e., pro-eating disorder) content and communities. Particularly, the interpretation that r/ProED violated "our Content Policy, [due to], specifically, the posting of content that encourages physical harm" (Reddit). Reddit's decision is not an anomaly; the ban of r/ProED is part of a larger movement on Reddit and other social media platforms to remove 'harmful' content from their online spaces (Blackwell et al., 2017; Chancellor et al., 2016b,d; Chandrasekharan et al., 2017; Crawford, 2019; spez, 2020; Staff, 2012). However, I argue that the ban of r/ProED resulted in a number of serious, unintended, and *harmful* consequences. I introduce my dissertation work by sharing this important, yet neglected, point of view. My aim throughout my research is to foreground this – and other – perspectives and experiences that are underrepresented in current media, industry, and academic discourses surrounding mental illness content and communities online.

1.1. Andrea's Story

The ban of r/ProED was a stark, distressing event for its 32.1k registered members. These individuals lost not only a space for expressing themselves and communicating shared experiences, but also for providing and receiving abundant mutual support. To illustrate the consequences of r/ProED's ban, I turn to one of my participants. Andrea[1] is a 23-year-old white woman who *"had undiagnosed bulimia since I was 14 to probably the age of 18 and then I had undiagnosed binge eating disorder from 18 to about 21. And then I developed anorexia, which was diagnosed. I went to treatment when I was 22 for three months. And I'm still just in recovery for that right now."*[2] Like many others, Andrea was a member of r/ProED. I spoke with her seven months following the ban of the community by Reddit. Andrea's story, presented alongside how Reddit described r/ProED (i.e., encouraging physical harm), is unexpected. Through Andrea, we come to understand r/ProED as an important communal space where she and others gathered to support one another and share experiences, such as difficulties during recovery, that other people – people without eating disorders – often don't accept or even empathize with (Puhl and Suh, 2015; Roehrig and McLean, 2010). Andrea's personal narrative challenges Reddit's classification and subsequent decision; instead, showing us how r/ProED was a space for community, understanding, and healing.

Andrea joined the r/ProED subreddit when she was *"at the tail end of [her] bad experience."* On r/ProED, Andrea was able to find *"a lot of support. Just reading people's posts and feeling validated. As the anorexia got worse, I started posting my own stuff.*

---

[1] Not her real name. Pseudonyms are used through this dissertation, unless otherwise noted.

[2] Quotes throughout the introductory chapter are modified for clarity and concision.

Mostly vent-type things. Like, how hard the day was because of food. Or a comment that someone made that made my life feel hard." Though r/ProED provided a space for Andrea to find support and validation, as well as vent about her own struggles related to living with anorexia, it was not solely an online space where she went to receive support. An important and meaningful part of Andrea's interactions within the community involved caring for and helping other people. She explained, *"Also, I like to really support people online 'cause I appreciated it a lot at the time, too. I would frequently post on people's posts. I comment on people's posts saying things like, 'You're totally valid' or 'Hey, let me know if you need support' or 'I feel that, too.'"* For Andrea, r/ProED was a community of collective, mutual aid (Piepzna-Samarasinha, 2018) – not a place that encouraged harm. On r/ProED, Andrea found a supportive space for her own, at times difficult, experience with anorexia and, in turn, provided care and support to others.

When r/ProED was banned, Andrea was *"already in recovery,"* but *"was still really active on that site in terms of posting about recovery, and how hard it was, and supporting people who are going through it. I specifically have this one article that I like to share with people because it really helped me."* The article, Andrea said, *"is about how your body changes. I think it's called 'The Body's Response to Adequate Fuel in Anorexia Recovery.' And it's an awesome article. It just talks about how, like, immediately your belly gets huge and that's, I think, the part that's really hard for a lot of people going through this."* In the following excerpt, Andrea talks about her reaction to Reddit's decision to ban the community. Though Reddit's ban of r/ProED hinged – publicly, at least – on the classification of r/ProED as a community that encouraged physical harm, Andrea's experience with r/ProED as a space for healing and recovery support troubles this oversimplified account and calls attention to the consequences of Reddit's decision. She said:

> *"It felt like no one was trying to understand what the purpose of the community was. It felt really, really unfair and confusing. And, also, very helpless. Like, there's nothing we can do about this. They took away our content. Something great about r/ProED is that it was so old that you could search things. Like, if you needed help or something or if you remembered a post that was helpful or something, you could search for that. And then all of that content – like so much good, helpful content – was removed."*

When asked to elaborate on the good, helpful content that was removed, Andrea explained:

> *". . . it was stuff that I would go and read if I was having a hard day. Like, someone had posted what to do if you feel like you're going to binge or what to do if you feel like you can't eat today. I would go and read that actively, like, 'Oh, here's some reminders for myself,' and now it's gone and I can't access that."*

The specific details of Andrea's narration call into question Reddit's decision to ban r/ProED across several interrelated points. First, Andrea's firsthand account illustrates how disruptive it was for her when r/ProED was banned, as well as how the ban of this community involved the loss of an archive of good, helpful content. Through Andrea, we see how the ban of r/ProED resulted in a complex and traumatic experience of loss. Firsthand accounts, stories, and perspectives like these are largely neglected when we think about questions related to harm, social media, and governance. Second, the ban of r/ProED shows us how platform decisions intended to reduce harm for the ‘greater good’ have serious consequences. With Reddit’s ban of r/ProED, we see how oversimplified, inaccurate classifications of what constitutes harmful content and communities can – and does – result in actual consequences that hurt people, like Andrea. Specifically, Andrea’s story illustrates how, in trying to reduce harm, Reddit actually caused it. Finally, we see that Reddit’s description of r/ProED is only part of the story – though, arguably, it is one of the most visible components. When we introduce Andrea, and many others like her, we see that the situation is much more complicated.

Reddit’s ban of r/ProED was not a malicious or isolated incident. It’s a single moment within a long history of social media moderating content and regulating the ways people post about their bodies and experiences with mental illness – or neurodivergence, madness, craziness, and psychosocial disability – more broadly. For Andrea and others, the ban of r/ProED is just one instance of oppression. This event is set within a backdrop of other types of violence and control that individuals with eating disorders experience as members of a marginalized group. The decision Reddit made to ban r/ProED and other eating disorder support subreddits echoes and is reflected back across ecologies of social media platforms and online communities. It’s reminiscent of the ways that Instagram removes and censors content related to self-harm and how people – other members of platforms like Reddit and Tumblr – can leave cruel, demeaning messages on posts about mental illness, repeatedly report an experience with mental illness that another person is trying to post, and unintentionally trigger someone through a careless, well-intentioned comment.

Andrea's story highlights three important components of my dissertation. In particular, Andrea's participation on Reddit shows us how: the use of this social media platform was part of her everyday experience of life with and in recovery from an eating disorder; classification of content related to mental illness online is complex and embedded within broader contexts of accounts, communities, and social practices surrounding online participation; and moderation of content related to eating disorders online, though often taken for granted, is a harmful and traumatic practice. My empirical chapters build from these insights: in Chapter 4, I argue that mental illness should be approached and understood as an everyday experience; in Chapter 5, I critique current mental illness classification systems and the ways they're operationalized online; and, in Chapter 6, I challenge how platforms aim to control social media use and representations of mental illness online. Following, in Chapter 7 I illustrate how my empirical work provides a foundation for reconfiguring the relationships between researchers and members of research populations as well as the ways we design social media platforms alongside people with mental illness.

1.2. Dissertation Overview

My dissertation research attends to the ways people with mental illness have, historically and contemporarily, been 'othered' and marginalized. By centering people with mental illness, I push back against systems of exclusion, displacement, and defamiliarization. I demonstrate how online spaces, a relatively new way for people to express themselves and gather with others, can operate to reinforce and perpetuate existing stigma, oppression, and violence toward people with mental illness. Through my empirical research, grounded in more than two and a half years of digital ethnography, including online observation, online participation, and interviews, I illustrate the diverse ways in which people with mental illness express their experiences online; resist controlling sociotechnical systems working to restrict their content, practices of communication, and ways of gathering online; and care for themselves and one another. My research highlights inequalities in online participation for people with mental illness and addresses how contemporary social media and research practices contribute to these, and other, forms of oppression. Following, I detail how the remainder of my dissertation is organized across six chapters.

For a brief historical background of madness and mental illness, and the ways psychiatrized people have been stigmatized and mistreated, medically and socially, see Chapter 2. This historical background grounds my decisions to represent people with mental illness as members of a marginalized group and to address mental illness from the perspective of disability, rather than through a medicalized or psychiatrized lens. Given this context, when I talk about online practices of inequality and oppression, I am doing so in conversation with the past. In addition to setting this historical stage, Chapter 2 provides related work. I address social media research practices and recommendations with respect to mental illness online. I illustrate how these practices typically exclude firsthand accounts and extract content 'explicitly' related to mental illness from other types of content that provide important context for understanding the specifics, subtleties, and subjectivities of how people living with mental illness use and experience online spaces.

In Chapter 3, I detail my research approach, which involves a multi-year and multi-site digital ethnography. My approach to digital ethnography throughout my empirical chapters involves two core components: interviews and online observation. In this chapter, I describe data collection and analysis, as well as how my populations and methods have changed over the years. I provide behind-the-scenes insight into decisions, including those concerning self-disclosure, research population, and ethnographic method, that have shaped the work I've conducted. My aim is to illustrate the development of my scholarship and how I've grappled with different – and difficult – ethical decisions during my program of ethnographic research.

Chapter 4 is the start of my empirical work, where I address the question: **How do individuals with mental illness express and communicate about their experiences online?** I illustrate that people living with mental illness post about mental illness as an everyday experience and aspect of life. This perspective provides an alternative to medicalized and deviant representations of mental illness content and communities, such as those prevalent throughout related literature (Chancellor et al., 2016b, 2017). I argue that posts about mental illness should not be extracted from accounts and other contexts in which they are produced. My aim is to normalize – albeit not trivialize – and complicate representations of mental illness online by contextualizing related content within the everyday as it is described through interviews and the ways people post online. I extend this thread of inquiry in Chapter 5 by addressing how classification of mental illness content online neglects a number of experiences that cannot be readily collected through computational means or understood through dichotomous categories that juxtapose mental illness and mental health. Lastly, attending to how individuals with mental illness post online necessitates examining audience reactions to content. In this chapter, I begin to describe how audience members – both human and computational – influence the ways people with mental illness post online. I revisit one particular facet of sociotechnical control, content moderation, in detail in Chapter 6.

Building on my work in the previous chapter, Chapter 5 examines the following question: **What types of individual and collective experiences with mental illness and mental health are missed and misinterpreted by human and computational actors operating within sociotechnical classification systems?** Algorithms on social media platforms frequently crawl and review content, supporting human actors and processes, such as content moderation, that are overwhelmed by the sheer quantity of content produced and shared every second, minute, and hour. Though computational approaches to online content and people are integral to platform operations, they are embedded within and contribute to systems that repeatedly and systemically disadvantage and, at times, oppress members of marginalized and minoritized groups (Bowker and Star, 2000; Cheney-Lippold, 2018; Noble, 2018). My findings demonstrate how computational approaches to mental illness online neglect content that is not explicitly (i.e., often this means hashtagged) related to mental illness; misinterpret mental health and mental illness content that share similar representations; and make several assumptions regarding how mental illness content is generalizable, indicative of what is happening beyond the screen, and static over time.

Though not the focus of Chapters 4 and 5, examples of content moderation are described in each. These examples, albeit briefly described, provide the motivation for my third empirical chapter. In this chapter, Chapter 6, I ask: **What is the impact of content moderation on people with eating disorders?** I emphasize content moderation as enacted by social media platforms, which often involves entangled processes, such as the provision of support resources. I describe how people with eating disorders experience content moderation as a type of loss; resist unjust practices, such as through the maintenance of multiple accounts and creation of new online community spaces; and work together to practice community norms intended to support access and navigation through spaces that can be both harmful and supportive. I argue that processes of content moderation, though integral to what platforms do (Gillespie, 2018), are ultimately contemporary forms of oppression and control that restrict how people post about their experiences with eating disorders online, enacting a particular type of conformity with respect to bodies and body management practices. Rather than err toward conformity, even if inadvertently, I discuss how platforms could design for a multiplicity of experiences with mental illness – and madness, neurodiversity, and disability – online.

Finally, in Chapter 7, I take a moment to reflect across the three empirical studies I've conducted. I use this chapter to highlight the primary contributions of my dissertation, detailing the ways these contributions cut across my body of empirical research. Building from these contributions, I discuss how power flows through and consolidates within certain ways of studying mental illness online. Though my examples are brief, I use them to illustrate how the prevalence of certain practices, as they exist now, unintentionally limits the versions of mental illness produced through the human-computer interaction (HCI) and computer-supported cooperative work (CSCW) research enterprise. Limited representation has serious consequences, including epistemic violence and oppressive care. After discussing these consequences, I move to consider potential future work for social media and online community research on mental illness. I discuss how we can reconfigure our relationships with our research populations by drawing on reflexive practices (e.g., similar to other work within and outside of HCI and CSCW) in order to recognize and appreciate how knowledge is co-constructed and affectively entangled. I end by detailing a new approach to design with and alongside people from marginalized groups: anti-oppressive care.

1.2.1. Dissertation Contributions

My dissertation makes five contributions to the HCI and CSCW research communities. First, though digital ethnography has had a home within these communities for decades (see, for example: Ames et al. (2018); Blomberg and Karasti (2013); Irani and Silberman (2013)), mine is one of the first – if not the only – to engage with people with mental illness as they participate online. This contribution provides the foundation for two more. Namely, my research involves centering the experiences, perspectives, and concerns of people with mental illness. This standpoint is underrepresented in current work, which, though involving people with mental illness, tends to center the perspectives and concerns of social media platforms and healthcare practitioners. Additionally, through digital ethnography involving online observation and participation as well as interviews, my work contributes to the diversification of mental illness online; particularly in that my work extends conversations of online depictions of mental illness beyond hashtags and observable or explicit signals.

Beyond the contributions of my empirical work, in Chapter 7 I make two additional contributions. The fourth contribution of my dissertation involves researcher self-disclosure as it relates explicitly to online research examining mental illness. Researcher self-disclosure as a practice is not new (see, for example: Ames et al. (2018); Bardzell (2010); Bardzell and Bardzell (2011); DeVito et al. (2018); Rode (2011); Schlesinger et al. (2017)). However, it is under practiced in social media research examining mental illness content online. Building from a discussion of my own apprehension in Chapter 3, in Chapter 7 I revisit researcher self-disclosure and reflexivity, making a case for writing researcher disclosure as relational (i.e., specific to a researcher and the population they research). I end Chapter 7 by introducing a novel approach to research and design with marginalized groups: anti-oppressive care. Anti-oppressive care is a reflexive approach to research and design that involves thinking-with people with mental illness, disclosing researcher-research population relationships, and centering logics of care that, as much as possible, are not harmful or oppressive. I introduce and start to develop this approach here, intending to build and extend this work moving forward.

In the following paragraphs, I detail the above contributions and how they are addressed throughout the following dissertation:

Foremost, my research approach through digital ethnography provides the flexibility and the means through which people's experiences can be heard, examined, and understood within their interdependent sociotechnical contexts. Through digital ethnography, including online observation, online participation, and interviews, my dissertation examines how people with mental illness use and navigate online spaces, including online forums and social media platforms. This approach involves close digital proximity and a commitment to individuals with mental illness. Proximity, as I mean it here, involves observing and hanging out in digital spaces (i.e., rather than physical ones), such as ones for eating disorder support, and inviting individuals to participate in interviews. Digital ethnography is an immersive, interconnected, emotional approach that supports embedded understanding of how individuals, as well as the ways they narrate their experiences online, and communities change over time. It supports attention to nuanced social dynamics and tensions, such as those between specific individuals and within and between online communities. In Chapter 7, I detail how digital ethnography provides spatially and temporally flexible data collection (e.g., through observation, participation, and interviews drawing on contemporary and historical events), insight into social dynamics between various online actors, and affective understandings of experiences and practices. This approach to examining how people with mental illness use social media is uncommon in the HCI and CSCW literature. Yet, it compliments existing work by attending to significant gaps with respect to first-person accounts, informant insights, and content that can't be detected or readily classified by current computational approaches.

Throughout my dissertation work, I aim to enrich our understanding of how people with mental illness use online spaces by centering their perspectives and experiences – rather than what we think we know about mental illness online (e.g., with respect to what is good or bad, helpful or harmful). Centering is a multi-dimensional process through which I listen to and engage with the needs and concerns of people with mental illness (e.g., such as letting concerns about content moderation shape Chapter 6), recognize and appreciate that individuals living with mental illness are the experts of their own experiences, and position myself as a learner within interviews and online observation and participation. This approach also aims to normalize experiences with mental illness, particularly with respect to the variety of ways in which these experiences are expressed online. Particularly, because experiences are centered here, they are not presented on the margins or as ‘the other’. This perspective has supported a new understanding of what is ‘good’ and ‘bad’ for people with mental illness in online spaces, as well as recognition of the labor people perform within these online spaces, such as to resist unjust practices and help one another navigate online content that may simultaneously be harmful and helpful. These perspectives are valid and pressing, but currently underrepresented.

My dissertation also contributes to diversifying experiences with mental illness in the literature by interviewing people and observing accounts and communities, rather than relying on thousands of single posts collected through hashtags. In much current work, particularly work relying on hashtags for data collection, experiences may need to be explicitly signaled in order for algorithms to capture them during data collection. My work illustrates how many people express their experiences with mental illness online without making any explicit signals or claims. These experiences and ways of representing mental illness online are largely absent from related work; which, additionally, may remove posts related to mental illness from surrounding content during data collection and analysis. In my work, I aim to keep this context intact when collecting and analyzing data. I argue that context is vital for understanding the diversity of experiences with and representations of mental illness online, as well as for understanding nuanced similarities and differences between mental illness and mental health.

In Chapter 3 and, again, in Chapter 7, I discuss how my own feelings and practices around researcher self-disclosure can be described through ‘apprehension’. In the majority of my empirical work, I did not self-disclose my position as a researcher with an eating disorder due to discomfort and concern of stigma. As such, I was complicit in producing research that sees everywhere and everything while coming from nowhere (Haraway, 1988). This type of work is common with respect to social media research studying mental illness. Rarely do researchers practice self-disclosure or consider their relationships with participants and research populations (i.e., including when mediated by content produced by individuals online). In this respect, we are behind other HCI and CSCW research domains, where self-disclosure and reflexivity are important parts of research processes. In Chapter 7, I discuss what reflexivity can contribute to research on mental illness online. Namely, I discuss how incorporating a reflexive practice – one that involves varying elements of researcher self-disclosure – supports examining researcher-research population relationships in order to acknowledge and uphold co-constructive, affective elements of knowledge and care.

I end my dissertation by introducing a new direction for social media research and design: anti-oppressive care. Anti-oppressive care involves thinking-with people with mental illness, disclosing researcher-research population relationships, and centering logics of care. This approach spans all aspects of research and design, including initial inquiry, study development, recruitment, representation, and so on. I did not begin my dissertation work with anti-oppressive care. This approach is something I’ve developed throughout my empirical work and dissertation writing. As such, it’s very contextualized in my research over the past several years. I aim to outline anti-oppressive care, including its background and important elements, to provide a basis for myself and other scholars conducting research with members of marginalized groups.

CHAPTER 2

Background and Related Work

Studying mental illness as digitally mediated through social media involves examining how people communicate about their own experiences, participate and gather in various online spaces, and respond to sociotechnical systems of control. Throughout my dissertation, I use ‘sociotechnical’ to mean the interconnected and intertwined social and technical dimensions of various processes, such as communicating, gathering, and participating within online spaces. Sociotechnical relates to the interplay in which society shapes technology and, simultaneously, technology shapes society. Sociotechnical systems and processes are political (Winner, 1980). This dimension is particularly important for my work, as the politics of sociotechnical processes often disproportionality impact marginalized groups, of which individuals with mental illness are part.

My approach to this area of study involves centering people with mental illness, such as through digital ethnography and a sincere commitment to listening (Williams and Boyd, 2019). In discussing ‘centering’ throughout my work, I commit to attending to the positions and standpoints of individuals who are frequently relegated to the margins. Centering this particular group involves recognizing that people live with mental illness, madness, neurodivergence, and psychosocial disability amidst harmful and oppressive societal stigma and practices; violent practices that are amplified for individuals who encounter oppression multiply, including across gender, race, and sexuality (Clare, 2017; Crenshaw, 1990; Piepzna-Samarasinha, 2018; Schlesinger et al., 2017; Ymous et al., 2020). With the HCI and CSCW research community, the idea of ‘centering’ is pervasive, usually applied within the context of user-centered or human-centered research and design. However, as Emma J. Rose and colleagues note (Rose et al., 2018), research and design within these fields consistently center certain types of people and experiences (e.g., white, heterosexual, cis, male, and abled) over others. While ‘centering’ people with mental illness is not without its own problems or potential for harm, as I address in Chapter 7, my intention here is to use this practice of ‘centering’ to normalize marginalized people and groups as well as amplify and engage with their needs and concerns.

Societal attitudes toward people with ‘othered’ bodyminds (Clare, 2017) – in this instance, particularly those which are disabled, mad, and ill – are dangerous and traumatizing. Take, for example, historical and contemporary practices of forced institutionalization (Amador and Johanson, 2000; Foucault, 2003; Scull, 2015), as well as pervasive microaggressions and discrimination – individual and structural – within and beyond clinical practice (Corrigan et al., 2004; Gonzales et al., 2015; Green and Ubozoh, 2019; Holley et al., 2016; Peters et al., 2017). Taken together, these vast and far reaching practices of oppression impact lives, including the decisions that people make with respect documenting and communicating their personal experiences, as well as, when desired, seeking support and healing from peers and mental healthcare practitioners (Chancellor et al., 2016b). Though I frequently employ the use of ‘mental illness’ in the following chapters, which is a medicalized way of talking about certain experiences, my point of view is informed jointly by Disability and Mad Studies literature and firsthand accounts of mental illness, madness, and psychosocial disability –

not to mention the myriad of other ways in which people communicate about and reclaim these experiences. Medicalized language related to illness and disorder permeates non-medicalized spaces, including online forums and social media platforms. My use of ‘mental illness’ is not intended to psychiatrize individuals and their lives, but, rather, to illustrate how many individuals within my dataset (e.g., interview participants, online posters) narrate and construct their own experiences, and connect with others.

Normalized acts of societal oppression and interpersonal violence contribute to the ways in which people living with mental illness use online spaces. Specifically, online spaces are some of the few in which people can gather without treatment as some sort of requisite (e.g., group therapy sessions, inpatient facilities). Social media and online forums provide digital affordances for communicating about mental illness and participating in networks and communities with people who have similar experiences (Andalibi et al., 2017; Bartlett, 2015; Blair and Abdullah, 2018; Chancellor et al., 2018; De Choudhury and De, 2014; De Choudhury et al., 2017; Li et al., 2016; Zhang et al., 2018). Specifically, social media and online community forums present opportunities for individuals to connect and gather with others, seek and give support, and communicate about their experiences with mental health and mental illness. Additionally, some technical affordances associated with anonymous accounts on social media platforms and online forums facilitate communication and other types of participation divorced from concerns of being outed or encountering offline repercussions, such as loss of employment or forced treatment. Another appeal of online spaces relates to the ways in which social media provide opportunities for individual and group interactions connected through shared experiences rather than geographic proximity. In these instances, social media platforms and online forums can provide spaces for people to exist and be heard *just as they are* (Piepzna-Samarasinha, 2018), wherever they are in their particular experience with mental illness.

As researchers, we still have so much to learn about the role that social media and other digital spaces play in the experience of living with mental illness. In the research presented throughout my dissertation, I use digital ethnography to examine how people with mental illness use social media and to interrogate the sociotechnical systems and cultural norms implicit in these practices and surrounding interactions. Below, I provide a brief historical context to help situate my understanding of mental illness and a review of the literature as it relates to the study of mental illness online and the ways in which marginalized groups use online spaces.

2.1. Historical Context

The contemporary experience of living with mental illness cannot be understood without considering its historical context. Historically, individuals living with mental illness have encountered stigma, social ostracization, and forms of oppression, including forced institutionalization and other forms of treatment (Foucault, 2003; LeFrançois et al., 2013; Scull, 2015). Madness and insanity, though linguistically reclaimed within the last few decades by mad and disabled activists, originated centuries ago as ways to describe variances within human experience. The history of these terminologies, as well as the ways in which they intersect with the comparatively young concepts of mental health and mental illness (Bertolote, 2008; Scull, 2015), shapes current practices and perceptions of individuals who are neurodivergent and/or otherwise living with mental illness, madness, and psychosocial disability. For example, there has long been an interplay between religious and medical interpretations of lived experience with madness. These entanglements, though now leaning heavily toward medicalized versions of experience, are present in stigma and the ways in which people with mental illness are still marginalized (e.g., including how individuals are not necessarily viewed as the experts of their own experiences, how people with mental illness are stereotyped and represented as weak and/or wicked, and how people can be forced into institutional settings).

Marginalization of people with mental illness, in part, may stem from quantified versions of mental illness which, like other types of disability, have been used to bolster ableism, racism, and misogyny (Davis, 2006). Specifically, scholars have addressed the ways that the scientific revolution in psychiatry contributed to the quantification of experience and emphasis on biomedical diagnosis and pharmaceutical treatment (Lewis, 2006). Encapsulating and following World War I was an emphasis on national interests, productivity, and eugenics in Western culture. Developing within this environment, modern medical labels and practices picked up further ground following the Second World War, in which psychiatry rose to prominence as an authority of mental health and mental illness in all aspects of life (LeFrançois et al., 2013; Scull, 2015). Derivations from ‘the norm’, whether due to disability, mental difference, or other human attributes, were viewed as contributing to the disease of the nation (Davis, 2006). As I use it here, ‘the norm’, or normativity, references an evaluative standard to which everything and everyone is compared. In the United States, normativity is deeply, irreparably entangled with whiteness, maleness, and able-mindbodiedness. As Patty Berne describes, people who are part of the "subjugated other" not only differ from an implicit societal norm but are also positioned subordinate to it (Berne as quoted by Leah Lakshmi Piepzna-Samarasinha in *Care Work: Dreaming Disability Justice* (Piepzna-Samarasinha, 2018)):

> "...able-bodied supremacy has been formed in relation to other systems of domination and exploitation. The histories of white supremacy and ableism are inextricably entwined, both forged in the crucible of colonial conquest and capitalist domination. One cannot look at the history of US slavery, the stealing of indigenous lands, and US imperialism without seeing the way that white supremacy leverages ableism to create a subjugated 'other' that is deemed less worthy/abled/smart/capable..."

In the excerpt above, Berne addresses how the construction of normativity is founded through a history of white supremacy and patriarchy. These systems of oppression aim to classify and control, subjugate and exterminate brown, black, femme, disabled, ill, and mad bodyminds. From this history, we see how madness, particularly as entangled with disability, became synonymous with depravity. 'Deficits' in mental behavior were associated with inefficiency, criminality, and promiscuity, particularly through association with "feeblemindedness" (Davis, 2006). These conceptual associations justified further oppression of psychiatrized individuals, including forced sterilization and, following deinstitutionalization, the rise of social control through medication (LeFrançois et al., 2013; Scull, 1991).

Though "feeblemindedness" has been replaced by clinical terminology through the dissemination of the Diagnostic and Statistical Manual of Mental Disorders (DSM)

(American Psychiatric Association, 2013) and similar Western texts, the stigma and stereotypes solidified during this era of American history still underlie national attitudes toward mental illness. State asylums contributed to coercive and inhumane treatment, including electroconvulsive therapy and lobotomy, for decades (LeFrançois et al., 2013; Scull, 1991). The spectacle and sensation of the asylum saw Bedlam, along with other mental institutions of the time, open its doors to the profitable practice of asylum tourism, in which members of the public came to gawk at patients and other institutionalized individuals for shock and entertainment (Scull, 1991, 2015; Torrey and Miller, 2001) . Though treatment in asylums was challenged during the moral hygiene movement (Bockoven, 1972), the shockwave of its history (e.g., coercive practices, spectacle and romanticization, diagnostic labeling) reverberate into the current era.

Personal accounts of mental illness are important for understanding attitudes regarding lived experience (Green and Ubozoh, 2019; Kempe, 2000; LeFrançois et al., 2013; Mollow, 2013). However, they are also prone to psychiatric reinterpretation through applied labels of disorder and censorship through content moderation. Recent decades notwithstanding, the history of madness and mental illness was not passed down through the subjectivity or expertise of individuals with lived experience. Instead, many of the stories and knowledge that remain come to us from third party (i.e., outsider) interpretations and retellings (Foucault, 2003; Kleinman, 1989; Scull, 2015; Torrey and Miller, 2001). The autobiography of Margery Kempe is a well-cited exception (Kempe, 2000). In her writings, Margery Kempe details experiences with demonic visions and conversations with God. While some contemporary analyses situate this subjective narration in the context of modern diagnosis and psychosis (Delaney, 1975), others, instead, describe how Margery Kempe "transformed illness and health into adventure, devotion, and belief in a direct relationship with God" (Freeman et al., 1990). The differences in these interpretations call attention to how a clinical lens applied to subjective narration can reduce meaningful experiences and stories through pathologization. Perhaps unintentional, this practice is a subtle form of oppression in which personal experience is subjected to a psychiatric rewrite.

Contemporary narratives and movements related to mental illness and madness involve reclamation of identity and acceptance. Madness is a conceptual taking back of cultural, social, and personal identity that challenges practices associated with the categorization and pathologization of experience through psychiatric labels of mental illness (Costa et al., 2012; Poole and Ward, 2013). For example, the Mad Pride movement, originating in Toronto, Canada in the early 1990s, advocated for individuals to be proud of their madness. In part, it aimed to normalize experiences with madness and break down the stigma that mad individuals often encounter. To be clear: Being proud of a mad identity does not necessitate a rejection of wellness and healing – even medicalized versions of those experiences. Take, for example, the Fireweed Collective (Fireweed Collective), formerly known as the Icarus Project, which bridges this gap with their Wellness Map or Mad Map exercise (T-Maps Community; Werning, 2019); a "very practical document to be written in good health and shared with friends and loved ones" to determine what wellness looks like for a particular person and to help individuals care for one another (Du Brul, 2012).

As discussed, dominant societal narratives have long held counterproductive views of madness, which marginalize entire communities by casting them as deviant, outsider, and other and devalue diverse mental experience through biological reductionism and psychiatric diagnosis (Starkman, 2013). Mad Pride and related endeavors involve activism, social critique, and resistance. In particular, Mad Pride resists various forms of oppression, tracing its roots to the antipsychiatry, mental patients' liberation, and psychiatric survivor movements of the mid-nineteenth century (LeFrançois et al., 2013; Lewis, 2006; Scull, 2015). At the heart of these movements was a call to reject the oppressive nature of psychiatric labeling and treatment, which was particularly prevalent, and to undermine the prejudicial use of certain terminologies, such as "mad" (Abraham, 2016). Additionally, these movements worked to claim madness as an important and valid aspect of identity and to educate the public regarding the stigma against psychiatric survivors and mental illness (C/S Info Centre). These practices are reminiscent of how disability has moved in activist circles from a medical diagnosis or label to a social and political way of being (Linton, 1998). My work learns from and converses with these histories, aiming to engage with the concerns and needs of people with mental illness.

2.2. The Study of Mental Illness Online

Internet research focusing on mental illness has been around for over a decade (Baker and Fortune, 2008; Brotsky and Giles, 2007; Fox et al., 2005; Juarascio et al., 2010; Shade, 2003; Sharkey et al., 2011). Early work often investigated websites and online forums – and, at times, social networking sites – as spaces where individuals talked about topics such as self-harm, depression, and eating disorders (Bartlett, 2015;

Csipke and Horne, 2007; Juarascio et al., 2010; Westerlund, 2012). With few exceptions, including Baker and Fortune (2008) and Csipke and Horne (2007), much of this early work involved website and online forum analysis without incorporating firsthand accounts or perspectives from the people with mental illness producing certain content and gathering in particular online spaces. Though I vaguely refer to ‘particular’ spaces, what I mean to say is that much work focused on content perceived to be ‘pro’ or promoting suicide and self-injury, including cutting and eating disorders. If we look at how contemporary research approaches mental illness online, we see that, in some regards, not much has changed. For example, research without firsthand accounts is more common than research with firsthand accounts, and there is still an emphasis on communities that researchers interpret as promoting or glorifying behaviors that are risky or dangerous to individual and community health. Recently, however, researchers have started to diversify these approaches and topics (Chancellor et al., 2018; McCosker and Gerrard, 2020; Pater et al., 2019a).

Within HCI and CSCW, the study of mental illness online coincides – and continues to be entangled – with the rise of online communities and, especially, social media platforms, including Facebook, Instagram, Reddit, and Tumblr (Andalibi et al., 2017; Blair and Abdullah, 2018; Chancellor et al., 2016a,c,d, 2017, 2018; De Choudhury, 2015; De Choudhury and De, 2014; De Choudhury et al., 2013b, 2014, 2016, 2017; Li et al., 2016; Pater and Mynatt, 2017; Pater et al., 2016a,b, 2019a,b; Tsugawa et al., 2015; Zhang et al., 2018). Researchers early to these online spaces favored computational tools for data collection and analysis; in particular, using these computational, quantitative approaches to characterize and predict certain types of content (De Choudhury, 2015; De Choudhury et al., 2013a,b,c, 2014). Much of this early work set the agenda and the stage for studying mental illness online. For example, Munmun De Choudhury pioneered the application of predictive computational techniques with respect to studying mental illness, initially post-partum depression (De Choudhury et al., 2013a), on social media. Along with many collaborators, she extended this approach and thinking to work related to experiences such as depression (De Choudhury et al., 2013c), suicidal ideation (De Choudhury et al., 2016), and eating disorders (Chancellor et al., 2016a,c; De Choudhury, 2015). Many others have followed suit, extending research on mental illness online to a variety of platforms and experiences, as well as ethical considerations of these particular practices (Chancellor et al., 2019b). Due to the ways in which researchers have characterized certain intersections of experiences and online spaces, we've given legitimacy to certain people and practices, excluded others, and, ultimately, can now control that which we 'see' (Bowker and Star, 2000; Scott, 1998), such as through recommended design interventions leveraging systems of classification. As of yet, classification is not informed by or produced alongside people living with mental illness. However, though classification (as related to detection and prediction) is certainly a common approach to online spaces, it is certainly not the only one. Researchers may also study how people communicate and disclose about sensitive mental health and mental illness experiences (Andalibi, 2020; Andalibi et al., 2017; Blair and Abdullah, 2018; Chancellor et al., 2018; De Choudhury et al., 2017; Ernala et al., 2017; Manikonda and De Choudhury, 2017); react to platform-imposed interventions, such as content moderation (Chancellor et al., 2016b,d; Gerrard, 2018); and use online spaces for support and empowerment (Kushner and Sharma, 2020; Li et al., 2016; Pruksachatkun et al., 2019; Zhang et al., 2018).

Research incorporating interviews, including Pater et al. (2019a), Zhang et al. (2018), and Li et al. (2016), illustrate how firsthand accounts complicate certain well-held empirical findings related to how mental illness is expressed online. For example, these works examine tensions between different actors online and describes how people with mental illness may not post explicitly about their experiences or about mental illness at all. In contrast, much of the work that does not elicit firsthand accounts relies on overt, culturally prominent and, at times, stereotypical representations of mental illness (e.g., #proana, a black and white image of an arm with scratches and cuts) in order to classify content. This body of work, because it does not engage with content posters, cannot develop interpretive insight that takes into account the very individual and nuanced ways that people post about their experiences – including, for example, how people might not engage in obvious forms of representation. For example, similar to my work in Chapters 4 and 5, Jessica A. Pater and colleagues describe how three individuals in recovery from eating disorders participated online in ways that would not have been classified as related to living with a mental illness, even though these individuals were clinically diagnosed with eating disorders (Pater et al., 2019a).

As mentioned, much work in this domain does not directly engage with individuals living with mental illness who use online spaces (e.g., such as through interviews, surveys, and hanging out with members of the research population online). Many researchers, instead, collect and analyze online content. In doing so, researchers may extract content related to mental illness from other aspects of community and individual life (Chancellor et al., 2016a,c; De Choudhury, 2015; Manikonda and De Choudhury, 2017; Pater et al., 2016a). In data collection, this translates to reliance on a strict selection of hashtags through which to collect, organize, and examine specific practices. Separation of mental illness from other parts of life, both online and off, neglects a holistic examination of the ways in which context plays a role in understanding illness and the subjective experience of being ill (Andalibi et al., 2017). As other researchers note, mental illness does not exist in a vacuum (Pater and Mynatt, 2017). For this reason, throughout my digital ethnography, though, at times, I certainly foreground mental illness, I aim to keep experiences contextualized.

In addition to the extractive nature of much work, including aspects of my own digital ethnography, many researchers bring into their interpretation of mental illness a clinical point of view. For example, researchers may reference DSM descriptions of eating disorders and self-harm in their related work, work with clinicians during data analysis, and develop or recommend clinical uses for social media content (Andalibi et al., 2017; Li et al., 2016; Manikonda and De Choudhury, 2017; Pater and Mynatt, 2017; Pater et al., 2016a, 2019a; Tsugawa et al., 2015). One benefit of partnering with clinicians means potentially connecting individuals with digitally mediated forms of treatment and care when in-person services are not wanted or readily accessible. However, I write this with a note of caution: The possibility for digital spaces to provide access to medicalized forms of treatment should not be a stand-in for physical locations and services. As researchers who work to develop accessible health technologies, we also need to fight for policies that provide mental health funding, especially to remote areas and underserved groups of people.

Despite benefits of partnering with clinicians, there are drawbacks, as well. For example, not all experiences with mental illness are easily categorized or psychiatrized – nor should treatments and interventions always be the end goal (Ymous et al., 2020). For example, not all individuals view clinical treatment or recovery as desirable or even possible. While joining with mental health practitioners in research does invite a particular set of interpretations and solutions for a specific subset of individuals with mental illness, it does not cater to everyone or every problem within this domain. Emphasis on clinical definitions of mental illness may neglect the holistic experience of having or having had a mental illness and marginalize the perspectives of individuals who do not abide by socially normative illness narratives or trajectories (e.g., such as individuals living with a chronic mental illness and individuals who experience cycles of recovery and relapse).

As is common in much HCI and CSCW work, researchers set out to solve problems. In this domain, the predictive power of computational techniques with respect to mental illness online, in particular, helps researchers characterize online spaces – though in very specific, limited ways – and address risky content and perceived behaviors through interventions, many of which are intended to reduce the prevalence of certain content, such as pro-eating disorder content, online (Chancellor et al., 2016d, 2017; De Choudhury, 2015; Pater et al., 2019a). While some research does recommend promoting positive health content and behaviors (Chancellor et al., 2016c; De Choudhury, 2015;

Gerrard, 2018), this is driven by researcher and, often, clinical practitioner expertise, rather than through engagement with the needs and concerns of individuals living with mental illness and posting or viewing certain ‘risky’ content. Though these approaches are developed out of concern for individuals using online spaces in particular ways, they problematize certain experiences with mental illness and privilege others. This means that this type of work may miss aspects of individual experience where risky mental illness content is not the problem. For example, individuals with mental illness may encounter other problems due to experience with mental illness (e.g., stigma, oppression, invalidation, targeted forms of harassment), as well as other problems and disruptions in life (e.g., break-ups, failed tests, global pandemics), some of which may be triggering or can exacerbate dangerous aspects related to living with a mental illness.

Across my dissertation, I argue for increased nuance in order to understand other disruptions, as well as the diverse ways that people communicate about mental illness online. Here, I’ll end by calling attention to research on self-harm and suicide websites from the 2000s conducted by Darren Baker and Sarah Fortune (Baker and Fortune, 2008). Baker and Fortune make several statements that it feels like we’re just now coming back around to in HCI and CSCW social media research on mental illness. Specifically, findings from this interview study illustrate how the authors engage with the tensions of online spaces where people discuss self-harm and suicide. Explicitly, they mention, “If health professionals and researchers hope to understand people who use self-harm and suicide websites, and engage them in their services, they must take a more balanced view and not focus solely on the possible risks associated with using such sites." While current research may, at times, engage in this balance, such as acknowledging recovery content in pro-anorexia spaces, we still see dichotomies in how these different content types and communities are positioned in relation to one another. My aim is to move beyond this interpretation by "[abandoning] simplistic classifications" of online spaces as good or bad, risky or supportive (Baker and Fortune, 2008).

Though I do not contest that certain practices associated with mental illness, such as purging, may be harmful, dangerous, risky, or destructive, I do question the appropriateness of technological solutions and recommendations that aim to remove, reduce, or otherwise control how people post about these practices online. Design recommendations for mental illness content online should not limit the personal agency of individuals with mental illness, who are often subject to policies put in place to 'help' them, but that are not informed *by* them (cf. Park et al. (2013); Pater et al. (2016b)). Researchers have begun to examine the ethical and social ramifications of technology in this space (e.g., censorship on marginalized communities (Pater et al., 2016a) and the ethics of online inference (Chancellor et al., 2019b)). In my work, I advance perspectives on 'helpful solutions' by foregrounding the experiences and concerns of people with mental illness, rather than traditional societal perceptions (i.e., surrounding mental illness, deviance, and treatment) and the importance of individual differences. Additionally, throughout my dissertation, my aim is to normalize communicating about mental illness online and to engage with certain groups in order to understand how to better develop technical systems that support their wellness and ways of interacting online.

2.3. Marginalized Communities Online

Individuals labeled as outsiders to social and cultural mental health norms have endured stigma, social ostracization, and oppression (Clare, 2017; Foucault, 2003; LeFrançois et al., 2013; Piepzna-Samarasinha, 2018; Scull, 1991, 2015). As such, an important vein of work in HCI and CSCW involves examining how members of marginalized groups find and build communities online (Haimson and Hoffmann, 2016; Haimson et al., 2015a, 2016; Renninger, 2015). In this work, I join with other mental health and mental illness and social media researchers in considering the experiences of people with mental illness through a history of oppression and marginalization (Li et al., 2016; O'Leary et al., 2017; Pater et al., 2016a; Zhang et al., 2018). Examining how members of marginalized groups participate in online spaces involves reflecting on research practices and ethics (Barron, 1999; Haimson et al., 2015b). For example, other researchers have argued that scholars should be mindful of the ways in which researcher norms, as well as societal ones, differ from those of the 'subculture' being studied (Bruckman, 2006; Haimson et al., 2015b). Situating the experiences of these individuals in the context of marginalization helps us attend to power dynamics and differentials, acknowledge labor practices, and contribute to a growing body of literature that examines the marginalization of groups and designs for more equitable online experiences (Blackwell et al., 2017; Haimson et al., 2016; Irani and Silberman, 2013; Renninger, 2015; Scheuerman et al., 2018).

People with mental illness may seek online socialization and support for a number of reasons (Andalibi et al., 2017; Bowler et al., 2012; Eikey, 2016; Keski-Rahkonen and

Tozzi, 2005; Li et al., 2016; Zhang et al., 2018). Online spaces can help reduce feelings of loneliness and isolation derived from stigma and also connect individuals with people and communities where experiences are shared and understood (Chang and Bazarova, 2016; Eikey and Booth, 2017; Keski-Rahkonen and Tozzi, 2005). Similarly, beyond research examining mental illness, there is a body of work that examines how individuals with other 'non-normative' identities and experiences engage and participate online (DeVito et al., 2018; Dym et al., 2019; Haimson et al., 2015a; Liu et al., 2017; Mehra et al., 2004). Research within this corpus often addresses the everyday lives of people from marginalized and minoritized groups. For example, Bharat Mehra and colleagues examine internet usage by low-income families, sexual minorities, and African-American women (Mehra et al., 2004), finding that the internet can effectively operate as an instrument of empowerment. Other research foregrounds the importance of having safe and supportive online spaces, particularly for opportunities related to self-disclosure and identity work (Andalibi and Forte, 2018a; Andalibi et al., 2017; DeVito et al., 2018). Collectively, this scholarship documents the benefits and detriments of participation online and discusses how to improve online spaces for individuals at society's margins.

Though there are benefits to online participation for members of marginalized groups, there are also an array of harms. For example, women, people of color, members of the LGBTQ+ community, and individuals with mental illness all encounter disproportionate and targeted forms of harassment online (Duggan, 2017; Feuston and Piper, 2019; Lenhart et al., 2016). Ongoing research aims to address problems with harassment, such as through work with social organizations, communities, and platforms, including

Hollaback (Dimond et al., 2013) and HeartMob (Blackwell et al., 2017). Social media platforms are also invested in understanding and solving problems related to online harassment (landoflobsters, 2019). However, as Tarleton Gillespie describes, platform efforts related to reporting and mitigating harassment can themselves contribute to the problem (e.g., such as when individuals organize to use reporting features to flag or report a specific user – or group of users – who they do not agree with or like) (Gillespie, 2018). Throughout my dissertation, I consider how features designed for good (i.e., moderation to support positive experiences and health) can work to exclude individuals with mental illness and contribute to the oppression of a marginalized group online.

CHAPTER 3

Methodological Overview

In this section, I describe my approach to research and provide an overview of the methods used throughout my empirical work. The specific details of methods and datasets are described in the following chapters (i.e., rather than here), where they are most relevant. The various methods used across my empirical research are deeply entangled with one another, connected and extended over years of work. They are not intended to be viewed as compartmentalized or isolated sets of information. Additionally, note that the chapters are arranged outside of chronological order. This decision reflects my intent to, first, describe how mental illness is a familiar and, at times, routine part of everyday life for people living with mental illness. Following, my work continues to present empirical findings related to mental illness and mental health online, while also working to critique certain approaches, including those related to computational and categorical research methods as well as how online social media platforms approach community and site governance, that aim to exclude facets of mental illness from the online everyday.

3.1. Researcher Self-Disclosure

I am white, queer, cis, and female; was raised upper middle class in suburban Pennsylvania; and have an eating disorder. Until this point, I've been hesitant to explicitly include my positionality in publications of my empirical work. Across Chapters 4, 5, and

6, for example, these important considerations are largely implied or absent. However, in the reflexive undertaking of writing this dissertation, I've taken a breath to pause and consider my stance with respect to writing on my own, multiple positions. Though apprehensive, here I aim to make my position clear in order to situate my research approach and decisions related to how I represent my research population.

The context of my own experience, particularly as it relates to having an eating disorder, is important for all aspects of my work, including times when my approach to digital ethnography has been 'problematic', such as through covert observation or lurking. Claiming an eating disorder is difficult. This difficulty, and the apprehension I mention earlier, stems from multiple sources, including concerns around academic ableism and the potential it has to invalidate my work (Dolmage, 2017) as well as, and, perhaps, more pressing, personal experiences of living with – and concealing – a mental illness. As someone who's been able to pass (i.e., as not having a mental illness), talking so openly about my own experiences is uncomfortable and feels, in some ways, unnecessarily risky. However, my experiences greatly influence how I conduct research, including my decisions regarding the populations I study and the ways I approach this work, which is why I have decided to write them here. I continue this conversation in Chapter 7 by reflecting on researcher self-disclosure in HCI and CSCW research, describing tensions between self-disclosure and outing, and discussing how emphasizing researcher-research population relationships may provide an alternative form of disclosure that does not necessarily entail firmly locating or 'outing' every researcher position (i.e., particularly those that may inadvertently put the researcher at risk).

3.2. Research Approach

In studying how people with mental illness participate online, my aim is to foreground the point of view of the population and specific subgroups that I work with. In saying that I foreground this point of view, or that I am committed to this population, I mean that, as much as I am able, I engage with the needs and concerns of my research population and, more specifically, the people I encounter through interviews and online spaces. Though the decision and responsibility of which voices to include – and, by association, exclude – is ultimately mine, I aim to co-construct knowledge as much as possible, rather than strictly apply my own interpretive lens. By including these voices, I hope to reduce perceived 'otherness' of people with mental illness by presenting expression of mental illness – and mental illness itself – as a familiar and normal part of everyday life. This means finding a balance between dis/encouraging harmful behaviors. While I certainly don't want people to harm themselves, I completely understand needing a space to talk about ugly, shameful, and scary experiences – particularly without the judgement of those who may not understand. In positioning mental illness and talking about mental illness as normal, I simultaneously seek to illustrate how 'normative practices', such as content moderation and the provision of support resources, are strange. This isn't to diminish their potential to help, but, instead, to call attention to the ways in which they can harm people with mental illness. Ultimately, my approach flips dominant paradigms of mental illness in social media research upside down and inside out in order to center people with mental illness who use online spaces and to question the role of sociotechnical systems in these practices of care and togetherness.

An important analytic lens that cuts across the various empirical work of my dissertation relates to anti-oppressive design (Smyth and Dimond, 2014). I approach research and design for good "as that which strives to end one or more forms of oppression" (Smyth and Dimond, 2014), and attend to the ways in which marginalized communities negotiate and resist socially oppressive forces and care for one another online. This relates to viewing mental illness through the sociocultural and political lens of disability, rather than as medically-defined diagnostic criteria or even individualistic, medicalized disability. My work not only centers people living with mental illness, but also positions them as experts with respect to their own experiences. This view is significant, as expertise is typically withheld for medical practitioners and HCI and CSCW researchers, rather than people living with mental illness, madness, and psychosocial disability.

3.2.1. Research Population

When I began this research in November 2017, I was primarily interested in examining how people living with mental illness – then phrased as 'mental health conditions' – expressed relevant experiences online; specifically, on Instagram[1]. I approached mental illness broadly at this point, aiming to be in conversation with prior work and to not

---

[1] Instagram (www.instagram.com) is a social media platform released in 2010 that prioritizes visual types of content. For example, account holders can upload pictures and videos with – and without – a number of different in-platform filters. Posters may also use captions when sharing content. On Instagram, captions are varied, including text and emojis as well as links to other accounts (i.e., tagging using the '@' symbol) and content groupings indicated by the '#' symbol (Messina, 2007). As a platform, Instagram is interactive. Account holders may follow and direct message one another. Followers of certain accounts, as well as other account-holding audience members, may like, comment, share, and report posts. Additionally, people can view content on a curated content feed comprised of posts from accounts they follow, search for specific types of content (through the use of hashtags as search terms), and receive content recommendations through the Explore tab, which generates content through an algorithm that involves the types of posts and accounts a person likes and follows.

exclude individuals who did not have a diagnosis. The use of ‘mental health conditions’ in the recruitment materials for my first two studies was reflective of what I encountered in existing literature and was appropriated in order to ‘tread lightly’. At the time, I felt this approach did two things. First, it avoided stigma associated with ‘mental illness’ and, second, it distanced my work from clinical relevance, which helped when requesting approval from Northwestern University’s Institutional Review Board (IRB).

The participants in my first two empirical chapters, Chapters 4 and 5, include individuals with and without clinically diagnosed conditions, as well as individuals with a broad spectrum of experiences, such as anxiety, depression, and disordered eating. Interviews focused on these experiences and mental health, generally. Meaning, for example, that in addition to talking about how participants posted about mental illness online, I also asked about posting practices related to mental health. The online data I collected from Instagram at this time also included content about mental illness and mental health. Considered together, the majority of experiences within this dataset (i.e., interviews and Instagram content) are related to anxiety, depression, and eating disorders. With respect to my online collection, these experiences were likely most represented due to the set of hashtags I used to start my inquiry. Though as part of my ethnography I extended my collection beyond hashtags, there is no doubt that the ones I selected as a starting point shaped what I was able to observe.

Inclusivity was an important part of my initial research. Though anxiety, depression, and disordered eating are vastly different experiences, people share a number of commonalities, including how those experiences are viewed and treated by the general public as well as academic researchers. As I continued throughout my work, however, I came to recognize important differences in content that I inadvertently obscured by working through such a broad approach. I decided to change research populations, working, instead, specifically with individuals with eating disorders and disordered eating – arguably, still a very large and heterogeneous group of individuals. However, many people with eating disorders gather in similar spaces online and encounter similar concerns, such as harassment and platform moderation. I decided to work with people who have eating disorders largely because of my own experiences. In particular, being an 'insider' researcher, at least in some respects, with a population living with mental illness isn't a common position in HCI and CSCW research. I felt it could add a valuable dimension to the field.

3.2.2. Digital Ethnography

Digital ethnography has a foothold in a number of disciplines, HCI included (Dourish, 2006; Pink et al., 2016; Rode, 2011). My approach to digital ethnography has grounding in the 'ethnographic turn' in HCI (Dourish, 2006), which Paul Dourish, in part, contextualizes in relation to the history of ethnographic research in anthropology and related fields (e.g., studies of colonized peoples, research examining particular subcultures and 'outsider' groups). Set against this backdrop, my understanding of ethnography has also been informed by a number of studies within HCI, CSCW, and adjacent communities. While much of this work is firmly situated within an ethnographic realm, work at the boundaries – particularly community-based and participatory design work – has also provided insights for navigating participation and ways of engaging with communities online. For example, immersive, ethnographic fieldwork by scholars – such

	Beyond the Coded Gaze	Everyday Experiences	Conformity of Eating Disorders
Instagram	3,143 posts	6,223 posts	6,223 posts
Reddit			208 threads
Tumblr			160 posts
Interviews: Expressions of Mental Health and Illness	14 adults	18 adults	
Interviews: Content Moderation			20 adults

Table 3.1. This table is an overview of the data – interview and online – I collected for my three empirical studies. I present this information chronologically (i.e., in the order it was collected). However, please note that, for the upcoming chapters, I have decided to place Everyday Experiences (Chapter 4) before Beyond the Coded Gaze (Chapter 5). While this table does account for posts that I actively collected and saved, either to spreadsheets and/or to folders on my computer, it does not account for the hours of online observation or the posts that were not formally recorded in a quantifiable way.

as Morgan G. Ames, Silvia Lindtner, Shaowen Bardzell, Jeffrey Bardzell, Lilly Nguyen, Syed Ishtiaque Ahmed, Nusrat Jahan, Steven J. Jackson, and Paul Dourish (see Ames et al. (2018) for a collaboration) – studying practices related to making and hacking; research examining labor and crowdsourcing, such as work by scholars Lilly Irani (Irani and Sengul-Jones, 2015; Irani and Silberman, 2013) and Mary L. Gray (Gray and Suri, 2019); and community-based design approaches, such as those practiced by Amanda Lazar (e.g., research with individuals with dementia (Lazar et al., 2017)), Christina Harrington (e.g., research with older adults in Chicago's South Side (Harrington et al., 2019)), and Christopher A. Le Dantec (e.g., research with members of Atlanta's urban homeless population (Le Dantec, 2012)). Additionally, with respect to the predominantly digital, danah boyd's study of Friendster (boyd, 2007, 2004), Kathryn E. Ringland's multi-year ethnographic work on Minecraft with children with autism (Ringland, 2019; Ringland et al., 2016), and Devin Proctor's inquiry within the Otherkin community on Facebook (Proctor, 2018) have provided valuable opportunities to understand the diversity of ethnography as it relates to, broadly, sociotechnical spaces, concerns, and opportunities.

In addition to the above, my approach has been significantly influenced by ethnographer danah boyd (boyd, 2010, 2014) and sociologist Dhiraj Murthy (Murthy, 2008, 2011). Through boyd's ethnography of the social lives of teenagers in a networked era we see a commitment to immersion, observation (e.g., digitally-mediated and physical), and in-person interviews (boyd, 2014). While drawing on these practices, I also turn to Murthy, who writes extensively on ethics and describes digital ethnography as "ethnography mediated by digital technologies. [. . .] As this definition suggests, digital ethnographies can be ethnographic accounts of both offline and online groups. The 'digital' in this mode of ethnography stems from the methods rather than merely the target ethnographic object" (Murthy, 2008). Drawing together these varied approaches, I've come to understand digital ethnography as an approach to research that effectively supports a range of practices and forms of engagement with research populations. Meaning, there is no one correct way to conduct ethnography, only the context of research (e.g., involving the researcher and research population) and decision-making logics of particular scholars (Varis, 2016). Sans unity of method, a common thread across ethnographies, then, is a commitment to the standpoint, position, or perspective of members of the research population (Dourish, 2006; Varis, 2016). Though, of course, this is not standardized in meaning or in practice.

Like traditional, non-digital ethnography, the aim of digital ethnography is to understand the experiences of a particular population through the perspectives and standpoints of members of that population (i.e., in the context of their everyday lives and interactions) (Brown, 2013). As Paul Dourish writes, "ethnography argued that, through daily participation in everyday life, one could come to understand what members of those cultures *experienced* through their actions" (Dourish, 2006). As such, ethnography supports the examination of practices, social relations and roles, and power dynamics within and surrounding particular groups. The ways these various aspects are examined include participant observation and participation (e.g., within the research population's setting or environment), copious fieldnotes, and interviews. Ethnography – digital and otherwise – necessitates a reflexive research practice, where the position of the scholar within society and in relation to members of the research population is not taken for granted or assumed (England, 1994; Pacheco-Vega and Parizeau, 2018; Rode, 2011). Reflexivity is vital, though not always practiced within certain HCI and CSCW domains, because it locates and situates the researcher, rather than presenting the individual and their scholarship as all-knowing, all-seeing (Haraway, 1988). Importantly, ethnographers, even though we are particularly immersed within our research population and setting, are still restricted through our own situated positions and knowledges. As I discuss in Chapter 7, knowledge is produced in the betweenness of the researcher, research population, and methodological approach (England, 1994). Within this context, my aim during this ethnography has been to, as much as I can, support the voices of the particular individuals and groups I work with and to foreground their points of view and experiences. Though I do not claim to speak for anyone or any particular group, this approach has been an important way of engaging with the needs and concerns of my research population and representing an underrepresented perspective.

My research does not include a conventional field site or sites, such as geographic locations or collocated communities. Though traditional ethnographies of the past typically focused on one site, modern ethnographies often span multiple (Marcus, 1995). My work follows in the footsteps of this younger tradition. The posts, accounts, and communities I observed online, as well as the participants I interviewed, are geographically distributed. They gather online through mediating technologies and the boundaries they produce, including online forums and networks of hashtags and people. As such, my ethnography is better understood through the mediating technologies of a networked world (boyd, 2010). Digital ethnography provides a way to approach this networked world, through participation online and traditional ethnographic methods, and supports observation not only in the moment, but also over time. Additionally, this approach fosters opportunities to connect with, learn from, and give back to members of particular communities and, more broadly, members of a marginalized group.

As with many ethnographies, my digital ethnography is not apolitical or disengaged with ethics or researcher responsibility. I aim to use this ethnography to foreground and center the experiences and perspectives of people living with mental illness. In particular, the experiences I highlight trouble dominant and normative views of mental illness online. I address and challenge research and design that positions certain mental illness communities and content, such as pro-eating disorder, as deviant and that promotes content moderation as a partial-but-necessary salve for online contagion, bad actors, and inappropriate behavior. I emphasize challenges and consequences that people with mental illness encounter online, as well as the ways that people support and care for one another, to address relevant gaps in the literature (i.e., with respect to first-hand accounts and interpretation of online content) and to critically examine the structures of social media platforms and online communities that, at times, support power dynamics and unequitable forms of online participation.

Below, I detail two specific components of my digital ethnography: online observation and participation, and interviews. As the specifics for each study are housed within their empirical chapter, I only provide an overview of the practices in the text below. Table 2.1 shows the number of posts collected and interviews conducted for each study to provide a sense of how each aspect of my digital ethnography built from what came prior.

3.2.2.1. *Online Observation and Participation*

Social media present a shift in how content is produced, consumed, and circulated. This, in part, comes from the specific technical ways in which communication and other social interactions and relationships are flattened and performed online (e.g., upvoting, downvoting, liking, commenting, posting a status update) (boyd, 2007). People establish themselves online by leveraging the multimodal nature of most online sites, as well as through the common practice of maintaining multiple accounts in tandem. In my dissertation work, I attend to, though not exclusively, captions, comments, usernames, liking a post, hashtags, profiles, private messages (i.e., when shared by participants), and images online. In November 2017, while waiting for approval of the interview component of my research protocol, I started my online observation on Instagram. I chose Instagram because of its prevalence in other academic literature studying posts and comments about mental illness. Additionally, I was interested in Instagram due to its multimodal posting affordances and norms (i.e., such as privileging visual content).

Throughout most of my digital ethnography – that is, my practices spanning all three empirical chapters – I primarily participated online through lurking. With respect to ethnographic practice, this method is more than a little contentious (Garcia et al., 2009; Murthy, 2008; Pacheco-Vega and Parizeau, 2018), though not uncommon (see, for example: Csipke and Horne (2007); Denzin (1999); Schaap (2002)). Ethical concerns around lurking are many, including extensions of issues related to 'covert' ethnographic research practices (Murthy, 2008; Pacheco-Vega and Parizeau, 2018), such as harm to members of the research population (Clark, 2004; Hine, 2000; Steinmetz, 2012; Sveningsson, 2004), and lurking as epistemically insufficient (i.e., researchers cannot develop 'complete'[2] knowledge about a particular group or community without interaction) (Beaulieu, 2004; Hine, 2000, 2008; Steinmetz, 2012). Ethnography, as set apart from other qualitative and immersive methods, involves researcher practices of participation. Within digital spaces, this is often discussed in the context of creating accounts, filling out profiles, posting and commenting, liking posts, and messaging with members of the research population. Proctor goes as far as to say that "lurking is a great practice for preparatory research, but it is a poor ethnographic methodology" (Proctor, 2020).

---

[2] To be clear, researchers never – or rarely – develop 'complete' knowledge about a particular group.

Frankly, I disagree. The stance Proctor describes operates to restrict who can practice ethnography, as well as the ways in which ethnographic methods can be realized. Though lurking, like other forms of ethnographic method, is not unproblematic, here I want to focus on its legitimacy within ethnography, rather than its particular flaws.

In discussing the legitimacy of lurking, I do so within the context of my own position; namely, a researcher with an eating disorder. For others, mileage may vary and necessitate different decisions and considerations with respect to engagement. Within many online spaces, lurking is a valid form of online participation. For example, many of the individuals I've spoken with over the past few years lurk online, including on eating disorder communities and networks. Lurking as a practice supports individuals in making sense of their experience (i.e., understanding that their practices of eating and body management may be interpreted within the category 'eating disorder'), effectively operating as what Michel Foucault would call a technology of the self (Foucault, 1988). Additionally, lurking can mitigate concerns around privacy, and may also be entangled with people's experiences with relapse and recovery, particularly as individuals move within and between various online spaces.

With respect to my own approach and as a member, partially, of the group I research, lurking provided a certain amount of protection. It was a way to participate and engage while mitigating and, perhaps, controlling the risk I was introducing to my own, at times difficult and nebulous, recovery. I was cautious against adopting a more – or, perhaps, differently – immersive research practice due to my experiences with mental illness, not wanting to potentially exacerbate them or unintentionally trigger myself. This concern was not just grounded in the content I was observing, such as posts related to depression and disordered eating, but, also, the ways people interacted with that content and the commercial practices prevalent on Instagram – for example, advertisements and influencer recommendations for weight loss teas and other explicit calls for thinness. Ethnography – and, truly, any research practice – should not put the researcher at risk, particularly risk an individual is not wanting or able to accept. To be clear, this should not discount people, such as myself, from utilizing an ethnographic approach. Simply, it means that my approach – and the approaches of similar and adjacent scholars – involves attending to different sensitivities and ethical considerations; specifically, navigating and balancing these in relation to the researcher's own position and health.

All of this said, in these days of lurking I did participate in other small-yet-significant ways: I reported content. Written into my first research protocol was the following sentence:

> "If any of this public content seems to suggest an individual is suicidal or practicing self-harm, I will report it to the online media host (e.g., Facebook, Twitter, Instagram). These organizations have teams that reach out to users in these situations."

At the time, I did not want to be complicit in watching people harm themselves. While I did not report the majority of content I observed, I did flag posts where individuals seemed in acute distress (e.g., suicidal ideation and some instances of self-harm that looked particularly at risk for infection). My intent was for Instagram to send these individuals support resources. However, I didn't understand my own potential to cause harm through this interaction. As I became more familiar with the types of practices on

Instagram, such as posts decrying content moderation and new accounts created when old ones were banned, I realized that I had contributed to a very real problem that people were having. I stopped reporting what I observed. My only other interaction during this time involved reaching out to Instagram accounts – and, later, moderators on Reddit[3] and other online forums – to recruit for the interview components of my work. For the most part, my initial online observation was only immersive with respect reading content. However, though I read all of the content I collected, I did not engage with the people or communities within my digital field sites (outside of potential recruitment).

This changed in December 2019. Though none of my empirical work in this dissertation is impacted by this change in ethnography, as everything was submitted for publication prior to this date, I make room to include this experience because it has influenced my dissertation writing. In December 2019, I created a Reddit account to participate in several eating disorder subreddits. However, anxiety and apprehension around participation, disclosure, and my own experience with mental illness stayed my hand. It wasn't until March 2020, when Stay-at-Home orders were issued across the United States in response to the COVID-19 pandemic, that I began to interact beyond

---

[3] Reddit (www.reddit.com) is a social media platform where social news and other types of content are aggregated into several content feeds, such as r/all and r/popular. Though content is generated and uploaded by account holders, much of what Reddit features (i.e., what gains popularity) is originally sourced from other content production sites online, such as news production websites. However, this isn't to dismiss the personal types of content that individuals share. In addition to having content feeds that aggregate popular content across Reddit, Reddit is organized through smaller communities, called subreddits. Subreddits allow Reddit account holders to join and often, though not always, post and comment on content. Subreddits are moderated by subreddit-specific moderation teams, which develop specific rules for each community. These rules are developed and housed within Reddit platform rules. On Reddit, account holders, in addition to posting comments, may interact with posts and other comments through upvote and downvote mechanisms, as well as saving and reporting certain content. Account holders may also message one another and connect through individual and group chat features.

reading posts and comments, such as through liking, commenting, messaging, and reporting content to subreddit moderators (i.e., rather than Reddit). My Reddit profile has always included my position as a researcher; additionally, in several posts and messages I've discussed this part of my biography. Participating within various subreddits has been eye opening – more so than I initially expected. I've gained a different, more emotional, and, perhaps at times, 'behind the scenes' understanding of social interactions and roles within these communities. For example, I've participated on a Reddit chat with other community members to discuss concerns with moderation on one specific eating disorder support subreddit. Through participation in this particular chat, I also observed and was part of the creation of a new subreddit, one made specifically to field critical posts about perceived disordered eating in a number of diet communities on Reddit.

As the above examples illustrate, I now hang out and take part in ways that I did not before (Salvador et al., 1999). And, again, while 'hanging out' has not influenced the empirical chapters here, it is relevant for the discussion and for my work moving forward. Providing these examples isn't to discount my earlier work or to say that practices of lurking are invalid or 'poor ethnographic practice'. Rather, approaching my digital field sites initially through lurking was necessary for mitigating the risk this research domain poses to my personal health and wellness.

3.2.2.2. *Interviews*

Interviews have always been a core component of my research approach. Soliciting firsthand accounts of people living with mental illness (with respect to HCI and CSCW social media research) is not a common approach, though it has gained some prevalence since I first started this work in 2017. Interviews are important because they can help us understand why people gather, communicate, lurk, and participate in the ways they do, as well as other aspects of life that are concealed or not shared online. In my work, interviews have been particularly helpful in expanding my understanding of how people post about mental illness and mental health without overt signals, such as hashtags. Though not without critique (England, 1994), interviews are important for providing opportunities to directly engage people and collectively interpret experiences within the timespan of the conversation. Interviews provide opportunities to learn from members of the research population and to engage with their concerns and needs. Like my online corpora, my interviews are a multiyear practice that build on one another. While most of my interviews have been over the phone, several have been in person and over email. Even though text-based interviews make research more available for some individuals (e.g., for people who do not feel comfortable talking about mental illness over the phone, do not have time during the day to take an interview call), I did not include them in my last study, Chapter 6. My experience with them during my earlier empirical work was not overwhelmingly positive. For example, participants responded with one sentence answers and did not reply to follow-up questions. Though, certainly, that is not the case with all text-based interviews – and, in other studies, I have been able to develop more of a connection – I felt that my conversations were richer with people in person and over the phone.

I also found that the richness of conversations was partially entangled with my recruitment practices. For my first two studies I tried to recruit on Instagram, without success, and ultimately interviewed people who replied to my recruitment advertisement on Craigslist (and/or who heard about the study from another participant). Obtaining access to populations with mental illnesses can prove difficult and ethically ambiguous. In my initial interviews, I faced hurdles to recruitment, with only one out of every 10 respondents following through for an arranged interview. While many of these interviews were illuminating, others were not. For example, some participants were reluctant to talk about how they expressed their experiences with mental illness and mental health on Instagram, even though those topics were clearly stated in the recruitment calls for the first two empirical studies (Chapters 4 and 5). Of course, talking about these topics can be difficult. Many people that reached out to me from Craigslist ultimately decided to not go through with the interview due to concerns around their experiences and the privacy afforded through the research process.

I've recently had more success recruiting participants from online forums and communities, such as specific subreddits. Rather than reaching out directly to members of these communities or posting my recruitment materials, I have, first, reached out to moderators, asking for permission to share my study's information. Recruiting people from the spaces that are directly relevant to my research is preferable. For example, before I started doing this, when I was recruiting from Craigslist for individuals with eating disorders who had content moderated, I connected with people who had legitimate – but vastly different – experiences than what I was hoping to examine. One participant with an eating disorder participated in various diet Facebook groups, where she had content banned for being too “intense”. By recruiting from eating disorder subreddit and online eating disorder forums, I was able to connect with members of the specific population I was hoping to address.

Though my recruitment practices have changed, in conversations, I’ve always aimed to position the participant as an expert who I am learning from. With respect to interview styles, this most closely falls into ‘supplication’, which involves “exposing and exploiting weaknesses regarding dependence on whoever is being researched for information and guidance” (England, 1994). While supplication does aim to address power asymmetry in research, it, of course, cannot overcome the privileged position of the researcher. For example, across my empirical work, I have sent participants mental health resources prior to our conversation and informed them about their access to our study’s mental health consultant. I’ve also been the one paying people for their time.

In considering how people would respond to being interviewed about topics related to their experiences with mental illness, I also initially excluded people experiencing suicidal ideation and practicing self-harm. This constrained my potential sample and, ultimately, marginalized individuals with those experiences. When I decided to interview people about content moderation, I removed this exclusion criteria. People experiencing suicidal ideation and self-harm were included in this work – largely because the content moderation interviews did not aim to directly discuss those topics. Participants were informed that they could be referred to our practitioner if they seemed to be talking about currently distressing experiences with self-harm and suicidal ideation. Throughout my work, I have also only included individuals able to consent for themselves. This means that I excluded adolescents from participation in my interview cohorts. As others have noted, adolescents contribute to mental health and illness-related content on Instagram and other social media platforms (Moreno et al., 2016). Although I did not interview these individuals in my current study, I believe it is important to engage with these populations, so as to not neglect or misrepresent the experiences of these groups.

Research in this space often lacks diverse representation. My samples, not uncommon in this type of work, are predominantly female and white. This should not be interpreted as a propensity for white women to have mental illnesses, or as an absence of mental illness in other populations. Instead, we should consider how cultural values and norms shape how mental health and illness are conceptualized, shared, and expressed, and impact participant self-selection. It is our responsibility to reach out to communities and work toward building trust and reciprocal relationships.

3.3. Researcher Responsibility and Ethical Considerations

From building a research team to producing and communicating results, every decision made by researchers involves ethics. Though our decisions aren't always obvious, even to ourselves, we privilege certain people and experiences over others, and choose to represent our research population in certain ways. Here, I walk through some of the ethical decisions made throughout the course of my multi-year digital ethnography.

I received approval for this research from Northwestern University's IRB. In working with the IRB for this approval, I included certain provisions for working with people living with mental illness. These provisions were intended to mitigate the potential for harm in interviewing people about experiences that can, at times, be difficult and sensitive. Throughout my work, these provisions have been twofold: a mental health practitioner and a document of mental health support resources. I included the healthcare practitioner's name and contact information on the consent document and sent participants a document of support resources (i.e., including several phone and text-based helplines) prior to the interview. Note, the mental health practitioner was available as a resource for participants and did not participate in data collection or analysis. When I first started this work, healthcare practitioners, when involved in social media research, were likely to work alongside researchers in initial analysis. However, by including a practitioner as a resource, rather than an analytic lens, my aim was to have a healthcare expert available for participants, should they need or want to speak with someone. This provision was also intended to distance myself from the role of a healthcare practitioner or therapist. Though I specifically – and several times – mentioned my role as a graduate HCI researcher, having a mental healthcare practitioner as part of the team helped set those roles and responsibilities apart. While the healthcare practitioner was available as a resource, I did inform participants that, should they describe current experiences with suicidal ideation or self-harm, the interview would immediately end, and their contact information would be shared with the practitioner, who would then reach out.

Ethnography – digital and otherwise – involves a commitment to examining the relationship between the researcher and the subject, or the population, at hand (England, 1994; Murthy, 2008, 2011; Pacheco-Vega and Parizeau, 2018; Pink et al., 2016). This type of reflexive examination contributes to the production of knowledge. As part of my method, I am mindful of how the analytic frame of marginalization requires accounting for and reflecting on how my expectations, values, and norms as a researcher, and as an individual within society, differ from those of my participants and online posters (Bruckman, 2006; Haimson et al., 2015b). My work is shaped by my personal experiences, including my own individual experience with an eating disorder and my experiences alongside close friends, family members, and colleagues with eating disorders. My personal experience inherently shapes my approach and perspective, including my interest in this line of inquiry, interpretation of the data, and conscious commitment to foregrounding perspectives and experiences that are largely unrepresented in current HCI and CSCW research. Though this experience does not make me any more or less eligible to conduct research in this domain, it does provide a perspective that is both inside and outside the studied population – a tension that I've had to grapple with throughout my digital ethnography.

I view myself as having an obligation to the communities and people I research (Murthy, 2008; Pacheco-Vega and Parizeau, 2018; Schrum, 1995). In part, this means engaging with the needs and concerns of my research population and foregrounding experiences without sensationalizing them. As such (i.e., not wishing to sensationalize images of harm or illness), I do not include any images or video screenshots in my dissertation. Visual modes of communication, in particular, when taken out of context can make certain topics or practices seem unfamiliar or othered. When used in prior work to illustrate types of content, images are used most often with implied consent (Pater et al., 2016a), which is not consent at all, and frequently feature bare skin and bodies in lingerie – which is not only sensationalist, but also at risk of 'fetishizing the other' (England, 1994). While other researchers have taken a similar stance, they have also found creative ways to display visual content within a paper, such as through their own forms of remix (Andalibi et al., 2017). I opt, instead, to describe visual content and style with text, and to paraphrase from captions and comments (Chancellor et al., 2017). By not including images, I aim avoid sensationalizing images of mental illness, potentially capitalizing off of novel or 'shocking' visual content and rousing emotional distress in readers, as well as associations with mental illness as a kind of aesthetic.

Just as I do not share images because I do not have the consent of the individuals I observed online; I do not share unmodified text excerpts (Fiesler and Proferes, 2018). All captions, hashtags, and account names are paraphrased or have the majority, if not all, of the words substituted to preserve the anonymity of people who, though posting publicly, did not consent to participate in this study and, ultimately, do not participate online so that their content can be used for research. Further, I also modified textual information from posts that participants in my interview studies did not agree to share with the greater public. The text and images described in the following chapters are phrased to best maintain privacy and the integrity of the post, including affect and language. Additionally, I decide not to name active online communities (i.e., those currently existing online that have not been removed or deleted by platforms or other providers). This is to avoid unwarranted attention. As one participant offhandedly mentioned, it's their belief that Reddit's r/ProED subreddit was banned, in part, due to a research publication that explicitly named – and outed – the subreddit. Whether or not that is the case, researchers should not cause that type of harm or disruption to a community of individuals. It's important to be mindful of how we represent the individuals and groups of people who we study.

CHAPTER 4

Everyday Experiences: Small Stories and Mental Illness on Instagram

For people living with mental illness, mental illness is a part of their everyday experience. As obvious at this sounds, current approaches to studying and intervening on mental illness on social media extract certain content from the context of everyday life and online practices, highlighting a narrow vision of what people with mental illness post about online. In this chapter, I address how individuals living with mental illness express their experiences online – including experiences not related to mental illness. Here, I make three contributions: 1) I center people living with mental illness and their everyday lives and social media practices, of which mental illness is part; 2) I write about mental illness as an everyday, albeit not insignificant, part of life; and 3) I discuss how existing features, such as content moderation, on social media platforms reinforce power differentials that ultimately serve to harm a marginalized group of people.

An earlier version of this chapter is published with Anne Marie Piper in the Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems:

> Jessica L. Feuston and Anne Marie Piper. Everyday experiences: small stories and mental illness on Instagram. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019.

4.1. Introduction

Social media platforms and online forums provide digital spaces where people with mental illness can gather and communicate about their experiences. Given the prevalence of these online spaces and the academic interest in mental illness, over the past decade, research on mental illness online has increased exponentially. Though individuals do not post online to have their content observed or analyzed by researchers – an ethical tension we must grapple with while conducting this work – it's undeniable that digital spaces provide ways for researchers to examine individuals, communities, and practices that are traditionally difficult to reach.

Online spaces that people with mental illness participate within hold traces of subjective experience. Intentionally or not, individuals archive aspects of life with mental illness through posts, likes, comments, and accounts. Though mental illness is but one aspect of life, a condition or experience that people are living through and with, prior work often centers 'mental illness' explicitly (Andalibi et al., 2017; Chancellor et al., 2016a,c; De Choudhury et al., 2013b, 2014), relying on data collection methods that extract posts related to mental illness with surgical precision. In extracting this content, researchers sever mental illness from other aspects of life and experience, such as school and hobbies, that can provide insight into how people live and participate online with mental illness. In this chapter, I illustrate the importance of context in understanding mental illness online and, more broadly, in understanding mental illness as part of the everyday.

My aim in this chapter is to extend prior work on mental illness and social media by contributing a perspective where mental illness is represented as an integrated part of – rather than apart from – everyday experience. This perspective illustrates how, for people living with mental illness[1] on social media, mental illness is not always in focus; but may, instead, be sidelined or entirely absent. Though mental illness is still featured in this work, it is not to the detriment of other experiences and content, which are kept throughout data collection and analysis, or an individual, more generally. I aim to foreground and describe people living with – and, more specifically, people using social media with – mental illness. I use small stories (Georgakopoulou, 2015) as an analytic framework to understand how people 1) communicate about mental illness as an aspect of their lives and/or condition through which they live and 2) navigate ways of communicating – or telling – about mental illness on Instagram, a picture sharing-based social media platform, where narration is shaped by sociotechnical processes of control. Small stories research emphasizes understanding non-linear events, world-making (i.e., "ordinary, everyday events" (Georgakopoulou, 2015)), story recontextualization, and story co-construction. In this chapter, I argue that small stories as an analytic frame can help us understand how people communicate about mental illness and related, everyday experiences online.

In this work, I draw on interviews with 18 people who self-reported having, or having had, experience with mental illness, as well as public accounts on Instagram which, in some capacity (e.g., living or coping with illness, recovery, relapse), included posts

---

[1] The published version of this chapter frequently uses the phrase "living with or posting about mental illness." In differentiating these two conditions, my aim was to avoid inadvertently medicalizing an account holder or poster's experience. For example, posting about mental illness does not necessarily indicate an off-screen experience of living with mental illness. However, when I made this decision I was grappling with the nature of mental illness – as something social and experiential, but also heavily psychiatrized. In reflection, many of my participants do not fit medical diagnosis of mental illness and/or have not sought a clinical diagnosis. Yet, these individuals still use that social category when communicating about certain experiences. As such, I've decided, in this version, to simply write about people living with mental illness.

about mental illness. Using small stories research as an analytic framework, I demonstrate how mental illness is contextualized within everyday experience specifically *because* it is a familiar part of everyday experience for those living with it. This perspective has implications for understanding timescales, disruption, and the ways individuals must strategically navigate and negotiate processes of social control. I argue that the pervasive focus on 'deviant' mental illness content within HCI and CSCW literature sensationalizes and decontextualizes certain facets of experience, and completely dismisses and ignores others. Selective attention on 'deviant' content, as well as extracting 'mental illness' from surrounding context, inadvertently contributes to stigmatization and the continued marginalization of individuals with mental illness. Additionally, I discuss how social interactions between individuals with mental illness, audience members, and Instagram create new instantiations of traditional forms and processes of social control.

I make three contributions through this work. First, I extend prior work by centering my analysis on personal experience with mental illness through semi-structured interviews and online observation of public Instagram accounts. Though increasing in frequency, first-person accounts of mental illness in conjunction with social media analysis are still uncommon. Second, my work situates mental illness as part of everyday and ordinary, albeit not insignificant, experience. This position stands in contrast to prior work in which deviance and medicalized conceptions of disorder frame analysis or are otherwise conceptually invoked. I argue that an alternative, non-deviant, and non-medicalized view of mental illness can help us to modify current research practices and recommendations, particularly those that inadvertently contribute to harmful stereotypes, and expand the design space related to supporting individuals living with mental illness. Finally, I discuss the ethical implications of research and design with respect to social media and mental illness. Certain Instagram features, such as content moderation and helpline resources, reinforce power dynamics that marginalize and unjustly disadvantage certain populations. By attending to how individuals with mental illness communicate about their lives online, I contribute new understanding related to how mental illness is represented and negotiated within the everyday, including how individuals posting about mental illness are subjected to social control through audience member interactions and Instagram as a platform.

4.2. Related Work: Health and Illness Narratives in HCI

Interpersonal, everyday stories are a large part of how we communicate with others, construct identity, and understand experience – both our own and others' (Greenhalgh, 2016; McAdams, 2008; Rappaport, 1993; Riessman, 2003). In studying illness, scholars often use narrative as a means of understanding the subjective experience of being ill. Yukari Seko and Stephen P. Lewis's study of self-injury narratives on Tumblr (Seko and Lewis, 2018), as well as Su Holmes's analysis of anorexia narratives on YouTube (Holmes, 2017), are two recent examples that demonstrate how scholars leverage components of narrative research to examine mental illness and the online construction of narrator identity. Though I discuss these narratives in the context of illness (i.e., here and in Chapter 6), as much prior work does, understanding the subjective experience of mental illness may, actually, be more closely related to disability narratives, which are "the stories people with disabilities tell about their lives and experiences that can highlight issues of disability identity" (Dunn and Burcaw, 2013).

Few researchers within HCI and CSCW have drawn on narrative research or analysis to examine lived experience with illness. A recent exception includes Jordan Eschler et al.'s examination of how cancer survivors strategically use tattoos to recover from trauma associated with cancer diagnosis and the subsequent treatment process (Eschler et al., 2018). Similarly, Lydia Michie et al.'s research on reproductive rights (Michie et al., 2018), though examining a highly stigmatized health decision rather than illness, draws on certain pro-life narratives (e.g., trauma, regret) and the ways that pro-choice stakeholders come to challenge them. Additionally, recent studies examining mental illness online within CSCW (Andalibi et al., 2017; Zhang et al., 2018), though not calling upon narrative research explicitly, incorporate notions of storytelling to describe how individuals share their experiences with others. Here, I aim to make the connection to narrative and storytelling more explicit. This serves to highlight subjective experiences that mental illness is part of, as well as to help us draw connections between mental illness narratives and other types of illness narratives, such as those related to living with and healing from cancer.

To date, scholars studying illness online have predominantly focused on the content of personal illness narratives. I extend this research by using small stories as an analytic framework for exploring how narrative activity is accomplished through interaction (Georgakopoulou, 2015; Phoenix and Sparkes, 2009). Developed in opposition to 'big stories' (i.e., stories often narrated in the retrospective, in which individuals describe prominent life events or their life stories for the purpose of connecting particular components into a descriptive whole (Bamberg, 2006; Phoenix and Sparkes, 2009)) typical of narrative research, small stories prioritize interactional aspects of the everyday. Specifically, small stories “refer to stories told during interaction, generally within everyday settings, about very mundane things and everyday occurrences” (Phoenix and Sparkes, 2009). Small stories research provides an opportunity to examine commonplace, interactional exchanges, such as those shared via social media; ongoing events or stories that may unfold across the span of multiple tellings, without a particular beginning or end; and co-constructed experiences and identities that demonstrate the influence of social and power dynamics. Additionally, in thinking about narrative for this particular chapter, I consider narratives to be “located in particular times and places” (Riessman, 2003). This means that storytelling is not only influenced by individual and collective experience, but also “cultural evaluation of what can be narrated and how that can be narrated” (Atkinson, 1997).

4.3. Method

To examine the subjective experience of life with mental illness, I conducted digital ethnography combining semi-structured interviews and online observation of public Instagram accounts. Interview participants, as well as accounts included in my Instagram corpus, shared content related to various facets of mental illness, including living or coping with mental illness, recovery, and relapse. My analysis draws on small stories research (Georgakopoulou, 2015) and multimodal discourse analysis (Constantinou, 2005; Kress and Van Leeuwen, 2001) to understand how individuals express mental illness on Instagram.

4.3.1. Interviews

I conducted semi-structured interviews with 18 adults (ages 18 – 36; M = 29 years; 14 female)[2] who post about mental illness on Instagram. For this study, I was interested in understanding experiences with mental illness. However, due to the colloquial equation of mental health with mental illness, participants frequently used these terms interchangeably (i.e., mental health conditions). Therefore, my interpretation of mental illness derives from participant references to low mood, named disorders (e.g., anxiety, depression), and aspects of recovery, such as clinical treatment (e.g., medication, therapy) and self-care. Though all of the participants self-reported experiences with anxiety, depression, and eating disorders, only several indicated receiving a diagnosis. Per my IRB protocol, I did not interview individuals who self-reported currently practicing self-harm or experiencing suicidal ideation. However, I did analyze public Instagram accounts with posts related to these topics. In the instance that any participant had self-reported current experiences with self-harm or suicidal ideation, the interview would have ended and this study's mental health consultant would have reached out in accordance with my IRB protocol and her own set of practices. The study's mental health consultant (i.e., an art therapist and licensed counselor by training) was an important member of the study team who was available as a participant resource. For example, this team member's information was included on consent documentation. I also informed participants that they could contact her at any time. Additionally, in accordance

---

[2] I present the method in this chapter as it appeared in publication. Additional details about this approach and dataset appear in the next chapter.

with this study's IRB, I emailed a document of mental health resources (e.g., helplines, practitioner locators) prior to the interview.

I recruited participants through websites, such as Instagram and Craigslist, and word of mouth. While I did reach out to Instagram accounts that had public posts about mental illness, none of those individuals responded to my recruitment message. I only interviewed individuals living in the United States, though the Instagram account corpus includes posts and accounts from other countries. I conducted 30-minute semi-structured interviews through phone calls and text-based mediums, including email and Skype. Text-based interviews typically lasted longer than 30 minutes due to the asynchronous back and forth nature of question and answer. These interviews allowed me to connect with people who, for whatever reason, did not prefer to talk over the phone or in person. To reduce potential confusion between a research interview and talk therapy, I clarified my role as human-computer interaction researcher several times prior to the start of the interview.

My interview protocol was specific to Instagram. I asked participants about posts, including images, captions, hashtags, and comments, as well as Stories, a feature on Instagram that allows individuals to share content for up to 24 hours. Though I focused on Instagram posts that included some element of mental illness or mental health, I also asked participants to describe the types of content they typically and recently shared. This helped me gain a base understanding of how the participant used their Instagram account or accounts. For this study, my analysis centers people's experiences with mental illness, of which mental health may be part, rather than focusing explicitly on mental health.

Prior to the interview, and when possible, participants shared their Instagram accounts with the study team. Some participants did not feel comfortable sharing their accounts, even when public, and opted to share screenshots of particular posts or describe the posts during the interview. Given that my inquiry aimed to examine particular Instagram posts, many of which included pictures, I used photo elicitation as a way of having an artifact present to support recall of particular posts and experiences (Harper, 2012). When shared with me, these artifacts also provided a point of reference that my participants and I could use throughout our conversation. I reviewed three to seven posts per participant and asked questions pertaining to how and why they posted about experiences with mental health and mental illness, as well as the ways others responded to this content. Interviews were audio recorded and transcribed for data analysis. Participants received a $20 Amazon gift card for their participation.

4.3.2. Instagram Account Corpus

Data collection practices influence researcher interpretation, and vice versa. By conducting online observation on public accounts, rather than on a random amalgam of posts, we are able to examine the temporal, and often non-sequential, ordering of events as presented on Instagram. This provides a different perspective of online expressions of mental illness. I identified content and accounts through data collection in November 2017, March 2018, and June 2018, where I used hashtags as search terms (e.g., #depression, #ednos, #secretsociety123, #bipolar, #anorexia (Andalibi et al., 2017; Chancellor et al., 2016b,d)). Many of the hashtags I selected to use were described in previous research studies on mental illness and social media. As such, other researchers had also used them as a way to start collecting content about mental illness on social media platforms, such as Instagram, Twitter, and Tumblr. I started my collection with five particular hashtags. I manually (i.e., without a web crawler) collected 6,223 unique mental illness-related posts by 2,188 users. This post corpus provided an initial area of inquiry. I also extended these hashtag-based collections to include accounts and posts initially excluded (i.e., those using different hashtags or no hashtags at all) and analyzed 15 public accounts in detail, such as looking through the entirety of posts and comments, to get a deeper understanding of individual posting practices that, at times, mentioned or referenced mental illness. These fifteen accounts were selected through purposive sampling (Palinkas et al., 2015) and represented a number of different ways that people communicate about mental illness online.

4.3.3. Data Analysis

I used an inductive, qualitative approach to develop themes from the data. This involved iterative coding, memo writing, and constant comparison of data and initial concepts to one another (Charmaz, 2014). My analytic approach involved simultaneous review of interview data and content from Instagram. Meaning, I did not analyze these data separately or view them as distinct. I used small stories research (Georgakopoulou, 2015) as a guiding framework during my analytic process. Using small stories as a lens supported foregrounding mundane events told through interaction. Specifically, I attended primarily to account and post composition as well as audience reaction as a way of grounding myself in everyday, interactional storytelling. However, though my analytic interest foregrounded small stories, I also collected personal mental illness narratives from my participants (e.g., big stories explaining personal mental illness trajectories and attitudes). These 'big stories' were leveraged to situate my understanding of an individual's overall lived experience and contextualize the small stories they shared during interviews.

To call attention to a minute-yet-significant detail, I conducted research with small story narratives rather than on them (Bamberg, 2012). Specifically, I used small stories and interview retrospectives (i.e., about specific experiences and Instagram posts) to understand and examine the ways people post about and live with mental illness (i.e., in contrast to generating various small story typologies). This approach familiarizes mental illness by aiming to keep it in context. My use of small stories in analysis allows me to center the lives and experiences – as we can understand them through online social media usage – of people with mental illness.

Finally, small stories are not necessarily grounded in the oral or textual practices of traditional narrative research. I want to clearly state my commitment to examining the ways in which multiple modalities shape how mental illness is communicated on Instagram. As definitions of narrative shift and expand to accommodate new media forms (Lundby, 2008; Underberg and Zorn, 2013), it becomes increasingly necessary to incorporate a variety of modalities, such as images and videos, into narrative analysis. In this paper, I include captions, comments, usernames, liking a post, hashtags, profiles, private messages, offline discussion, interview responses, videos, and images in analysis. Through this approach – specifically, one that incorporates social media content and firsthand interview accounts – I call on several modes of expression (Kress and Van Leeuwen, 2001) to understand the inherent complexity of how mental illness as part of everyday life is communicated online.

4.4. Findings

Relying on social media hashtags as the primary means of data collection and filtering, as researchers often do, means extracting mental illness from other content and performing analysis absent the context in which mental illness exists (Feuston and Piper, 2018; Gerrard, 2018). In this chapter, I examine online content related to mental illness in conjunction with an individual's broader experience (i.e., in part, what Nazanin Andalibi et al. refer to as 'contextual expression' (Andalibi et al., 2017)). By carefully attending to where and how mental illness is present and, in many instances, absent, I prioritize the everyday experience of living with a mental illness. This includes examining mental illness when it is featured and when it is not, as well as examining content in which no mention of mental illness is communicated at all. This approach stands in contrast to research practices that recontextualize mental illness by extracting it from surrounding context. Here, I aim to keep this context intact as a way to foreground how people live with and post about mental illness. I describe how Instagram account holders communicate their experiences with mental illness online and navigate how they post – or narrate – about their lives through sociotechnical systems that aim to control them.

4.4.1. Everyday Life with Mental Illness

The research in this chapter details how experiences shared on Instagram include mental illness, often bringing it alongside other everyday activities and posting routines. Specifically, on Instagram, people communicate about their everyday experiences, in which mental illness may be part (Andalibi et al., 2017); post with respect to their own timescales (e.g., idiosyncratic practices and mental illness experiences); and reframe what it means for something to be disruptive.

My interpretation of 'mental illness as part of everyday life' is grounded in a new way of looking at and thinking about mental illness. Rather than look at mental illness when it has been decontextualized, I aim to understand mental illness as part of life. This framing is subtle and, at times, difficult to articulate. However, in this chapter, my aim is to center the lives of people with mental illness, rather than to emphasize mental illness alone. Through this new way of looking, mental illness becomes ordinary precisely because of how people communicate it online. The incorporation of small stories as an analytic lens supports thematic development grounded in interactional and ongoing tellings where mental illness is brought alongside or featured, as well as communicated explicitly or implicitly. Examining how mental illness and other elements of life entangle supports understanding parts of lived experience with mental illness that are overlooked and backgrounded. In the context of this analysis, mental illness is rendered ordinary and familiar, rather than deviant or other, due to how deeply ingrained it is within the fabric of daily life and Instagram posting practices.

4.4.1.1. *Daily Routines, Cultural Trends, and Mental Illness*

Studies of mental illness on social media often analyze data absent daily experiences and activities. Here, I extend research on contextual expression (Andalibi et al., 2017) by attending to how people with mental illness post about a variety of topics on Instagram, including school and hobbies. At times, mental illness is brought alongside these other topics. This illustrates how mental illness is not always central or foregrounded within the lives or posting practices of people living with it. In this section, I show how content related to mental illness becomes part of the ways that people post on Instagram.

Members of my interview cohort, as well as account holders within my Instagram corpus, shared posts about classes, pets, food, exercise, artwork, and vacation that – at times but not always – included explicit and implicit references to mental illness. By keeping this variety of topics throughout my analysis, I call attention to how mental illness is one of many components of everyday life that an individual may choose to narrate. One account holder in my Instagram corpus frequently posted content related to school and entertainment. Periodically, these posts, where school and entertainment take the center stage, referenced mental illness. In a post about school (i.e., *"back to school after break, got a 65 on my test...yayy"*) - part of a series of a related, ongoing telling about exam - the poster also wrote, *"zero days clean."* This reference to self-injury occurs alongside and as part of everyday life, which involves returning to school and failing tests. Further, the mention of mental illness is not prominently featured. In fact, given the positioning of the text at the bottom of the caption, as well as the casual phrasing, it's almost an after-thought; a small reminder of this individual's experience with mental illness and how it intertwines with schoolwork.

Keeping these topics together during analysis allows us to examine mental illness as it occurs in the interactional everyday, which is ongoing and, at times, mundane. We can start to approach mental illness differently by understanding the ways in which it is relevant – or not – for an individual. For example, while many posts made by the account holder in the previous paragraph reference mental illness and low mood alongside other topics (e.g., *"hi, feeling bad, so i'm gonna watch harry potter. i hope ur all fine"*), many do not; dealing with, instead, other documentation of the day to day (e.g., LGBTQ+ pride, music). Still yet, this individual posted content where mental illness is distinctly featured. Take, for example, the ways in which they draw upon black and white imagery, text overlay (e.g., *"warring against myself"*, *"angry, depressed, and ugly in & out"*), and captions mentioning depression. In posts such as these, we see mental illness featured – with and without other topical inclusions. These examples show us that mental illness is communicated in a variety of ways online. For many individuals, these ways of communicating about mental illness are everyday, familiar, and mundane.

Posters within my Instagram corpus and participants in my interview cohort relied on multiple modalities, including images, videos, captions, and hashtags, to cast mental illness as a part of daily life. For example, in one post, an account holder shared a picture of a bright puzzle laid on a wooden surface. Yellow, blue, red, and green pigment punctuate the image, drawing together a kaleidoscope of spheres interrupted only by several missing pieces. The caption of this post references the puzzle on the table (i.e., *"Missing a few bits"*) and a long workday. The hashtags include general social media tags (e.g., #instadoll, #follow4follow) as well as those explicit to mental illness (e.g., #edrecovery, #depression). This individual carried forward similar hashtags in a variety of posts, including those about inpatient hospitalization, recovery-related eating practices, and studying. In one school-related post, in which a laptop, notebook, and pen take the visual center stage, the caption reads, *"Ate so much, but my brain isn't letting me study. Why is my mind so foggy?"* Elements related to everyday activities and circumstances co-exist with mental illness content (e.g., eating disorder and depression-related hashtags), which itself is rendered routine by association.

Within the past several years, HCI and CSCW scholars have incorporated cultural practices in analyses of mental illness online (Li et al., 2016; Zhang et al., 2018). Oftentimes, however, this research emphasizes deviant or 'othered' aspects of specific conditions rather than examining how the appropriation of cultural or community practices might embed mental illness within ordinary practices and liken it to the behavior of non-deviant others. By appropriating popular forms of communication and online participation, such as memes and selfies, individuals assimilate experiences with mental illness into mainstream ideologies (Strike Your Note, 2014). Several members of my interview cohort, including P1 and P4, discussed reposted memes during photo elicitation as a way to post about their experiences with mental illness through culturally relevant and common artifacts (e.g., Kermit the Frog (Know Your Meme, 2015)) that are communicative and relatable. Similarly, individuals regularly entangled mental illness with selfies (Andalibi et al., 2017). Accounts within my Instagram corpus frequently did so in conjunction with hashtags (e.g., #eatingdisorder, #skinny, #depression, #edrecovery);

however, members of my interview cohort often did not include any overt signals to mental illness (Ames and Naaman, 2007). P7, for example, posted a selfie without any hashtags, though she did caption it, *"Not feeling great. Doubt has crept in and my eyes feel heavy with sadness. Thanks for listening. Feeling better already knowing I can put this out there and acknowledge my hurt and heartache."* Appropriation of community practices and cultural artifacts demonstrates how individuals combine elements of mainstream digital culture to communicate personal experiences with mental illness.

4.4.1.2. *Posting about Mental Illness on Individual Timescales*

Understanding how mental illness is an integrated part of everyday experience involves attending to archival aspects of Instagram as a platform for ongoing narration. Here, I build on work featuring time as an integral component of understanding mental illness phases or trajectories (Chancellor et al., 2016c; De Choudhury et al., 2016) with the caveat that time is relative and intimately individual. Though Instagram is archival, the goal is to understand timescales without implying linearity in content due to practices involving 'throwback images', account maintenance (e.g., individuals selecting to removing content), and Instagram reporting features (i.e., in which posts are forcibly removed by the platform).

While certainly permissible to post once with respect to mental illness and never again, my sample comprises of accounts and participants who shared about mental illness repetitively. Though many of these accounts did not feature or reference mental illness in every post, I observed accounts where mental illness was a steady constant. Some individuals, such as P16, shared content that included components related to mental illness every day, whereas others, such as P3 and P18, incorporated mental illness into their posts only every few months. With respect to periodically including mental illness in posts, P15 talked about how *"over time, my posts will vary with how disordered and negative they are. Right now, everything's pretty peachy...well, it's not as bad, so I'm not going to be posting about that stuff as much."* It was common for the frequency of content containing mental illness to change over time and vary with how salient an individual thought it was to their daily experience.

During my online observation, I observed accounts that contained long breaks or pauses in content related to mental illness. By examining these 'gaps' in posting at a longer timescale, we can attend to the ways that mental illness ebbs and flows across narration of lived experience. Furthermore, we can attend to times that individuals may stop posting entirely. Gaps in posting were related to a variety of events, including those specific to mental illness (i.e., *"I'm going to try recovering"*) and those related to other aspects of life (e.g., vacation, school). Occasionally, individuals addressed their absence; however, much of what I observed was individuals stepping back from accounts without explanation, sometimes surfacing months later or not at all. Activity timescales and periodicity for generating new posts vary for each individual and, while sometimes consistent within a relative window of observation, are likely to change. Understanding this context is vital to examining how individual lives and posting habits change, as well as the ways in which mental illness is, at times, more relevant or necessary to communicate.

Just as I observed changes in posting activity, I also saw changes in content. Here, I focus on changes in content related to mental illness, to show how mental illness is not a constant experience. One poster observed during my digital ethnography only began to incorporate #edrecovery in posts within the past six months, despite an account history that dates to early 2017 and account content, consisting of recipes, that has remained constant. With respect to content related to mental illness across time, another poster, documenting their eating disorder, previously shared pictures of their body, face blurred or otherwise obscured, with captions including, *"Huge. Super fat. Gross #ednos #ana,"* and *"Lost weight but still fucking disgusting #500caloriesaday."* At the time of analysis, these selfies are nonexistent; their prevalence subsumed by screenshots depicting caloric intake through a common fasting application. Similar to practices of decreased posting, individuals sometimes announced transitions through mental illness (e.g., maintenance, recovery, relapse) with content and, specifically, captions, including, *"Recovery again. But I'll probably relapse... my weight is terrifying."* By attending to content variation across individual timescales, we begin to understand the different ways in which mental illness can both remain consistent and fade in and out of online expression of lived experience as narrated through Instagram.

4.4.1.3. *Reframing Disruption within Individual Experiences*

For individuals living with mental illness, mental illness may be an entirely common part of their everyday life. On Instagram, individuals express their experiences with illness as ordinary through entanglement with other components of daily life and certain posting practices, such as posting about mental illness every day or consistently only every once in a while. This perspective demonstrates the importance of different types of context and challenges societal perceptions of mental illness as deviant (i.e., people with mental illness are not necessarily expressing mental illness as deviant). However, it does not suggest that mental illness is not disruptive. For example, one account observed during my online observation recently shared content related to an eating disorder relapse and subsequent hospitalization. In one particular post the individual wrote, *"My eating disorder wants me to restrict and to continue losing even more weight. This is ridiculous. I didn't come [to the hospital] to keep losing weight. It'd be really dumb to waste this treatment opportunity."* This caption suggests that elements of mental illness (e.g., restrictive behaviors, losing weight) are disruptive to everyday life during healing and recovery, such as attempts to gain weight. These findings corroborate traditionally held views with respect to mental illness as a disruptive force (Pater and Mynatt, 2017). However, examining mental illness as an ordinary part of an individual's life, rather than a sensationalized extreme, calls into question how we think about disruption.

In narrative research, illness is frequently framed as a disruptive event (Riessman, 1990). Individuals who are ill are often relegated to suffering or to 'overcoming' their illness (Frank, 1995; Kleinman, 1989), sentiments we see carried through work in HCI and CSCW. Bias toward diagnosis and treatment neglects that certain individuals may not want to recover, that mental illness is a form of neurodiversity and psychosocial disability, and that recovery, though often aligned with 'good', may be disruptive to daily routines established while living with mental illness. I observe that what constitutes disruption is based on the daily narration of experience. Many posts within my dataset suggest that individuals are reframing what it means for mental illness to be 'disruptive'.

When people live with mental illness, disruptive events are sometimes the antithesis of this everyday experience. Specifically, changes to daily life and routine, such as being forced into treatment or attempting various practices linked with recovery (e.g., eating certain foods, not self-harming), become disruptive. One account within my online observation often posts from inpatient treatment facilities, writing about the disruption of inpatient treatment and recovery. In one post, complete with a mirror selfie from a treatment restroom and hashtags including #picoftheday, #psychward, and #eatingdisorders, the individual writes, *"Why can't I go home? I promise I'll eat. I promise. No one understands. They're not supportive, they only hinder my progress."* In another post – a black and white image with overlaid text (i.e., *"I hate myself so much"*) – the caption reads, *"I have gained a lot of weight. I fucking hate myself, all I want to do is eat. I am so fucking depressed. This place has ruined me."* While a desirable clinical outcome of recovery may involve placing mental illness "under the person's control or [to where it] is at least no longer intrusive or disruptive" (Davidson et al., 2005), I find that recovery itself can be a disruptive process. Just as contextualization demonstrates the ordinary, familiar lives of people living with mental illness, so can it depict how certain practices related to mental illness, including recovery, can be disruptive. Individuals must be foregrounded, so that their experiences with mental illness are contextualized within their own life and not the lives, or randomized posts, of others. By challenging traditional notions of mental illness as an inherently disruptive event and other activities, such as recovery, as non-disruptive (or desirable) events, I confront the appropriateness of contemporary technological 'solutions' employed by many social media sites, Instagram included, to address the 'problem' of some content related to mental illness.

4.4.2. Negotiating the Telling of Experience

In line with prior work, many participants in my interview study described how social components of Instagram contributed to their motivations for participation, such as sharing experiences (without the necessity of reaching out personally or in person), receiving support and validation, connecting with specific communities (e.g., eating disorder communities), and providing support for others (Blair and Abdullah, 2018; Chancellor et al., 2018; De Choudhury and De, 2014; Zhang et al., 2018). Given the social nature of Instagram, expressing mental illness within everyday life is not straightforward nor only individual (Frank, 1995). Individuals must navigate the sociotechnical structures that shape personal narration of everyday experiences, including those related to mental illness. Through the use of Instagram, individuals implicitly and, sometimes, explicitly invite collaborative storytelling and reinterpretation of experience by others (i.e., an audience).

While audience members may provide a range of reactions (Andalibi et al., 2017; Brown et al., 2018; Ernala et al., 2018; Zhang et al., 2018) to tellings of mental illness, many of which are positive, researchers, with few exceptions (e.g., Pater et al. (Pater et al., 2016a)), have yet to thoroughly engage with how these interactions, compounded with features of social media platforms, influence the content individuals share and how they select to share it. Here, I turn to social control as a way of understanding how other people and the Instagram platform participate in narration of mental illness. Social control involves the "ways in which a society tries to prevent and sanction behaviors that violates norms" (Barkan, 2011). All audience reactions, including a range of positive (e.g., approval, reward) and negative (e.g., disapproval, punishment) (Young, 1949), have implications for the ways in which social control plays out online and how individuals must negotiate the telling of their experience. Though some forms of social control may occur through subtle means of interaction, such as brief praise or a 'like' on Instagram, others are more overt. Historically, individuals with mental illness have been subject to these direct forms of social control (Horwitz, 1982; Opalić, 2007; Scull, 1991). I find this to be the contemporary case, as well. Individuals must navigate interaction narration of mental illness, even when told as and within their everyday, liable to be rewritten, reported, and removed.

4.4.2.1. *Narrative Alignment and Validation of Experience*

For some of my interview participants, co-constructed storytelling and interpretation of experience by audience members was of little or no concern. P1, for example, mentioned that followers of his private account were free to *"derive whatever meaning they wanted to from"* his posts. P2, similarly, explained how *"people can get their own interpretation. You can read the same thing. Everybody can read the same thing and get a different meaning."* Oftentimes, reinterpretation by audience members was a benign or positive activity, particularly when it aligned with the expression shared by the poster. For example, in one post shared by a public account within my Instagram corpus, an individual discusses how their bipolar disorder contributes to *"what's wrong with me."* Coming from a place of support, another individual commented, *"I've seen these comments that say nothing's wrong, but we know that's not true. I'm bipolar too and there's something wrong. Do what's best for you and ignore the ignorant..."* Validation of experience, a form of support and care (Naqshbandi et al., 2016; O'Leary et al., 2017), plays an important role in motivating some individuals to share about their everyday life with mental illness. P5, for example, described how her desire to share *"came from a place of validation. And I think that's a lot of what Instagram is grounded on."* Validation of experience extends beyond social support. It is confirmation of the way – and worldview – through which someone experiences having a mental illness.

In addition to providing validating reinterpretations of mental illness, audience members may react positively to posts through textual praise (e.g., comments, direct messages) and 'liking'. For example, a post by P7 alluding to her experiences with depression received audience responses such as *"Thank you"* and *"respect you for sharing."* In addition to comments of appreciation gently rewarding this particular act of sharing, P7's post received 30 likes. Another member of my interview cohort, P17, discussed how positive reactions motivated him to post about his life with mental illness: *"It's like a positive assurance. Positive words. It's nice to have people who kind of give you advice, give you support. It's kind of, like, motivational, I would say."* While many positive reactions are online and public (provided the account is public), others are not. Audiences may engage with individuals through direct messages and, when possible, offline means, suggesting that mechanisms of social control flow through a variety of channels, not all of which are explicitly observable.

4.4.2.2. *Narrative Conflict and Erasure of Experience*

Not all narrations of mental illness are socially acceptable. At times, members of the audience reinterpreted expressions of daily life with mental illness in ways that conflicted with, and often erased, the original telling (i.e., within the Instagram post). This type of reinterpretation is particularly salient in relation to socially unacceptable expressions of mental illness, including posts about self-harm and self-critique (e.g., an individual with anorexia captioning a selfie with *"Yuck. What a whale"*). At times, audience reinterpretation, though conflicting, meant to offer support. For example, on a post P15 shared after feeling *"hella bad about drinking"* a Sprite, a friend commented, *"Bitch shut up your body is perfect."* Publicly, P15 'liked' this reaction. However, during the interview she described how *"someone telling me that my body is perfect doesn't mean that my body is perfect. Because it's not. No one's body is perfect and – it's like, even as nice as it feels that someone cared to comment like that, it's just annoying because it's obviously not true. It's obviously just meant to make me feel better and not actually a real thing."* Support is frequently tied to positive audience reactions and mechanisms of social control that reward certain posting practices. However, as demonstrated here, social support may have a negative, invalidating effect when it conflicts with certain experiential aspects of living with mental illness.

Audiences may also respond with conflicting reinterpretations of experience that shame or harass the original poster, potentially to push the original poster toward a more socially acceptable posting norm. For example, an audience member, perpetuating a harmful stereotype, referred to one poster in my Instagram dataset as an *"attention whore"* for uploading content, including pictures, depicting self-injury. Similarly, on a suicide helpline post shared by another poster in my online corpus, another audience member commented, *"u r not alone ur just fucking stupid. ur emo community is a fucking sickness im done with seeing ur stupid shit."* These hostile comments may be reminiscent of personal attacks, where audience members attempt to control the posting practices, and, implicitly, the underlying behaviors of particular Instagram posters through aggressive and, even, hostile reinterpretations of content. In addition to conflict, these comments demonstrate attempts to rewrite and erase the context of the original posts by providing new, albeit harmful, perspectives. Individuals for whom mental illness becomes part of daily expression on Instagram must negotiate reactions such as these, where audience members interactively participate in constructing, and attempting to control, personal narration and ways of experiencing life with a mental illness.

Though audience members may attempt to exert control on a poster's narration of experience, comments themselves are not mechanisms through which total control can be exerted. To this effect, Instagram has several means of reporting and moderating content. Instagram's reporting feature allows an audience's negative reaction to content (e.g., reporting a post) to impose upon the daily routine of individuals living with or posting about mental illness. Reporting practices such as these (e.g., an Instagram message saying, "We're Reaching Out to Offer Help") may attempt to alter or 'fix' the posting or underlying behavior of a person by providing them with nonspecific resources, which are sometimes unnecessary and unwanted, through the guise of care and support. One account holder in my online corpus posted their reactions to receiving these notices and helpline resources. In one caption, they mention, *"Again, fuck off,"* while, in another, they describe the practice as *"super counterproductive it's stupid."*

While some reporting practices may be limited to the imposition of unsolicited help, others silence and erase the subjective expression of an individual already marginalized by society. The vague description of removal (i.e., "We've removed your post for not following our Community Guidelines") provides little opportunity for recourse or transparency with respect to who – or what (i.e., an algorithm) – reported the content. It is nearly impossible for people to have posts or accounts reinstated. Though researchers have addressed instances in which content moderation might be appropriate (Wohn et al., 2017), my analysis provides an alternative view: processes of content moderation can ultimately oppress marginalized groups and exert paternalistic social control through platform features favoring 'mainstream' reactions over the expression of individuals living with mental illness.

4.4.2.3. *Reclaiming Control through Multiple Accounts*

In the current implementation of Instagram, processes of social control related to the reporting or removal of posts outweigh platform mechanisms that may be used to indicate positive audience reaction. Specifically, individuals with mental illness are subject to social interactions that may reinterpret or erase personal expression. One way that individuals exercise control over the telling of their own experience is through the practice of creating multiple accounts. While other researchers suggest that the maintenance of multiple accounts may facilitate sharing secret or sensitive information (Andalibi et al., 2017), here I suggest a new interpretation: multiple accounts are a necessity based on the culture in which we live, where individuals with mental illness experience intolerance, misunderstanding, and oppression. Maintenance of multiple accounts provides opportunities to interact with different audiences and to share content where it will receive the most validating, least judgmental response.

Several members of my interview cohort described their preference for including posts featuring or referencing mental illness on their spam or Finsta account (or accounts, in the case of P16) rather than on their main account. As P10 and P15 explained, their decisions to post about mental illness were grounded in trust. For these individuals, their spam accounts were shared with fewer and closer friends. Like P10 and P15, P16 did not post about mental health or mental illness at all on her main account. Instead, she shared *"mental health posts such as things I see in papers about mental health or suicides and what i can do to prevent it"* on her popular spam account and, on her less followed spam account, *"put more personal ones...because the followers on that account are ones i can trust."* The strategic use of multiple accounts demonstrates how perceptions of audience, and the trust an individual has with this audience, influence where and how narration occurs.

I also observed how the need for privacy and like-minded community influence the maintenance of multiple accounts. P15, for example, discussed previously maintaining *"a special Instagram account for my eating disorder and related things"* where she would only *"follow other people with eating disorders. Like, none of my friends. I didn't post my name or my face."* Multiple accounts need not exist in tandem, though they can. Rather, individuals may leverage account creation features as a means of circumventing certain forms and processes of social control. For example, the first post on an account including graphic images of self-harm referenced a previous account that had been removed by the platform. Despite attempts to circumvent platform social control, similar to lexical variations of hashtags (Pater et al., 2016a), new and multiple accounts maintained by individuals living with mental illness are still subject to intense scrutiny, surveillance, and control.

4.5. Discussion

In this chapter, I describe how researchers and other technologists can better understand the daily lives of individuals living with mental illness, as well as how stories told from group members are co-constructed by various outside actors – both human and computer. I demonstrate how mental illness is a familiar part of the fabric of everyday life for some people. Additionally, I show how individuals with mental illness negotiate online instantiations of social control in order to express and communicate their experiences with mental illness. The use of small stories research provides an analytic lens for prioritizing interactional, ongoing, and everyday stories, such as those occurring on Instagram. By attending to narratives within specific interactions, such as posts and comments, rather than broader retrospective accounts or societal perceptions of mental illness, I offer counter, but complimentary, interpretation to understanding mental illness online. In the following sections, I discuss the implications of examining mental illness within the context of an individual's daily experience or as an ordinary, albeit not trivial (Michie et al., 2018), part of their life; the ways in which current implementations of features on Instagram may facilitate oppression of a marginalized group; and how, as researchers and designers, we contribute to processes of social control.

4.5.1. Contextualization through an 'Everyday Lens'

Certain methods of data collection (e.g., the use of hashtags and key terms to identify mental illness content (Chancellor et al., 2016d; Manikonda and De Choudhury, 2017)) prioritize the explicit expression of mental illness and neglect how mental illness is, at times, more or less salient within everyday life (Andalibi et al., 2017; Feuston and Piper, 2018; Gerrard, 2018; Pater et al., 2016a). Following from these practices, researchers often recontextualize mental illness within a narrow window of interpretation, not taking into account a wide possibility of meaning. This narrow interpretation, which is often not engaged with subjective experience, ultimately aims to produce generalizable findings. These findings have important implications for understanding community and cultural norms (Chancellor et al., 2018) and contributing insights to diagnosis and treatment of mental illness online. However, design recommendations that center on broad, population-level approaches to mental illness, do not fully appreciate the individual impact of life with mental illness and of the potential impact of technological 'solutions'. In this context, mental illness becomes decontextualized from the everyday, lived experience in which it occurs. Decontextualization has the potential to sensationalize certain facets of experience by highlighting them absent other aspects of life and time. Incorporating the subjective experience of mental illness is an important component of understanding and including the idiosyncrasies and forms of expression produced by members of a marginalized population. Prioritizing the individual and the everyday means shifting focus from a generalized population to a person; a practice in line with traditional forms of psychological treatment (e.g., talk therapy).

The experience of mental illness is different for everyone. In this paper, I present mental illness not only as the topic of narration, but as the condition through which narration occurs (Frank, 1995). Here, I extend an agenda put forth by other HCI and CSCW researchers studying mental health and mental illness (e.g., Jessica A. Pater and Elizabeth D. Mynatt (Pater and Mynatt, 2017), Nazanin Andalibi et al. (Andalibi et al., 2017)) that argues for a more diverse understanding of mental illness and online spaces. By emphasizing an everyday context, we can better understand normalcy and disruption for the individual and attend to the ways that design for certain populations fails. By prioritizing individuals and their small stories, we have an opportunity to explore and disseminate alternative narratives of mental illness, such as those demonstrating ways that individuals with mental illness live and thrive. This reframing is reminiscent of the "living with AIDS" movement, which reframed life with serious illness when mainstream narrative focused on dying (Grover and Crimp, 1987). Examining the holistic experience and attending to how people narrate mental illness presents an opportunity to support and validate individual interpretations, rather than conflict them. Further, an 'everyday lens' can be helpful for thinking about marginalized groups consistently 'othered' due to association with deviance, whether historically or presently, such as individuals previously incarcerated and LGBTQ+ community members.

4.5.2. Technological Responses to Mental Illness

Instagram includes a range of technical features and process, including comments, reporting, and content detection algorithms. These features, through use, produce new instantiations of social control, which enable audiences to influence the content that people generate, including how individuals express their experiences and communicate about mental illness online. Though platform features are not inherently coercive or restrictive, the ways they are used contribute to a socially regulated environment in which members of a marginalized group (i.e., individuals with mental illness) are subject to various forms of oppression, including surveillance and erasure. Platform features, including the option to report "posts encouraging or promoting self injury, which includes suicide, cutting and eating disorders," as described by Instagram, are founded in mainstream perspectives and calls to action; rather than, for example, the perspectives and concerns of people with mental illness. I revisit this discussion in Chapter 6, where I examine how platform processes of content moderation work to enforce a particular type of conformity with respect to body image and body management practices online.

Dominant colloquial and medical attitudes often conceptually entangle mental illness with deviance, such in that mental illness is viewed as outside the 'norm' of socially sanctioned behaviors and ways of being. This view of mental illness as deviant, and the stigma associated with such attitudes, means that people with mental illness are not just viewed as dangerous to themselves, but also to others – even others online (i.e., contagion (Carlyle et al., 2018)). As I demonstrate through my analysis, however, mental illness is ordinary or unusual, disruptive or not in the context of an individual's life, rather than any particular societal perspective. For example, recovery, frequently favored and privileged as a desirable outcome, can be disruptive – and, at times, even traumatic (Green and Ubozoh, 2019). Automatically providing certain resources (e.g., helplines, websites) to targeted posters, though framed within a commitment to social good, may actually be harmful and "counterproductive," as one poster wrote. We must reflect on the clinical and societal assumptions embedded in technology, which stand to affect marginalized groups the most. To be clear, this doesn't involve abandoning recovery-related resources or outcomes. As I discuss in Chapter 6, it involves, instead, understanding how people respond to current practices of support on social media platforms and designing for a multiplicity of experiences with mental illness.

Community awareness concerning algorithmic bias around gender, race, and age is growing (Diaz et al., 2018; Hamidi et al., 2018; Schlesinger et al., 2018). Additionally, researchers have begun to address the ways algorithms, often relying on strict classification, contribute to unjust social structures (Fox et al., 2017). As I'll detail in the next chapter, Joy Adowaa Buolamwini talks about these particular types of algorithmic practices and bias through her coined term 'the coded gaze' (Buolamwini, 2017). This way of thinking about bias in sociotechnical systems enables us to closely examine the many algorithmic applications involved in the maintenance of online forums and social media platforms (i.e., particularly those relying on categorical approaches to content). These algorithmic practices have implications for content recommendation, organization, moderation, and censorship; for example, the detection of harassment and toxic content (Bretschneider et al., 2014; Chen et al., 2012). Censoring content, despite necessary and beneficial applications (e.g., detection of cyberbullying (Ashktorab, 2016; Nandhini and Sheeba, 2015)), restricts online participation and, as Lindsay Blackwell and colleagues describe, has unintended repercussions related to the potential erasure and invalidation of 'atypical' (i.e., not easily classified) experiences with online harassment (Blackwell et al., 2017). Given the potential for some content related to mental illness to be categorized as toxic, harmful, and deviant, my work demonstrates a blatant, albeit different, form of discrimination and restriction – one that must be addressed.

As I'll revisit and more thoroughly detail in Chapter 6, content moderation (e.g., through various means, such as audience reporting features and algorithmic detection) perpetuates a surveillance culture that disproportionately impacts certain individuals with mental illness – those who already have a history of marginalization and subjugation. Practices associated with content moderation, more than just producing a chilling effect (Chancellor et al., 2016b), are oppressive; used to control behavior and the ways people participate online (Pater et al., 2016a). Processes of content moderation favor mainstream ways of being and participating online as well as 'treatment' of individuals with mental illness involving lack of transparency and recourse. Individuals who have been reported or had content removed cannot face an accuser – audience member or algorithm – or easily petition to have posts or accounts restored. This approach to mental illness online, void of broader life context and unproven with respect to the provision of help, can silence those who need a space to explore their experience – including the worst aspects of it – and connect with others without fear of judgment. Social media should not obsessively focus on 'fixing' certain experiences with mental illness (e.g., providing mental health resources) or removing them from online spaces. Instead, platforms can consider ways to support everyday experience and to assist people with mental illness in negotiating forms of social control enacted by certain audiences.

4.5.3. The Role of Researchers and Designers

By contextualizing mental illness within the everyday and calling attention to the ways in which social control manifests online, my aim is to raise awareness of the sensationalism of mental illness content and the paternalistic underpinnings of technological solutions. Researchers and designers, though often inadvertently, perpetuate stereotypes, stigma, and systems of social control, such as through highlighting certain types of mental illness content and labeling this content as deviant or subversive (or comparing it to other undesirable types of online content, such as cyberbullying (Pater et al., 2016a). Representations of mental illness in research can sensationalize certain behaviors and content (e.g., self-injury) by drawing attention to components of mental illness absent the vast majority of lived experience as presented online. Given this view, it is unsurprising that the design of technology often focuses on identifying and mitigating behaviors and content associated with certain types of mental illness content and communities (e.g., pro-eating disorder, pro-ana). Ultimately, current research practices neglect much of the richness and diversity of online representation of mental illness. For example, current approaches to studying mental illness online have no means of collecting or interpreting content without captions or hashtags, something I'll discuss in the following chapter. We must consider life (i.e., online and offline) with mental illness as a cohesive whole, rather than as a convenient sample of content. This can, for example, help us differentiate between people using online spaces to cope versus individuals who are actually and actively seeking to cause others harm (e.g., people without eating disorders who offer to be eating disorder coaches).

The experiential turn in HCI (e.g., Dourish (2006); Wright and McCarthy (2010)) suggests a more humanist agenda and emphasizes research and design practices that enhance lived experience. Without careful reflection, research framing and design solutions may reduce individual agency and exert control over marginalized populations, who are living with the histories and present realities of enduring controlling institutions and relationships. By acknowledging how researchers and designers contribute to processes of social control that are enacted through technology, we can reevaluate what it means to design for 'social good' (Smyth and Dimond, 2014). In part, this involves increased reflexivity (Bardzell, 2010), wherein we recognize that certain goals and values, particularly those developed and promoted by outside actors, with respect to mental illness can have unforeseen consequences. For example, the goal of helping society by reducing 'contagion' of mental illness content can inadvertently punish those in need of a space for coping, venting, understanding, and shared experience.

Additionally, practices of experience-centered design and reflexivity create space for us to consider the role of capitalism in the design of technology, such as social media. With respect to the capitalist structure of social media sites, Instagram included, companies benefit from the creation of multiple accounts (e.g., increased advertising revenue) and content, even if this content is at odds with community guidelines and terms of service. When corporate goals aim to drive traffic and usage, focusing on satisfaction of the 'general user population' takes priority and, arguably, renders certain regulatory practices (e.g., removing accounts that post about mental illness in certain ways) ineffective. Platforms, however, are responsible to all who use them. While mainstream sentiment stigmatizes and misunderstands mental illness, capitalism and individual expression of mental illness will remain at odds. In Chapters 6 and 7, I briefly revisit capitalism in online spaces as it relates to mental illness through Jessa Lingel's work on digital gentrification (Lingel, 2019) and Mark Andrejevic's research on digital enclosures (Andrejevic, 2008). Here, however, I question the role and response of technology companies to content related to mental illness online, just as I question the role of researchers and designers. Instead of falling in line with dominant societal norms and values, we should closely consider our influence and how we might use our position to support increased agency for individuals with mental illness.

4.6. Conclusion

Online communities and social media provide opportunities for members of marginalized groups to find and build spaces in which they belong. Platform features, however unintentional, may inadvertently restrict expression and contribute to the marginalization of certain individuals and groups of people. Here, by considering an individual's everyday experience with mental illness, I begin to challenge societal notions of deviance and sociotechnical processes of control, such as 'universal' helpline resources and content moderation. Through the use of small stories research, I have illustrated how individuals living with mental illness narrate their experiences as part of their everyday lives and circumvent social control (e.g., through the use of multiple accounts) that influence and, at times, inhibit expression. By attending to the individual experience of mental illness, in which mental illness as contextualized alongside other aspects of everyday life may, at varying times, be ordinary, unusual, and disruptive researchers and designers have opportunities to reflect on their impact with respect to the implementation of controlling technologies and to explore alternative ways of approaching design of technology for people with mental illness.

CHAPTER 5

Beyond the Coded Gaze: Analyzing Expression of Mental Health and Illness on Instagram

In the previous chapter, I detailed how mental illness is part of everyday life for those living with it. As such, mental illness isn't necessarily narrated medically or as deviant (e.g., rule breaking). Emphasizing personal narratives and tellings of mental illness can enable us, collectively as researchers and designers, to better understand what is normal, or not, and what is disruptive for people with mental illness – including the ways in which platforms handle certain type of content and send support resources. Additionally, as a part of everyday life, mental illness is surrounded by, brought alongside, and entangled with a rich and diverse variety of contexts. However, as I argue in the previous chapter, many approaches to examining content about mental illness online neglect to keep context intact. This results in a distorted view of mental illness online, which ultimately impacts practices associated with research and design.

Building from Chapter 4, here I examine the types of individual and collective experiences with mental illness that are missed and misinterpreted by human and computational actors operating within sociotechnical classification systems. To illustrate the importance of context, I bring mental health into my examination, demonstrating how the ways people communicate about mental illness and mental health share a number of similarities that are difficult to tease apart. My empirical work forms the foundation for my critique of current computational and classification systems and extends my argument for the necessity of context in designing online spaces for people with mental illness.

Here, I make four contributions: 1) Through my analysis I combine firsthand accounts of people living with mental illness and data collected from public Instagram accounts; 2) I attend to the ways in which mental health and mental illness are interrelated, co-existing concepts; 3) I take an interdisciplinary approach to social media analysis, including works from visual sociology, cultural studies, and multimodal discourse analysis; and 4) I highlight ethical considerations for future social media research in the context of mental illness content.

This chapter is published with Anne Marie Piper in the Proceedings of the ACM on Human-Computer Interaction:

> Jessica L. Feuston and Anne Marie Piper. 2018. Beyond the Coded Gaze: Analyzing Expression for Mental Health and Illness on Instagram. In Proceedings of the ACM on Human-Computer Interaction, Vol. 2, No. CSCW, Article 51, November 2018. ACM, New York, NY. 21 pages. https://doi.org/10.1145/3274320

5.1. Introduction

Representations of mental health and, particularly, mental illness have long contributed to the cultures of Western societies. In art and literature, these artifacts come to us through religious paintings and accounts of demonic possession and divine punishment, as well as secular interpretations of and responses to suicide[1] and depression[2]. Traditionally conveying a grander message of cultural values (e.g., morality), representations of mental health and mental illness have come, in recent centuries, to depict personal and deeply emotional experiences.

Art and religion are neither alone in representing mental health and illness nor influencing perceptions of particular disorders and those living with them. News and advertisement media, as well as certain medical disciplines, including neurology and psychiatry, have impacted the ways mental health and mental illness are represented (Gilman, 2018; Wahl, 1995). Across domains, technology plays an increasingly influential role. From the camera to the phone application, technological developments shape how mental health and illness are portrayed, as well as who has the opportunity to portray them.

In CSCW and HCI, scholars have also contributed to representations of mental health and illness, such as through graphs and charts generated through mental health and mood applications (Bardram et al., 2013; Sonderegger et al., 2016) and analysis of mental health and illness-related posts on social media platforms (Andalibi et al., 2017; Chancellor et al., 2016d; Tsugawa et al., 2015). On platforms such as Instagram and Tumblr (Chancellor et al., 2017; Manikonda and De Choudhury, 2017; Pater et al., 2016a), scholars have attended to understanding content and style of mental *illness* (Manikonda and De Choudhury, 2017), particularly eating disorders (Chancellor et al., 2017; Pater et al., 2016a) and self-harm (Pater and Mynatt, 2017; Seko, 2013), as well

---

[1] The Suicide of Dorothy Hale (Frida Kahlo, 1938).

[2] Portrait of Dr. Gachet (Vincent van Gogh, 1890).

as content moderation (Chancellor et al., 2016b, 2017), behavioral intervention (Pater and Mynatt, 2017), disclosure (Andalibi et al., 2017), and predictive modeling (Chancellor et al., 2016a,c). Though this research provides insight into the representation of mental illness on social media, it's built on a foundation of descriptive statistics and generalizable trends, which prioritize population-level practices and rigid classification that may disguise uncommon, albeit meaningful, practices and attitudes.

Although we know social media provide an opportunity for self-expression, sharing personal experiences with mental health and illness (Andalibi et al., 2017), and empowerment (Ackermann et al., 2009; Ammari and Schoenebeck, 2015; Blackwell et al., 2016), we have a limited understanding of how and why people craft, share, and disseminate representations of mental health and illness on these platforms. In this domain, social media posts are often analyzed in absence of interviews with community members, which implicitly emphasizes the significance of observable artifacts (e.g., posts), or representations, rather than broader expressive processes (Constantinou, 2005), such as practices associated with posting (e.g., selecting an image). Though an element of expression, representation offers a narrow, captured view. By incorporating first-person narratives in my data, I expand my analysis to consider produced artifacts and surrounding processes, which I refer to as expression.

In this paper, I build upon prior work through digital ethnography, joining manual collection of Instagram posts with semi-structured interviews and photo elicitation to understand how posters (i.e., people posting content) on Instagram express attitudes and share experiences related to mental health and mental illness. In contrast to prior work, I argue that content on social media benefits from an analytic approach recognizing that all methods of interpretation are situated and partial (Harper, 2012). Here, I turn to Joy Adowaa Buolamwini's concept of the coded gaze (Buolamwini, 2017) to address theoretical assumptions underpinning current practices in data analysis and how these assumptions are supported by related activities, such as data collection. The coded gaze refers to algorithmic bias present in artificial intelligence due to human values and priorities (Buolamwini, 2017). Though coined to address individual and collective values, I employ the coded gaze through its capacity as an analytic lens. Given similarities between machine methods of analysis and other statistical methods (i.e., content analysis (Bell, 2001)) employed in social media analyses of mental health and mental illness, I extend the coded gaze to address methods that share similar theoretical assumptions. I argue that the coded gaze restricts our ability to understand how individuals negotiate claims to mental health and mental illness, such as through visibility and signaling, and participate in community practices, such as remix. I also discuss how assumptions underlying the coded gaze dehumanize posters through the application of population-level technological solutions that neglect individual experience.

I make four primary contributions. First, my analysis, which combines manual online data collection with semi-structured interviews, extends prior work by placing the poster's experience at the center of analysis. First-person accounts of mental health and mental illness on social media are rare in the literature. Second, my work attends to mental health and mental illness as co-existing, flexible, and interrelated; offering a richer view than previous attempts to classify 'user' behavior. My findings challenge the dichotomy applied to mental health and illness by providing an analysis of visual and textual content, as well as poster accounts. Third, my method involves an interdisciplinary approach to social media analysis that incorporates techniques and theoretical perspectives from visual sociology, cultural studies, and multimodal discourse analysis. Finally, I discuss how my approach and findings highlight broader ethical concerns for research in this space.

5.2. Background

In this paper, I confront the relationship between mental health and illness. This relationship has, at times, been described as a spectrum (Payton, 2009) and at others a strict dichotomy (Boruchovitch and Mednick, 2002). I, instead, approach this relationship through a modified version of the Dual Continuum Model of Mental Health and Mental Illness (Keyes, 2014)[3]. This model demonstrates how "mental health and mental illness belong to two separate but correlated dimensions in the population" (Keyes, 2014). Instead of adopting the definitions of mental health and mental illness used in the standard version of this model, I adopt the framings described below.

---

[3] The Dual Continuum Model of Mental Health and Mental Illness was initially helpful for conceptualizing the relationship between mental health and mental illness. However, it no longer accurately represents how I understand these two social constructs. In particular, this is a medical model for understanding mental health and mental illness. Though I did not intend to use it as such – for example, pairing this model with work by Thomas Szasz – it nevertheless produces a medicalized lens for thinking about mental health and mental illness. Over the course of my dissertation research, I've taken to considering mental illness through disability and madness. In part, this has involved thinking about people with mental illness as members of a social group; particularly one that resists normative and psychiatrized ways of being and shares histories of marginalization, stigma, and oppression. This approach radically attends to the ways that people construct their own experiences and narratives with mental illness, disability, and madness, such as through various forms of online participation. Considering mental health, then, involves understanding how disabled and mad folks construct practices of healing and wellness. This latter consideration is not within the scope of my dissertation, though I intend to build on this thinking in future work.

Current paradigms of mental health (Manderscheid et al., 2010) position the concept as more than the absence of mental illness; situating it, instead, within a continuum of wellbeing and mental functioning. Mental health is associated with ‘flourishing’ (e.g., positive mood and unimpaired functioning) and its absence is associated with ‘languishing’ (i.e., negative mood and impaired functioning) (Keyes, 2002). Silvana Galderisi et al. (Galderisi et al., 2015) contend that this continuum restricts and excludes a “variety of emotional states,” as well as “imperfect functioning”. By contextualizing functioning and emotions within specific situations and experiences, Galderisi et al. reposition mental health as “close as possible to human life experience, which is sometimes joyful, and at other times sad or disgusting or frightening” (Galderisi et al., 2015). I adopt this framing to acknowledge that flourishing and languishing mental health are situational and dependent on context.

Mental health and related concepts, such as mental illness, are deeply entangled with cultural values and priorities. In CSCW and HCI, the disciplines of clinical psychology and psychiatry are influential. These approaches to mental illness presume a biological (e.g., neurological) or biopsychosocial etiology for illnesses (Sullivan, 2016), which manifest through symptoms of behavioral deviance from an implicit social norm. Writing in critical dissent, Szasz reframes mental illnesses as ‘problems with living’ rather than actual diseases, which have biological markers (Szasz, 1961). This theoretical approach to mental illness promotes flexible interpretation and change over time. For these reasons, my own understanding of mental illness is grounded in Szasz’s seminal work (Szasz, 1961), where mental illnesses are reflective of deviance from a narrow window of societal standards. Mental illness is produced, in part, from the responses of people to it, including actions and attitudes that stigmatize and marginalize people living with mental illness. To maintain inclusivity of the various ways individuals with mental illnesses conceptualize their experiences, I extend my definition to include clinically diagnosable mental illnesses, as well as self-identified mental illnesses.

5.3. Related Work

Scholars from multiple communities, including human-computer interaction, sociology, communications, and medicine, have studied representations of mental health and illness. The present analysis brings together these diverse perspectives applied to analyzing content on Instagram.

5.3.1. Institutional Representations of Mental Illness

Absent personal experience, many Americans receive information about mental illness through the news media (McGinty, 2017). Mass media representations of mental illness are particularly vital due to the potential reach of their messaging. Individuals with mental illnesses are frequently portrayed by the media as dangerous, violent, criminal, and "fundamentally different from others" (McGinty, 2017; Stout et al., 2004; Wahl, 1995), though anyone may experience mental illness. These unfavorable and inaccurate representations of psychiatric disorder and neurodivergent experience contribute "to the stigma and discrimination that represent formidable barriers to treatment and recovery" (Wahl, 1995). Representations of mental illness produced by the news media often exclude points of view emphasizing how individuals live with mental illness, navigate recovery, and manage challenges that arise from stigma.

In addition to the media, other institutions have shaped public attitudes toward mental health and mental illness. Traditional medical techniques as well as the development of modern technologies have impacted representations of mental health and mental illness. For example, mental illness may be represented by the countenance (Bersani et al., 2013), diaries produced by patients (Almeida, 2005), and images of the brain (Charney et al., 2013; Drysdale et al., 2017). Drawing from psychological practices, in particular, researchers within HCI have contributed to representations of mental health and mental illness. Contributions largely take the form of health applications utilizing self-report or unobtrusive monitoring to track medication adherence, sleep, mood, steps, and location (Chancellor et al., 2016a; Lee and Hong, 2017; Wang et al., 2016, 2017a) in order to produce graphs, charts, and avatars (Bardram et al., 2013; Matthews and Doherty, 2011; Matthews et al., 2015). These examples demonstrate how mental health and mental illness may be represented, as well as who may represent them. Often, representations are crafted by particular institutions (e.g., media, medicine) acting within a set of rules, practices, and values, rather than individuals living with mental illnesses (c.f., Snyder et al. (2019)).

Art presents an opportunity for individuals with mental illnesses to participate in personal practices of expression and representation. Art brut, more commonly known as ‘outsider art’, provides additional context to the analysis presented in this paper. This genre of art is often created by individuals considered within or beyond the margins of traditionally trained artists (Maclagan, 2010). Underlying this genre is the implication of the artist as ‘the other’, individuals unable or unwilling to abide by societal norms (Parr, 2006) (e.g., institutionalized individuals (Beveridge, 2001; Schrift, 2006; Wilson, 1992)). The value of art brut comes, in part, from the juxtaposition of art producers and consumers. Consumers from the mainstream populace may fetishize artists due to their identities. This can perpetuate stigma surrounding mental illness, disability, and marginalized communities. I draw parallels between outsider art and social media, which changes the nature of who can participate in creating these representations as well as how they are crafted, shared, and interpreted.

5.3.2. Interdisciplinary Approaches to Social Media Analysis

Social media platforms present a shift in how content is produced, consumed, and circulated. Specifically, these platforms provide flexibility in the modes, such as still images, music, video, and writing, through which posters are able to represent and communicate various content (Johansson and Sternudd, 2015; LeVine and Scollon, 2004a). Nevertheless, traditional methods of social media analysis foreground written modes above all others (Jones, 2004), thus privileging aspects of written representation and people employing this specific mode, as well as distorting analysis through a conceptual lens that views content as monomodal rather than multimodal (LeVine and Scollon, 2004b). By expanding "beyond linguistic structures" (Jones, 2004), the context for analysis begins to encompass other modes frequently incorporated within posts on social media sites, such as layout (de Saint-Georges, 2004; Jewitt and Oyama, 2001) and images. These modes are not restricted to the digital or physical reality surrounding online content (Jones, 2004). They are also inclusive of private online and offline interactions, as well as the personal and cultural lenses through which individuals inevitably operate (Collier, 2001), and may, at times, contradict one another (Van Leeuwen, 2001).

To understand identity within this multimodal context, I draw from media and cultural studies (Lister and Wells, 2001). Though not always explicitly multimodal, scholars from this discipline incorporate various modes in data collection and analysis. In doing so, they consider how expressive (i.e., “those that take place as we live” (Lister and Wells, 2001)) and representational (i.e., those aimed at producing artifacts) practices intersect and articulate identity. A particular benefit to this approach is the notion that identity is unfixed and fluid. Further, identity, such as of individuals with mental illness communicating on Instagram, is established through intertextuality (Van Leeuwen, 2001), as well as multiple modes of representation and expression. In this work, these modes include, but are not limited to, captions, comments, usernames, liking a post, hashtags, profiles, private messages, offline discussion, and images. By joining my analysis of Instagram posts with participant interviews, I combine several modes of expression to understand the richness surrounding mental health and illness on Instagram.

5.4. Method

This research draws on posts and account information collected manually from Instagram, as well as semi-structured interviews conducted with members of Instagram who share posts related to mental health, mental illness, and related topics (e.g., self-care, awareness, recovery). My analysis incorporates visual methodologies (Harper, 2012; Rose, 2016) and semiotics (Hodge et al., 1988; Van Leeuwen, 2001) to examine how individuals express experiences with mental health and mental illness on Instagram.

5.4.1. Instagram Post Corpus

I assembled an initial dataset over a period of seven days in November 2017, in which I used Instagram's search feature to collect more than 3,000 Top Posts and Most Recent posts from five mental illness-related hashtags (i.e., #depression, #anorexia, #anxiety, #bipolar, #mentalillness) on Instagram. These hashtags were selected due to their prevalence or association with hashtags used in related literature (Andalibi et al., 2017; Chancellor et al., 2016b,d). After accounting for duplicates, my dataset comprised of 2,102 unique posts. I did not remove any posts from this dataset and I also did not seek, for example, to validate these posts with mental healthcare practitioners, as others have done (Andalibi et al., 2017; Manikonda and De Choudhury, 2017). Instead, I viewed all posts in this corpus as potentially relevant to mental illness in a certain context. All non-English posts were translated approximately using Google Translate.

Due to the impact of circadian rhythms associated with certain mental illnesses (Abdullah et al., 2014; Germain and Kupfer, 2008), I collected data at four points throughout the day: early morning (12AM to 5:59AM), morning (6AM to 11:59AM), afternoon (12PM to 5:59 PM), and evening (6PM to 11:59PM). To capture the variability of posts, I randomized each data collection period within these time ranges. During each data collection period, I collected nine Top Posts, as well as the 15 Most Recent. At the time of data collection, all Instagram posts included an approximate timestamp as part of the post (e.g., 33 seconds ago; April 24; July 21, 2017). Given the popularity of all hashtags incorporated in my data collection, Recent Posts were collected 5 minutes within the collection time, and often less. Instagram's Top Posts for each hashtag are determined by a proprietary algorithm. As such, I cannot determine what a 'Top Post' truly entails, aside from a rough approximation of popularity, and, instead, used these posts as an opportunity to collect additional data. Though these posts were collected within specific time ranges, Top Posts, unlike Most Recent posts, may remain for several days and were, therefore, potentially posted prior to my week of collection. Posts were collected manually, in line with Instagram's Terms of Service, as well as their platform API's policy at the time of data collection. Manual collection provided an opportunity to review each individual post in the data set. This meant that data collection and initial memoing occurred simultaneously, with the identification of themes beginning the first day of data collection. I iterated on these themes throughout data collection and when revisiting posts.

This online corpus provided an initial area of inquiry for my observation of the presentation of mental illness and related topics on Instagram. Over six months, I expanded upon these initially collected posts by examining co-occurring hashtags (e.g., #prorecovery, #secretsociety123, #mentalhealthawareness) and accounts within my dataset, including those of individuals who shared posts, as well as those who commented on and liked posts. I continued in this way, traversing through accounts, posts, and comments that were not collected during the initial data collection period. I decided during this time, as well, to include hashtags, posts, and accounts who actively promoted mental health and wellbeing to broaden my scope from mental illness alone. Collection was not intended to be exhaustive or comprehensive. Themes identified during this process (e.g., remix and recycling of images, similarity of visual styles between mental health and mental illness, perceptions of support) guided participant interviews, which in turn guided further data collection and iterative development of the findings discussed in this paper.

5.4.2. Interviews and Photo Elicitation

After beginning analysis of the online corpus, I conducted semi-structured interviews with 14 adults who post about experiences related to their mental health and mental illness on Instagram (ages 22 – 36; M = 27 years; 11 female). During the time of the interview, several participants were attending counseling sessions with a mental health practitioner or taking prescribed medication. Others had, in the past, received care from a physician to address aspects of mental health or mental illness. Still others described experiencing a mental illness (i.e., depression) without seeking care. Participants indicated experiences with anxiety, including panic attacks, depression, and eating disorders. At the time of the interview, no participants self-reported practicing self-harm or experiencing suicidal ideation. Per my approved research protocol, I did not interview individuals who self-reported currently practicing self-harm or experiencing suicidal ideation, though posts and accounts related to these topics were analyzed as part of my larger Instagram dataset.

The team for this research study included a mental health consultant (i.e., art therapist and licensed counselor by training). This team member's information was included on the consent documentation and participants were informed they could contact her at any time. Were any of the participants to start discussing present practices around suicidal ideation or self-harm, the interview would have promptly ended, and the study's mental health consultant would have reached out to them in accordance with her own set of practices and my IRB protocol. Additionally, per my IRB, I sent a document of mental health resources (e.g., helplines, practitioner locators) and the study's consent documentation to participants prior to the interview.

Following the identification of themes of interest from the post dataset and subsequent online observation, I recruited participants through websites, such as Instagram, Craigslist, and word of mouth. Though I recruited from Instagram, to my knowledge, none of my participants came from posts or accounts in my online dataset. Like many qualitative studies, self-selection was a limitation to my sample. However, I do not believe this reduces the importance of this work. It has provided an opportunity to understand Instagram and mental health and mental illness from an angle that is underrepresented. Though my corpus contains content from people around the world, I only interviewed individuals currently living in the United States. Participants received a $20 Amazon gift card following the interview.

I conducted semi-structured interviews through phone calls and text-based mediums (e.g., email). Though not preferable, interviews through text allowed me to reach people who were otherwise put off by in-person interviews and phone calls about a high personal topic. Prior to the start of the interview, the researcher clarified their role as a human-computer interaction professional and not a mental health practitioner. During the interview, I focused on how participants used Instagram to express experiences with mental health and mental illness. I did not delve into specific experiences with mental illness or reasons for those experiences; instead, I asked participants to provide me with context regarding their relationship with mental health and mental health 'conditions'. My interview protocol was tailored to Instagram posts, including images, captions, hashtags, and comments, as well as Stories, a feature on Instagram that allows people to share content for up to 24 hours. When possible, participants shared their Instagram accounts prior to the interview. Barring this, due to concerns of privacy and feelings of anxiety or discomfort, participants shared screenshots of specific posts or simply described mental health and mental illness related posts during the interview. I used photo elicitation to review individual posts (3-7 per participant) and asked questions in the context of each post (Harper, 2012). I asked participants about the types of mental health and mental illness expression they engaged with, including mediums, modes, and processes (i.e., how posts were produced), and audience. Interviews lasted approximately 30 minutes and were audio recorded and transcribed for data analysis.

All participants used Instagram regularly, though some preferred to share their experiences through Instagram Stories, rather than traditional posts. None of the participants interviewed had an Instagram account dedicated solely to mental health or mental illness. Most maintained personal accounts that included posts about events (e.g., concerts), scenery, and family, as well as topics related to mental health and mental illness. When participants did share about their experiences with mental health and mental illness, they represented their experiences in a multitude of ways, including sketches, memes, inspirational quotes, and music. My participants maintained public and private accounts. When the account was public, the participant would typically share the name, so, during the interview, both researcher and participant could review content at once. Several participants gave consent for the researcher to temporarily follow a private account. The researcher stopped following the account immediately after the interview to maintain the participant's privacy.

5.4.3. Data Analysis

Data analysis followed a constructionist grounded theory approach (Charmaz, 2014), where I worked with my co-author to develop emergent themes through iterative coding, memo writing, and constant comparison of data to emerging concepts. Themes were first identified through Instagram posts and accounts acquired during data collection in November 2017 and subsequent online observation. These themes were used to inform interviews, which, in turn, informed additional online observation and Instagram data collection. The research team met frequently to discuss emergent concepts in Instagram and interview data. The interview protocol was also iteratively revised to guide sampling, as well as refine particular concepts of interest. My analysis incorporates practices within critical visual methodologies (Harper, 2012; Rose, 2016), semiotics (Hodge et al., 1988; Van Leeuwen, 2001), and multimodal discourse analysis (LeVine and Scollon, 2004a,b) to examine how modes of expression (e.g., visual, textual, and oral), as well as social interactions, context, and practices of content production, affect and are affected by community norms and attitudes surrounding mental health, mental illness, and related experiences.

5.5. Findings

My analysis illustrates that individuals use a variety of practices and the features of Instagram to make experiences with mental health and illness visible to others. I find that mental health and mental illness are interrelated and have flexible, ill-defined boundaries. Further, I observe the emergence of practices and norms around content reposting and remix.

5.5.1. Visibility and Signaling of Mental Health and Illness

Central to my analysis is the question of how individuals make their experiences with or attitudes toward mental health and mental illness visible. Participants in my interview cohort and individuals within my Instagram post corpus use features associated with Instagram posts (e.g., pictures, captions) and account pages (e.g., profile information) to make their experiences with mental health and mental illness visible, as well as signal these experiences to particular audiences.

I find that self-expression involves recording components of mental health and mental illness (e.g., digital traces of behavior, physical changes and scars, emotion) and re-instantiating these elements in new ways. While some aspects of mental health and illness are inherently visible (i.e., weight loss or gain, scars from self-harm), others are not (e.g., depression, anxiety). Individuals make invisible or overlooked aspects of mental health and mental illness visible through different modes; frequently visual and textual. Some posters within my Instagram corpus, for example, posted screenshots from fasting applications, such as Vora, to demonstrate time without eating. Other individuals within my online corpus, as well as interview informants, posted inspirational quotes and uncaptioned visual content. P2 shared an uncaptioned picture with the text, *"No matter what knocks you down, get up and keep on going."* There is no explicit mention of mental health or mental illness within this post; even the visual content is sparse, displaying the silhouette of a hiker in gold and grey. Yet, P2's description of the post is thoughtful and deeply personal: *"I stopped talking to people. I just isolated myself, but then, you know, when I started feeling good about myself. . . . I still keep going no matter how I feel these days. . . So, even if I'm feeling really sad and don't feel like doing* *anything, I still get up, you know, and keep going. And just always hope for a better day."* P2 demonstrates how the visual content of this post was used to make certain invisible facets of mental health and mental illness visible.

Similarly, P5, who discussed mental health primarily through self-care and wellbeing, shared a post in which she had just emerged from the ocean in winter, *"freezing but fulfilled."* The picture was taken on the last day of a wellness retreat, where she was *"trying to test the boundaries between self-care and pushing yourself out of your comfort zone a little bit."* There is no explicit link between the post itself and mental health or mental illness, demonstrating that these concepts cover a broad spectrum of content. However, even when individuals represent their experiences with mental health and mental illness, their relationship to these concepts may not be apparent to an external audience or followers on Instagram.

Interview participants reported infrequently including captions or hashtags to contextualize visual content or signal a direct relationship to mental health or mental illness, typically due to internalized stigma surrounding undesirable attention-seeking behaviors. Many posts within my Instagram corpus, however, did include textual content related to mental health or mental illness, such as captions, usernames, hashtags, and profile information. In a post intended to *"build awareness of post-partum depression,"* one individual in my post corpus shared two images. In the first picture, she's smiling. Her baby's head rests on her shoulder. In the second picture, the composition is much the same. However, in this picture, the poster isn't smiling. Her face is straight, conveying, perhaps, sadness. The first picture is what she *"wanted to post on Instagram. . . but the next pic shows how I really feel."* In this post, the individual addresses the invisible nature of postpartum depression, making her experience visible through images, captions, and hashtags. Through the use of two disparate and juxtaposed images, as well as #postpartumdepression, she potentially signals to others her belonging within a specific community or association with a specific experience (American Psychiatric Association, 2013; Andalibi et al., 2017). Making elements of mental health or mental illness visible is not necessarily dependent on any single component of a post, but, rather, the specific construction of elements.

I learned of other factors that make visible and signal elements of mental health and mental illness, such as offline social contacts' personal knowledge about an individual and an individual's commenting behavior. P4, who experiences anxiety, discussed a picture of herself on the beach, commenting, *"I was able to go to California, and took a picture of myself feeling completely serene at the beach. It was awesome. And I actually had people that know me that were like, 'Oh my gosh, you look so much more relaxed.'"* The picture records a specific moment showing a woman on a sunny beach. P4's experience with anxiety relief and wellbeing was only visible to individuals who know her well. Personal experience and attitudes are also revealed through what individuals comment on others' pages. For example, commenting on a post in my Instagram corpus where an individual talks about their experience with overbearing parents and bipolar disorder, one individual said,*"Early prevention is so important. I had to deal with helicopter parents and bipolar disorder, as well, and it was frustrating at times."* Though this individual mentioned personal information in this comment, they do not make this experience with mental health or illness visible – at least to me, as a researcher-outsider – on their own account.

Across my interview data and online observation, I found that individuals constantly negotiate the visibility of their experiences with mental health and illness in their posts (and other platform contexts, such as profile information), particularly given negative societal perceptions of mental illness and the features and norms of this particular technical platform (e.g., ability to report or block posts and accounts). These examples begin to draw out how socially held views of mental health and mental illness affect the ways in which individuals create and share content.

5.5.2. Claiming a Relationship with Mental Health and Mental Illness

Rather than presuming a connection with mental health and illness or viewing visibility and signaling as self-disclosure (Andalibi et al., 2017; De Choudhury et al., 2017), I analyze the ways people position themselves and their expressions with respect to mental health and mental illness, thus 'claiming' – or not – mental health and mental illness as part of their identity.

For many individuals, establishing an 'objective' or explicit relationship to mental health and mental illness is not necessarily a goal. P2 mentioned, *"people can get their own interpretation. You can read the same thing. Everybody can read the same thing and get a different meaning."* This participant, like others in my sample, did not frequently caption pictures or use hashtags to contextualize or signal the relationship of a post (e.g., text or visual content within a post) to mental health or mental illness. Another participant, P13, described her expression related to mental illness as "suggestive," due to her tendency to err on the side of distancing her experiences with anxiety and Major Depressive Disorder from content shared. For example, in a post on her account page, P13 shared a picture of a bouquet of bright, yellow sunflowers. In addition to the picture, she wrote, *"Hey I completely forgot world mental health day but here's some flowers for all u out there struggling with annoying inner demons (fuck them). It's okay if you're not okay. You're awesome and loved."* There is a connection to mental health within the caption (i.e., support for others), as well as mental illness (i.e., *"annoying inner demons"*), but P13 does not use this post to claim a personal relationship with mental illness.

Frequently, the textual content of a post may clarify a post or individual's relationship or claim to mental health or mental illness. However, it can also have the opposite effect (Van Leeuwen, 2001). In the literature, scholars discuss how visual content, captions, and hashtags may be incongruent (Pater et al., 2016a). While this is one interpretation, my data offers another: individuals carefully craft posts to present a particular relationship with respect to mental health or mental illness. This is a strategic part of negotiating online spaces. P3, for example, does not make any explicit claims regarding her personal experience with mental illness in her Instagram posts despite living with anxiety and depression. She, instead, crafts her posts using humor, an upbeat tone, and nonsensical hashtags to present a relationship with mental illness that is *"kind of humorous... like, not take my posts too seriously."* In one post, P3 described how she *"had so many things to do, and I just did not feel good, and I wanted to do nothing at all, and just kinda be depressed all day. But I picked up my guitar and that helped me so much."* In the post's video, P3's sitting in front of the camera, playing her guitar, she sings, *"I'm getting tired; getting the soul sucked out of me."* Though she made her experiences with mental illness visible as a way to help her *"get in touch with my feelings,"* her personal relationship with mental illness is presented to others within the context of a cheerful caption (i.e., *"I wanted to hide in my blankets because of the amount of things that need to be done. But then i picked up my guitar. Good therapy session!"*) and hashtags (i.e., #haha).

Relationships with mental health and mental illness are not only constructed through Instagram posts, but also the surrounding context (e.g., followers, additional posts, profile information). Previous work on Instagram and mental illness tends to focus on the post as the unit of analysis, potentially missing this additional information. For example, an account in my post corpus wrote in their profile that they are *"RECOVERING since"* the early 2000s. Another individual from my Instagram corpus, in their profile, mentions, *"Don't report, just block. I'm receiving help. Not supporting eating disorders."* Context surrounding the person, rather than just a particular post, contributes insight to their relationship with mental health and mental illness. Further, many individuals within my Instagram post corpus incorporate words related to mental health and mental illness, such as "bipolar", "depressed", and "recovery", within their username (e.g., bipolar.pixie.dreamgirl[4], depressed_daveth). This seems to directly claim mental health or mental illness as a part of identity or, at the very least, online presentation. Understanding usernames, in addition to how individuals carry blocks of hashtags through posts, may aid in understanding the different ways people on Instagram claim a relationship to mental health or mental illness.

Although individuals carefully construct their posts and their relationship to mental health and illness, the presentation of their expression may be misinterpreted. One

---

[4] Anonymized usernames. At the time of writing this paper, neither example username was in use on Instagram.

individual in my Instagram corpus, for example, posted a greyscale image of an individual pulling down their shirt, their legs covered in scars. This image depicts self-harm without advocating for or claiming the activity. Accompanying the visual content is a caption that reads, *"My old account was deleted. It was apparently violating terms by promoting self-harm, which I would never do. I post pictures to vent. Whenever I get sent pictures or anyone asks how to do it, I tell them to stop cutting. Then I tell them how to prevent infection. If you don't like my posts, just block me. Don't report me #cutting #depression #selfharm #suicidal #ana."* This example is rare; many of the posts shared by this individual only include hashtags. One interpretation might be that this individual is self-harming or pro-self-harm, but that may not be the case. Bound up in the interpretation of content is an assumption that an individual's posts are indicative of current practices and thus current mental health status. Individuals who have practiced self-harm in the past, however, may post pictures of those past activities. While this might indicate current negativity in mental state, it does not necessarily indicate current physical harm. Nevertheless, in analyzing these images, there is a tendency to assume a temporal relationship between the content and an individual's current state: what the individual is posting is a reflection of what they are experiencing in the here and now. This may or may not be the case. We see this assumption, along with the implicit classification of an illness state, at work when systems react to an individual's post with an offer of help and mental health resources, or by removing the post and, potentially, disabling the account.

5.5.3. Mental Health and Mental Illness as Interrelated Concepts and Experiences

In Western thought, binary opposition refers to related concepts that, while theoretically opposed, support the context of meaning for one another (Bracken and Thomas, 2010). For example, in colloquial dialogue as well as academic discourse, illness and health are often juxtaposed. This binary opposition further invokes others (e.g., deviant and normal, bad and good) that carry through in the way we discuss these concepts in the literature. Many studies of mental health and mental illness on Instagram, such as those related to eating disorders, emphasis mental illness. While prior work acknowledges recovery and associated facets of health (Chancellor et al., 2016c; Manikonda and De Choudhury, 2017), there have been few studies that fully integrate mental health and mental illness in a holistic analysis. Instead, a common convention is to categorize communities in ways that position one group in illness (e.g., 'pro-harm') and another in health. For example, prior work examining eating disorders online often discusses communities in the context of 'pro-ED' (eating disorder) and 'pro-recovery'. Even when researchers acknowledge content similarities between these two groups, the stereotypical differences are still foregrounded. I have observed these communities and types of content and view them as more entangled than classification permits. Specifically, I observed content within my Instagram corpus and from informant reports that challenge these traditional dichotomies and presents, instead, blurred boundaries between health and illness, and related concepts.

As a prime example, content related to food (i.e., food diaries) has been linked to individuals with eating disorders (Manikonda and De Choudhury, 2017; Pater et al., 2016a) but also to individuals currently navigating recovery and maintaining certain wellbeing or fitness practices (Chancellor et al., 2016c; Manikonda and De Choudhury, 2017). While some individuals in my post corpus provided additional context in relation to food, such as captions discussing recovery or restricted caloric intake, others simply wrote, *"Breakfast."* One individual, with a sparse comment of "breakfast" and no accompanying hashtags posted a picture of a bowl of sliced bananas and peaches. The bowl, tinted turquoise, but otherwise transparent, rests on a white table. Without additional context, this post may not seem related to eating disorders or recovery from eating disorders. However, other posts shared by this person offer insight. As indicated by a post shared four days following the one about breakfast, the individual mentioned feeling *"so mad"* because they were *"doing good, 600 calories a day for the past week, but I binged yesterday."* Between these two posts are not only four days, but also three other posts, all of which contain captions regarding caloric intake, although only one displays actual food. Without a deeper understanding of the individual, here provided by captions on other posts, the post about breakfast seems unrelated to mental illness and difficult, when standing alone as an individual post, for even human analysts to interpret as healthy or illness-related eating practices.

My analysis calls into question whether strict classification should be the prevailing empirical goal for research in CSCW. As illustrated by the breakfast example, mental health and mental illness on Instagram share a surplus of similarities. For example, body comparison posts are found within eating disorder and recovery content, as well as fitness accounts. (Pater et al., 2016a) notes that this "ED journey" archetype is particularly popular as a mechanism to track progress toward a goal weight. However, in my sample, I also note that this trend is visible in discussions of recovery.

In addition to thematic and compositional similarities, I observed that style often carries across posts related to mental health and mental illness. For example, while (Pater et al., 2016a) describes the high occurrence of black and white imagery in eating disorder communities, and (Manikonda and De Choudhury, 2017) comments on high contrast as a visual signature of mental illness-related content, individuals discussing their experience with recovery, as well as people just following a certain visual aesthetic, also make use of similar visual styles. One individual in my post corpus, for example, posted a black and white image of her wrist; curving, black text creating a tattoo of her children's initials. In the caption, she discusses being *"nearly recovered from postpartum depression."* Posts related to recovery portrayed in greyscale are not uncommon on this individual's account. Just as community boundaries are porous (Andalibi et al., 2017), so are the means of visual communication and expression.

Posts may also straddle and deconstruct the conceptual line between mental illness and health. In one post, a solid stack of five pancakes rises on a white plate. Each layer of pancake is separated by an additional layer of sliced strawberries and syrup. On top of the stack, a generous heap of whipped cream, chocolate shavings, and additional strawberry compote. In the caption, the poster discusses the concept of partial or half recovery. They write, *"If you're recovering within the safety of your eating disorder, you're not actually recovering. This is a state of half-recovery. We can look like we're recovering, even when we're counting our calories and unwilling to eat 'bad' foods. #anorexiarecovery #fightingmentalillness #foodisfuel #eatingdisorder, #mentalillness."* Supporting this concept of half-recovery, another poster shared a post picturing a large plate of penne pasta. While the image of the post, in conjunction with the hashtags (e.g., #recovery), underlie an assumption that the food will be eaten, the caption mentions otherwise, indicating that the hurdle overcome was ordering the food, rather than eating it. Further yet, profiles from multiple accounts include lines such as *"attempting recovery,"* and *"relapsing. . . trying to recover."*

These examples demonstrate how mental health and mental illness co-exist within experiences. In some instances, mental health actually seems to mediate experience with mental illness (e.g., during recovery and relapse, in particular); illustrating the interconnectedness of these experiences. Additionally, the posts above demonstrate how recovery is not a straightforward, linear process. The concepts of recovery and relapse assume directionality and binary states, moving from worse to better or better to worse. In my data, I found a richer negotiation between the experiences of health and illness, recovery and relapse. As I argue throughout this section, mental health and mental illness can be difficult to distinguish from one another. At times, the goals of distinguishing and categorizing may also be unproductive. Instead, attending to the emergence of visual patterns and content over time provides insight into cultural norms among these communities, as well as what is meaningful for people who post about experiences with mental health and mental illness online.

5.5.4. Reposting and Remix as Community Participation

Prior analyses of mental health and mental illness on Instagram call attention to the particular attributes and aesthetics of visual content, aiming to visually distinguish and cluster content based on categories of mental illness grounded in clinical interpretations of experience. Yet, little is known about the emergent norms of how people generate and share this content. Within my interview cohort, many of my participants described practices that are reminiscent of recycling and remixing, which others have studied with respect to online, creative communities (Fiesler et al., 2015) and plagiarism detection (Gruner and Naven, 2005). Understanding Instagram data as recycled and remixed has implications for how we analyze online interactions. In particular, these practices imply an information flow from one person or community to the next. This means we cannot simply treat an individual post as belonging to a single individual and we cannot assume the content, particularly visual content, originated with that poster. The notions of recycling and remix are also important because they are a form of community participation, in which members abide by certain social norms (e.g., content attribution, expected post novelty).

While reposting images from Instagram and other platforms was a common practice within my dataset, carrying over associated hashtags and captions was not. This may be a limitation of platform sharing features or indicative of the online sites that individuals tend to use for content appropriation (e.g., Google Images, Reddit). Additionally, in the context of the images selected, many of my participants preferred for their reposted content to standalone (without caption or hashtags) and speak for itself, potentially due to the text embedded in many of these reposted pictures.

In my Instagram post corpus, however, I observed ways that reposted content can acknowledge an earlier source (e.g., tagging another account in a post, linking to the individual's account in the caption), though not all reposted content is clearly attributed as such. For example, participants in my interview sample did not tend to attribute visual content to a source. Some, such as P12, went to lengths to avoid attribution, lest they repost from an account they did not know personally or an intimate moment with which a friend did not want to be publicly associated. Several accounts I observed during my digital ethnography, particularly those posting thinspiration (Ghaznavi and Taylor, 2015; Lewis and Arbuthnott, 2012), did not attribute sources to reposted content. They did, however, frequently include a blanket statement in their profile to indicate that pictures were not theirs or of them (e.g., *"no pics are mine"*, *"not mine"*, *"not me"*). One account in my Instagram dataset reposts images of thin, feminine individuals, frequently in bathing suits and lingerie. While some pictures include faces, many posts feature tight shots on specific body parts, such as abdomens and thighs. The only text associated with each post is a block of hashtags (e.g., #skinny, #pretty, #anorexiarecovery, #fat, #hate, #like4like) that differs little from post to post, and does little to claim or deny personal association with or ownership of any of the visual content. However, similar to other accounts that I saw on Instagram during my digital ethnography, this individual's profile mentions that *"None of these pictures are mine."* The phrase comes just before request to *"just block"* rather than report, and a goal weight of 46 kilograms (approximately 100 pounds). I noticed similar language on other accounts, as well; potentially indicating how these statements may be used by members of a community to negotiate how a relationship with mental illness is claimed.

Cultural artifacts, such as advertisements and internet memes, also influence how people make their experiences visible, creating a form of remixing and appropriation of visual content, aesthetics, and practices. Several of my participants described creating their own memes, though most tended to strictly repost this type of content, if they found it relevant. Some accounts collected with my Instagram corpus, however, subsisted entirely on memes. One individual within this post corpus created many of the memes displayed on their page, as indicated by several captions, including, *"I spent a few hours making mental illness memes. Here you go,"* and *"This is one of the new memes I created today."* As another practice, posts associated with eating disorders (e.g., living with and recovery from) in my Instagram post corpus often include pictures showcasing side-by-side comparisons of an individual's body. Typically, these posts are done in a before and after format, where the left side of the picture incorporates a picture of a previous body, and the right side of the picture is demonstrative of the current one. Many types of advertisements, from wrinkle creams to diet plans, employ this side by side comparison. However, embedded within this structural trope in the context of bodies is not only an element of time and progress, but a value judgement, where the after portion of the image is framed as desirable. This type of format is not uncommon for individuals with eating disorders, as well as those recovering.

P01 relied on a similar facet of remix when creating one of his Instagram Stories. P01 created, *"one of those lifeline, I'm not sure, medical images,"* which resembled a heart-rate visualization to convey his experience across a particular morning (i.e., *"But, it was inspired because I was actually having a really great morning, and, midday, I got a text from my mom that my uncle passed away"*). Remix does not necessarily involve digital or physical materials. It may, instead, involve concepts, compositions, and styles (Feuston and Lindley, 2018). For example, participants discussed using Instagram filters to imitate styles they felt could express mental health and illness. P13 used the

Instagram application to put a blue filter over a picture of herself because she associated blue with depression, and P12 talked about how the black background of one repost was stylistically reminiscent of struggle. Based on my analysis, I cautiously write that certain visual signatures (Manikonda and De Choudhury, 2017) do seem to exist within mental health and mental illness visual content. I am hesitant to include an emphasis on aesthetics, however, because mental illness is not a style (Tatum, 2013). In considering aesthetics, we also need to be aware of how style and content evolve over time through community practice, draw on other cultural artifacts, and may or may not implicate a relationship to mental health or illness. These later aspects are those we should aim to center in future work.

5.6. Discussion

Interpretation of mental health and mental illness is shaped by lived experience and intertwining influences of traditional, current, and cross-cultural artifacts, media, and technology. Without reducing the importance of understanding how individuals express mental health and mental illness, I address and reflect on the theoretical underpinnings implicit in interpretation of these expressions. How we interpret content affects the research enterprise. Though themes and styles may cross domains (e.g., religion, art, medicine, media), they do not necessarily carry with them consistent meaning. A visual representation of an apple might reference certain cultural symbolism, such as ‘the fall of man’ or the apple of discord, or simply record a type of fruit. Below, I discuss insights from my analytic approach to inform mental health and mental illness research conducted on social media, as well as ethical considerations for this avenue of scholarship.

5.6.1. Beyond The Coded Gaze: Insights for Research

Berger, in *Ways of Seeing*, states: “The way we see things is affected by what we know or what we believe” (Berger, 2008). For this reason, my approach to examining expression of mental health and mental illness on Instagram frames social media content as a site “at which many gazes or viewpoints intersect” (Lutz and Collins, 1991). This intersection of gazes complicates interpretation of expression. There are many ways to interpret the same post. However, this diversity of interpretation is not sufficiently addressed in the current literature. Mental health and mental illness scholars often study Instagram and similar platforms from a biomedical perspective, which assumes objectivity and erases the researcher as an important conduit through which analysis occurs. A similar objectivity is applied through content analysis and machine learning algorithms, which emphasize representative domains of content (i.e., classification and frequencies of values) over the potential multiplicity of content meaning (Van Leeuwen, 2001; Van Leeuwen and Jewitt, 2001).

To revisit the approach to mental health and mental illness within CSCW, I return to the coded gaze. I describe how the coded gaze takes a particular approach to interpreting mental health and mental illness-related content, and how this approach influences technological development and practices, such as those related to content moderation (Chancellor et al., 2016d) and health interventions (Chancellor et al., 2016a; De Choudhury et al., 2016). In my discussion of the coded gaze, I incorporate research from media and cultural studies, which recognizes “that ‘looking’ is always embodied and undertaken by someone with an identity. In this sense, there is no neutral looking” (Lister and Wells, 2001). In addition to reflecting researcher and dataset biases, the code gaze, as an analytic tool, carries certain systematic expectations, analytic commitments, and supportive practices (i.e., data collection). Here, I call attention to three assumptions underlying the coded gaze and ways to address them.

5.6.1.1. *The Assumption of Generalizability.*

Methods of data collection impact how we come to interpret data. Though these methods are not necessarily analytic, they support the analytic practices that follow. The prevalence of certain collective practices advances the agenda of machine learning techniques and content analysis through unquestioned systematic prioritization of large, and largely independent, quantities of data (i.e., millions of Instagram posts) and strictly observable attitudes and behaviors. A common practice amongst CSCW and HCI researchers, including the present study, involves relying on keywords (e.g., #depression, #secretsociety, #eatingdisorder) to collect data. Collecting data through this method assumes that hashtags, more so than any other element of a post, account, or person, index a relationship to mental health and mental illness. My findings demonstrate that expression of mental health and mental illness is a more nuanced practice. An individual's relationship with mental health and illness can be tenuous, ill-defined, and under active negotiation. Sampling on certain hashtags or keywords excludes individuals, such as many of my interview participants, who do not engage with certain modes of expression or signaling practices on Instagram, and inadvertently inscribes a non-relationship with mental health and mental illness to their online content. We must be mindful and accountable to the ways our practices of data collection imprint relationships and impact post inclusion. Practices of data collection must consider the unit of analysis; most standardly, the post. Though this includes multiple modes of expression, it assumes independence from other data points, including those produced by the same account, and disregards an individual's grander context, such as profile information and surrounding posts.

Assumptions during data collection carry through to data analysis. The coded gaze frequently examines posts against a checklist of relevant codes or features. This practice neglects the idiosyncrasies of a post. Instead, a limited number of elements are absorbed into representative domains of content, which implicates a relationship between frequency of observable behavior and importance of meaning and dehumanizes individuals in favor of population-level descriptions. These population-level descriptions are then applied to entire communities of individuals online in the form of technological 'solutions', including content moderation. Data collection techniques that shift the unit of analysis from the post to the account, such as the digital ethnography employed in this work, or to the community, as well as methods that prioritize a multimodal approach during data collection, provide complementary perspectives to existing literature that could support new understanding of this topic and avenues for technological development.

5.6.1.2. *The Assumption of an Objective Record.*

At present, the theoretical assumptions underlying the coded gaze do not support understanding how the meaning of recorded content may change over time or differ depending on the timeline of posting or viewing. Additionally, individuals' relationships with mental health and mental illness may not be observable through recorded, public activities, such as posting, commenting, and liking. Though analytic methods drawing from the coded gaze might recognize content related to self-harm, such as pictures and associated hashtags, machine learning and content analysis are less helpful in examining meaning and significance, particularly given how these concepts shift over time and across social and cultural contexts. For example, individuals may post previous instances of self-harm or repost content related to the practice to manage their own impulses (e.g., as a coping mechanism), rather than to indicate current practices or support for harmful behaviors. Additionally, individuals employ multiple and 'secret' accounts to negotiate how they express and claim (e.g., signal, make visible) their relationship with mental health and illness (Andalibi et al., 2017). Whether people operate accounts in parallel or create new ones following the ban of an old one, use of these accounts merits further consideration.

The coded gaze predominantly relies on content as it has been recorded, rather than how it may be interpreted in a variety of ways. Given the frequent discussion of health interventions as a desirable outcome of research involving social media and mental illness, there is an implicit assumption that social media content is an objective record of truth. However, a long history examining the interpretability of written texts and, in particular, images demonstrates otherwise (e.g., (Goodwin, 1995; Harper, 2012; Van Leeuwen and Jewitt, 2001)). In my findings, I describe how posts can seem to record one thing (i.e., inspiration, motivation), but actually represent a different concept entirely (i.e., loneliness, challenges to mental health, experiences with mental illness). The coded gaze disregards that elements of a post may not represent what they seem to record and, as such, does not support interpretability of content or acknowledgement of intersecting viewpoints. By considering multiple viewpoints or gazes during data interpretation, such as those provided by posters (i.e., via interviews, photo elicitation), the assumptions implicit in the coded gaze may become less central to how data in this space are analyzed and used to inform technology and platform policies.

5.6.1.3. *The Assumption of Classification.*

The coded gaze assumes that data is classifiable, and that classification is a desirable empirical outcome, within a determined set of researcher- or machine-created labels. Though beneficial in certain contexts, classification reduces the holistic nature of the data being analyzed, and creates bounded, potentially artificial ways of seeing content, behavior, and practices. In this paper, I describe ‘mental health and mental illness’ to equally acknowledge and emphasize the complexity and multiplicity of human experience. Individuals move between mental health, mental illness, recovery, and relapse in parallel and with fluidity. My findings demonstrate that individuals practicing self-care, raising awareness of mental health and mental illness, trying to recover, and living with mental illness share similar visual content, composition, and style, as well as captions and hashtags. Black and white filters, employed by individuals posting about self-harm, are also used by those discussing recovery from post-partum depression.

Classification is applied to certain facets of content, such as ‘deviant behavior’, account trajectories of mental illness (Chancellor et al., 2016a; Tsugawa et al., 2015), likelihood of recovery (Chancellor et al., 2016c), and practices of disclosure and support (Andalibi et al., 2017). Outcomes of this research have many valuable applications.

However, as researchers, we must consider the ways that we exploit public content to demonstrate that certain people, such as those sharing mental illness-related content online, are ill or somehow deviant from the norm (Lazar et al., 2018; Parr, 2006). Classification of content may not align with how an individual conceptualizes and experiences their own identity. Individuals sharing eating disorder content online do not necessarily identify as pro-eating disorder. While group distinctions support our ability, as a community of researchers, to discuss mental health and mental illness, they implicitly assume categorical associations that may be inaccurate or harmful when perpetuating stereotypes. We should, instead, step back a moment to understand how individuals negotiate expressing experiences, signaling experiences to others, and claiming relationships with mental health and mental illness.

5.6.2. Ethics and Responsibility

As other researchers have addressed, there are ethical concerns involving the use of public social media data as well as interviewing individuals who may have mental illnesses (Andalibi et al., 2017; Chancellor et al., 2017; O'Leary et al., 2017). I reflect on these issues here as a way of continuing this discussion within CSCW.

Reviewing content previously shared to Instagram can be evocative and emotional. However, this content formed an essential part of my methodology for informants and for researchers. As a researcher, I was mindful of potential feelings of anxiety or discomfort, or triggering experiences, when I asked informants to review their Instagram account and recent posts. Although participants tended to have positive discourse, some addressed feelings of anxiety and depression, as well as hesitancy to share posts or account information. In addition to participant comfort and wellbeing, my co-author and I attended to how the research team managed the emotional aspects of this research. Reflexivity is central to an interpretivist approach, and through this process, researchers were attuned to their own emotional state (Moncur, 2013; Wolters et al., 2017). As part of a research team, I also discussed, reflected on, and memo-ed regarding how my experiences and attitudes toward mental health and mental illness influenced my interpretation of the data. For example, when one researcher sees emotional pain in a particular post or account, another may, instead, see objectified bodies (Fredrickson and Roberts, 1997). The reflexive practice of qualitative research resonates with scholarship in critical algorithm studies (Diaz et al., 2018; Friedman et al., 1996), in which positionality and interpretation are central to questions of how algorithms are designed to interpret human experience.

As a research community, we must also consider the ways in which we represent the individuals who contribute to our data analysis. We should question the ethics involved in using personal content to draw conclusions about an individual's mental health or mental illness without their expressed knowledge or consent. By classifying individuals, we may apply incorrect and harmful labels that reinforce negative community stereotypes and introduce the potential for new risks. Drawing on the history of outsider art, which includes many artists with mental illness, we can observe how sensationalizing artifacts because the creator's experience lies outside of the mainstream can ultimately reinforce stigma and negative views of mental illness. However, the tension between communicating research and avoiding sensationalizing these individuals' experiences begs the question: what does it mean for a paper reliant on visual content and associated practices to not include images? This is an ethical and methodological question that, as a community, we have yet to resolve or standardize in terms of practice.

Finally, it is worth noting that the findings and recommendations detailed in research on mental health and mental illness may be put into practice by other academics, corporations, and professionals. In light of abuse and misuse of machine learning and artificial intelligence applications (e.g., facial recognition (Bowyer, 2004; Brey, 2004), criminal sentencing (Angwin et al., 2016; Crawford and Calo, 2016)), particularly against communities already facing systematic oppression, we must be aware of the potential for harm posed by similar applications in this domain. Though the issue of classification has implications for content moderation and other types of intervention (Chancellor et al., 2016b), it could also be employed by law enforcement agencies to identify and monitor certain individuals due to the misinformed belief that individuals with mental illness have a propensity for violence (Link et al., 1999). Classification, in practice, only interprets a narrow slice of data and, as I describe in my findings, does not always have a clear moniker. The relationship between mental health and mental illness is thin, porous, and flexible, and not fully observable through behaviors on social media. Additionally, there are thematic and stylistic, visual and textual similarities between mental health and mental illness. By moderating content or intervening on individual practices through a one size fits all approach, we may cause harm to certain individuals and ultimately reduce potential for self-expression, leading to a chilling effect (Chancellor et al., 2016b). Current practices around content moderation seem to engender system workarounds and frustration (Chancellor et al., 2016d), rather than systemic change in behaviors. As researchers, we must rigorously consider the tradeoffs between censorship and expression, as well as the benefits and potential consequences of interpreting content through the coded gaze. In the next chapter, I extend this thinking through examining the impact of platform content moderation – which involves classification by human and computational actors – on people with eating disorders. In part, what I detail in the following chapter illustrates how moderation of certain related to mental illness is oversimplified, inaccurate, and harmful.

5.7. Conclusion

Expression of mental health and illness is an active negotiation of content production and circulation. Individuals on Instagram make their experiences visible through sharing original, reposted, and remixed content without necessarily making claim to a particular relationship with mental health or illness. The ways that individuals choose to make their experiences salient is influenced by platform practices and social norms, as well as broader cultural representations of mental health and illness. HCI and CSCW scholars frequently employ the coded gaze to analyze online content related to mental health and illness. Though beneficial for a variety of reasons, such as the ability to recognize depression and predict suicidal ideation, application of the coded gaze assumes generalizable findings that de-emphasize individual experience; a static, objective record of content that does not explore alternative interpretations; and an implicit goal of classification. I argue that the coded gaze restricts opportunities for understanding the intricacies of how individuals claim, signal, and make visible experiences with mental health and mental illness. By diversifying methodological approaches in this space, we can open up new avenues for research and support complimentary ways of understanding how individuals express mental health and illness online.

CHAPTER 6

Conformity of Eating Disorders through Content Moderation

In Chapter 5, I illustrated how algorithmic bias in computational and sociotechnical systems – the coded gaze – impacts the ways researchers approach, study, interpret, and understand content related to mental illness and mental health online. For example, I showed how posts on Instagram without hashtags are often excluded; meaning, that people and representations of mental illness online sans hashtags do not exist in many datasets and are, therefore, not accounted for in models related to the detection, classification, or prediction of mental illness content online. By expanding data collection beyond hashtags, such as I do in my work, I've been able to interact with a broader diversity of content related to mental illness. Part of this diversity involves a topic that has been present, albeit backgrounded, in the past two chapters: content moderation.

For many social media to work as they do, content moderation is a necessary part of platform organization and governance. However, processes of content moderation involve classification, which, as I've discussed, can be biased against certain types of people and posting practices with respect to mental illness. In Chapter 4, I began to discuss processes of content moderation, including the provision of support resources, as a 'counterproductive' force. Here, I continue this discussion by bringing another interview study into my ongoing digital ethnography. In this chapter, I address the impact of content moderation on people with eating disorders.

My work in this chapter makes three contributions: 1) I provide, through my empirical work, a detailed account of how members of a marginalized group experience online content moderation; 2) I show how content moderation, as a process of social control, works to enact a specific type of conformity as related to body image and management; and 3) I discuss what it means to design for a multiplicity of eating disorders – and experiences with mental illness – online.

This chapter has been published with Alex S. Taylor and Anne Marie Piper in the Proceedings of the ACM on Human-Computer Interaction.

> Jessica L. Feuston, Alex S. Taylor, and Anne Marie Piper. "Conformity of Eating Disorders through Content Moderation." Proceedings of the ACM on Human-Computer Interaction4.CSCW1 (2020): 1-28.

6.1. Introduction

Participation of diverse groups on social media platforms, such as Facebook, Instagram, and Reddit, occupies a large contingent of work in Computer-Supported Cooperative Work (CSCW). Research addresses the proliferation of networks and communities across these platforms as well as the content of discussions and practices of sharing (Andalibi, 2019; Bazarova et al., 2015; Fiesler et al., 2017; Jeon et al., 2016; Kumar and Schoenebeck, 2015; Olteanu et al., 2015). Emergent within this literature is an emphasis on understanding the practice of content moderation and associated experiences. As Tarleton Gillespie writes, content moderation is central to what online platforms do (Gillespie, 2018). Moderation of participation and discussion has been studied within general contexts, such as Reddit (Jhaver et al., 2019a,c), as well as specific ones, including examination of hate speech and online harassment (Chandrasekharan et al., 2017; Phadke et al., 2018; Vashistha et al., 2019; Vitak et al., 2017). Much discussion in this domain involves identifying specific topics of conversations (Chang and Danescu-Niculescu-Mizil, 2019; Vitak et al., 2017), determining which topics should be encouraged or removed (Croeser, 2016; Franks; Goldman, 2010; Seering et al., 2017), and understanding interactions between manual and automated forms of regulation (Jhaver et al., 2019b; Seering et al., 2019). In the CSCW- and Human-Computer Interaction (HCI)-related literature, as well as publicity from large tech firms (Plumb, 2019), the moderation of individuals and groups has been treated as a necessity.

In this chapter, I aim to approach the topic of platform content moderation from another perspective. Platform moderation involves various configurations of human and algorithmic activity (e.g., flagging content, content removal). Here, I closely attend to how moderation happens and what the consequences of moderation are for members of marginalized groups expressing non-dominant narratives. I argue that the relations between the social and technical (i.e., the *sociotechnical*) afforded on social media platforms exert an active force, producing and reproducing a conformity to particular norms and values. Conformity, therefore, is not only established through the formally documented rules in a platform's standards and guidelines (Fiesler et al., 2018; Gillespie, 2018; Pater et al., 2016b). Similarly, it is not solely dependent on the technological features and underlying structures of a platform. Rather, it emerges as a tacit set of norms and values through the interplay between a platform's features and how a group's members come to actively moderate talk and interaction. My interest, then, is in how the distinctive features of social media platforms interweave with the social practices of moderation and how such sociotechnical relations serve to sustain and amplify certain norms and values that often exclude or marginalize non-dominant narratives.

I examine these social and technical practices of content moderation on social media platforms as they relate to individuals with eating disorders. The work that follows is grounded in two years of digital ethnography, most recently focusing on the experiences of individuals with eating disorders across an ecosystem of social media platforms. In addition to analyzing online content, I interviewed 20 individuals with eating disorders who reported having content removed from social media platforms, including Facebook, Instagram, Reddit, Tumblr, and Twitter. Through a constructivist grounded theory approach to analysis (Charmaz, 2014), I show that the pressures of moderation can have damaging consequences, especially for marginalized groups. These consequences include loss, labor, and displacement, as well as wider processes that reinforce ideas around which versions of mental illness are sanctioned as 'normal' and 'acceptable' in online spaces. I will show that resistances arise in response to these many consequences and to the effects of being marginalized. Individuals and groups find ways to work around platform processes through the creation of different user accounts and establishment of splinter communities forged through ingroup, grassroots processes of community moderation.

I make three primary contributions. The first is a detailed account of how members of a marginalized group—individuals with eating disorders—experience content moderation, extending prior work in this space (Chancellor et al., 2016d; Jhaver et al., 2019a,c; Myers West, 2018). Although content moderation is typically conceptualized as necessary for the greater good of online communities (e.g., preventing harassment, protecting individuals from graphic or triggering content), its potential harms are not well-understood or documented. My analysis reveals the ways in which content moderation has consequences, sometimes severe, for people with eating disorders. These consequences include loss of personal content (e.g., used for self-reflection) and community support, as well as the creation of additional restorative work for people who have been subject to moderation.

Second, I turn to conformity as a way of understanding the broader social and technical practices of content moderation. In this chapter, I view these practices as mechanisms of social influence and control. Conformity, in this context, is simultaneously a particular configuration of norms and values and an active process in which people with differing norms and values are pressured to assimilate or comply (Cialdini and Goldstein, 2004). Conformity is central to many social processes. My aim is not to contend the importance of conformity or to call for its eradication. Rather, I use this chapter as a space to question and call attention to a particular practice of conformity—content moderation as it relates to eating disorders online—that has become pervasive and taken-for-granted across many online spaces. I discuss how content moderation contributes to wider processes of conformity, set within historical and contemporary contexts, where particular versions of mental illness are legitimized and others are rejected.

Finally, as a counterpoint to conformity, I reflect on what it means to design for multiplicity in online social platforms. Drawing from Annemarie Mol's work (Mol, 2002), I discuss how eating disorders are enacted differently across various sociotechnical configurations of online spaces and actors. These differing, multiple versions of eating disorders are simultaneously performed and entangled within different platforms, communities, accounts, and people (Haraway, 2013; Swanson, 2018). Here, I use multiplicity to focus on the many different versions of body image and body management as they are performed online with respect to eating disorders and, more broadly, mental illness. In this context, multiplicity helps us attend to the range of norms and practices within eating disorders online and the restrictive impact of platform content moderation. I articulate directions for future work aimed at creating more diverse and equitable online spaces.

6.2. Background

Eating disorders – and even the concept of *disordered* eating – are relatively new ways of constructing meaning around certain practices related to food consumption and body management, including practices of self-starvation, bingeing, purging (e.g., for example, vomiting, exercising, and laxatives), and eating non-foods (e.g., iron nails, human body fluids). As such, in providing a historical view, brief and Westernized as this is, I do not conceptualize these practices in their past contexts as disordered. Rather, I talk about them as non-normative eating practices or practices of food consumption to illustrate how, in varying degrees, they were not necessarily common, mainstream, or entirely socially sanctioned within their particular historical era. Take, for example, the overeating and vomiting said to have been practiced by wealthy Romans (Miller and Pumariega, 2001) (though they did not have special rooms for vomiting (Pappas, 2016)). While these behaviors were not socially approved, they were certainly not viewed as disordered or through any other medicalized frames. Similarly, throughout Europe's Middle Ages there are records of individuals, often sanctified and holy women, practicing, most frequently, fasting as a form of protest or religious significance[1](Bemporad, 1997; Bynum, 1988; Harris, 2014).

While researchers have reframed historical practices through the lens of contemporary disorders (Forcen and Forcen, 2015; Harris, 2014), 'anorexia nervosa' wasn't actually coined until 1873 by Sir William Gull (Gull, 1874). In 1952, it was the first and only eating disorder included in the first version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-I) (Dell'Osso et al., 2016). Other disorders, including bulimia nervosa and binge eating disorder, don't have the same medical-historical legacy. Bulimia nervosa, for example, was first written about with respect to diagnostic terminology in 1979 by Gerald F. M. Russell (Russell, 1997). However, it wasn't significantly recognized until 1987, when it was legitimized as a diagnosis through its inclusion in the DSM-III-R. Binge eating disorder, conversely, wasn't included in this manual until the DSM-V, published in 2013 (American Psychiatric Association, 2013), though it was first identified in 1959 as Night Eating Syndrome by Albert J. Stunkard (Brewerton, 1997; Stunkard, 1959). Inclusion in the DSM is a powerful form of validation for specific clinical interpretations of experience, particularly with respect to

---

[1] That much of this history discusses European women and restriction illustrates a preoccupation with certain bodies and ways of being embodied, practices of food consumption, and subject positions. While consequences of this narrow focus may impact contemporary research inquiries, such as focus and funding for particular experiences, it similarly influences my own historical review. For example, similar practices of restrictive food consumption practiced by religious or holy men are not typically included within historical perspectives on eating disorders, though self-abnegation, asceticism, and other practices could involve restriction, bingeing, and eating non-food items (Diamond, 2004).

insurance coverage and the generation of treatment plans. However, clinical criteria often change (Goyal et al., 2012) and, with respect to the exclusivity of the DSM, there are psychiatrized versions of eating published outside of this manual; including, for example, orthorexia (Arguedas, 2020; Dunn and Bratman, 2016). I mention these medical developments not because I explicitly use them in my own work, but because they help situate and contextualize this chapter.

Though my intent is to understand eating disorders as experienced and described by people living with them – undiagnosed or otherwise – there is no doubt that medical interpretations of experience bleed into colloquial, everyday forms and practices of narration and identity work (Foucault, 1988), such as through medical publications and consultations, media discussions of certain groups of people and practices of food consumption, and social media communities and hashtags. Psychiatrized language is so pervasive that, even without a clinical diagnosis, people might self-describe, self-label, or self-diagnose their experiences as related to anorexia, bulimia, and EDNOS (Eating Disorder Not Otherwise Specified in the DSM-IV). These clinical frameworks, even when used in non-clinical ways, can help individuals make sense of their experiences. For example, all of the online spaces observed during my digital ethnography use ‘eating disorder’ or more specific eating disorders, such as anorexia, bulimia, and binge eating, to describe themselves, which is a type of boundary work that excludes individuals with shared eating practices, such as fasting, that may not be interpreted as disordered. In considering how clinical classifications influence individual narration and meaning-making, take as an example the case of orthorexia. While not established within the DSM, orthorexia has been written about in medical texts, picked up by the media, and used by individuals with certain eating practices as a way of describing what they are doing (Fixsen et al., 2020) – even though the clinical diagnosis is not widely accepted (e.g., hence current exclusion from the DSM) and has, at times, been problematized.

Modern eating disorders are, at least in the United States of America, situated within and alongside certain pervasive cultural practices, such as fad diets, that comprise what Michael Pollan has called "our national eating disorder" (Pollan, 2004). Notably, this national eating disorder entails not only practices of food consumption, but also food production, distribution, and the various meanings constructed around these particular practices (e.g., morality and food, such as notions of certain foods as clean or dirty). Though many of these practices are specific to the United States, globalization operates to create similar pressures in other nations and cultures. For example, with respect to my own digital ethnography, I have observed certain eating disorder and diet communities online, where membership is international, rather than specific to any one country or locality[2]. With respect to the United States, food production, distribution, and consumption are deeply entangled with systemic inequalities, such as those related to gender, race, and socioeconomic status. For example, food insecurity has been linked with binge eating, overeating, and other practices of disordered eating (Lydecker and Grilo, 2019; Stinson et al., 2018; Tester et al., 2016). Though my work does not currently engage with larger processes related to food production and access, it is, undoubtedly, entangled with eating disorders and the ways people participate online (e.g., through

---

[2] However, as researchers note, different cultures do have different ways of constructing meaning around food and practices related to eating (Rozin et al., 2006).

posting ‘safe foods’ and experiences related to compulsive food ordering through Uber Eats) and should set the stage for other lines of inquiry moving forward.

6.3. Related Work

To situate the current inquiry around eating disorders and content moderation, I briefly address other research related to content moderation on social media as well as, more specifically, content moderation on social media as it relates to eating disorders.

6.3.1. Content Moderation on Social Media

A large body of work within CSCW- and HCI-related literature examines content moderation in the context of social media and online communities (Blackwell et al., 2017; Chandrasekharan et al., 2017, 2018; Gillespie, 2018; Jhaver et al., 2018, 2019a,b,c; Jiang et al., 2019; Lampe and Resnick, 2004; McGillicuddy et al., 2016; Myers West, 2018; Seering et al., 2017, 2019). Practices of moderation aim to facilitate quality content, civil discussion, and, generally speaking, online spaces where individuals can engage and participate without overt fear of abuse, harassment, or accidental viewing of violent, illegal, or triggering activities (Lampe and Resnick, 2004; Lampe et al., 2007). Throughout this chapter, when I refer to content moderation, I refer to “the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse” (Grimmelmann, 2015). What I call platform moderation, others have termed commercial content moderation (Roberts, 2019). This practice of moderation involves the organized ways in which content produced by social media users is subject to surveillance, report, review, and removal (Myers West, 2018). These practices often rely on decisions passed down by dispersed groups of outsourced laborers (Gray and Suri, 2019; Roberts, 2019).

Though mechanisms behind content moderation are largely proprietary and private (i.e., a black box (Jhaver et al., 2019b)), some researchers have illuminated the underpinnings of these sociotechnical processes (Gillespie, 2018; Gray and Suri, 2019; Myers West, 2018; Roberts, 2019). Broadly, content moderation may involve automated systems, community flagging and reporting (Crawford and Gillespie, 2016), and outsourced labor (Gray and Suri, 2019; Roberts, 2016, 2019). Several social media platforms, including Reddit and Facebook (e.g., subreddits, Facebook groups), also rely on community moderators—at times, with automated systems—to manage groups of individuals with similar interests, as well as transient visitors (Jhaver et al., 2019b; Jiang et al., 2019; Kiene et al., 2019; McGillicuddy et al., 2016; Seering et al., 2019). I distinguish this instance of moderation, in which moderators and other members of communities engage in shaping (i.e., moderating) particular forms of participation online, from platform moderation. However, as I argue, practices of platform and ingroup community moderation are entangled.

Given the pervasiveness of content moderation, a growing area of interest involves understanding the experience of being moderated (Jhaver et al., 2018, 2019a,c; Myers West, 2018). This research thread speaks to the frustration and, at times, confusion of having content removed. Though marginalized communities and groups of people are not highlighted currently in this body of work, researchers have suggested that content moderation may have more detrimental effects on their members (Jhaver et al., 2018; Myers West, 2018). The present study helps bridge this gap in the literature by engaging with a particular marginalized group (i.e., individuals with eating disorders) through digital ethnography, including online observation and interviews. In addition to demonstrating the harms of content moderation in this context, I animate its role in constituting eating disorders and, as I detail in the discussion, illness narratives online.

6.3.2. Moderating Eating Disorders Online

Researchers have also studied content moderation as it relates to eating disorders. This work typically engages with 'deviant' (i.e., rule-breaking) content from pro-eating disorder (pro-ED) communities. Research in this domain has used machine learning techniques to characterize types of content removed (Chancellor et al., 2016b, 2017) and behavioral responses to moderation, including the ways that individuals use platform features, such as hashtags, to circumvent banned content (Chancellor et al., 2016d; Gerrard, 2018). Findings from these works provide valuable insight into how the practices of platform moderation (i.e., particularly the banning of hashtags) amplify existing challenges to moderation and may inadvertently overlook others. For example, Stevie Chancellor and colleagues (Chancellor et al., 2016d) found that attempts to moderate certain types of eating disorder content through hashtag bans resulted in a broader diversity and lexical variation of hashtags. The increased lexical variation of eating disorder hashtags resulted in additional challenges to moderation conducted via hashtags. Ysabel Gerrard (Gerrard, 2018), similarly, detailed limitations to practices of hashtag-based moderation, including the ways in which recommender systems can actively circulate pro-ED content. Due to these pitfalls of platform moderation, researchers note that alternatives are necessary (Chancellor et al., 2016d; Feuston and Piper, 2019; Gerrard, 2018). In this chapter, I extend these prior works through an empirical study of the experience of content moderation and a subsequent discussion detailing new avenues for design.

Content moderation is not, of course, the lone interest for researchers examining eating disorders online. Prior works detail a large and diverse spectrum of inquiries, including characterizations of content (Borzekowski et al., 2010; Cavazos-Rehg et al., 2019; De Choudhury, 2015; Ging and Garvey, 2018; Juarascio et al., 2010; Pater et al., 2016a, 2019b; Wang et al., 2017b), information-seeking behaviors (Bowler et al., 2013; Fleming-May and Miller, 2010; Lewis and Arbuthnott, 2012; Oh et al., 2013), recovery likelihood (Chancellor et al., 2016c), and ethical concerns, including those related to censorship (Shade, 2003). Across this body of research, we see a commitment to supporting people with eating disorders and understanding the complexities of eating disorders in digitally-mediated spaces. Speaking to this complexity, Elizabeth V. Eikey, across a number of collaborations (Eikey and Booth, 2017; Eikey and Reddy, 2017; Eikey et al., 2017), describes how technologies, including social media platforms and weight loss applications, can be simultaneously beneficial and negative for individuals with eating disorders. Similarly, Pamara F. Chang and Natalya N. Bazarova's examination of disclosure-response sequences on Pro-Ana Nation, an online forum, demonstrates how community-provided support within pro-anorexia spaces can be detrimental to health (Chang and Bazarova, 2016). These examples highlight tensions within online spaces that individuals with eating disorders frequent. Specifically, as these spaces are not inherently or wholly positive or negative, benefits and consequences are entangled in the ways that people use them.

Given the complexities of technology use in this domain, computer-mediated support for individuals with eating disorders presents a challenging area for research. With respect to research focusing on pro-ED content and communities, recommendations tend to settle within a narrow window of approaches. These approaches may include novel forms of moderation (Chancellor et al., 2017; De Choudhury, 2015), such as automated systems to assist human moderators, and health interventions (Chancellor et al., 2016b). These design recommendations may benefit a number of people. For example, they aim to reduce the prevalence of triggering content online, and its potential for contagion (Carlyle et al., 2018), and provide recovery support to individuals posting about certain topics. However, little is known or understood about their potential for harm. Understanding conflicting and negative effects from well-intentioned and health-minded design can support academic and industry professionals in developing equitable approaches to eating disorders online that work to mitigate unintentional harm and oppression. Additionally, in considering prior works, many studies do not engage directly (i.e., through interviews) with the communities they observe and plan to serve. First-person accounts are vital to better understand the complexity of eating disorders online. With this chapter, I build on these earlier studies with interviews and attention to the consequences of content moderation for individuals with eating disorders.

6.4. Method

A two-year digital ethnography, including online observation and interviews, grounded my understanding of the experience of content moderation for individuals with eating disorders. During this digital ethnography, I studied topics related to mental health and mental illness across several social media platforms and online communities. Most recently, I integrated my on-going digital ethnography, particularly those data collected through online observation, with 20 semi-structured interviews. In these interviews, I spoke with individuals who have or have had eating disorders and experienced content moderation online.

6.4.1. Online Observation

During the third week of November 2017, I conducted online observation on Instagram that resulted in an initial corpus of 2,102 posts once duplicates were removed. Initially, my inquiry focused on understanding multimodal expressions of mental health and illness (LeVine and Scollon, 2004b). To build this corpus, I used five hashtags (i.e., #anorexia, #anxiety, #bipolar, #depression, #mentalillness) that had been validated in previous research (Andalibi et al., 2017; Chancellor et al., 2016b,d). I collected posts four times a day by manually saving (i.e., hand-scraping) nine Top Posts and nine Most Recent posts for a given hashtag. During this preliminary data collection period, in addition to saving posts, I also spent time memoing on the content I observed. This involved copying and pasting URLs and taking screenshots of images, videos, captions, and comments to incorporate in documents. Juxtaposed with this online content, I wrote extensively during my initial interpretive work.

Following this one-week data collection period, I used my initial corpus as a starting point for continued online observation. Between November 2017 and September 2018 I collected a total of 6,223 Instagram posts (n=2,188 unique users) by tracing through accounts of individuals who had posted, commented, and liked posts in my corpus. This traversal included full accounts, posts without hashtags, and posts without any text. Though I was interested in understanding how people posted about mental health and mental illness on Instagram, I noticed a number of instances related to content removal. Specifically, I observed a number of posts in which individuals who had content removed, including content related to self-starvation and self-harm (e.g., cutting), shared about their frustrations with content removal, reporting features, account terminations, and Instagram as a platform.

In addition to Instagram, I conducted data collection on Reddit and Tumblr[3] . I did this to expand my corpus beyond a single social media platform. Broadening data collection provided a more nuanced view of the ecosystem (Burgess et al., 2019; DeVito et al., 2018) in which people with eating disorders interact, and how this ecosystem changes and is disrupted by platform moderation of content. This ecosystem view is important for a number of reasons. Namely, people who use social media or online communities often belong to a number of different online spaces, rather than just one (DeVito et al., 2018). By examining experiences across an ecosystem, rather than a

---

[3] Tumblr (www.tumblr.com) is a social media blogging platform where account holders can post multimedia blog entries, such as those including text, picture, and video. Account holders can customize the look and feel of their specific Tumblr blog or blogs, follow others, share posts created by others, and comment on posts, such as through reblogging with caption commentary. Tumblr has several other interactive features, including chat and the ability to send questions or ‘asks’ to other account holders. Individuals using Tumblr may also connect with other people or content through tagging posts with specific topics, activities, and other words. These tags, like on other social media platforms, such as Instagram and Twitter, are searchable and can be used to organize and navigate content.

specific platform, we can better understand what is common and what is extraordinary. Understanding the systemic reach of common practices, as I do here (i.e., platform content moderation), can support critical interrogation of activities or features that are concealed or taken for granted.

My online observation of Reddit began during November 2018, when Reddit issued a series of bans to communities such as r/ProED, r/ProEDMemes, and r/ProEDAdults. Following this incident, I observed how a number of banned subreddit members joined other social media and online communities. At this time, and continuing throughout the following months, I collected public posts on Reddit discussing the platform's decision to ban these subreddits. I also gathered relevant content from other online spaces, including online communities, individual blogs, and social media.

When I began interviewing for this study, several participants described content removal and account bans on Tumblr. As such, I decided to collect posts from Tumblr beginning in June 2019 to supplement my understanding of how people with eating disorders use the platform and respond to content moderation. I began my manual crawl through Tumblr using several eating disorder search terms, such as those used and validated in previous research (i.e., anorexia, proana, proED, bulimia, eating disorder) and others occurring alongside these hashtags on posts and within accounts (Pater et al., 2016a). I included posts and accounts in my analysis when I observed a mention of moderation. However, though many posts and accounts were not included in analysis, they did support my understanding of Tumblr as a platform, as well as the ways in which people with eating disorders use it.

I included an additional 208 threads started by 103 unique users from Reddit and 160 posts from 23 accounts on Tumblr in my analysis. These data were used to inform my line of questioning for interviews and supplement analysis. When presenting these data, as well as the content from Instagram and other online spaces (e.g., blogs, online communities) mentioned above, I alter the wording of posts so that they are not easily searchable or identifiable. Additionally, my online observation involves currently active and quarantined subreddits, as well as several smaller, online communities that are not housed on social media platforms. To preserve the privacy of these communities and their members, I do not name them. However, in this chapter I do refer to banned eating disorder support communities on Reddit by their names. I do this as a form of activism to raise awareness about the termination of communities that provided support for a marginalized group.

6.4.2. Interviews

I conducted semi-structured interviews with 20 adults (ages 18 – 57; M=29) with eating disorders who had content related to their disorders removed from online communities and social media platforms. Though eating disorders can impact anyone (Pater et al., 2019b), regardless of any particular facet of identity, only three participants in this study identified as male (17 female). This is not to suggest that eating disorders are more prevalent or significant for women, only that my methods of recruiting did not adequately reach out to or engage with other individuals. With respect to race, eating disorders often run the risk of being associated predominantly with white women

(Lafayett, 2017). While the majority of my participants were white (n=12), six were African-American, one was Hispanic, and one identified as multi-ethnic.

Eligibility for this study was not contingent on a diagnosis. However, barring diagnosis, participants were required to identify as having an eating disorder. I invited individuals living with and in recovery from eating disorders to participate. As such, there is a broad spread of experiences represented by my participants. For example, several participants described being in recovery, while others were relapsing at the time of the interview or had grown accustomed to living with their disorder. Many individuals in my study described specific categories of eating disorders (e.g., anorexia, binge eating disorder, bulimia, other specified feeding or eating disorder), even when they had not received a diagnosis. Others, however, identified through particular activities, such as experiences with self-starvation and binge eating. These various participant experiences are difficult to neatly categorize. Many participants self-described having multiple experiences, such as with diagnosis and with disordered eating practices. I interviewed individuals who were members of pro-ED communities, as well as individuals who were members of pro-recovery or diet communities. Many individuals were members of multiple communities. The thread connecting participants were their experiences, even those in the past, with content moderation. The content removal experienced by participants included posts, accounts, and communities.

I recruited participants from an online eating disorder support community (n=4), Reddit (n=3), Craigslist (n=12), and snowball sampling (n=1). I issued a pre-interview phone screener, where I called participants to verify their age, eating disorder status, and experience with content moderation. Interviews lasted an average of 45 minutes and were held over the phone (n=18) or in-person (n=2). During the interview, we discussed topics related to experiences with online eating disorder accounts and communities, content removal, reactions to content removal, support resources, and opportunities for platform redesign. Interviews were audio recorded and transcribed for data analysis. Participants received a $30 Amazon gift card or $30 in cash. When referencing participants throughout the chapter, I use pseudonyms.

6.4.3. Data Analysis

My approach to data analysis follows a constructivist grounded theory process, which involves theme development through iterative coding, memo writing, and constant comparison of data to developed concepts (Charmaz, 2014). The process of memoing, as it relates to online observation, involved manually saving links to spreadsheets and taking screenshots of images and text to move into digital documents (i.e., Microsoft Word, Microsoft OneNote). I then wrote memos directly juxtaposed with the online content I was observing. With respect to videos, which were rare in my dataset, I watched the video online and took a screenshot of the thumbnail, as well as anything visually relevant to the current inquiry (i.e., content moderation). I did not discretely memo around online observations and interviews. Insights, quotes, and images from these various collection methods were entangled in my memos, where they co-informed interpretations of one another and of the thematic development.

Preliminary categories included types and motivations for posting content that was eventually moderated, receipt of news (i.e., how participants came to know their content had been moderated), sensemaking around moderation, consequences of moderation, workarounds and resistance, and tensions with coexistence (i.e., how individuals navigate eating disorder communities that may include triggering content). Through my analysis, I began to understand the ways in which harm can be caused by good intentions (e.g., content moderation and support resources), as well as how individuals push back on oppressive practices and participate online in ways that support the diversity of experiences with eating disorders.

6.5. Findings

Through my analysis, I show how content moderation involves the interplay of social norms and technical features of a platform that work to silence individuals and remove support, create new labor by encouraging responses and resistance, and shape community-led practices of moderation. To set the scene for my findings, I first walk through a case with one of my participants that illustrates how platform content moderation works in this context.

Dani, now 20 years old, has participated on social media and online communities for nearly a decade. Though her personal experience with eating disorders was not the only content she shared online, it did specifically result in account bans on both Tumblr and Instagram. With respect to Tumblr, prior to the termination of her account, Dani used a number of strategies to manage her public eating disorder blog and limit unwarranted attention. For example, she avoided using features that could establish links to other content or aid in platform search and providing tips or advice to other users (i.e., *"telling people you should do this"*). Despite these strategies, Dani felt like she was *"walking on eggshells"* whenever she posted. Her sensitivity to the workings of Tumblr (e.g., its capacities for linking and connecting content) was motivated by wanting to maintain a highly personal blog detailing her own sense of self and body image, while, at the same time, wanting to escape criticism and platform moderation. Specifically, she *"didn't want people to come crucify me because I was talking about, you know, the part of eating disorders that nobody wants to see. That nobody wants to hear."*

Despite Dani's strategic use of Tumblr, her eating disorder blog attracted attention. A year into managing this blog, Dani received an "aggressive" anonymous message asking her to delete an unspecified post about body image. *"I didn't know exactly which post they were talking about,"* she said. *"[T]hat wasn't the first time I posted about me not liking the way I looked... So, for a moment I sat and stared at the [message], and I was like, 'What? Which one?'"* Rather than remove any posts, Dani sent a message back to the anonymous user, telling them to "just block" her. Shortly after, Dani's account was terminated by Tumblr. An email from Tumblr's support team notified her the eating disorder blog had been deleted for *"violating their terms"* and, though it invited appeal, Dani's efforts to receive an explanation and reinstate her content remain unanswered. Though what triggered the ban is unclear, Dani placed blame with the anonymous user who messaged her earlier in the day. However, it may have been another, or even an automated content reporting system, that was ultimately responsible. Despite being subjected to regulation, Dani resisted the ban on Tumblr by creating a new account and, ultimately, finding new online communities, including those off of social media, to join.

Even with new accounts and online spaces, Dani's experience of being banned shaped her future interactions online, including practices of participation. She explained:

> *"I'm not as talkative anymore... I just kind of lurk... I know there's still people posting about eating disorders on there, but, when I see a post from them, I immediately get nervous saying, you know, if I interact with this person...someone is going to find my account and find a reason to make me disappear."*

In Dani's case, we see how a range of sociotechnical mechanisms and practices can work together to monitor and regulate content. Specifically, we see that moderation is made possible through the tight coupling of social interactions and the underlying technical structure of a platform (i.e., how the platform makes possible specific moderation practices, such as reporting and removal). These entangled relations—the interplay between the social and technical—do not only influence what and where some individuals post, but also shape appropriate or acceptable versions of having an eating disorder online. For Dani, we see that content moderation has serious consequences, including reduced social engagement and online expression. Additionally, we see how moderation and its consequences, including the possibility of further sanction, serve to amplify Dani's sense of being subject to control and surveillance.

As this example begins to show, individuals experience a number of serious consequences following from moderation. These consequences may lead many to react against and resist platform moderation. However, as I will show, moderation is not simply an external force. It is also an interactive process that shapes how groups of individuals with diverse and varied experiences of eating disorders establish their own community-led moderation practices as part of engaging and participating within online spaces.

6.5.1. Experiencing Content Moderation as Loss

Throughout my data, and exemplified in Dani's case, I learned of many unintended consequences of moderation, including reduced online engagement and loss of community. Marie, discussing an experience with account termination, addressed how, for her, moderation *"was kind of embarrassing."* She *"felt like I was being told I was wrong. Or getting punished when I hadn't done anything. I felt like I hadn't done anything wrong and I was angry about that, as I felt it was unfair."* The initial anger and confusion associated with moderation, as Marie and others in my dataset described, have been detailed in prior research (Jhaver et al., 2019a). These—often strong—emotions are entangled with the ways that individuals learn about and make sense of the experience of moderation, which can be confusing due to the lack of transparency and consistency. Marie's comment, in addition to describing her embarrassment and anger, speaks to recent findings detailed by Shagun Jhaver and colleagues (Jhaver et al., 2019a). Notably, that many individuals who have been moderated feel that they were done so unfairly. Here, rather than focus on perceptions of fairness or emotional responses to moderation, I attend to the various losses, including personal content for reflection and community support, that moderation entails.

Loss of content is central to the experience of being moderated. Platform moderation often involves unsolicited removal of personal posts and accounts, which are maintained by and for the individual. As most participants were not in the habit of saving content to multiple locations, their content was lost entirely. By removing or deleting this personal content, platforms effectively remove certain experiences and prevent opportunities for reflection and catharsis. Andrea and Dani both equated aspects of their online content with "diary" entries. This perspective shows how online content related to eating disorders is not merely a snippet of conversation or the representation of an experience. Rather, it plays into how people think of themselves and gives shape to an archive where posts can be revisited and reflected upon. Specifically, content in aggregate becomes a resource for reflection in the short and longer term (Li et al., 2010). While access to online content, particularly content functionally similar to a diary or journal, is valuable at any point during an individual's experience with an eating disorder (Linville et al., 2012), Andrea talked about how rereading her earlier posts was beneficial during recovery. She said:

> *"I remember I used to post a lot of intrusive thoughts and then, going through recovery, I started having a lot fewer of those. And then there's a lot of elements where you're like, 'Oh, am I in a really bad place?' And then you go back and look at it and you're like, 'Oh, I'm not having 50 obsessive thoughts today about needing to weigh myself...' I can actively see how it's changed or even like at the time too, seeing how it got worse. That was really helpful to me right when I started recovery..."*

Content removal as a practice of moderation can suggest that certain experiences with mental illness are unwelcome and unworthy (Feuston and Piper, 2019). Notably, this interpretation coexists alongside the view of content removal as beneficial due to, for example, the reduction of potentially triggering imagery and text. In the instances I articulate here, moderation can feel like a loss of personal voice or silencing of experience. While many participants shared content related to living with an eating disorder, Grace discussed how posts on her Instagram account centered on *"trying to be healthy"* and *"trying to gain my weight back."* Despite this recovery context, Instagram removed a selfie that Grace shared because she looked too thin. The removal of her post from Instagram left Grace feeling sad, ashamed, and *"unworthy to be seen."* This example demonstrates how various types of eating disorder content, such as recovery imagery and thinspiration, can share similarities. These similarities speak to the difficulties of classifying mental health and illness content on social media (Feuston and Piper, 2018), as well as the ways platforms may inadvertently delegitimatize experiences while aiming to provide certain protections or support (e.g., helping people avoid triggering content).

Another form of loss that individuals experience as a result of moderation involves loss of community and social support. When platforms moderate content, they may *"[take] away a support system,"* Christy explained. Loss of community, such as through practices related to account and community bans, can lead to social isolation, particularly for individuals who *"don't have anywhere else to go,"* one former member of the now banned r/ProED wrote. As another former member described, the subreddit ban was *"extremely upsetting. So many people used this [subreddit] for help and support. We can't always find that support offline."* Social isolation due to practices of moderation can affect health. For example, Dani had a few helpful *"people [on Tumblr] that would tell me, you know, 'You're not alone. I'm here to talk,' and stuff like that."* Following the ban of her Tumblr account, Dani lost these meaningful connections, which caused her to feel *"depressed, 'cause I didn't have anyone to talk to."* In addition to depression, I observed instances in which the experience of moderation led to dangerous or unhealthy offline behaviors, including purging. A former member of r/ProED wrote, *"I was really trying to recover... I don't know what to do now. I really feel like purging everything. This is so stupid."* In attempting to remove content classed as non-normative and harmful, platforms can create a downward stream of negative consequences, including loss of social support that, at times, amplifies illness.

Content removal is not the only practice of moderation that results in loss of community. For example, on Reddit, the practice of quarantining effectively isolates certain communities and their members from the larger Reddit community. In particular, quarantine suggests that, while certain subreddits are "not prohibited," they are, nevertheless, not normative or socially sanctioned. Quarantine on Reddit is established in several ways. Take, for example, the community that Morgan moderates. At the time of the interview, the subreddit had been under quarantine for several months. Functionally, this means that visitors receive a warning screen prior to viewing the subreddit. This warning screen includes the following message:

> *"Are you sure you want to view this community? This community is quarantined. If you or someone you know is struggling with an eating disorder, there are resources that can help. Visit the National Eating Disorders Association website or contact their telephone helpline at 1-800-931-2237 for more information. Are you certain you want to continue?"*

The visitor is then presented with two options: "no thank you," a button that is emphasized through the blue of its background, or "continue," in gray. Should the visitor continue to the subreddit, a similar warning message (i.e., “This community is quarantined”) is displayed at the top of the page. In addition to these warning messages, quarantined subreddits are excluded from non-subscription feeds (e.g., r/All, r/Popular) and search results, making them more difficult to find.

Prior to the quarantine of her subreddit, Morgan described how she revived the community to the point where hundreds of people subscribed every few weeks. Following quarantine, new subscriptions to the subreddit, as well as member engagement, have slowed to a halt. Quarantine, Morgan said, *“severely affects the subscribers”* of a community. *“It also makes you not want to talk, really. It kind of feels like you’re under watch. Like, the thing you say, that’s going to be the next – that’s going to be the thing that makes you get banned.”* As this example illustrates, content moderation through quarantining can result in loss of participation and constrained expression due to its surveillant property, which ultimately works to constitute which versions of eating disorders are permitted online.

Loss of community, particularly with respect to the removal of community spaces and content from online platforms, also involves the loss of a shared archive of resources. Andrea discussed how the loss of community resources on r/ProED *“totally sucked, because it was stuff that I would go and read if I was having a hard day. Like, someone had posted what to do if you feel like you’re going to binge or what to do if you feel like you can’t eat today.”* Rather than cultivate community-provided resources, when certain content related to eating disorders is moderated, many social media platforms share support helplines—namely, the National Eating Disorder Association (NEDA) helpline. When Marie was provided the NEDA helpline following the termination of her MyFitnessPal account, an account she had used for nearly a decade, she felt “insulted.” Vehemently, she said:

> *“People will just be, like, ‘Here’s the NEDA helpline. Hail Mary full of grace. The Lord is with thee.’ Because they just don’t know what else to do. They don’t know what else to say. You just sort of start to feel, like, here’s the NEDA helpline. Now please go away... Stop having an eating disorder.”*

Helpline resources such as these can be beneficial for individuals wanting to recover or learn more about their particular experience. However, these resources may feel *“unrealistic and unfair”* to many because of how they push recovery. Upholding recovery as an ideal neglects a multitude of differing experiences and needs. For example, recovery-related resources may not support individuals who do not feel ready to recover. These resources may also inadvertently exclude people who feel “kind of stuck in that revolving door of treatment,” as Amy described. These are, for example, individuals who cycle between recovery and relapse. Important to note is that helpline resources promote a limited range of norms and values with respect to eating disorders. They ultimately operate to reproduce a particular version of body image and body management online. Provision of these resources in tandem with moderation also suggests that platforms are forcibly creating loss—by removing opportunities for reflection, spaces for expression, and online networks for support and connection—and filling that absence with a resource list and phone numbers. The compounded losses experienced by individuals with eating disorders lead many to develop strategies of resistance that aim to circumvent or push back on oppressive platform practices.

6.5.2. Responding to and Resisting Content Moderation

Given the significant and even traumatizing effects of content moderation, individuals with eating disorders respond in a variety of ways. Here, I address responses to moderation through the lens of resistance. By emphasizing individual and collective action, I acknowledge the labor performed by individuals with eating disorders. Much of this labor relates to the ways that individuals resist oppressive sociotechnical practices in order to raise awareness about and appeal decisions of moderation, rejoin online platforms by creating new accounts and communities, and engage on platforms by mediating the types of content they decide to share.

Because the removal of content through platform mechanisms often plays out in the background, individuals with eating disorders must work to raise awareness about their experiences. For example, when personal content and accounts are deleted, only the individual who posted the content or who owned the account is notified by email. The constant stream of content via personalized social media feeds, such as on Instagram, Twitter, Facebook, and Reddit, makes it difficult for other users, even those within the same communities or networks, to notice that reblogged or reposted content has been banned or that accounts have been muted or removed. To foreground practices of moderation, raise awareness, and confront other posters, particularly those implicated in practices of moderation, individuals with eating disorders may post about their experiences with moderation online. One Instagram user, for example, captioned a post, *"Why are you reporting me? Why do you want to delete my posts? It makes me feel bad. Seriously, just block me."* Dani similarly addressed her Instagram followers, via a secondary account, when her primary account, a private account, was banned:

> *"I hopped over to my second account and said, 'Hey, guys. Someone reported me' ... I made a post saying, 'Hey, guys. I got my account deleted. I don't know which one of you did it, but gee thanks. That really did - that did me a great favor.' Like, 'Thank you so much because that made my day so much better.' I was furious and I could not for the life of me figure out who it was."*

As we see here, rather than change any personal content or settings, such as post or account privacy, individuals may confront their followers and those who come across their account, holding these other social media posters accountable for outcomes of moderation and requesting they cease and desist. Similarly, many former members of r/ProED posted to other subreddits and online communities about their frustration, anger, upset, and outrage at the ban of their support community. In these examples, we see how individuals use social media, sometimes the very same ones from which they had content removed, as platforms to speak out. This suggests that individuals are highly attuned to how social media can be used for activism (Anderson et al., 2018; Halupka, 2014; Hemphill and Roback, 2014; Lee and Hsieh, 2013) and the ways in which other posters contribute to practices of platform moderation (e.g., such as via reporting posts).

Individuals use the technical features they have at hand to raise awareness and respond to what they view as unjust and unfair practices. For example, several participants used platform appeal features provided within email messages detailing content bans. However, many participants spoke to the idiosyncratic and opaque nature of appeal. Out of all of my participants, only Grace had her Instagram account restored—and on the stipulation that she remove all posts in violation of community guidelines and discontinue her prior posting practices. Other participants described appealing content moderation through forms of collective action, including community-led petitions and surveys. As a key instance, former members of r/ProED created a survey to submit to Reddit admin to *"tell them how most people found the [subreddit] to be really helpful. How it made us feel less alone, like we had people who understood. I hope they listen and unban it. If not, at least we can speak out."* Favorable outcomes to these individual and collection actions were rare, however. Appeals, though offering some mitigation to oppression, are still sanctioned and overseen by platforms. Meaning, appeal processes remain inherent within, rather than as a check and balance to, platform moderation.

Of course, platform-sanctioned appeals are not the only ways that people return to platforms following account and community termination. Sasha, for example, found it easier to create a new Instagram account, which she did "instantly" after her first one was banned. She was then faced, however, with the task of regaining followers and connecting with individuals from other communities of which she was part. For some, such as those on Tumblr, it is common to create a new account following a ban and request for others to not only follow, but to circulate new blog information via reblogging features. In a post on Tumblr, one person wrote, *"Hey, it's [former blog name]. I got deleted, but I had 800+ followers. Can you share this to help me get them back?"* On most social media platforms, following an account ban, there are few, if any, actual barriers to account creation. Platforms, after all, want new users. As such, it is relatively straightforward for individuals to use platform features in unintended ways as a form of resistance and to rejoin spaces from which they have been forcibly removed and displaced. Similarly, it is just as straightforward to join new online spaces and communities. By allowing for these practices of resistance, which, ultimately, platforms do through their technical affordances, platforms offload the labor and burden of content, account, and community recreation to the individuals themselves.

Despite practices of resistance, such as the creation of new accounts or community spaces, platform moderation still exists. This means, for some individuals, continued engagement and participation involves changing content-sharing practices and internalizing the norms that practices of platform moderation aim to establish and enforce. Andrea, for example, described hesitancy around posting on a new eating disorder support subreddit following the ban of r/ProED. *"There was definitely something I wanted to post,"* she said. *"And it was, like, I don't know, I feel like it had specific numbers or I was complaining about not being small enough... And I felt like I couldn't post it and I felt like that was frustrating."* This hesitancy and, ultimately, assimilation of platform standards through self-censorship can have, as I show, a chilling effect on behavior (Marder et al., 2016; Penney, 2017). Sasha, following the ban of an Instagram account, created a new account where she posted content that was *"still in the same arena, just not as intense."* Christy, similarly, described how she stopped posting thinspiration. She explained, *"I just save it or archive it now. So, I don't want to risk, like, getting anything banned. So, I just save, archive stuff that you can find it on, like, various eating disorder websites or on Instagram."* Internalizing platform standards to mitigate risk of moderation conceals experience in a way that is similar to how individuals with eating disorders may hide certain behaviors from friends, family, and social others (Vandereycken and

Van Humbeeck, 2008). Concealing mental illness, eating disorders included, has serious consequences (Pachankis, 2007). However, in order to exist and participate within certain online spaces, this is exactly what must be done. Notably, even though individuals assimilate to online norms, they do not necessarily change their offline behaviors. As Marie explained, *"[An account ban] isn't gonna make me stop having an eating disorder."*

6.5.3. Establishing Norms through Community-Led Moderation

An everyday part of engaging with online eating disorder spaces involves brushing elbows with a variety of individuals and content—which can challenge online participation and result in practices of ingroup community moderation. As Marie described, *"different types of people exist in the same spaces. It's muddled at times."* Among interview participants, individuals described varied diagnoses (e.g., binge eating disorder, anorexia, bulimia), both clinically provided and self-applied, and relationships with eating disorders, including those related to recovery, relapse, and living with a disorder.

Beyond diagnosis or experience, participants also engaged across a number of platforms and types of communities, including those described in prior research as pro-ED (De Choudhury, 2015; Ging and Garvey, 2018; Pater et al., 2016a). Pro-ED has a long history of negative publicity and association. However, as my data suggest, contemporary usage by community members has reconfigured pro-ED to refer to groups of individuals who support people with eating disorders (i.e., *"in favor of—or pro—people with eating disorders"*), rather than being supportive of the disorder itself. For Marie, pro-ED meant, *"I'm dealing with a disorder and I don't want help right now. And I want a place to vent about that. And it's not so much as being, like, give me tips, give me tricks on* *how to be skinny... It's more just the support."* My understanding of pro-ED is not solely grounded in interview data, but also in the types of content that come to be socially sanctioned on social media and online platforms by members of pro-ED communities (Howard and Irani, 2019). For example, content related to the difficulties of having an eating disorder, attempts to recover, and replacing certain harmful practices with less harmful, or safer, ones can exist in the same spaces alongside thinspiration, food diaries (i.e., in the context of an eating disorder and recovering from one), and body checks (e.g., progress pictures of weight loss or gain). Given this diversity of content and experiences, individuals work to establish what is 'normal' and what is not through community-led practices of moderation.

Community-led practices of moderation develop, in part, through individual reflection and action related to the ways in which diverse, heterogeneous groups of people can coexist (e.g., pro-ED communities and the broader ecosystem of eating disorders online). For example, sharing content related to the reality of living with an eating disorder may upset or unintentionally trigger others. This includes certain experiences related to recovery, where individuals may keep detailed food diaries and share successes related to enumerated weight gain. Individuals are aware of the complex relationships between content and people. As one Tumblr poster wrote:

> *"I feel bad about running an ED blog. Does anyone else ever feel that way? Like, just kind of guilty. At least a little bit. This blog is for me to vent and cope and meet other people with the same issues. But, like. I'm really nervous that how I express myself is going to mess up some other kid."*

This commentary demonstrates tensions between wanting to engage online (e.g., to vent, cope, and connect) and a deep concern regarding the potential to negatively affect others. This concern influences how people post (i.e., how people self-moderate and self-censor), as well as the ways communities self-govern.

Although informants were willing to engage in community-led moderation and self-governance, social media platforms, by the very implementation of their features, present a number of challenges to these practices. Marie, for example, explained how *"everyone uses the same, like, 12 [eating disorder] tags on Tumblr for everything. So, everything bleeds together."* While this overlap of hashtags blurs boundaries (e.g., between individuals, content, and the potential for classification), it also presents a number of risks, including the overlap of content in detrimental ways. For example, in one Tumblr post, a user wrote, *"I've recently seen a bunch of recovery tags in non-recovery spaces! Do NOT post recovery tags with thinspo!"* Awareness of the ways in which content bleeds together, as well as its potential risks, is not enough to establish or enforce norms around hashtag use. In part, this is due to colloquial usage of hashtags to broaden audience—and, with it, the potential for followers and likes—as well as the decentralized forms that groups and communities of people with eating disorders take on social media.

Nevertheless, even in these muddled, entangled online spaces, individuals can be mindful of one another. Rose, currently in recovery, talked about how she's able to safely access and participate within one of Twitter's eating disorder communities due to the ways that she, and others, make use of content warnings—labels within posts that are separate from hashtags. According to Rose, on Twitter, content warnings are *"when someone posts something, for them to actually put up a content warning on top of* *[the tweet]. So, just to say, like, eating disorder or food or weight or, you know, whatever."* In Rose's community, the use of content warnings are *"kind of an unspoken rule."* Though these warnings are not necessarily standard within or across platforms, their presence in Rose's community enable her and others to *"safely and, in a positive way, access Twitter—is if I have those warnings, so that I can scroll past, if I need to without being triggered to start doing unhealthy behaviors."* Other online communities, such as various subreddits, may also have community practices around flairing or labeling posts. Similarly, many smaller online communities for individuals with eating disorders build community spaces on traditional forums, such as those that allow for category-specific subforums (e.g., Anorexia, Recovery, Thinspiration). In these instances, content has designated spaces. As my participants shared, some online communities are successful at upholding the organization of these spaces, both in part due to an active moderation team, as well as the willingness of members to post in appropriate spaces and call attention to those who do not. Community-led approaches, when successful, *"makes you feel safe,"* Marie mentioned. Ingroup community moderation can facilitate safety and a diversity of content in ways that practices of moderation enforced by platforms do not.

Nevertheless, practices of community-led moderation interact with practices of platform moderation as members work to establish and enforce community norms. Here, I provide the example of harm reduction to illustrate differences between the ways that community-led and platform moderation interact to regulate content. Harm reduction refers to materials or resources that help individuals take care of themselves while living with an eating disorder. For example, harm reduction involves reminding individuals to hydrate during episodes of self-starvation and to not brush their teeth immediately following a purge. As Christy mentioned, *"I think harm reduction is great. I love—because I purge. And, if it weren't for harm reduction, I think I would've fucked my teeth up so much more than I have."* Harm reduction provides resources for individuals who have an eating disorder, but cannot or will not recover, to stay safe and informed. Despite benefits, harm reduction resources are treated differently across eating disorder spaces online. While some communities freely permit them, others, such as one of the active subreddits in my digital ethnography, have moderation teams dedicated to removing posts related to tips or advice and carefully overseeing content related to harm reduction. These community-led practices differ from their historical precedent, in which harm reduction was not liable for removal or modification at the discretion of a moderation team. This example illustrates an easy to miss point. Harm reduction, previously unregulated, is now subject to new, restrictive practices of community-led moderation. These new practices are grounded in past and present interactions with platform moderation. As such, underlying these new community-moderated restrictions is the ever-present awareness of surveillance and potential threat of subreddit quarantine or removal from Reddit. Even in spaces intended to be welcoming and supportive for people with eating disorders, practices of moderation (i.e., interaction between platform and community-led moderation) can enforce certain versions of body image and body management and contribute additional labor to ingroup moderators (Dosono and Semaan, 2019; Wohn, 2019).

6.6. Discussion

In the following sections, I turn to conformity as a way of understanding content moderation and the way it contributes to broad social and structural effects on marginalized groups. Conformity can help us to examine and address the pressures that individuals from marginalized groups experience online. As I describe here, conformity through content moderation is a form of social control that influences the ways people participate online. Though I am not arguing against conformity or content moderation in all of their forms and applications, I discuss how moderation *in this specific context* works to establish and enforce a particular conformity to body image and management, particularly among people with eating disorders, as well as, more broadly, people with mental illness. I discuss its consequences for members of marginalized groups. In contrast to conformity, I then discuss how platforms can design for a multiplicity of experiences online.

6.6.1. Content Moderation as Enforcing an Order of the Normal

People with eating disorders have historically been subjected to processes of conformity that aim to dictate overarching norms and values, particularly with respect to the enforcement of certain ideas regarding body image and body management. Processes of conformity are not only projected onto particular groups or communities (e.g., through guidelines, codes of conduct, or diagnostic manuals), but, in practice, come to be enacted through unfolding relations between varied actors and sociotechnical structures. My claim is that through practices of content moderation these platforms are, in effect, enforcing an *order of normal* that restricts and aims to influence the ways people can participate online. In short, content moderation enforces a particular range of norms and values.

Though describing this ‘particular range of norms and values’ is beyond the scope of the current chapter, I briefly speak to the broad diversity of bodies and body management practices online. Body image and body management practices online are often found within topics related to fitness, health and wellness, fashion, sports, and the everyday (e.g., selfies, hobbies and events, meal preparation). Content within these domains may be ‘acceptable’ for a number of reasons, many of which are assumed rather questioned. For example, within the health and fitness spheres, several diets, including intermittent fasting and one-meal-a-day (OMAD), recommend that individuals restrict their eating throughout the day. These body management practices, as well as other types of socially acceptable content (e.g., fitspiration), share a number of similarities with content that is moderated in online spaces (Boepple and Thompson, 2016; Chancellor et al., 2018; Tiggemann and Zaccardo, 2018). The acceptability of content and, therefore, the perception of content as adhering—or not—to a social norm is contingent on context. As I show here, content related to body image and management is less acceptable when it is posted within the context of life with an eating disorder.

Following from the empirical materials presented above, I argue that social media platforms play a part in the wider sociotechnical processes of conformity. Specifically, across social media platforms, conformity to versions of body image and body management is established and enforced through platform features and the capacities for interaction that are afforded through them (e.g., commenting; labeling value through ‘likes’; content promotion and demotion via features, such as ‘upvotes’ and ‘downvotes’ and algorithms that prioritize content). With respect to my findings, I recognize that it is altogether too easy to attend to specific instances within my data, such as the particular wordings of a comment, the reporting of a specific post, or the deletion of an individual's account. However, here, I shift my attention to what is happening across these cases and across an ecosystem of online spaces. This supports attending to the structural forces at play. Key to this structural framing is that individual instances of moderation, moderation that occurs within and among community groups, and distinctive interactive features of the platform must be understood together. Making sense of the platform altogether, rather than through specific cases in isolation, has been my means of understanding what content moderation across social media platforms is doing.

Consider, for example, the reporting features on social media platforms. By approaching their design and use in terms of the wider structural practices of content moderation, we get a clearer picture of how conformity and an order of normal are enforced. As we have seen, the threat of being reported regulates online behavior with respect to what people are willing to say about eating disorders and their own actions and beliefs. The power of moderation in this context is not in the reporting *per se*, but in its perpetual threat. Reminiscent of Michel Foucault's well-rehearsed reading of the panopticon (Foucault, 2012), surveillance, rather than punishment, is the primary means of control. The Panoptic qualities of a platform (Wood, 2003), or more broadly its structural configuration designed to support content moderation, controls users and, in the case of eating disorders, regulates content so that it adheres to a norm. What the structural mechanisms of content moderation serve to do, then, is actively delineate the boundaries of what is acceptable within particular online spaces. They come to constitute a 'structural machinery' that sanctions some bodies and forms of mental illness, while simultaneously casting others as other-than-normal or deviant. Classifications of wrongness and deviance are amplified by practices of moderation that target certain versions of eating disorders, at times removing them from social media platforms and, therefore, the ability to participate in constituting broader versions of mental illness online.

Consequences of conformity clearly resonate with feminist and biopolitical accounts of bodies (Braidotti, 2000; Butler, 1993; Fuchs, 2013; Malson, 1999). These accounts detail various ways in which bodies are constructed and regulated in postmodernity. In particular, I attend to the ways social media platforms exert structural forms of control on how bodies are performed (Dejmanee, 2013). Underlying these structural forms and related research recommendations (e.g., such as health interventions) is the assumption that social media users should internalize, rather than question or counter, mechanisms of social control. Current platform practices, as well as research and design suggestions that call for additional, albeit different, forms of moderation and intervention, may result in new aggregates of human and technical control that work to establish and enforce existing norms and values (Chancellor et al., 2016d; De Choudhury, 2015; Jhaver et al., 2019a,c; Myers West, 2018). This applies, even, to recent research on content moderation that suggests a shift toward an educational paradigm rather than a punitive one (Myers West, 2018). Existing norms and values constitute only certain versions of body image and body management. They neglect the broader range of bodies and practices. I argue platforms and researchers should explore alternative ways to support individuals with eating disorders, rather than impose an order of normal that further marginalizes and subjugates certain experiences.

6.6.2. Consequences of Conformity for Marginalized Groups

Specific moderation practices on any one platform might not pose a problem in isolation. However, the ways in which platforms operate in similar ways with respect to content related to eating disorders contributes to systemic discriminatory practices and displacement of individuals on the margins. By addressing the consequences of conformity, as I do here, I contribute to a developing body of work that examines how the sociotechnical machinery of platforms and algorithms (e.g., on social media and elsewhere) exclude non-normative identities and forms of expression and interaction (Bivens, 2017; Bivens and Haimson, 2016; Haimson and Hoffmann, 2016; Hamidi et al., 2018). Platform content moderation, as I describe here, operates to enforce a particular conformity to body image and body management online. These localized findings connect with broader processes of conformity, including those related to mental illness and marginalization. Here, I take a moment to discuss how conformity, as enacted by dominant groups and processes, has consequences for individuals with eating disorders and other members of marginalized groups.

Posts about body image and body management as they relate to eating disorders on social media can be interpreted as a type of illness narrative (Feuston and Piper, 2019). Illness narratives provide opportunities for individuals to work through, reflect on, and make meaning of the subjective experience of being ill or living with an illness (Frank, 1995; Kleinman, 1989). Personal narratives about mental illness share much in common with other types of illness narratives, including cancer and chronic illness (Sannon et al., 2019; Yang et al., 2019). For example, these narratives are highly idiosyncratic and describe experiences with suffering, coping, and healing. Despite these similarities, narratives of mental illness online are more likely to be subjected to moderation and removal. As I have shown, individuals may respond to practices of moderation by concealing the full extent of their eating disorder on social media. This poses a problem, particularly for the many individuals for whom social media platforms may be the only spaces in which they feel comfortable disclosing and discussing their experiences. Without those spaces, and without others elsewhere in their lives, these individuals are at risk of psychological consequences related to hiding stigmatized experiences (Pachankis, 2007).

Processes of conformity also operate to displace individuals and communities. Displacement may be likened to a form of digital gentrification (Lingel, 2019), in which marginalized groups are forcibly removed from platforms to benefit a majority. As Jessa Lingel describes, gentrification involves power and control. In particular, it involves the ways that corporations, including social media platforms, increasingly shape sociotechnical relations in digitally-mediated spaces (Lingel, 2019). Given how platforms operate, there are a number of motivations for 'gentrification' and the displacement of certain individuals and groups of people—many of which are grounded in concerns of revenue and legality. For example, displacement of individuals with eating disorders and other marginalized groups might be conducted to present a vision of an advertisement-friendly social media. Alternatively, platforms may encounter additional expenses when working to manage a particular population, such as those costs involved with the development and deployment of support resources and health interventions.

Nevertheless, displacement in the context of this study creates inequalities with respect to content production and which voices are permitted online. Displacement often occurs alongside moderation and simultaneous provisions of helpline resources. Though these resources may provide valuable and informative assets for individuals with eating disorders, as well as their family members and friends, they do not—and cannot—replace support networks and opportunities for socialization and disclosure. Further, when support resources are limited, such as to the NEDA helpline, they present a bounded interpretation of life with an eating disorder. These constraints ignore personal histories and experiences with eating disorders, including recovery, relapse, and management, demonstrating how a blanket solution (e.g., providing the same resources for everyone) may exclude many.

Finally, in considering regulation and subsequent responses, it is vital for platforms to acknowledge the labor they create for individuals with eating disorders, a group of people who face marginalization and stigma in online and offline spaces (Livingston and Boyd, 2010; Rüsch et al., 2005). As other reports have shown (Online Censorship, 2019), social media platforms can negatively impact marginalized groups (e.g., Rohingya people, Native Americans, Black Americans) through practices of content moderation. Though I have also found consequences, I additionally show how individuals can use certain platform features to resist control and moderation. These forms of resistance, which others have aligned with practices of civil disobedience (Myers West, 2018), share their spirit and their histories with other forms of social activism, including those related to Mad Pride and #MeToo (LeFrançois et al., 2013; Mendes et al., 2018). However, resistance is itself a burden. It requires individuals with marginal status to go beyond typical platform interactions just to enjoy the same access. Furthermore, 'going beyond' involves using platform features in unintended ways (e.g., immediately creating a new account after another has been banned). Rather than necessitate additional labor, particularly when it is at odds with a platform, technologists can reconsider how to design for online social spaces that are host to multiple and differing versions of experiences.

6.6.3. Supporting a Multiplicity of Eating Disorders

Important for rethinking design are cases where we see moderation and conformity operating in constructive ways. Several participants discussed practices of moderation (e.g., forum organization, content warnings) that resulted in a conformity of body image and body management tailored to respect the standards of their ingroup community. These grassroot processes of moderation and the type of conformity they enact are not a panacea. They operate in a similar way to other online platforms and social media sites (e.g., through sociotechnical relations) and can restrict versions of body image and body management. However, my findings suggest that certain community-led practices of moderation can be practiced in ways that support different, multiple experiences of eating disorders. This speaks to the complexity of moderation as it relates to conformity and highlights promising possibilities for supporting the multiplicities inherent in individuals, communities, and platforms (Haraway, 2013; Mol, 2002; Swanson, 2018).

6.6.3.1. *Accepting Multiple Versions of Eating Disorders*

As my participants discussed, many experiences with eating disorders are difficult—recovery included. Rather than casting content as non-normative or deviant (i.e., such as through its removal or by other technical configurations that set it apart), social media platforms should reconsider the ways in which marginalized experiences and illness, of any type, are addressed and moderated online. As I describe above, illness narratives can be productive ways to document, reflect, and share the experience of being ill. Yet, these personal narratives are at risk when they do not adhere to the order of normal enforced on social media platforms and online communities. Rather than constraining experiences, such as through practices of content moderation disproportionately impacting non-normative types of content, platforms should reconsider ways to rework interactions surrounding narratives of illness. To this effect, support resources could represent more than the ideal of recovery. For example, they could also include community-curated posts and articles and content related to harm reduction, which can help people living with an eating disorder stay safe by avoiding particularly risky or dangerous behaviors. Additionally, rather than reporting features or educating those who have been subject to moderation (Jhaver et al., 2019a; Myers West, 2018), platforms could provide educational resources to others (i.e., people reporting content) about the experience of living with mental illness and importance of online disclosure.

6.6.3.2. *Coexisting through Strategic Content Practices*

Designing for multiplicity also means mindfully attending to the ways that different types of people coexist with one another. Certain ways of posting about eating disorders, including experiences related to recovery, can be triggering or upsetting for others. Content removal is one common way of addressing this type of content. However, as I have shown, it is not without consequence. Instead, I turn to the practice of content warnings. As my informants described, these warnings provide an alternative way to coexist and safely access online spaces. However, content warnings are not universal—even on a single platform. Additionally, the labor of including a content warning falls to the individual who is posting or, in some instances, a small, community-led moderation team. Platforms, therefore, have a design opportunity to simultaneously support expression and disclosure, as well as safety and access. For example, Instagram has a content warning feature that blurs a 'sensitive' image until an individual makes the decision to view it. However, this type of content warning is an external force. It's applied by the platform following a report of content that "some people may find offensive or disturbing." My data suggest that individuals with eating disorders seem likely to appropriate similar features as a way of online self-preservation, community sustainment, and consideration of others. As part of reworking content warnings, we should consider how platform features could support awareness of potential audiences and draw attention to the impact of viewing certain content. For example, including numbers (e.g., calorie counts, weight) in a post can be triggering for some people. Through new and improved mechanisms for self-moderation, platforms could make available technical affordances that maintain diverse forms of expression *and* help individuals navigate content.

6.6.3.3. *Reconfiguring Power Dynamics*

Another way that platforms and online communities can move toward a more equitable and just experience online is to shift power dynamics embedded within content moderation features and practices. In particular, platforms could provide additional control and agency for individuals who have been reported. Currently, power is in the hands of other platform members, as well as computational actors and contracted laborers, who are able to report, flag, and remove content anonymously. Oftentimes, this results in the person who had content moderated not being able to face the person or system who was responsible—or even know how or what happened when content was reported or removed. This lack of transparency and accountability contributes to the marginalization that individuals with eating disorders face on a day-to-day basis. Additionally, the emotional burden and labor of restoring activity online, including finding a space to exist, is shouldered by those who may need support the most. Productive design changes may include increasing transparency with respect to moderation and its motivation (Jhaver et al., 2019c); temporarily archiving an individual's account or content during a process of deliberation; and turning moderated content over to individuals to restore their control over their personal data and its usage (e.g., such as for self-reflection).

6.7. Conclusion

Practices of content moderation are integral to what social media platforms do. However, they are far from perfect and increasingly difficult to get right. Despite good intentions, practices of moderation have consequences for individuals with eating disorders and other members of marginalized groups. These consequences include loss of personal content and community support and labor associated with practices of resistance. In this chapter, I examine the experience of content moderation and how, in particular, mechanisms of content moderation (e.g., the sociotechnical relations involved in content removal, quarantine, helpline resources) work to establish and enforce a conformity to body image and body management on social media. I argue that processes of conformity as reproduced through sociotechnical structures afforded by platforms work to exclude people with eating disorders and other non-normative identities and experiences. Additionally, my work raises the question of who or what gets to decide which experiences and narratives are acceptable for online participation. Rather than design for restrictive content moderation practices, I suggest that platforms consider supporting a diversity of eating disorder and illness experiences by designing for multiplicity.

CHAPTER 7

Discussion

Let's briefly return to Andrea – my participant from Chapters 1 and 6. Before we ended our call, I asked Andrea what, if anything, she would like to say to Reddit regarding the ban of r/ProED. She said:

> *"[I would tell them] about how harmful it is for them to just take away community without even attempting to understand the purpose. And, honestly, if they did, then it's pretty shitty that they didn't have the empathy to understand the purpose of the community. And the other thing that I think about is that people outside of the eating disorder community don't understand how few safe spaces there are for that. There isn't a lot of safe space for open dialogue about your experiences and having a space to connect about what's going on. So, for them to just either A) not attempt to understand it and remove it is shitty or B) to attempt to understand it and then not have the empathy to try to give them, give us, a space to do better or maintain. I think that is their biggest misstep."*

Andrea's response illustrates how she feels misunderstood, unsupported, and unheard. Though her experience is, perhaps, part of an extreme event (i.e., the ban of an entire community), it resonates across the small stories and everyday experiences of people with mental illness who participated in my interviews and who I observed online. Throughout my dissertation work, my aim has been to center people with mental illness and their experiences, particularly in instances where people are underrepresented and underserved by existing social media platform and research practices. This perspective is relatively absent from HCI and CSCW social media research on mental illness. As I've illustrated, however, it's vital that we make room for these stories and experiences. In foregrounding people like Andrea, my work prioritizes the experiences and recognizes the expertise of people with mental illness who use social media. In the following sections, I return to the primary contributions of my work and discuss new directions for social media research on mental illness.

7.1. Revisiting Dissertation Contributions

Here, I connect the work across my specific empirical studies to address three of my primary contributions. First, I discuss an approach to studying mental illness online involving digital ethnography, which has not been widely adopted within this domain (i.e., mental illness and mental health) in HCI and CSCW. Even as a researcher with an eating disorder, digital ethnography has been – and continues to be – a deeply immersive, informative, and insightful experience. Second, I address what it means to center people with mental illness in my program of research. Currently, in related scholarship, researchers and designers frequently apply a psychiatrized lens to framing, method, and analysis. In this body of work, other forms of expertise (e.g., researcher expertise, clinical expertise) are privileged over the expertise of lived experience. I argue for centering lived experience and expertise to develop new research trajectories, such as those examining the needs and concerns of people with mental illness who participate online (i.e., from the perspectives of those individuals). Finally, I describe the importance of context in understanding the diversity of mental illness representations and experiences online. While much of the current literature emphasizes certain linguistic and visual cues and styles, my research illustrates how many people with mental illness do not engage with these practices of expression and communication. I argue for supporting and understanding the diversity of mental illness content and communities online by maintaining the context of everyday life and other online practices.

7.1.1. Digital Ethnography

Within the HCI and CSCW fields, my dissertation represents one of the first, if not the only, digital ethnographies of how people with mental illness participate online. My ethnographic practice is an everyday, or nearly every day, engagement. I have spent, at the very least, hundreds of hours online, digitally surrounded by people with mental illness. Daily, I aim to spend two hours – if not more – online in the context of this work and research approach. At the start of my ethnographic practice, I tried to collect all of the posts I observed. For example, I saved the links of Instagram posts to spreadsheets and, at times, Word documents. The goal was to literally account for everything I had seen. However, much of my recent ethnographic practice does not involve a quantified collection of posts. Instead, I prioritize time spent participating and memoing. Though other qualitative approaches are similar to certain ethnographic methods, the ongoing commitment (e.g., time, multiple locations), practices of participation, and reflexivity of my ethnography make this approach distinct. Meaning, ethnography, as I've practiced it, compliments existing research practices, but, importantly, also provides new and important insights.

Though not all ethnographic practices span multiple years, many (not all) do involve a lengthy time commitment to participating within a specific group, community, or research population. This approach, particularly contextualized within my own practice, decidedly differs from short term research projects – even related projects that occur intermittently throughout the years. My approach to digital ethnography involves continual, albeit at times shifting, engagement. For example, while I have spent the majority of my time on Instagram and Reddit, I've also observed accounts and content on Discord, Tumblr, Twitter, and several online community forums to contextualize information and stories shared by informants, interview participants, and online posters. Data collection is not strictly bounded by time, communities, or a certain number of posts. Every day, I participate online in spaces where people with mental illness gather and memo on what I see and do. As such, data collection is not necessarily extremely distinct from analysis. Similarly, within the work I've done, various paper submissions have bled into and built off one another as I've learned more about individual and community needs and concerns. The immersion and entanglements of this approach provide certain insights, scholarly development, and ways of thinking are grounded in the flexible, continuous, and fluid approach of ethnography.

Ethnographic practices of participation, such as online lurking, posting, and commenting, provide opportunities for understanding and examining individual and group social practices and dynamics. Through participation within a particular eating disorder support subreddit, for example, I've been able to understand reactions and responses to new community and platform policies and actions, such as Reddit's recent push to ban hate subreddits. Additionally, closely interacting with others in these spaces, such as through interviews, commenting, direct messaging, and chats, has informed my understanding of how content and moderation practices within particular communities has changed – and not necessarily for the better, as described by my informants. The social dimension of ethnography is uniquely bolstered due to ethnographic emphasis on participation in various forms. Rather than researcher interpretation influenced by personal expertise or external literature, participation through digital ethnography provides a means to check interpretations of content with informants and participants on a regular basis. This means that interpretations of content and practices are co-informed, involving the work and understanding of the researcher and members of the research population.

Participating within online spaces with members of a research population contributes, in part, to the emotional investment and labor associated with digital ethnography. This is, at times, exhausting. For example, for two years I followed an account on Instagram where an individual with anorexia posted about her experiences, including school and inpatient treatment stays. One day, I realized the account had been removed from Instagram. This was a terrible, anxious realization. Though I'll never know what happened, or how that individual is now, I do still wonder. Similarly, content on a particular eating disorder support subreddit has, more and more frequently, romanticized eating disorders and related practices and provided tips, such as how to conceal eating disorders from others. While this seems to stem from a lack of active community moderators, individuals (e.g., informants and interview participants) are concerned this subreddit will go the way of r/ProED (i.e., banned). I'm concerned, as well. However, anxiety and concern are not the only emotional attachments associated with my ethnographic practice. Some content and individuals are deeply funny. Others are gently supportive and caring. Digital ethnography has provided an emotional appreciation of how people with mental illness participate online, bringing to my interpretation and analysis an additional, affective layer of understanding.

7.1.2. Centering People with Mental Illness

Stigma surrounding mental illness contributes to how people decide to express, communicate, and share – or not – their experiences (Chang and Bazarova, 2016; Corrigan et al., 2004; De Choudhury and De, 2014; Livingston and Boyd, 2010). During my digital ethnography, I've seen and spoken with individuals hesitant to talk about their experiences with friends, family members, and physicians. I've witnessed stories told by a number of people regarding the lack of validation and belief from healthcare providers; which, in turn, can result in exacerbation of disordered behaviors and further concealment. Invalidation of experience, along with other forms of stigma, can reduce a person's proclivity for telling others, even *trusted* others, about mental illness. Ultimately, this can act as a barrier to healing and receipt of treatment, when desired. Research within HCI and CSCW does little to break down the stigma that people with mental illness encounter (Chancellor et al., 2019a). For example, recommending certain types of content and intervening on others creates situations where a narrow range of particular experiences are valued and privileged. Ultimately, this serves to reify existing stigma toward mental illness and the people living with it. Additionally, though current work does, arguably, foreground mental illness, it's frequently viewed from the perspectives of platforms and healthcare practitioners or related documents, such as the DSM. Meaning, certain content and experiences related to mental illness are othered and treated as non-normative.

With respect to my own position and understanding, I argue that research does not center people with mental illness if it does not engage with the priorities, values, and needs of relevant individuals and communities (Pacheco-Vega and Parizeau, 2018). For example, common practice in current, related literature involves the development and recommendation of design implications and interventions by a privileged few (i.e., HCI and CSCW researchers, clinical practitioners, industry technologists). Though these individuals conduct research and design *for* people with mental illness, members of the research population are notably absent. Similarly, given dominant framings of current research, scholars do not often engage with or represent the needs and concerns of members of the research population. Instead, they present the concerns and needs of researchers, social media platforms, and clinicians. Representing content, experiences, and people with mental illness, and designing systems for people with mental illness, without including these individuals, these stakeholders, can be harmful and othering. In my work, by centering people with mental illness, I aim to engage with the needs of this population and normalize how people with mental illness participate online.

Centering is an integral part of my ethnographic approach. In addition to attending to the values and practices, needs and concerns of people with mental illness, it also involves recognizing that people with mental illness are not a homogenous group. Understanding similarities and differences, such as between people and communities, involves a commitment to listening. Partly, this means acknowledging and embracing the role of the researcher as a learner. Throughout my digital ethnography, I have continually learned from people I've interviewed, as well as individuals I've observed and encountered online. For example, when I first started my ethnography, I reported graphic self-harm content to Instagram. My intention was to support people who seemed distressed. However, through observation, followed by an interview study, I learned that my actions of 'help' were, at times, potentially harmful, oppressive, and counterproductive. My openness to sincerely learning from my research population provided the basis for this point of view and insight. As a practice, centering does not mean that researchers will always get it right. However, it does involve openness, sincerity, humility, and willingness to learn and grow.

7.1.3. Diversity of Experiences and the Importance of Context

The significance of approaching mental illness online through digital ethnography also involves collecting and analyzing content that does not explicitly engage with linguistic and visual cues related to mental illness. Recall, for example, participants from Chapter 4 who posted inspirational quotes and memes, as well as beach selfies. These representations are not necessarily hashtagged, captioned, or otherwise explicitly related to an experience with mental illness (e.g., anxiety, depression). Currently, representations of mental illness such as these are underrepresented in the literature (c.f., McCosker and Gerrard (2020); Pater et al. (2019a)). Specifically, many current practices of data collection cannot account for posts related to mental illness that are not visibly linked or claimed through hashtags or other types of signifiers. This ultimately works to obscure the diversity of mental illness online in a number of ways. As my work illustrates, other forms of data collection, such as interviews, promote a broad understanding of how people post about their experiences with mental illness online.

Even when collected, posts related to mental illness may be obscured, flattened, and misrepresented. This is due to data collection and cleaning processes, which often remove posts about mental illness from surrounding context. As I've illustrated across my dissertation work, context is vital to understanding an individual's experience with mental illness. For example, in Chapter 4, I illustrate how mental illness is better understood when contextualized within an individual's everyday life and posting practices. This work argues for keeping content related to school, entertainment, cars, and other topics as part of data collection and analysis. Similarly, in Chapter 5, I demonstrate how posts related to mental illness and mental health share a number of similarities. Recall, for example, the Instagram post captioned "Breakfast." Without context, posts like these are indecipherable; they hold too many interpretive potentials to make any analytic sense, particularly with respect to human and computational practices of classification.

Though it's important to include context in data collection and analysis, whether through surrounding posts, the entirety of an account, or a complementary interview study, when I talk about keeping context intact, I don't mean attending to every single element in an individual's account or online history. Rather, keeping context intact is about, at the very least, bringing other types of content alongside. This involves making the explicit decision to not remove any part of an individual's online account or posts (i.e., within a particular space or beyond). Even when context is kept there are ways to center or foreground an individual's life with mental illness, such as I've done in Chapter 4.

As demonstrated across my body of work, I am committed to working with people with mental illness in order to listen to their experiences and perspectives and engage with their needs and concerns. Here, I've revisited the primary contributions of my empirical chapters, including my digital ethnography, centering of people with mental illness, and representation of diverse experiences and context. Building on this foundation, I detail current HCI and CSCW research practices with respect to mental illness online and discuss consequences associated with dominant paradigms.

7.2. Examining Current HCI and CSCW Research Practices

In this section, I attend to the ways in which people are produced in social media research on mental illness within HCI and CSCW. This particular investigation has roots across my empirical work, where I have aimed to understand several lines of inquiry related to how people post about their experiences with mental illness online. For example, in Chapter 4 I illustrated how mental illness is depicted within and surrounding other everyday experiences, such as school, work, and hobbies, as well as the ways in which other actors participate in shaping how mental illness is narrated. With respect to this latter instance, Chapters 5 and 6 extend this work by describing how human and computational actors (mis)interpret certain experiences and operate to enforce a certain type of conformity around how people with mental illness post and gather online.

In this section, I use research on eating disorders to discuss how people (with eating disorders) are produced through the sociotechnical structures of social media and the god trick of ‘objective’ science (Haraway, 1988). The collective entanglement of these various regimes works to enact certain versions of people online *and* within the academic literature. While commercial aspects of platforms also have relevancy here, such as the ways in which platforms generate revenue and retain ‘users’, here I center HCI and CSCW research practices in order to work through a particular case of how power moves and consolidates within this specific research enterprise (for a related example in education see: Vossoughi and Escudé (2016)). I locate my examples within research on pro-anorexia and pro-recovery communities, some of which stem back to the mid-2000s, to illustrate how the technical specificities of digital platforms support certain opportunities for inquiry and contribute to how people with eating disorders are then reproduced within academic research projects. My aim is to call attention to how power is consolidated across the HCI and CSCW research enterprise, which has implications for the versions of people that are produced. Following a discussion of how consolidated power leads to narrow worlds and consolidated, inaccurate versions of people, I address two consequences associated with these practices: epistemic violence and oppressive care.

7.2.1. Producing Eating Disorders through the Research Enterprise

Social media platforms have everything to do with how people are enacted online and, subsequently, how people are produced through and within social media research. As illustrated across my empirical work, people interact online through various modes of communication and engagement, including accounts, friend lists, images, hashtags, comments, and likes. The specifics of these sociotechnical machinery vary from platform to platform. For example, Instagram uses an image of a heart to indicate 'likes', whereas Reddit uses pictures of arrows to indicate 'upvotes' and 'downvotes'. These technical specificities produce different versions of people across online spaces – not to mention the ways in which social norms and practices vary from platform to platform. The versions of people we see online did not exist prior to social media platforms and do not exist independently of them.

For example, while studying pro-anorexia on Instagram now might involve collecting hashtags, account profile information, post media (e.g., images, videos, text), and comments, studying pro-anorexia websites in the late 1990s and early 2000s was vastly different (e.g., collecting characteristics related to website provider cost, date of last update, information about the webmaster, and website content (Norris et al., 2006). In addition to differences in technical affordances, the social specifics of pro-anorexia differed in the 1990s and early 2000s, an era that I've seen called 'the first wave of pro-ana' during my online observation. All of this is to say that the ways in which pro-anorexia and people communicating about pro-anorexia are produced depends on the sociotechnical affordances and practices of specific online platforms at specific times. Reminiscent of Annemarie Mol's examination of atherschlerosis in The Body Multiple

(Mol, 2002), this means that the versions of people produced in online spaces are technospecific and situated. By attending to differences, as well as the ways in which these multiple versions cohere, we can gain insight into how power flows through the HCI and CSCW research enterprise.

Researchers are reliant on specific human-machine assemblages to conduct research and produce scientific knowledge. Take, for example, my empirical work in Chapter 6. I started recruitment on Craigslist and moved to a specific subreddit and an online eating disorder support forum, once moderator permission was granted. In comparison to the subreddit and eating disorder support forum, which very explicitly had many members with eating disorders who decided to gather around that facet of experience, the flyer on Craigslist was available to a much broader population. Several of the individuals who met the inclusion criteria for my study from this particular recruitment approach were individuals with eating disorders who participated in diet groups, such as those on Facebook. These individuals were very different than the population I had in mind. For example, one individual spoke about how her restriction-based intensity on a diet Facebook group ended in a ban when she suggested multiple dieting practices at once. Similarly, another spoke about how Instagram suspended her account for sizeism and fatphobia, even though it was directed toward herself. These experiences are valid and connected to how these women experienced moderation in relation to their eating disorders. However, neither of these experiences are those in which the eating disorder is explicit or even visible to an external audience. Meaning, it’s likely the experiences with content moderation here were due to other processes of conformity around body image and body management practices, rather than those explicitly related to policing how people with eating disorders communicate about their experiences with eating disorders online. (Though, of course, the end result is similar.)

Moving away from the specifics of my interview recruitment methods, I want to pivot to discuss hashtag-based data collection, which is very prevalent within social media research on mental illness. Hashtag-based data collection, the particulars of the practice as well as the specific hashtags used, illustrates how power and knowledge are entangled, as well as how they consolidate in order to produce certain versions of eating disorders within the HCI and CSCW literature. As I move through the ways in which certain versions of people with eating disorders are consolidated through the research enterprise, my aim is not to police or gatekeep any particular ways of approaching mental illness – or eating disorders, more specifically – on social media. Rather, I want to illustrate how certain practices and versions of people become more prevalent than others.

Though not the first empirical chapter of my dissertation, Chapter 5 represents my initial foray into understanding and examining how people with mental illness use social media. In that chapter, I describe how I compiled an initial dataset – an entry point into mental illness on Instagram – using a set of five hashtags, all of which had been, at the very least, mentioned in some capacity in prior work. Though hashtags have a number of limitations when used to collect social media data (Gerrard, 2018; McCosker and Gerrard, 2020), I used them in order to reference related literature; aiming to make my work credible by association. At the time, as a very junior researcher – even more so than now! – using hashtags, as well as Instagram, which had been studied in prior work on mental illness, was a vital part of my research strategy. Without getting into all of the hashtags here, I want to solely focus on #anorexia and take a step into the history of this hashtag's use, as it relates to the two papers I referenced for this part of my method: *"This Post Will Just Get Taken Down"* (Chancellor et al., 2016b) and *#thyghgapp* (Chancellor et al., 2016d).

With respect to #anorexia, both of the above publications cite the following work: Anorexia on Tumblr (De Choudhury, 2015). In this publication, De Choudhury collected 18,923 unique blogs and "first manually examined a number of Tumblr blogs mentioning common eating disorder and anorexia symptomatology tags" (De Choudhury, 2015). It is unclear how these initial blogs were selected, however, because we're not given a sense of how the author arrived at the "common eating disorder and anorexia symptomatology tags." Given the prior work cited in this publication, we can imagine at least one reason: some of these tags were grounded in content analysis and linguistic studies of blogs and other online pro-anorexia and pro-eating disorder spaces (Borzekowski et al., 2010; Wolf et al., 2013; Yom-Tov et al., 2012). Many of these earlier studies set the stage for the described dichotomy between pro-anorexia and pro-recovery online spaces and content, which may recognize, but not certainly not foreground, the many similarities and people across these groups. All of this history is folded into something as deceitfully simple as using #anorexia as part of a hashtag-based data collection. This brief example illustrates how certain hashtags or words gain footing and legitimacy when studying certain topics. Using the same hashtags repeatedly, even just as entry points into additional data collection, consolidates the possible versions of eating disorders we see online; therefore, consolidating what we produce through our research.

Hashtags are, of course, not the problem here, though they help to illustrate some of it. Consolidation of power through certain research practices is endemic to the entire endeavor (Bowker and Star, 2000) – though we can see the ways in which certain works aim to push back (e.g., such as in my data collection including interviews and content without hashtags). Nevertheless, even work that pushes back in one way may be complicit in another. Interviews and content analysis also work to constrain the ways in which people are produced, such as through researcher-generated questions and classifications. Once work has been legitimized, such as through peer-reviewed publication or receipt of funding, there's a certain momentum that carries practices forward. These operate to create a type of homogeneity with respect to people and eating disorders online.

However, we know that people with eating disorders aren't homogenous. Even the community groups of pro-anorexia and pro-recovery are misleading, as the individuals moving between these content and spaces are similar and numerous. The world, eating disorders included, is multiple, messy. Dominant research practices may inadvertently conceal this mess, producing a certain type of 'sameness' that is not indicative or accurate with respect to the many multiple versions of people with eating disorders online. As others have written and demonstrated through their research commitments, in aiming to examine particular facets of the world, it's important to lean into the messiness, the ways in which the world and its various actors are interconnected (Baumer and Brubaker, 2017; Haraway, 2013; Law, 2004; Piepzna-Samarasinha, 2018; Taylor, 2020). After all, "the openness and culturally constructed nature of the social world, peppered with contradictions and complexities, needs to be embraced, not dismissed" (England, 1994).

7.2.2. Consequences of Consolidating Research Approaches

There are consequences for the ways that power consolidates across the HCI and CSCW research enterprise, particularly with respect to social media research on mental illness. Here, I highlight two consequences: 1) epistemic violence and 2) oppressive care. Epistemic violence and oppressive care are harmful, traumatic experiences for individuals with eating disorders and other types of mental illness, madness, and neurodivergent experiences studied in social media research. I discuss theses consequences in the context of research on eating disorders, as well as my own empirical work, in which I aim to address the harm that we, as researchers, inadvertently cause through our practices of constructing knowledge and caring for others.

7.2.2.1. *Epistemic Violence*

With respect to social media research on eating disorders, few researchers incorporate the perspectives of people with eating disorders in their work. While my work is one exception, recent work by Jessica A. Pater and colleagues illustrates the experiences of three *patients* in recovery from eating disorders, including the ways in which individuals may use digital content to self-harm (Pater et al., 2019a). I emphasize 'patients' here to make a clear point: Much of the work within HCI and CSCW on eating disorders participates in the medicalization of these experiences. While interviews with patients are a very important perspective to incorporate, not all individuals with eating disorders are patients or want to become patients. Even the use of patients suggests a researcher-designated hierarchy and framing of experience. So, while Pater et al.'s work does invite firsthand accounts and tellings of online experience and mental illness, it further consolidates those experiences within a medical realm of interpretation and classification that still suggests group, such as pro-eating disorder (as mentioned in the publication), homogeneity. Another interpretation: the inclusion of patients in this work means that the words of people with eating disorders are appropriated in order to reinforce existing processes of domination. I include this example to illustrate that it's not only firsthand accounts that are necessary, but a variety of accounts that engage with multiple, including non-medicalized, versions of eating disorders.

Exceptions aside, much of the social media research on eating disorders does not include firsthand accounts or collaborative, co-constructions of knowledge. Meaning, individuals with eating disorders are not frequently permitted to produce knowledge about themselves within HCI and CSCW research. This is evident in the ways that we, as a collective of researchers, often devalue expertise derived from lived experience; instead relying on ourselves and clinical practitioners as intellectual gatekeepers. In this context, exclusion of individuals with eating disorders is epistemic violence, a "violence against one's status as a knower; one's role as a creator and communicator of knowledge" (Ymous et al., 2020). In *Can the Subaltern Speak?* (Spivak, 2003), Gayatri Chakravorty Spivak uses epistemic violence to describe how knowledge is produced about the colonized other by the dominant social group. The question inherent in this article, more than mirroring the title, is, rather, "When the subaltern speak, will anyone listen?" With respect to people living with madness, mental illness, and other forms of neurodivergence, the answer frequently feels like a resounding, “No.” After all, as I address in Chapter 2, personal accounts of lived experience were all but nonexistent until the 15th century. Though communicating about mental illness is more readily available today than ever before, particularly due to a variety of digital spaces, knowledge production about mental illness is still largely relegated to clinical actors and those who assume authority by association, such as researchers in HCI and CSCW. Ultimately, we are not listening.

Recently, Anon Ymous and colleagues addressed epistemic violence in disability related technology research. The authors make a number of poignant points, illustrating how:

> “. . . disabled people are often persistently excluded from making meaning about technologies presumably for them. Such epistemic exclusion allows the circulation of implicit bias against disabled subjects as producers of knowledge. Privileging, for example, neurotypical perspectives on neurodivergent needs, demarcating the space for technologies almost exclusively to domains of cure, therapy and intervention – all of which are tied to rhetorics removing the intricacies of disabled lives from the range of human experiences” (Ymous et al., 2020).

In this article, the authors discuss how disabled scholars, themselves included, working and living within the academy and industry are subject to various processes of exclusion and control. These processes include, for example, who gets to produce knowledge and conduct research, as well as the ways in which the lives and experiences of disabled people are represented and framed. Ultimately, disabled people – of which people living with mental illness are part – are excluded from knowledge production and subject to ableist interventions, which I discuss in the following section.

7.2.2.2. *Oppressive Care*

In addition to epistemic violence, people with mental illness who use certain online spaces, such as social media platforms, can be subjected to interventions that impose upon their daily routine. For example, during interviews, individuals have discussed deep frustration with post removals and accounts bans, as well as with the provision of certain support resources, such as the NEDA helpline. Though research recommendations don't always materialize on social media platforms, they are, of course, implicated, as well. Interventions on people with mental illness, particularly those individuals posting or viewing certain content, are often developed without the input of those very people. While, on one hand, this is another example of the epistemic violence mentioned above, here I want to slightly pivot. Social media platforms and researchers develop and implement interventions and recommendations from, I argue, a place of care. Essentially, interventions are how researchers and platforms care for people with mental illness. However, the type of care we see produced by 'experts' in this space isn't caring with or even supporting how people with mental illness care for one another; it's oppressive.

Mad folks and individuals living with mental illness have a collective history of centuries of oppressive care, including forced institutionalization and coercive treatment (e.g., in various forms, such as medication, electroshock therapy, and lobotomy) – including from medical practitioners, religious and spiritual leaders, and family members. Oppressive care is nonconsensual, an imposition. It involves acts of caring for an individual where the individual's agency and, at times, personhood is denied or ignored. Oppressive care is a type of violence. It's controlling. Often, it's paternalistic in assuming that an individual needs a specific un-asked for type of care, support, treatment, or healing. Oppressive care is done to a person, rather than with them or alongside them.

Digital interventions, including content moderation and the provision of support resources, are types of oppressive care. The historical past I draw from provides additional context for what is going on. It situates new ways in which people with mental illness may encounter oppression under the guise of care. Many online interventions, though this is not their intention, have ableist undertones and reinforce stigma against people with mental illness by targeting certain content. For example, recovery is promoted as an end goal or ideal and deeply personal content about living with a mental illness, such as anorexia, may be removed due to its potential to harm others. My digital ethnography over the past few years has illustrated that many people with mental illness are not producing content or communicating in certain ways to harm others. People simply want places where they can openly discuss all aspects of their experience with an eating disorder – even (and maybe especially) the really 'messed up' parts.

Interventions, though they attend to societal notions of designing for 'good', miss the mark because of how they can continuously harm the people who encounter them (Smyth and Dimond, 2014; Ymous et al., 2020). For example, though people resist interventions, such as by not using hashtags or by creating multiple and/or new accounts (e.g., following the ban of another), there are no systems in place to ensure these interventions, as processes of control, won't happen again. Here, I want to reframe interventions as oppressive care. As I discuss in Chapter 4, understanding disruption in the life of someone with mental illness depends on context. My argument, specifically, is that mental illness is not always a – or the only – disruptive experience or force in an individual's life. As such, there are instances when oppressive care, such as digital interventions, are the disruption to an individual's everyday. This is clear throughout Chapter 7, when I discuss how content moderation – which, here, I've readdressed through the lens of oppressive care – has consequences for people with eating disorders who have had posts, accounts, and communities impacted. For a specific example, recall from Chapter 7 the individual who discussed purging to cope with the ban of r/ProED.

The examples I address above with respect to oppressive care are the tip of the iceberg. These forms of care implicate more than just consequences to people with mental illness, but an entire system of surveillance, which, in part, relies on outsourced and underpaid laborers. Mark Andrejevic writes that "entry into the digital enclosure carries with it, in most cases, the condition of surveillance or monitoring" (Andrejevic, 2008). By participating online, people with mental illness are already subject to surveillance – often this type of surveillance is linked to how platforms make money. However, when posting certain content related to mental illness, individuals may find themselves in a position where they are hyper-surveilled, running the risk of being reported by other platform members or flagged by machines. Surveillance, particularly the ways in which it entangles with certain systems of classification and moderation, operates to enforce certain versions of mental illness online. This type of surveillance, and the classification systems entangled with it, form the basis for how people with mental illness are controlled (Bowker and Star, 2000; Scott, 1998). As researchers studying how people use online systems, we're implicated in these processes; such as through our design recommendations or the ways in which we observe – or surveil – people with mental illness online. I am not saying that researchers advocate for surveillance of people with mental illness. And, yet, we are certainly complicit in these processes. Rather than remain complicit, we need to design systems alongside people with mental illness that mitigate this type of harm and prioritize their concerns, expertise, and labor, as well as the ways in which people care for one another.

7.3. Future Work

People do not communicate and gather online for us to observe, collect, and manipulate their content or digital representations of themselves. To make this point clear, several informants from the eating disorder subreddits I participate within have discussed what I'll loosely refer to as the 'problem with researchers'. This problem is twofold: 1) individuals don't necessarily want researchers collecting data from the public spaces they gather and communicate about their experiences with eating disorders; and 2) individuals don't want these public spaces to become private because that would inevitably make them more difficult to find for people who need them.

Computational social scientists and digital ethnographers need to grapple with the ways in which we impact our research populations and, at times, specific individuals and communities. I don't expect or ask for a stop to these types of practices. Rather, we need to examine our relationships with research populations and develop a research practice dedicated to mitigating harm and supporting research populations, communities, and individuals. To put this another way, we need to "question both the potential harms (e.g. conflicts with the online group) and the benefits (e.g. legitimization of the group)" (Murthy, 2011). Additionally, rather than just question harms and benefits, we are accountable to working to reduce harm, including the ways we exploit communities and commit violence against them (e.g., epistemic violence and oppressive care).

As discussed above, we want more worlds and versions of eating disorders and other types of mental illness online, not less. We also want to attend to these worlds in such ways that the people within them are engaged with and their needs and concerns are valued and addressed. This doesn't mean discounting the dominant paradigm – which is starting to break apart – but broadening it out. In this section, I discuss how reconfiguring relationships and committing to an anti-oppressive care approach to research and design provide potential avenues for mitigating the harm we cause to our research populations in this domain.

7.3.1. Reconfiguring Relationships in Social Media and Mental Illness Research

> "Vision is always a question of the power to see – and perhaps of the violence implicit in our visualizing practices." – Haraway (1988)

Creating and maintaining multiple worlds with respect to social media research on mental illness necessitates a reconfiguration of relationships between researchers and research populations. In this section, I discuss how I've negotiated my research practices over the past several years, walking through the implications of my decisions and the ways these decisions have changed over time. Though I center my own experiences here, I also aim to 'read against the grain' of other relevant studies within HCI and CSCW to demonstrate the ways in which researchers already engage in these approaches to the study of mental illness online (Hustak and Myers, 2012). Reading against the grain involves what Carla Hustak and Natasha Myers describe as "reading with our senses attuned to stories told in muted registers" (Hustak and Myers, 2012). There are so many of these stories in work on social media and mental illness. I specifically attend to the stories where researchers are visible as part of this process. My aim, then, is to amplify relationships across this domain of work (Vossoughi and Escudé, 2016), and illustrate how attending to these relationships can support collective production of knowledge and epistemic justice, as well as affective entanglements that facilitate and recognize mutual aid and care.

7.3.1.1. *Incorporating Reflexivity in Online Mental Illness Research*

Throughout the next few sections, I address taken-for-granted and concealed decisions within research practices; namely, my own. To support this reflexive endeavor, I situate my work within feminist, queer, and other critical traditions. Though these literatures have a life beyond research in HCI and CSCW (Berger, 2008; Crenshaw, 1990; England, 1994; Haraway, 1988; Hustak and Myers, 2012; Pacheco-Vega and Parizeau, 2018), many researchers in these fields have worked to make these practices visible and approachable with respect to technology research, particularly as it is produced within many ACM venues. This includes, for example, works such as Feminist HCI (Bardzell, 2010; Bardzell and Bardzell, 2011), Intersectional HCI (Schlesinger et al., 2017), Queer(ing) HCI (Spiel et al., 2019), Critical Race Theory for HCI (Ogbonnaya-Ogburu et al., 2020), and Critical Disability Studies in HCI (Spiel et al., 2020; Williams and Boyd, 2019; Ymous et al., 2020). Beyond these articles, researchers also make their commitments clear through the ways in which they approach research populations and the production of knowledge, including publications and technological systems. Drawing from this body of work, I also bring alongside the analytic lens of doubly engaged ethnography (Pacheco-Vega and Parizeau, 2018), which resonates with my position as a digital ethnographer and, in many ways, talks across ethical issues that ethnographers and computational social scientists should aim to address in their research.

Doubly engaged ethnography, as described by the authors, “requires attention to the relationship between the researcher and their research community, and an acknowledgement of the power dynamics inherent in ethnographic research and writing” (Pacheco-Vega and Parizeau, 2018). As such, it shares much in common with the works I describe above, many of which focus on relations between researchers and research populations and critically engage with ethical dilemmas in research, including how power moves through various enterprises and configurations of sociotechnical spaces and actors. Like work within HCI and CSCW, doubly engaged ethnography argues for ethical considerations to be part of the entire research process. In part, this involves engaging with the needs and concerns of a research population and ensuring that the research process does not harm the wellbeing or livelihood of population members. Additionally, doubly engaged ethnography provides a clear way for researchers to think about their relationships with people, communities, and populations through the centering of three ethical issues: positionality, engagement versus exploitation, and representation. In this chapter, I focus on positionality.

Positionality – or considering the researcher or designer's position with respect to a particular research population and project – is something we see many researchers in feminist, queer, race, and disability scholarship discuss as part of their research practice (see, for example: Ames et al. (2018); DeVito et al. (2018); Thieme et al. (2018)). Engaging with positionality is not prescriptive and can involve a number of facets, including the researcher's position in the world, goals, and political beliefs (Bardzell and Bardzell, 2011). However, with respect to social media research on mental illness, positionality and researcher self-disclosure are rarely present in publications. Like other scholars within HCI and CSCW, those of us studying mental illness online should engage in processes of self-disclosure, regardless our research approach. Engaging with these processes allows us to reflect on our own, multiple positions and how those shape our work.

Below, I argue that we need to learn from other work, particularly adjacent work within HCI and CSCW, that engages with and grapples with researcher positions. I make this argument by very clearly showing the relevance of my positions to my research practice. My digital ethnography, and how it has changed across the years, involves a variety of techniques, some of which share similarities to computational approaches (e.g., the hashtag data collection mentioned previously). The hybridity of my own approach to studying how people with mental illness use online spaces provides a unique opportunity to engage with inescapable ethical dilemmas that researchers –

digital ethnographers, computational social scientists, and all those around and in between – encounter while working with specific online sites and configurations of people. I illustrate how considerations of researcher positions – in particular, the ways in which those positions constitute certain relationships with research populations – help us understand entanglements between researchers and members of research populations in order to discuss the implications for co-constructing knowledge and recognizing affective entanglements between ourselves and other actors within these online spaces.

7.3.1.2. *Locating Ourselves*

> “We argue that positions are fundamentally relevant in doubly engaged ethnography because they locate the researcher, subjects, and research site all within a system of knowledge production where there is a cohesive understanding of the need for scholarly engagement that actually benefits the community under study.” – Pacheco-Vega and Parizeau (2018)

In *Situated Knowledges*, Haraway writes, “Feminist objectivity is about limited location and situated knowledge, not about transcendence and splitting of subject and object. It allows us to become answerable for what we learn how to see” (Haraway, 1988). Due to the multiple positions we hold, and the ways in which these positions may change over time, we learn to interpret and act with the world in very specific ways. Ultimately, we are materialized, embodied entities and, as such, “we need to locate ourselves in our work and to reflect on how our location influences the questions we ask, how we conduct our research, and how we write our research” (England, 1994). While these reflexive considerations and the conversations they spark are prevalent within many HCI and CSCW research circles, such as those referenced above, they're noticeably limited within research on mental illness and online spaces.

Specifically, within social media research on mental illness there is a noticeable absence with respect to how researchers engage with their own positions and relationships to online communities. That absence relates to a number of positions, but, in particular, researcher experience with mental illness. To be clear, I'm complicit here, as well. While I argue for locating and situating our positions more visibly and explicitly – as well as carrying them through in our study designs, findings, discussions, and design recommendations – I recognize the difficulty of self-disclosure and potential for stigmatization as it relates to our experiences with mental illness, madness, psychosocial disability, and neurodivergence. Undeniably, it can be safer for those of us working within specific institutions to keep these positioned concealed (Clare, 2017; Ymous et al., 2020). As such, my intention is not for people to out themselves or for reviewers, when they enter into the process, to ask that people out themselves. There are other ways to detail positionality that focus on relationships, rather than all of the positions a researcher might hold. Here, I aim to illustrate, as many others have done, how our positionality is the foundation through which we work and view the world. Simply put, it impacts everything.

Not all researcher positions are absent from social media research on mental illness. Take, for example, instances when researchers refer to themselves as 'experts' or mention the training they have received in order to conduct their specific line of inquiry (Chancellor et al., 2017; Pater et al., 2019a). A researcher's position is also implicated by who they choose to work with, such as individuals who select to work with clinicians in some capacity (e.g., analysts, participant resources). Researcher positions are also, perhaps subtly, implicated in the interventions that people design, such as for handling pro-anorexia content (i.e., to moderate it or aim to share it with clinicians (De Choudhury, 2015)) or helping people respond to sensitive disclosures (Andalibi and Forte, 2018b). These particular positions, those directly addressed and those implicit, are, of course, only part of how people move through the world. Race, gender, class, and disability status – and the fluidity and flexibility of these positions for some – often go unmentioned.

When we don't provide insight into our own partially situated positions, we risk being complicit in "the god trick of seeing everything from nowhere" (Haraway, 1988). As I've mentioned, this practice of absolute, objective vision is violence. It's a practice that participates in constructing dominant, homogenous paradigms of what eating disorders – and other mental illness – are. Breaking away from this paradigm, without replacing it with another, involves acknowledging our embodied materiality and situatedness; arguing for more worlds. While this is a very easy argument to make, as my own apprehension in Chapter 3 shows, it can be a difficult argument to actualize. Across my body of work, I've discussed how researchers may inadvertently stigmatize and marginalize their research populations with mental illness. The impact of these practices is not constrained to research populations, however; they also impact the ways in which mental illness is stigmatized and marginalized within universities (Dolmage, 2017; Ymous et al., 2020). Given this understanding, as I mentioned previously, I am not aiming for us to engage in a culture of outing people with mental illness – or for us to necessarily out ourselves. Even though experiences with mental illness change across time, when disclosure is mentioned in a publication, it's frozen within that moment. It can, of course, be updated, but it can't necessarily be taken back. As such, we should be grateful when people engage with any of their positions in this type of research and aware that the positions that are made visible might, at times, conceal others.

While engaging with researcher identity situates research and supports engaging with reflexivity, there are other ways to incorporate this practice, as well. As I illustrate in the following section, highlighting relationships between researchers and research populations provides an additional way of writing about positionality that can involve (or not) researcher positions, as well as research practice.

7.3.1.3. *Examining our Relationships*

Engaging with positionality and reflexivity is vital for our research (Bardzell, 2010; Bardzell and Bardzell, 2011; England, 1994; Pacheco-Vega and Parizeau, 2018; Rode, 2011). It can help us attend to how power moves, such as through the consolidation of certain practices that impact how we know and construct people with mental illness, which I describe above. Identity-based positions and biographies are only part of what it means for researchers to engage with positionality. We should also work to include and highlight relationships between ourselves, including our research practices, and the individuals, communities, and populations we study (England, 1994; Pacheco-Vega and Parizeau, 2018; Toombs et al., 2017; Vossoughi and Escudé, 2016). By engaging with positionality and reflexivity in this way, we can aim be "more aware of asymmetrical or exploitive relationships" (England, 1994). Though we cannot always remove these relationships, we can be aware of them and work to reduce the harm we cause (England, 1994; Pacheco-Vega and Parizeau, 2018; Smyth and Dimond, 2014).

Research does not happen on the world or on people within the world, but rather, in the betweenness of our relationships with members of the populations we research; which are, in part, impacted by our positions, biographies, research sites, and methodological approaches (England, 1994). As Kim V. L. England writes, research does not involve examining uninterpreted lives, worlds, and social realities (England, 1994). Rather, it's an assemblage of multiple people, worlds, and practices. Below, I illustrate important relational aspects of research by discussing how knowledge is co-constructed between researchers and members of research populations. I end by illustrating how engaging with these relationships highlights affective entanglements between scholars and other actors within these assemblages, which can provide a basis for thinking with care (Puig de la Bellacasa, 2012).

7.3.1.4. *Co-Constructing Knowledge*

Research is structured by a number of actors. Here, I specifically want to address the ways in which knowledge is collectively constructed between researchers and research populations. In HCI and CSCW, many researchers and designers adopt participatory and co-design approaches to research that support collective knowledge production and interpretation (see, for example: Harrington et al. (2019); Lazar et al. (2017); Le Dantec (2012); Walker et al. (2019)). This work is vital, particularly as it relates to making research with marginalized groups equitable (Harrington et al., 2019). Participatory practices, under any guise, can address the potential for exploitation in research by bringing members of research populations into the research practices and representations in which they are implicated. Ultimately, it's important to offer various types of participation and to acknowledge the variety of ways in which knowledge is collectively produced.

However, not all work is or can be participatory. A good amount of work on social media research, particularly as it relates to mental illness, is not participatory. While this can lead to potentially exploitive relationships, there are ways to reduce exploitation. England, for example, discusses different researcher approaches to interacting with members of a research population, including positioning the researcher as a learner or supplicant (England, 1994; Pacheco-Vega and Parizeau, 2018), rather than as a knowledge holder or expert, in order to emphasis the researcher's dependency on the knowledge and expertise of the research population. In this way, online observation and other forms of data collection, such as web crawling, are *already* part of co-constructive practices. Rather than positioning the researcher as an expert, such as to justify particular methods and interpretations, we could collectively benefit by illustrating that, as researchers and scholars, we are also learners. This provides an opportunity to demonstrate how our research populations hold knowledge and educate us – even when our contact with them is mediated (e.g., through posts, comments, social media accounts). In this context, mitigating harm can mean acknowledging co-construction in research practices. This isn't to dismiss the ways in which researchers currently do this, such as speculating on reasons for hashtag variation or account creation following content bans. However, rather than interpretation presented by the expert researcher, we benefit from clear grounding in informant language, as constructed through online content, that supports the expert informant (i.e., as understood by the situated position of the researcher-learner).

By representing people with mental illness (e.g., informants, participants, and as produced on social media) as actors – as experts, as knowledge holders – we can work to reduce the effects of epistemic violence. Evelyn Wan et al. write about this as epistemic justice, the "proper use and allocation of epistemic power; the inclusion and balancing of all epistemic sources" (Wan et al., 2020). This is one way to create and maintain multiple, heterogeneous worlds, where we disrupt and resist the homogeneity of current social media and mental illness research practices and representations by incorporating more people, particularly those othered – even inadvertently so – in current research. We know that mental illness is a nuanced, intimately personal experience. Rather than focus on sameness, we can better account for difference and mess, as well as the tensions this difference can create. Additionally, by considering members of research populations as knowledge holders, we can also understand them through other positions of agency and power, including the labor that people perform within these online spaces.

7.3.1.5. *Affective Entanglements in Social Media Research*

Research is embodied, intersubjective, and interdependent. In the previous section, I illustrated what this means for the co-construction of knowledge; showing how, even when firsthand accounts and participatory methodology are not present, that academic knowledge is produced in the betweenness of researcher positions and biographies, relationships between researchers and research populations, and research approaches (e.g., online observation, data mining). An important element of co-construction is affect. Specifically, if we start to locate ourselves and our relationships, if we resist the god trick (Haraway, 1988), then we start feeling and attending to the affective dimensions of these spaces (Despret, 2013).

Within HCI and CSCW social media research on mental illness, we see instances where researchers discuss the emotional impact of studying mental illness online. Take, for example, Pater and colleagues' discussion of researcher wellbeing (Pater et al., 2016a):

> "The three researchers that coded this work often discussed coping mechanism used to ameliorate the impacts that immersing oneself into this type of data can have on an individual. While it is standard procedure for researchers in our field to protect the subjects within a research endeavor, it is far more rare to take into consideration the impact on the researcher(s). More discussions need to be had within our community as to what best practices and lessons learned could be shared by other disciplines that grapple with immersion into research areas that can negatively impact the researcher conducting the study."

The considerations above focus on the researcher and their wellbeing, briefly alluding to coping mechanisms practiced by Pater and colleagues to mitigate negative emotional aspects associated with examining content related to certain representations of mental illness online. This conversation is similar to many others within HCI and CSCW.

For example, a number of papers detail the potential emotional harms to researchers conducting certain types of work, advocating for researchers to be mindful and attentive to themselves during specific lines of inquiry (Moncur, 2013; Wolters et al., 2017). With respect to mental illness content, what we view online can, at times, be upsetting and at times triggering, even when it doesn't intersect with our own experiences with mental illness. Affective experiences, for example, motivated my movement away from content related to certain forms of non-eating disorder-related self-harm as well as from Instagram as a platform, where the rampant idolization of thin bodies extends beyond certain eating disorder networks and entangles with influencer culture and posts about health, beauty, and fitness.

Considering the affective impact of this work involves considerations beyond researcher reactions. I want to draw out the emotional character of these spaces, particularly as they relate to other types of affective entanglements, such as those apparent in how we represent people with mental illness and develop recommendations for social media design. Here, I want to think about care by thinking-with scholar Maria Puig de la Bellacasa (Puig de la Bellacasa, 2012). Puig de la Bellacasa writes that care involves "doings needed to create, hold together and sustain life's essential heterogeneity" (Puig de la Bellacasa, 2012). As I argue above, heterogeneity (i.e., difference and mess) should be an integral component to any agenda related to studying mental illness on social media. Rather than revisit the previous section through this particular thinking-with configuration, here, I want to consider how thinking with care, as an ontological practice (Puig de la Bellacasa, 2012), helps us understand the ways in which we are directly involved or, at the very least, implicated in the care webs and collectives that people with mental illness create, organize, and maintain (Piepzna-Samarasinha, 2018; Toombs et al., 2017).

Design recommendations are specific configurations of thinking and knowing – ones that often illustrate how researchers currently engage with care in this domain. With respect to online spaces where people with mental illness gather, we see a propensity for intervention in the context of content moderation and the provision of support resources. These approaches to care are prevalent and speak to the ways in which researchers and technologists, more broadly, are concerned for members of their research populations. Rather than engage with the needs and concerns of marginalized research populations with mental illness, however, researchers often opt to design for or recommend what they or other experts interpret as desirable outcomes. To be clear, desirable outcomes as determined by these experts do not typically engage with the needs and concerns of research populations. Instead, design recommendations and interventions are inadvertently entangled with an ableist type of care that, ultimately, can harm and oppress people with mental illness. As psychiatric survivor Imogen Prism writes, "where there is a helper and the helped, where one person is perceived as having a problem, and the other is perceived as being capable of fixing that problem. . . .well, this can lead to more problems" (Prism, 2017). It's time to rethink care as it relates to people with mental illness and how we can support individuals with mental illness in caring for one another. Whether we acknowledge it or not – and we should acknowledge it – we are part of the care webs and collectives of the populations we research (Piepzna-Samarasinha, 2018). In the following section, I briefly outline a new, alternative way of caring with people with mental illness: anti-oppressive care.

7.3.2. Anti-Oppressive Care

Building from my empirical work and writing across the last few sections, here I introduce and outline anti-oppressive care. Anti-oppressive care is a reflexive approach to thinking-with and designing-with people with mental illness that involves care at the core of its epistemic, political, and ethical intentions. The current section is a sketch; a start to imagining new practices of research and design with individuals with mental illness. For me, this type of imagining was not possible at the start of my dissertation work. Rather, it is situated in the specific context of the research I've conducted over the past several years. I've developed anti-oppressive care out of my commitment to the positions and standpoints of members of my research populations, as well as critical examination of how current sociotechnical approaches in this domain can result in a number of unanticipated consequences. The commitment to *anti-oppressive* care is significant. By positioning care here as explicitly anti-oppressive I mean to juxtapose this type of care with the oppressive versions and histories detailed above. Many of those versions of care are grounded in notions of good. However, as I have illustrated with respect to my own domain of study, 'good' can be insufficient when it leaves certain individuals behind. While I write about anti-oppressive care as it relates to individuals with mental illness, this perspective is valuable for all practices of research and design where care is involved or implicated in research and design for marginalized individuals and groups.

Anti-oppressive care is an opportunity to iterate on the 'centering' of people with mental illness that I describe throughout my dissertation. Though I draw from social justice work with respect to this practice (as illustrated in Chapter 2), centering is not unproblematic. For example, centering individuals with mental illness risks prioritizing one particular identity category and context over others. As I address in Chapter 4, mental illness should not be so heavily featured that it drowns out everything else about an individual and population. Rather than centering people, anti-oppressive care centers ontologies of care (i.e., ways of being with care) (Puig de la Bellacasa, 2012). So, rather than centering people with mental illness, an anti-oppressive care approach involves thinking-with people with mental illness, such as with respect to how content and experiences are interpreted and how problems, needs, and concerns are addressed. ‘Thinking-with’ in this context aims to displace or, at the very least, disrupt, contemporary practices of thinking for, of which much representation and intervention design is current part.

In addition to drawing from Puig de la Bellacasa’s scholarship on care, anti-oppressive care is informed by from Leah Lakshmi Piepzna-Samarasinha’s writings on collective care, mutual aid, and access within disabled, mad, sick, and neurodivergent communities – particularly care work and other types of ‘private and undocumented’ emotional and intellectual labor as practiced and written about by black and brown femmes (Piepzna-Samarasinha, 2018; Sins Invalid; Wong). An integral part of my own thinking is influenced by how Piepzna-Samarasinha describes care webs (Piepzna-Samarasinha, 2018); assemblages of access, healing, and support involving (potentially) may different actors, but led by the “needs and desires of the disabled people running them” (Piepzna-Samarasinha, 2018). In thinking alongside Piepzna-Samarasinha, my intention is to illustrate how social media researchers studying mental illness impose themselves within spaces and practices of collective care. Though this does not make researchers part of care webs, it does mean that they are implicated in how people with mental illness care for themselves and others online.

As I describe above, not all care is positive; it can be oppressive. Piepzna-Samarasinha discusses this as care that controls. To challenge and resist these practices of oppressive care, it's important that, in considering care, we also talk about how to practice an anti-oppressive approach to research. In particular, I aim to be in conversation with the anti-oppressive design framework (Smyth and Dimond, 2014) and other research calling for an end of oppression and harm (e.g., such as to livelihood and wellbeing (Pacheco-Vega and Parizeau, 2018)) to research communities (Diaz et al., 2018; Harrington et al., 2019; Lazar et al., 2017, 2018; Spiel et al., 2020; Ymous et al., 2020). Anti-oppressive design doesn't simply entail 'not being oppressive'. Rather, grounded in the anti-oppressive framework (Burke and Harrison, 1998), it is intentionally committed to ending oppression, such as through, in the context of anti-oppressive workplaces, "continual reflection, requests for outside help, and the implementation of genuine democracy" (Smyth and Dimond, 2014). My development of anti-oppressive care, though nascent, is where practices of care and anti-oppression intersect and entangle. These approaches contextualize my own work, which, in the following section, culminates in an overview of three core elements of anti-oppressive care.

7.3.2.1. *Caring Together: Practicing Anti-Oppressive Care*

Engaging with anti-oppressive care as an approach to research and design involves three core elements: 1) respecting and representing members of research populations, including people with mental illness, as actors; 2) disclosing researcher positionality, particularly as it relates to relationships with research populations; and 3) centering logics of care that, as much as possible, are not oppressive or harmful. Importantly, these 'principles' are not universal, but must be contextualized and situated between specific researchers, research populations, and methodological approaches.

To mitigate epistemic violence and its associated consequences, first, researchers must support the agency of members of their research population by considering people in all forms (e.g., participants in design interviews, human-machine assemblages through social media accounts) as actors. Actors, in varied forms of collaboration, participate in the co-construction of knowledge, rather than being subjected to researcher-driven interpretations that neglect epistemic agency. Positioning and writing about members of research populations as actors can also support considering the needs and concerns, expertise, and labor practices of these individuals. Here, I mean to emphasize the potential for thinking-with these individuals as we, researchers and designers, develop lines of inquiry, approach research questions through particular methods, and represent members of research populations in circulated articles. We have an opportunity to imagine care as integral (Puig de la Bellacasa, 2012), where thinking with care is necessary for work that embraces interdependent, collective, and multiple worlds – the types of worlds I've aimed to illustrate throughout my digital ethnography with people with mental illness online. As such, as much as possible, researchers should aim to directly work with members of their research populations. While this might involve formalized method, such as working with interview participants, it could also involve informal collectives where members of a research population are compensated for reviewing work (e.g., such as attending to representation) prior to its publication.

Second, anti-oppressive care is a reflexive approach to research and design that necessitates researcher self-disclosure. It is necessary for researchers and designers to acknowledge their positions and relations within these interdependent worlds. However, as I address above, with respect to mental illness, researcher self-disclosure may inadvertently pressure scholars into outing themselves as members of a marginalized and, often in academia, stigmatized group. With this consideration in mind, documenting the relationships between a researcher and research population provides an alternative to explicit researcher positionality. For example, researchers can discuss their reasons and motivations for working with a particular group, decisions behind making contact (or not), and affective entanglements with members of the research population or types of content produced by these individuals (Balaam et al., 2019; Toombs et al., 2017). A vital component in writing about these relationships is the understanding that they are founded within and through care. As such, in addition to the examples provided, scholars can also address disclosure through considerations of care – particularly, how their involvement does, or doesn't, impose upon the care webs and collectives of individuals with mental illness.

Finally, anti-oppressive care, involves – or at least provides a potential for – centering logics of care that are not oppressive or harmful. In centering these logics, the aim is to decenter particular identity categories, as well as dominant logics of authority and control. For example, centering logics of care could involve examining the work that individuals with mental illness perform in caring and providing mutual aid for (and with) one another (e.g., providing support and resources, making spaces accessible). Additionally, in writing this discussion, I've come to understand that, for my particular work, care is often juxtaposed with considerations of harm (e.g., content moderation of harmful content). Centering logics of care also requires reconsidering the meaning of harm; specifically, how harm is situated rather than universal. Current processes related to caring for people with mental illness, including content moderation, aim to reduce overall harm, such as by deleting content and banning accounts that are interpreted as generally harmful. We oversimplify sociotechnical spaces, however, when we generalize harm or aim to quantify it (i.e., moderation of certain eating disorder mitigates harm to more people than leaving that content online). There are certain experiences that current research fails to acknowledge and address. Though I am arguing for these experiences to be privileged, I recognize that this isn't without risk. Namely, that others could be harmed due to a lack of moderation of certain content. This is one of many tensions that we, collectively as technologists, need to address and grapple with; understanding that responses need to be technospecific and contextualized, and, as such, that there is no one or correct type of response.

In close, as I've begun to sketch it out here, anti-oppressive care is a reflexive approach to research and design that involves thinking-with people with mental illness, disclosing researcher and research population relationships, and centering logics of care that, as much as possible, are not harmful or oppressive. Even so, engaging with this approach does not mean that anti-oppressive care research will not be oppressive or harmful (Ymous et al., 2020). We all, at varying times, fumble through our scholarly work and grapple with difficult ethical decisions. Additionally, we all conduct research and design within a system that privileges certain people, practices, and values over others. To hold anyone to an unproblematic standard is unfair and unrealistic. Nevertheless, committing to an anti-oppressive care approach to research and design is one way we can work to change this system from within.

To draw this approach together in the following section I get speculative by outlining two concrete instances of what an approach through anti-oppressive care could do. These examples are not intended to be limiting, but, rather, to help move this conversation forward.

7.3.2.2. *Supporting the Work of Mutual Aid*

As I start to address in Chapter 6, individuals with eating disorders adopt a number of strategic practices in order to care for others within their communities. This involves (but is, of course, not limited to) self-moderating content, replying to posts with words – and resources – of recovery support and harm reduction, calling attention to the ‘bad’ or ‘unsafe’ behaviors of others (e.g., posting pro-anorexia content within pro-recovery hashtags), removing content from communities, and using trigger warnings in posts about specific types of content. Research often takes the labor of these practices for granted. However, “mutual aid is labor. Emotional labor” (Prism, 2017). As a community of HCI and CSCW technologists, we need to prioritize and value the work – including care work and other types of labors – that people with mental illness perform online.

In appreciating and learning from the types of work that people with mental illness do in online spaces, we can engage with people and communities in ways that support and privilege their needs and concerns. This may, for example, involve modifying existing computational recommendations to better suit the needs of people and communities. Rather than moderate content before it's posted, computational tools able to detect certain content could, instead, suggest the use of a particular trigger warning. This suggestion uses a computational practice to support an existing community and interpersonal strategy (Bennett et al., 2020). Including computational tools in care webs in such a way may also have the benefit of standardizing trigger and content warnings across an online space (e.g., by helping catch unintentional misspellings and typos). On some platforms, such as Instagram or Twitter, standardization could make it easier for people to block that specific content – which is, of course, helpful for navigating the space. Additionally, having a tool that supports community norms and practices may reduce some of the burden of education within these communities (e.g., this tool could provide educational information about why the inclusion of trigger warnings is important).

In considering how we can support the labor of people with mental illness online, we need to also be aware of how current approaches to interventions get in the way of what people are doing. Take, for example, the ban of r/ProED. A number of participants and other informants have mentioned how the ban of r/ProED came out of nowhere (i.e., there was no warning or option for a quarantined subreddit). The ban of this subreddit, then, directly neglects – and gets in the way of – the work performed by the moderators of this online space (e.g., removing posts and comments, responding to Modmail) as well as the other members (e.g., care work). Designing for anti-oppressive care means that we need to be mindful of how our involvement may destroy or otherwise unsettle existing practices. As Cynthia L. Bennett and colleagues discuss (Bennett et al., 2020), our aim should not be for artificial intelligence – or other technical features, systems, or processes – to step in completely or overtake the work that people are doing. Rather, we need to consider how technical systems can fit into and support existing practices and the work that people do together.

7.3.2.3. *Providing Resources*

As I address across my empirical research, people with many different types of experiences with mental illness gather within many of the same online spaces. Take, for example, an eating disorder support community on Reddit. This community houses individuals who are recovering, individuals who are relapsing, individuals with diagnosed disorders, and individuals who do not have access to healthcare. Additionally, people of many weights and disordered eating practices participate in this space together. The multiplicity of these gathering spaces has implications for how people care together. Even so, platform responses to support resources often draw the line at providing the NEDA website and helpline. On Reddit, this information is only provided on quarantined eating disorder support communities; however, on places like Instagram and Tumblr, the NEDA website and helpline, as well as similar resources, may be shared when searching for content using specific hashtags. While, again, the provision of these support resources is a form of care, care work means not leaving people behind (Piepzna-Samarasinha, 2018). As I've discussed, such as in Chapter 6, generalized, non-specific resources run the risk of being exclusionary, particularly when recovery is painted as the only way forward.

Individuals using online spaces, such as the eating disorder support subreddit mentioned above, provide many different types of information and resources for one another – collectively and at an individual level. For example, the eating disorder support subreddit I discuss in the previous paragraph has a designated link for community-compiled support resources. Additionally, individuals may create posts about books, articles, workbooks, recipes, and behavioral practices that helped them at various points in recovery (including resources related to harm reduction) or otherwise (e.g., staying well while living with a mental illness). Similarly, many people in this subreddit post about their experiences with eating disorders, such as difficulty eating, or specific problems they would like addressed, such as dental care. Other subreddit members are often quick to respond with support, resources, and anecdotes, which can be informative and comforting. Given the richness of experiences in these online spaces, the variety of ways in which people care with one another seems like commonsense. Just as there is no one way to have an eating disorder, there is no one way to stay well or heal with or from one.

Researchers have opportunities to leverage network and community-shared resources in order to support people with eating disorders and other mental illnesses with more varied and specific types of resources. While some researchers suggest directing people toward recovery communities or content (De Choudhury, 2015; Gerrard, 2018), I suggest, when possible, allowing people to opt in to seeing supportive, healing content within the communities or hashtags that they are already part of or searching through.

Even ED (eating disorder) Twitter – of which certain parts are known for hardcore content and gore – includes accounts that share tips for harm reduction. Recovery communities are not for everyone. Many of them are less active than other places for support (e.g., such as where recovery is not a requisite). Sending along recovery content and communities, in an existing ecosystem where these types of resources are already privileged, dismisses the importance of harm reduction and wellness while living with a mental illness. Researchers and designers have opportunities to work with people currently using these online spaces, in addition to people who have left (e.g., such as in order to recover), to provide resources that acknowledge many different ways of living with eating disorders and other forms of mental illness, madness, and psychosocial disability.

7.4. Concluding Remarks

Mental illness isn't going away. It – as well as madness, psychosocial disability, and related experiences – is entangled with a number of elements, including individual, collective, and generational experiences and trauma. Rather than solely leverage the expertise of certain fields, we need to diversify who is brought into the fold. We need to care together. As researchers, as people involved and implicated in the communities we study, and as people whose relationships are formed through caring with each other. This way of thinking has been integral to my dissertation, particularly as it has come together over the past few months. Rather than inadvertently restrict how versions of mental illness can be produced online and recommend interventions grounded in clinical understandings or desirable platform outcomes, we should aim to support a multiplicity of experiences. Undoubtedly, this is a difficult endeavor. However, individuals with mental illness and associated communities have strategies for helping one another navigate online content and experiences, such as through trigger warnings and community-led moderation (e.g., community member reports and moderator removal), as well as the sharing of anecdotes and particular resources. This labor, though acknowledged in other work, is often background – but it's exactly what we need to foreground, understand, and support. By pivoting to think of the work that people and communities are doing (i.e., the care they're providing for and with one another), as well as the role we have as researchers with respect to care, we can move forward to design *with* rather than for, and aim to end oppressive and harmful ways of caring.

References

Saeed Abdullah, Mark Matthews, Elizabeth L. Murnane, Geri Gay, and Tanzeem Choudhury. Towards circadian computing: " early to bed and early to rise" makes some of us unhealthy and sleep deprived. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, pages 673–684, 2014.

Amelia Abraham. Remembering mad pride, the movement that celebrated mental illness, 2016. URL `https://www.vice.com/en_uk/article/7bxqxa/mad-pride-remembering-the-uks-mental-health-pride-movement`. Accessed: 2019-01-18.

Edith Ackermann, Francoise Decortis, Juan Pablo Hourcade, and Heidi Schelhowe. Cultural coding and de-coding as ways of participation: Digital media for marginalized young people. In *Proceedings of the 8th International Conference on Interaction Design and Children*, IDC '09, pages 294–297. ACM, 2009. ISBN 978-1-60558-395-2. doi: 10.1145/1551788.1551864. URL `http://doi.acm.org/10.1145/1551788.1551864`.

David M. Almeida. Resilience and vulnerability to daily stressors assessed via diary methods. *Current Directions in Psychological Science*, 14(2):64–68, 2005.

Xavier Amador and Anna-Lisa Johanson. *I am not sick, I don't need help! Helping the seriously mentally ill accept treatment.* Vida Press, 2000.

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

Morgan Ames and Mor Naaman. Why we tag: Motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 971–980. ACM, 2007. ISBN 978-1-59593-593-9. doi: 10.1145/1240624.1240772. URL `http://doi.acm.org/10.1145/1240624.1240772`.

Morgan G. Ames, Silvia Lindtner, Shaowen Bardzell, Jeffrey Bardzell, Lilly Nguyen, Syed Ishtiaque Ahmed, Nusrat Jahan, Steven J. Jackson, and Paul Dourish. Making or making do? challenging the mythologies of making and hacking. Technical report, 2018.

Tawfiq Ammari and Sarita Schoenebeck. Networked empowerment on facebook groups for parents of children with special needs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2805–2814. ACM,

2015. ISBN 978-1-4503-3145-6. doi: 10.1145/2702123.2702324. URL `http://doi.acm.org/10.1145/2702123.2702324`.

Nazanin Andalibi. What happens after disclosing stigmatized experiences on identified social media: Individual, dyadic, and social/network outcomes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 137. ACM, 2019.

Nazanin Andalibi. Disclosure, privacy, and stigma on social media: Examining non-disclosure of distressing experiences. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(3):1–43, 2020.

Nazanin Andalibi and Andrea Forte. Announcing pregnancy loss on facebook: A decision-making framework for stigmatized disclosures on identified social network sites. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2018a.

Nazanin Andalibi and Andrea Forte. Responding to sensitive disclosures on social media: a decision-making framework. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(6):1–29, 2018b.

Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. Sensitive Self-disclosures, Responses, and Social Support on Instagram: The Case of #Depression. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 1485–1500, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4335-0. doi: 10.1145/2998181.2998243. URL `http://doi.acm.org/10.1145/2998181.2998243`.

Monica Anderson, Skye Toor, Lee Rainie, and Aaron Smith. Activism in the social media age. *Pew Research Center*, July 2018. URL `https://www.pewinternet.org/2018/07/11/activism-in-the-social-media-age/`.

Mark Andrejevic. Privacy, exploitation, and the digital enclosure. *Amsterdam LF*, 1:47, 2008.

Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine bias. *ProPublica, May*, 23:2016, 2016.

Amy A. Ross Arguedas. “can naughty be healthy?”: Healthism and its discontents in news coverage of orthorexia nervosa. *Social Science & Medicine*, 246:112784, 2020.

Zahra Ashktorab. A study of cyberbullying detection and mitigation on instagram. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*, CSCW '16 Companion, pages 126–130. ACM, 2016. ISBN 978-1-4503-3950-6. doi: 10.1145/2818052.2874346. URL `http://doi.acm.org/10.1145/2818052.2874346`.

Paul Atkinson. Narrative turn or blind alley? 7(3):325–344, aug 1997. ISSN 1049-7323. doi: 10.1177/104973239700700302. URL `https://doi.org/10.1177/104973239700700302`.

Darren Baker and Sarah Fortune. Understanding self-harm and suicide websites: a qualitative interview study of young adult website users. *Crisis: The journal of crisis intervention and suicide prevention*, 29(3):118, 2008.

Madeline Balaam, Rob Comber, Rachel E. Clarke, Charles Windlin, Anna Ståhl, Kristina Höök, and Geraldine Fitzpatrick. Emotion work in experience-centered design. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

Michael Bamberg. Stories: Big or small: Why do we care? 16:139–147, jan 2006. doi: 10.1075/ni.16.1.18bam.

Michael Bamberg. Narrative analysis. In H. Cooper, editor, *APA handbook of research methods in psychology*, volume 2, pages 77–94. 2012.

Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. Designing mobile health technology for bipolar disorder: a field trial of the monarca system. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2627–2636, 2013.

Shaowen Bardzell. Feminist HCI: Taking stock and outlining an agenda for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1301–1310. ACM, 2010. ISBN 978-1-60558-929-9. doi: 10.1145/1753326.1753521. URL `http://doi.acm.org/10.1145/1753326.1753521`.

Shaowen Bardzell and Jeffrey Bardzell. Towards a feminist hci methodology: social science, feminism, and hci. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 675–684, 2011.

Steven E. Barkan. Sociology: Understanding and changing the social world. 2011. URL `http://open.lib.umn.edu/sociology/`.

Karin Barron. Ethics in qualitative social research on marginalized groups. 1(1):38–49, jan 1999. ISSN 1501-7419. doi: 10.1080/15017419909510736. URL `https://doi.org/10.1080/15017419909510736`.

Jamie Bartlett. *The dark net: Inside the digital underworld*. Melville House, 2015.

Eric P. S. Baumer and Jed R. Brubaker. Post-userism. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6291–6303, 2017.

Natalya N. Bazarova, Yoon Hyung Choi, Victoria Schwanda Sosik, Dan Cosley, and Janis Whitlock. Social sharing of emotions on facebook: Channel differences, satisfaction, and replies. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pages 154–164. ACM, 2015.

Anne Beaulieu. Mediating ethnography: objectivity and the making of ethnographies of the internet. *Social epistemology*, 18(2-3):139–163, 2004.

Philip Bell. Content analysis of visual images. *Handbook of visual analysis*, 13, 2001.

Jules R. Bemporad. Cultural and historical aspects of eating disorders. *Theoretical Medicine*, 18(4):401–420, 1997.

Cynthia L. Bennett, Daniela K. Rosner, and Alex S. Taylor. The care work of access. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2020.

John Berger. *Ways of seeing*. Penguin uK, 2008.

Giuseppe Bersani, Elisa Polli, Giuseppe Valeriani, Daiana Zullo, Claudia Melcore, Enrico Capra, Adele Quartini, Pietropaolo Marino, Amedeo Minichino, and Laura Bernabei. Facial expression in patients with bipolar disorder and schizophrenia in response to emotional stimuli: a partially shared cognitive and social deficit of the two disorders. *Neuropsychiatric disease and treatment*, 9:1137, 2013.

José Bertolote. The roots of the concept of mental health. *World Psychiatry*, 7(2):113, 2008.

Allan Beveridge. A disquieting feeling of strangeness?: the art of the mentally ill. *Journal of the Royal Society of Medicine*, 94(11):595–599, 2001.

Rena Bivens. The gender binary will not be deprogrammed: Ten years of coding gender on facebook. *New Media & Society*, 19(6):880–898, 2017.

Rena Bivens and Oliver L. Haimson. Baking gender into social media design: How platforms shape categories for users and advertisers. *Social Media+ Society*, 2(4): 2056305116672486, 2016.

Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck. LGBT parents and social media: Advocacy, privacy, and disclosure during shifting social movements. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 610–622. ACM, 2016. ISBN 978-1-4503-3362-7. doi: 10.1145/2858036.2858342. URL `http://doi.acm.org/10.1145/2858036.2858342`.

Lindsay Blackwell, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. Classification and Its Consequences for Online Harassment: Design Insights from HeartMob. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW):24:1–24:19, December 2017. ISSN 2573-0142. doi: 10.1145/3134659. URL `http://doi.acm.org/10.1145/3134659`.

Johnna Blair and Saeed Abdullah. Supporting constructive mental health discourse in social media. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth '18, pages 299–303. ACM, 2018. ISBN 978-1-4503-6450-8. doi: 10.1145/3240925.3240930. URL `http://doi.acm.org/10.1145/3240925.3240930`.

Jeanette Blomberg and Helena Karasti. Reflections on 25 years of ethnography in cscw. *Computer supported cooperative work (CSCW)*, 22(4-6):373–423, 2013.

J. Sanbourne Bockoven. Moral treatment in community mental health. 1972.

Leah Boepple and J. Kevin Thompson. A content analytic comparison of fitspiration and thinspiration websites. *International Journal of Eating Disorders*, 49(1):98–101, 2016.

Evely Boruchovitch and Birgitte R. Mednick. The meaning of health and illness: some considerations for health psychology. *Psico-USF*, 7(2):175–183, 2002.

Dina L. G. Borzekowski, Summer Schenk, Jenny L. Wilson, and Rebecka Peebles. e-Ana and e-Mia: A Content Analysis of Pro-Eating Disorder Web Sites. *American Journal of Public Health*, 100(8):1526–1534, August 2010. ISSN 0090-0036. doi: 10.2105/AJPH.2009.172700. URL `https://ajph.aphapublications.org/doi/abs/10.2105/ajph.2009.172700`.

Geoffrey C. Bowker and Susan Leigh Star. *Sorting things out: Classification and its consequences*. MIT press, 2000.

Leanne Bowler, Jung Sun Oh, Daqing He, Eleanor Mattern, and Wei Jeng. Eating disorder questions in yahoo! answers: Information, conversation, or reflection? *Proceedings of the American Society for Information Science and Technology*, 49(1):1–11, 2012.

Leanne Bowler, Eleanor Mattern, Wei Jeng, Jung Sun Oh, and Daqing He. I know what you are going through: answers to informational questions about eating disorders in yahoo! answers: a qualitative study. In *Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries*, page 6. American Society for Information Science, 2013.

Kevin W. Bowyer. Face recognition technology: security versus privacy. *IEEE Technology and society magazine*, 23(1):9–19, 2004.

danah boyd. ”none of this is real”. identity and participation in friendster, 2007.

danah boyd. Social network sites as networked publics: Affordances, dynamics, and implications. In *A networked self*, pages 47–66. Routledge, 2010.

danah boyd. *It's complicated: The social lives of networked teens*. Yale University Press, 2014.

danah michele boyd. Friendster and publicly articulated social networking. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1279–1282, 2004.

Pat Bracken and Philip Thomas. From szasz to foucault: On the role of critical psychiatry. *Philosophy, Psychiatry, & Psychology*, 17(3):219–228, 2010.

Rosi Braidotti. Teratologies. In Ian Buchanan and Claire Colebrook, editors, *Deleuze and Feminist Theory*, pages 156–172. Edinburgh University Press, Edinburgh, April 2000. ISBN 978-0-7486-1120-1. URL `https://edinburghuniversitypress.com/book-deleuze-and-feminist-theory.html`.

Uwe Bretschneider, Thomas Wöhner, and Ralf Peters. Detecting online harassment in social networks. dec 2014. URL `https://aisel.aisnet.org/icis2014/proceedings/ConferenceTheme/2`.

T. D. Brewerton. Binge eating disorder: Recognition, diagnosis, and treatment. *Medscape Mental Health*, 2, 1997.

Philip Brey. Ethical aspects of facial recognition systems in public places. *Journal of information, communication and ethics in society*, 2004.

Sarah R. Brotsky and David Giles. Inside the "pro-ana" community: A covert online participant observation. *Eating disorders*, 15(2):93–109, 2007.

Barry Brown. Ethnographic approaches to digital research. *The SAGE handbook of digital technology research*, pages 189–202, 2013.

R. C. Brown, T. Fischer, A. D. Goldwich, F. Keller, R. Young, and P. L. Plener. #cutting: Non-suicidal self-injury (NSSI) on instagram. 48 (2):337–346, jan 2018. ISSN 0033-2917, 1469-8978. doi: 10.1017/S0033291717001751. URL `https://www.cambridge.org/core/journals/psychological-medicine/article/cutting-nonsuicidal-selfinjury-nssi-on-instagram/600ED6C6856EE21B7E875F08CB088DDB`.

Amy Bruckman. Teaching students to study online communities ethically. 15(2):82–95, 2006.

Joy Adowaa Buolamwini. *Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers*. PhD thesis, Massachusetts Institute of Technology, 2017.

Eleanor R. Burgess, Kathryn E. Ringland, Jennifer Nicholas, Ashley A. Knapp, Jordan Eschler, David C. Mohr, and Madhu C. Reddy. "i think people are powerful": The sociality of individuals managing depression. *Proc. ACM Hum.-Comput. Interact.*, 3 (CSCW):41:1–41:29, Nov 2019. ISSN 2573-0142/2019/November. doi: 10.1145/3359143. URL `https://doi.org/10.1145/3359143`.

Beverley Burke and Philomena Harrison. Anti-oppressive practice. In *Social work*, pages 229–239. Springer, 1998.

Judith Butler. *Bodies That Matter: On the Discursive Limits of 'Sex'*. Routledge, New York, 1993.

Caroline Walker Bynum. *Holy feast and holy fast: The religious significance of food to medieval women*, volume 1. Univ of California Press, 1988.

Kellie E. Carlyle, Jeanine P. D. Guidry, Kofoworola Williams, Ariella Tabaac, and Paul B. Perrin. Suicide conversations on instagram: contagion or caring? 0(0):1–7, mar 2018. ISSN 1753-8068. doi: 10.1080/17538068.2018.1436500. URL `https://doi.org/10.1080/17538068.2018.1436500`.

Patricia A. Cavazos-Rehg, Melissa J. Krauss, Shaina J. Costello, Nina Kaiser, Elizabeth S. Cahn, Ellen E. Fitzsimmons-Craft, and Denise E. Wilfley. "i just want to be skinny.": A content analysis of tweets expressing eating disorder symptoms. *PloS one*, 14(1): e0207506, 2019.

Stevie Chancellor, Zhiyuan Lin, Erica L. Goodman, Stephanie Zerwas, and Munmun De Choudhury. Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1171–1184, New York, NY, USA, 2016a. ACM. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819973. URL `http://doi.acm.org/10.1145/2818048.2819973`.

Stevie Chancellor, Zhiyuan (Jerry) Lin, and Munmun De Choudhury. "This Post Will Just Get Taken Down": Characterizing Removed Pro-Eating Disorder Social Media Content. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1157–1162, New York, NY, USA, 2016b. ACM. ISBN 978-1-4503-3362-7. doi: 10.1145/2858036.2858248. URL `http://doi.acm.org/10.1145/2858036.2858248`.

Stevie Chancellor, Tanushree Mitra, and Munmun De Choudhury. Recovery Amid Pro-Anorexia: Analysis of Recovery in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2111–2123, New York, NY, USA, 2016c. ACM. ISBN 978-1-4503-3362-7. doi: 10.1145/2858036.2858246. URL `http://doi.acm.org/10.1145/2858036.2858246`.

Stevie Chancellor, Jessica Annette Pater, Trustin Clear, Eric Gilbert, and Munmun De Choudhury. #Thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1201–1213, New York, NY, USA, 2016d. ACM. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819963. URL `http://doi.acm.org/10.1145/2818048.2819963`.

Stevie Chancellor, Yannis Kalantidis, Jessica A. Pater, Munmun De Choudhury, and David A. Shamma. Multimodal Classification of Moderated Online Pro-Eating Disorder Content. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 3213–3226, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4655-9. doi: 10.1145/3025453.3025985. URL `http://doi.acm.org/10.1145/3025453.3025985`.

Stevie Chancellor, Andrea Hu, and Munmun De Choudhury. Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 666:1–666:14, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3174240. URL `http://doi.acm.org/10.1145/3173574.3174240`.

Stevie Chancellor, Eric P. S. Baumer, and Munmun De Choudhury. Who is the" human" in human-centered machine learning: The case of predicting mental health from social media. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–32, 2019a.

Stevie Chancellor, Michael L. Birnbaum, Eric D. Caine, Vincent Silenzio, and Munmun De Choudhury. A taxonomy of ethical tensions in inferring mental health states from social media. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 79–88. ACM, 2019b.

Eshwar Chandrasekharan, Umashanthi Pavalanathan, Anirudh Srinivasan, Adam Glynn, Jacob Eisenstein, and Eric Gilbert. You can't stay here: The efficacy of reddit's 2015 ban examined through hate speech. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):31, 2017.

Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW):32:1–32:25, November 2018. ISSN 2573-0142. doi: 10.1145/3274301. URL `http://doi.acm.org/10.1145/3274301`.

Jonathan P. Chang and Cristian Danescu-Niculescu-Mizil. Trouble on the horizon: Forecasting the derailment of online conversations as they develop. *arXiv preprint arXiv:1909.01362*, 2019.

Pamara F. Chang and Natalya N. Bazarova. Managing stigma: Disclosure-response communication patterns in pro-anorexic websites. *Health Communication*, 31(2):217–229, 2016.

Kathy Charmaz. *Constructing Grounded Theory*. SAGE, mar 2014. ISBN 978-1-4462-9722-3.

Dennis S. Charney, Joseph D. Buxbaum, Pamela Sklar, and Eric J. Nestler. *Neurobiology of mental illness*. Oxford University Press, 2013.

Y. Chen, Y. Zhou, S. Zhu, and H. Xu. Detecting offensive language in social media to protect adolescent online safety. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, pages 71–80, sept 2012. doi: 10.1109/SocialCom-PASSAT.2012.55.

John Cheney-Lippold. *We are data: Algorithms and the making of our digital selves*. nyu Press, 2018.

Robert B. Cialdini and Noah J. Goldstein. Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, 55:591–621, 2004.

Eli Clare. *Brilliant imperfection: Grappling with cure*. Duke University Press, 2017.

David Clark. What if you meet face to face? a case study in virtual/material research ethics. In *Readings in virtual research ethics: Issues and controversies*, pages 246–261. IGI Global, 2004.

Malcolm Collier. Approaches to analysis in visual anthropology. *Handbook of visual analysis*, pages 35–60, 2001.

Odysseas Constantinou. Multimodal discourse analysis: Media, modes and technologies. 9(4):602–618, 2005. ISSN 1467-9841. doi: 10.1111/j.1360-6441.2005.00310.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1360-6441.2005.00310.x`.

Patrick W. Corrigan, Fred E. Markowitz, and Amy C. Watson. Structural levels of mental illness stigma and discrimination. *Schizophrenia bulletin*, 30(3):481–491, 2004.

Lucy Costa, Jijian Voronka, Danielle Landry, Jenna Reid, Becky Mcfarlane, David Reville, and Kathryn Church. “recovering our stories”: A small act of resistance. *Studies in Social Justice*, 6(1):85–101, 2012.

Angus Crawford. Instagram eating disorder content ‘out of control’, March 2019. URL `https://www.bbc.com/news/uk-47637377`. Accessed: 2020-07-15.

Kate Crawford and Ryan Calo. There is a blind spot in ai research. *Nature*, 538(7625): 311–313, 2016.

Kate Crawford and Tarleton Gillespie. What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, 18(3):410–428, March 2016. ISSN 1461-4448. doi: 10.1177/1461444814543163. URL `https://doi.org/10.1177/1461444814543163`.

Kimberle Crenshaw. Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stan. L. Rev.*, 43:1241, 1990.

Sky Croeser. Thinking beyond 'free speech' in responding to online harassment. *Ada: a journal of new media, gender, and culture*, 10:online–online, 2016.

C/S Info Centre. C/s info centre bulletin. URL `http://www.csinfo.ca/bulletin/Bulletin_374.pdf`. Accessed: 2019-01-18.

Emese Csipke and Outi Horne. Pro-eating disorder websites: users' opinions. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, 15(3):196–206, 2007.

Larry Davidson, Maria J. O'Connell, Janis Tondora, Martha Lawless, and Arthur C. Evans. Recovery in serious mental illness: A new wine or just a new bottle? 36 (5):480–487, oct 2005. URL `http://psycnet.apa.org/buy/2005-13212-004`.

Lennard J. Davis. *The disability studies reader*. Taylor & Francis, 2006.

Munmun De Choudhury. Anorexia on Tumblr: A Characterization Study. In *Proceedings of the 5th International Conference on Digital Health 2015*, DH '15, pages 43–50, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3492-1. doi: 10.1145/2750511.2750515. URL `http://doi.acm.org/10.1145/2750511.2750515`.

Munmun De Choudhury and Sushovan De. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *ICWSM*, 2014.

Munmun De Choudhury, Scott Counts, and Eric Horvitz. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 3267–3276, 2013a.

Munmun De Choudhury, Scott Counts, and Eric Horvitz. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, WebSci '13, pages 47–56. ACM, 2013b. ISBN 978-1-4503-1889-1. doi: 10.1145/2464464.2464480. URL `http://doi.acm.org/10.1145/2464464.2464480`.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*, 2013c.

Munmun De Choudhury, Scott Counts, Eric J. Horvitz, and Aaron Hoff. Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, pages 626–638. ACM, 2014. ISBN 978-1-4503-2540-0. doi: 10.1145/2531602.2531675. URL `http://doi.acm.org/10.1145/2531602.2531675`.

Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing*

Systems, CHI '16, pages 2098–2110. ACM, 2016. ISBN 978-1-4503-3362-7. doi: 10.1145/2858036.2858207. URL `http://doi.acm.org/10.1145/2858036.2858207`.

Munmun De Choudhury, Sanket S. Sharma, Tomaz Logar, Wouter Eekhout, and René Clausen Nielsen. Gender and cross-cultural differences in social media disclosures of mental illness. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 353–369, 2017.

Ingrid de Saint-Georges. Materiality in discourse: The influence of space and layout in making meaning. *Discourse and technology: Multimodal discourse analysis*, pages 71–87, 2004.

Tisha Dejmanee. Bodies of technology : performative flesh, pleasure and subversion in cyberspace. *Gender Questions*, 1(1):3–17, January 2013. ISSN 2309-9704. URL `https://journals.co.za/content/genderq/1/1/EJC167619`.

Sheila Delaney. Sexual economics, chaucer's wife of bath, and the book of margery kempe. *Minnesota Review*, 5(1):104–115, 1975.

Liliana Dell'Osso, Marianna Abelli, Barbara Carpita, Stefano Pini, Giovanni Castellini, Claudia Carmassi, and Valdo Ricca. Historical evolution of the concept of anorexia nervosa and relationships with orthorexia nervosa, autism, and obsessive–compulsive spectrum. *Neuropsychiatric Disease and Treatment*, 12:1651, 2016.

Norman K. Denzin. Cybertalk and the method of instances. *Doing Internet research: Critical issues and methods for examining the Net*, pages 107–125, 1999.

Vinciane Despret. From secret agents to interagency. *History and Theory*, 52(4):29–44, 2013.

Michael Ann DeVito, Ashley Marie Walker, and Jeremy Birnholtz. 'Too Gay for Facebook': Presenting LGBTQ+ Identity Throughout the Personal Social Media Ecosystem. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW):44:1–44:23, November 2018. ISSN 2573-0142. doi: 10.1145/3274313. URL `http://doi.acm.org/10.1145/3274313`.

Eliezer Diamond. *Holy Men and Hunger Artists: Fasting and Asceticism in Rabbinic Culture*. Oxford University Press on Demand, 2004.

Mark Diaz, Isaac Johnson, Amanda Lazar, Anne Marie Piper, and Darren Gergle. Addressing age-related bias in sentiment analysis. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 412:1–412:14. ACM, 2018. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3173986. URL `http://doi.acm.org/10.1145/3173574.3173986`.

Jill P. Dimond, Michaelanne Dye, Daphne LaRose, and Amy S. Bruckman. Hollaback!: the role of storytelling online in a social movement organization. In *Proceedings of* *the 2013 conference on Computer supported cooperative work*, pages 477–490. ACM, 2013.

Jay T. Dolmage. *Academic ableism: Disability and higher education*. University of Michigan Press, 2017.

Bryan Dosono and Bryan Semaan. Moderation practices as emotional labor in sustaining online communities: The case of aapi identity work on reddit. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 142. ACM, 2019.

Paul Dourish. Implications for design. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 541–550, 2006.

Andrew T. Drysdale, Logan Grosenick, Jonathan Downar, Katharine Dunlop, Farrokh Mansouri, Yue Meng, Robert N. Fetcho, Benjamin Zebley, Desmond J. Oathes, and Amit Etkin. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature medicine*, 23(1):28–38, 2017.

Sascha Altman Du Brul. Mad pride and spiritual community: Thoughts on the spiritual gift of madness, May 2012. URL `https://www.madinamerica.com/2012/05/mad-pride-and-spiritual-community-thoughts-on-the-spiritual-gift-of-madness/`. Accessed: 2020-07-15.

Maeve Duggan. Online harassment 2017. 2017.

Dana S. Dunn and Shane Burcaw. Thinking about disability identity, November 2013. URL `https://www.apa.org/pi/disability/resources/publications/newsletter/2013/11/disability-identity`. Accessed: 2020-07-14.

Thomas M. Dunn and Steven Bratman. On orthorexia nervosa: A review of the literature and proposed diagnostic criteria. *Eating behaviors*, 21:11–17, 2016.

Brianna Dym, Jed R. Brubaker, Casey Fiesler, and Bryan Semaan. “Coming Out Okay”: Community Narratives for LGBTQ Identity Recovery Work. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW):29, November 2019. doi: 10.1145/3359256. URL `https://doi.org/10.1145/3359256`.

Elizabeth V. Eikey. Privacy and weight loss apps: a first look at how women with eating disorders use social features. In *Proceedings of the 19th International Conference on Supporting Group Work*, pages 413–415, 2016.

Elizabeth V. Eikey and Kayla M. Booth. Recovery and maintenance: how women with eating disorders use instagram. *iConference 2017 Proceedings*, 2017.

Elizabeth V. Eikey and Madhu C. Reddy. “it’s definitely been a journey”: A qualitative study on how women with eating disorders use weight loss apps. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 642–654, 2017.

Elizabeth Victoria Eikey, Madhu C. Reddy, Kayla M. Booth, Lynette Kvasny, Johnna L. Blair, Victor Li, and Erika S. Poole. Desire to be underweight: Exploratory study on a weight loss app community and user perceptions of the impact on disordered eating behaviors. *JMIR mHealth and uHealth*, 5(10):e150, 2017.

Kim V. L. England. Getting personal: Reflexivity, positionality, and feminist research. *The professional geographer*, 46(1):80–89, 1994.

Sindhu Kiranmai Ernala, Asra F. Rizvi, Michael L. Birnbaum, John M. Kane, and Munmun De Choudhury. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–27, 2017.

Sindhu Kiranmai Ernala, Tristan Labetoulle, Fred Bane, Michael L. Birnbaum, Asra F. Rizvi, John M. Kane, and Munmun De Choudhury. Characterizing audience engagement and assessing its impact on social media disclosures of mental illnesses. In *Twelfth International AAAI Conference on Web and Social Media*, june 2018. URL `https://www.aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17884`.

Jordan Eschler, Arpita Bhattacharya, and Wanda Pratt. Designing a reclamation of body and health: Cancer survivor tattoos as coping ritual. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, pages 510:1–510:12. ACM, 2018. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3174084. URL `http://doi.acm.org/10.1145/3173574.3174084`.

Jessica L. Feuston and Siân E. Lindley. How social dynamics and the context of digital content impact workplace remix. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.

Jessica L. Feuston and Anne Marie Piper. Beyond the Coded Gaze: Analyzing Expression of Mental Health and Illness on Instagram. *Proc. ACM Hum.-Comput. Interact.*, 2 (CSCW):51:1–51:21, November 2018. ISSN 2573-0142. doi: 10.1145/3274320. URL `http://doi.acm.org/10.1145/3274320`.

Jessica L. Feuston and Anne Marie Piper. Everyday Experiences: Small Stories and Mental Illness on Instagram. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI ’19, pages 265:1–265:14, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300495. URL `http://doi.acm.org/10.1145/3290605.3300495`. event-place: Glasgow, Scotland Uk.

Casey Fiesler and Nicholas Proferes. “participant” perceptions of twitter research ethics. *Social Media+ Society*, 4(1):2056305118763366, 2018.

Casey Fiesler, Jessica L. Feuston, and Amy S. Bruckman. Understanding copyright law in online creative communities. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pages 116–129, 2015.

Casey Fiesler, Michaelanne Dye, Jessica L. Feuston, Chaya Hiruncharoenvate, Clayton J. Hutto, Shannon Morrison, Parisa Khanipour Roshan, Umashanthi Pavalanathan, Amy S. Bruckman, and Munmun De Choudhury. What (or who) is public?: Privacy settings and social media content sharing. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 567–580. ACM, 2017.

Casey Fiesler, Jialun "Aaron" Jiang, Joshua McCann, Kyle Frye, and Jed R. Brubaker. Reddit Rules! Characterizing an Ecosystem of Governance. In *Twelfth International AAAI Conference on Web and Social Media*, June 2018. URL `https://www.aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17898`.

Fireweed Collective. Fireweed collective website. URL `https://fireweedcollective.org`. Accessed: 2020-07-15.

Alison Fixsen, Anna Cheshire, and Michelle Berry. The social construction of a concept—orthorexia nervosa: Morality narratives and psycho-politics. *Qualitative Health Research*, 30(7):1101–1113, 2020.

Rachel A. Fleming-May and Laura E. Miller. I'm scared to look but i'm dying to know: information seeking and sharing on pro-ana weblogs. In *Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem-Volume 47*, page 61. American Society for Information Science, 2010.

Fernando Espi Forcen and Carlos Espi Forcen. The practice of holy fasting in the late middle ages: a psychiatric approach. *The Journal of nervous and mental disease*, 203 (8):650–653, 2015.

Michel Foucault. *Technologies of the self: A seminar with Michel Foucault*. Univ of Massachusetts Press, 1988.

Michel Foucault. *Madness and civilization*. Routledge, 2003.

Michel Foucault. *Discipline and punish: The birth of the prison*. Vintage, 2012.

Nick Fox, Katie Ward, and Alan O'Rourke. Pro-anorexia, weight-loss drugs and the internet: an 'anti-recovery' explanatory model of anorexia. *Sociology of Health & Illness*, 27(7):944–971, 2005. ISSN 1467-9566. doi: 10.1111/j.1467-9566.2005.00465.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9566.2005.00465.x`.

Sarah Fox, Jill Dimond, Lilly Irani, Tad Hirsch, Michael Muller, and Shaowen Bardzell. Social justice and design: Power and oppression in collaborative systems. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17 Companion, pages 117–122. ACM, 2017. ISBN 978-1-4503-4688-7. doi: 10.1145/3022198.3022201. URL `http://doi.acm.org/10.1145/3022198.3022201`.

Arthur W. Frank. *The Wounded Storyteller*. 1995. URL `https://www.press.uchicago.edu/ucp/books/book/chicago/W/bo14674212.html`.

Mary Anne Franks. The free speech black hole: Can the internet escape the gravitational pull of the first amendment? *Knight First Amendment Institute at Columbia University*. URL `https://knightcolumbia.org/content/the-free-speech-black-hole-can-the-internet-escape-the-gravitational-pull-of-the-first-amendment`.

Barbara L. Fredrickson and Tomi-Ann Roberts. Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of women quarterly*, 21(2):173–206, 1997.

Phyllis R. Freeman, Carley Rees Bogarad, and Diane E. Sholomskas. Margery kempe, a new theory: the inadequacy of hysteria and postpartum psychosis as diagnostic categories. *History of Psychiatry*, 1(2):169–190, 1990.

Batya Friedman, Eric Brok, Susan King Roth, and John Thomas. Minimizing bias in computer systems. *ACM SIGCHI Bulletin*, 28(1):48–51, 1996.

Christian Fuchs. *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*. Routledge, June 2013. ISBN 978-1-136-65527-2.

Silvana Galderisi, Andreas Heinz, Marianne Kastrup, Julian Beezhold, and Norman Sartorius. Toward a new definition of mental health. *World Psychiatry*, 14(2):231, 2015.

Angela Cora Garcia, Alecea I. Standlee, Jennifer Bechkoff, and Yan Cui. Ethnographic approaches to the internet and computer-mediated communication. *Journal of contemporary ethnography*, 38(1):52–84, 2009.

Alexandra Georgakopoulou. Small stories research. In Ana de Fina and Alexandra Georgakopoulou, editors, *The Handbook of Narrative Analysis*, pages 255–271. John Wiley & Sons, apr 2015. ISBN 978-1-118-45818-1.

Anne Germain and David J. Kupfer. Circadian rhythm disturbances in depression. *Human Psychopharmacology: Clinical and Experimental*, 23(7):571–585, 2008.

Ysabel Gerrard. Beyond the hashtag: Circumventing content moderation on social media. page 1461444818776611, may 2018. ISSN 1461-4448. doi: 10.1177/1461444818776611. URL `https://doi.org/10.1177/1461444818776611`.

Jannath Ghaznavi and Laramie D. Taylor. Bones, body parts, and sex appeal: An analysis of# thinspiration images on popular social media. *Body image*, 14:54–61, 2015.

Tarleton Gillespie. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press, 2018. ISBN 978-0-300-17313-0.

Sander L. Gilman. *Illness and image: Case studies in the medical humanities*. Routledge, 2018.

Debbie Ging and Sarah Garvey. ‘written in these scars are the stories i can’t explain’: A content analysis of pro-ana and thinspiration image sharing on instagram. *New Media & Society*, 20(3):1181–1200, 2018.

Eric Goldman. Unregulating online harassment. *Denver University Law Review Online*, 57:59, 2010.

Lauren Gonzales, Kristin C. Davidoff, Kevin L. Nadal, and Philip T. Yanos. Microaggressions experienced by persons with mental illnesses: An exploratory study. *Psychiatric rehabilitation journal*, 38(3):234, 2015.

Charles Goodwin. Seeing in depth. *Social studies of science*, 25(2):237–274, 1995.

Shrigopal Goyal, Yatan Pal Singh Balhara, and S. K. Khandelwal. Revisiting classification of eating disorders-toward diagnostic and statistical manual of mental disorders-5 and international statistical classification of diseases and related health problems-11. *Indian journal of psychological medicine*, 34(3):290, 2012.

Mary L. Gray and Siddharth Suri. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Eamon Dolan Books, New York, 2019.

L.D. Green and Kelechi Ubozoh, editors. *We’ve Been Too Patient: Voices from Radical Mental Health*. North Atlantic Books, 2019.

Trisha Greenhalgh. *ANNEX 3. A BRIEF HISTORY OF NARRATIVE RESEARCH IN HEALTH CARE*. WHO Regional Office for Europe, 2016. URL `https://www.ncbi.nlm.nih.gov/books/NBK391070/`.

James Grimmelmann. The Virtues of Moderation. *Yale Journal of Law and Technology*, pages 42–109, 2015. URL `https://heinonline.org/HOL/P?h=hein.journals/yjolt17&i=42`.

Jan Zita Grover and Douglas Crimp. Aids: Cultural analysis/cultural activism. 1987.

Stefan Gruner and Stuart Naven. Tool support for plagiarism detection in text documents. In *Proceedings of the 2005 ACM symposium on Applied computing*, pages 776–781, 2005.

W. W. Gull. Anorexia nervosa (apepsia hysterical, anorexia hysteria). 7:22–28, 1874.

Oliver L. Haimson and Anna Lauren Hoffmann. Constructing and enforcing” authentic” identity online: Facebook, real names, and non-normative identities. *First Monday*, 21(6), 2016.

Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. Disclosure, stress, and support during gender transition on facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW ’15, pages 1176–1190. ACM, 2015a. ISBN 978-1-4503-2922-4. doi: 10.1145/2675133.2675152. URL `http://doi.acm.org/10.1145/2675133.2675152`.

Oliver L. Haimson, Kathryn E. Ringland, and Gillian R. Hayes. Marginalized populations and research ethics online. In *CSCW Workshop on Ethics for Studying Sociotechnical Systems in a Big Data World*, CSCW Workshop on Ethics for Studying Sociotechnical Systems in a Big Data World, page 5, 2015b.

Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. Digital footprints and changing networks during online identity transitions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2895–2907. ACM, 2016.

Max Halupka. Clicktivism: A systematic heuristic. *Policy & Internet*, 6(2):115–132, 2014.

Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. Gender recognition or gender reductionism?: The social implications of embedded gender recognition systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 8. ACM, 2018.

Donna Haraway. Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist studies*, 14(3):575–599, 1988.

Donna J. Haraway. *When species meet*, volume 3. U of Minnesota Press, 2013.

Douglas Harper. *Visual Sociology*. Routledge, dec 2012. ISBN 978-1-135-27875-5. Google-Books-ID: cERcdCCZ4IcC.

Christina Harrington, Sheena Erete, and Anne Marie Piper. Deconstructing community-based collaborative design: Towards more equitable participatory design engagements. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–25, 2019.

James C. Harris. Anorexia nervosa and anorexia mirabilis: Miss k. r—and st catherine of siena. *JAMA psychiatry*, 71(11):1212–1213, 2014.

Libby Hemphill and Andrew J. Roback. Tweet acts: how constituents lobby congress via twitter. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1200–1210. ACM, 2014.

Christine Hine. *Virtual ethnography*. Sage, 2000.

Christine Hine. Virtual ethnography: Modes, varieties, affordances. *The SAGE handbook of online research methods*, pages 257–270, 2008.

Robert Hodge, , and Gunther R. Kress. *Social semiotics*. Cornell University Press, 1988.

Lynn C. Holley, Kyoko Y. Tavassoli, and Layne K. Stromwall. Mental illness discrimination in mental health treatment programs: Intersections of race, ethnicity, and sexual orientation. *Community mental health journal*, 52(3):311–322, 2016.

Su Holmes. ‘My anorexia story’: girls constructing narratives of identity on YouTube. *Cultural Studies*, 31(1):1–23, January 2017. ISSN 0950-2386. doi: 10.1080/09502386.2016.1138978. URL `https://doi.org/10.1080/09502386.2016.1138978`.

Allan V. Horwitz. *The social control of mental illness*. Academic Press, july 1982. ISBN 978-0-12-356180-0. Google-Books-ID: d_MgAQAAIAAJ.

Dorothy Howard and Lilly Irani. Ways of knowing when research subjects care. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 97. ACM, 2019.

Carla Hustak and Natasha Myers. Involutionary momentum: Affective ecologies and the sciences of plant/insect encounters. *differences*, 23(3):74–118, 2012.

Lilly Irani and Monika Sengul-Jones. Difference work: a conversation with lilly irani. *Catalyst: Feminism, theory, technoscience*, 1(1):1–14, 2015.

Lilly C. Irani and M. Six Silberman. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 611–620. ACM, 2013.

Grace YoungJoo Jeon, Nicole B. Ellison, Bernie Hogan, and Christine Greenhow. First-generation students and college: The role of facebook networks as information sources. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 887–899. ACM, 2016.

Carey Jewitt and Rumiko Oyama. Visual meaning: A social semiotic approach. *Handbook of visual analysis*, 134:156, 2001.

Shagun Jhaver, Sucheta Ghoshal, Amy Bruckman, and Eric Gilbert. Online Harassment and Content Moderation: The Case of Blocklists. *ACM Trans. Comput.-Hum. Interact.*, 2(CSCW):12:1–12:33, March 2018. ISSN 1073-0516. doi: 10.1145/3185593. URL `http://doi.acm.org/10.1145/3185593`.

Shagun Jhaver, Darren Scott Appling, Eric Gilbert, and Amy Bruckman. “Did You Suspect the Post Would be Removed?”: Understanding User Reactions to Content Removals on Reddit. *ACM Trans. Comput.-Hum. Interact.*, 3(CSCW):33, November 2019a. doi: https://doi.org/10.1145/3359294.

Shagun Jhaver, Iris Birman, Eric Gilbert, and Amy Bruckman. Human-Machine Collaboration for Content Regulation: The Case of Reddit Automoderator. *ACM Trans. Comput.-Hum. Interact.*, 26(5):31:1–31:35, July 2019b. ISSN 1073-0516. doi: 10.1145/3338243. URL `http://doi.acm.org/10.1145/3338243`.

Shagun Jhaver, Amy Bruckman, and Eric Gilbert. Does Transparency in Moderation Really Matter?: User Behavior After Content Removal Explanations on Reddit. *ACM*

Trans. Comput.-Hum. Interact., 3(CSCW):27, November 2019c. doi: https://doi.org/10.1145/3359252.

Jialun "Aaron" Jiang, Charles Kiene, Skyler Middler, Jed R. Brubaker, and Casey Fiesler. Moderation Challenges in Voice-based Online Communities on Discord. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW):23, November 2019.

Anna Johansson and Hans T. Sternudd. Iconography of suffering in social media: images of sitting girls. In *World suffering and quality of life*, pages 341–355. Springer, 2015.

Rodney H. Jones. The problem of context in computer-mediated communication. *Discourse and technology: Multimodal discourse analysis*, pages 20–33, 2004.

Adrienne S. Juarascio, Amber Shoaib, and C. Alix Timko. Pro-eating disorder communities on social networking sites: A content analysis. 18(5):393–407, sept 2010. ISSN 1064-0266. doi: 10.1080/10640266.2010.511918. URL `https://doi.org/10.1080/10640266.2010.511918`.

Margery Kempe. *The Book of Margery Kempe*. DS Brewer, 2000.

Anna Keski-Rahkonen and Federica Tozzi. The process of recovery in eating disorder sufferers' own words: An internet-based study. *International Journal of Eating Disorders*, 37(S1):S80–S86, 2005.

Corey L. M. Keyes. The mental health continuum: From languishing to flourishing in life. *Journal of health and social behavior*, pages 207–222, 2002.

Corey L. M. Keyes. Mental health as a complete state: How the salutogenic perspective completes the picture. In *Bridging occupational, organizational and public health*, pages 179–192. Springer, 2014.

Charles Kiene, Jialun "Aaron" Jiang, and Benjamin Mako Hill. Technological Frames and User Innovation: Exploring Technological Change in Community Moderation Teams. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):23, November 2019.

Arthur Kleinman. *The Illness Narratives: Suffering, Healing, And The Human Condition*. Basic Books, New York, reprint edition edition, October 1989. ISBN 978-0-465-03204-4.

Know Your Meme. Kermit the frog | know your meme, 2015. URL `https://knowyourmeme.com/memes/kermit-the-frog`.

Gunther Kress and Theo Van Leeuwen. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. {Arnold Publishers}, aug 2001. ISBN 978-0-340-60877-7. URL `http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0340608773`.

Priya Kumar and Sarita Schoenebeck. The modern day baby book: Enacting good mothering and stewarding privacy on facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1302–1312. ACM, 2015.

Taisa Kushner and Amit Sharma. Bursts of activity: Temporal patterns of help-seeking and support in online mental health forums. In *Proceedings of The Web Conference 2020*, pages 2906–2912, 2020.

Lakesha Lafayett. The Problematic Whitewashing of Eating Disorder Recovery, December 2017. URL `https://www.nationaleatingdisorders.org/blog/problematic-whitewashing-eating-disorder-recovery`.

Cliff Lampe and Paul Resnick. Slash(Dot) and Burn: Distributed Moderation in a Large Online Conversation Space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 543–550, New York, NY, USA, 2004. ACM. ISBN 978-1-58113-702-6. doi: 10.1145/985692.985761. URL `http://doi.acm.org/10.1145/985692.985761`. event-place: Vienna, Austria.

Cliff A. C. Lampe, Erik Johnston, and Paul Resnick. Follow the Reader: Filtering Comments on Slashdot. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 1253–1262, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-593-9. doi: 10.1145/1240624.1240815. URL `http://doi.acm.org/10.1145/1240624.1240815`. event-place: San Jose, California, USA.

landoflobsters. Changes to our policy against bullying and harassment. *Reddit*, September 2019. URL `https://www.reddit.com/r/announcements/comments/dbf9nj/changes_to_our_policy_against_bullying_and/`.

John Law. *After method: Mess in social science research*. Routledge, 2004.

Amanda Lazar, Caroline Edasis, and Anne Marie Piper. A critical lens on dementia and design in hci. In *CHI*, pages 2175–2188, 2017.

Amanda Lazar, Jessica L. Feuston, Caroline Edasis, and Anne Marie Piper. Making as expression: Informing design with people with complex communication needs through art therapy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2018.

Christopher Le Dantec. Participation and publics: supporting community engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1351–1360, 2012.

Kwangyoung Lee and Hwajung Hong. Designing for self-tracking of emotion and experience with tangible modality. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pages 465–475, 2017.

Yu-Hao Lee and Gary Hsieh. Does slacktivism hurt activism?: the effects of moral balancing and consistency in online activism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 811–820. ACM, 2013.

Brenda A. LeFrançois, Robert Menzies, and Geoffrey Reaume. *Mad matters: A critical reader in Canadian mad studies*. Canadian Scholars' Press, 2013.

A. Lenhart, M. Ybarra, K. Zickuhr, and M. Price-Feeney. Online harassment, digital abuse, and cyberstalking in america. data & society research institute. *Center for Innovative Public Health Research. Retrieved from: https://datasociety.net/pubs/oh/Online_Harassment_2016. pdf*, 2016.

Philip LeVine and Ron Scollon. *Discourse and technology: Multimodal discourse analysis*. Georgetown University Press, 2004a.

Philip LeVine and Ron Scollon. Multimodal Discourse Analysis as the Confluence of Discourse and Technology. In Philip LeVine and Ron Scollon, editors, *Discourse and technology: Multimodal discourse analysis*, pages 1–6. Georgetown University Press, 2004b.

Bradley Lewis. A mad fight: Psychiatry and disability activism. *The disability studies reader*, 3, 2006.

Stephen P. Lewis and Alexis E. Arbuthnott. Searching for thinspiration: the nature of internet searches for pro-eating disorder websites. *Cyberpsychology, Behavior, and Social Networking*, 15(4):200–204, 2012.

Guo Li, Xiaomu Zhou, Tun Lu, Jiang Yang, and Ning Gu. SunForum: Understanding depression in a chinese online community. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 515–526. ACM, 2016. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2819994. URL `http://doi.acm.org/10.1145/2818048.2819994`.

Ian Li, Anind Dey, and Jodi Forlizzi. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 557–566, 2010.

Jessa Lingel. The gentrification of the internet, March 2019. URL `http://culturedigitally.org/2019/03/the-gentrification-of-the-internet/`.

Bruce G. Link, Jo C. Phelan, Michaeline Bresnahan, Ann Stueve, and Bernice A. Pescosolido. Public conceptions of mental illness: labels, causes, dangerousness, and social distance. *American journal of public health*, 89(9):1328–1333, 1999.

Simi Linton. *Claiming disability: Knowledge and identity*. NYU Press, 1998.

Deanna Linville, Tiffany Brown, Katrina Sturm, and Tori McDougal. Eating disorders and social support: perspectives of recovered individuals. *Eating Disorders*, 20(3): 216–231, 2012.

Martin Lister and Liz Wells. Seeing beyond belief: Cultural studies as an approach to analysing the visual. *Handbook of visual analysis*, pages 61–91, 2001.

Fannie Liu, Denae Ford, Chris Parnin, and Laura Dabbish. Selfies as social movements: Influences on participation and perceived impact on stereotypes. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):72, 2017.

James D. Livingston and Jennifer E. Boyd. Correlates and consequences of internalized stigma for people living with mental illness: A systematic review and meta-analysis. *Social Science & Medicine*, 71(12):2150–2161, December 2010. ISSN 0277-9536. doi: 10.1016/j.socscimed.2010.09.030. URL `http://www.sciencedirect.com/science/article/pii/S0277953610006945`.

Knut Lundby. *Digital Storytelling, Mediatized Stories: Self-representations in New Media*. Lang, Peter, 2008. ISBN 978-1-4331-0273-8. Google-Books-ID: Sl_WM0tVV84C.

Catherine Lutz and Jane Collins. The photograph as an intersection of gazes: The example of national geographic. *Visual Anthropology Review*, 7(1):134–149, 1991.

Janet A. Lydecker and Carlos M. Grilo. Food insecurity and bulimia nervosa in the united states. *International Journal of Eating Disorders*, 52(6):735–739, 2019.

David Maclagan. *Outsider art: From the margins to the marketplace*. Reaktion Books, 2010.

Helen Malson. Womæn under erasure: Anorexic bodies in postmodern context. *Journal of community & applied social psychology*, 9(2):137–153, 1999.

Ronald W. Manderscheid, Carol D. Ryff, Elsie J. Freeman, Lela R. McKnight-Eily, Satvinder Dhingra, and Tara W. Strine. Peer reviewed: evolving definitions of mental illness and wellness. *Preventing chronic disease*, 7(1), 2010.

Lydia Manikonda and Munmun De Choudhury. Modeling and understanding visual attributes of mental health disclosures in social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 170–181. ACM, 2017. ISBN 978-1-4503-4655-9. doi: 10.1145/3025453.3025932. URL `http://doi.acm.org/10.1145/3025453.3025932`.

George E. Marcus. Ethnography in/of the world system: The emergence of multi-sited ethnography. *Annual review of anthropology*, 24(1):95–117, 1995.

Ben Marder, Adam Joinson, Avi Shankar, and David Houghton. The extended 'chilling' effect of facebook: The cold reality of ubiquitous social networking. *Computers in Human Behavior*, 60:582–592, 2016.

Mark Matthews and Gavin Doherty. In the mood: engaging teenagers in psychotherapy using mobile phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2947–2956, 2011.

Mark Matthews, Stephen Voida, Saeed Abdullah, Gavin Doherty, Tanzeem Choudhury, Sangha Im, and Geri Gay. In situ design for mental illness: Considering the pathology of bipolar disorder in mhealth design. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 86–97, 2015.

Dan P. McAdams. Personal narratives and the life story. In Oliver P. John, Richard W. Robins, and Lawrence A. Pervin, editors, *Handbook of Personality, Third Edition: Theory and Research*, pages 242–262. Guilford Press, aug 2008. ISBN 978-1-60623-738-0.

Anthony McCosker and Ysabel Gerrard. Hashtagging depression on instagram: Towards a more inclusive mental health research methodology. *New Media & Society*, page 1461444820921349, 2020.

Aiden McGillicuddy, Jean-Gregoire Bernard, and Jocelyn Cranefield. Controlling Bad Behavior in Online Communities: An Examination of Moderation Work. *ICIS 2016 Proceedings*, December 2016. URL `https://aisel.aisnet.org/icis2016/SocialMedia/Presentations/23`.

Emma E. McGinty. Mental illness and the media. *Community Mental Health: Challenges for the 21st Century*, pages 178–201, 2017.

Bharat Mehra, Cecelia Merkel, and Ann Peterson Bishop. The internet for empowerment of minority and marginalized users. 6(6):781–802, dec 2004. ISSN 1461-4448. doi: 10.1177/146144804047513. URL `https://doi.org/10.1177/146144804047513`.

Kaitlynn Mendes, Jessica Ringrose, and Jessalynn Keller. #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, 25(2):236–246, May 2018. ISSN 1350-5068. doi: 10.1177/1350506818765318. URL `https://doi.org/10.1177/1350506818765318`.

Chris Messina. Groups for twitter; or a proposal for twitter tag channels, August 2007. URL `https://factoryjoe.com/2007/08/25/groups-for-twitter-or-a-proposal-for-twitter-tag-channels/`. Accessed: 2020-08-20.

Lydia Michie, Madeline Balaam, John McCarthy, Timur Osadchiy, and Kellie Morrissey. From her story, to our story: Digital storytelling as public engagement around abortion rights advocacy in ireland. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 357:1–357:15. ACM, 2018. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3173931. URL `http://doi.acm.org/10.1145/3173574.3173931`.

Merry N. Miller and Andrés J. Pumariega. Culture and eating disorders: A historical and cross-cultural review. *Psychiatry: Interpersonal and biological processes*, 64(2): 93–110, 2001.

Annemarie Mol. *The body multiple: Ontology in medical practice*. Duke University Press, 2002.

Anna Mollow. 'when black women start going on prozac. . . ': The politics of race, gender, and emotional distress in meri nana-ama danquah's willow weep for me. *The disability studies reader*, pages 411–31, 2013.

Wendy Moncur. The emotional wellbeing of researchers: considerations for practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1883–1890, 2013.

Megan A. Moreno, Adrienne Ton, Ellen Selkie, and Yolanda Evans. Secret society 123: Understanding the language of self-harm on instagram. *Journal of Adolescent Health*, 58(1):78–84, 2016.

Dhiraj Murthy. Digital ethnography: An examination of the use of new technologies for social research. *Sociology*, 42(5):837–855, 2008.

Dhiraj Murthy. Emergent digital ethnographic methods for social research. *Handbook of emergent technologies in social research*, pages 158–179, 2011.

Sarah Myers West. Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, 20(11):4366–4383, November 2018. ISSN 1461-4448. doi: 10.1177/1461444818773059. URL `https://doi.org/10.1177/1461444818773059`.

B. Sri Nandhini and J. I. Sheeba. Cyberbullying detection and classification using information retrieval algorithm. In *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)*, ICARCSET '15, pages 20:1–20:5. ACM, 2015. ISBN 978-1-4503-3441-9. doi: 10.1145/2743065.2743085. URL `http://doi.acm.org/10.1145/2743065.2743085`.

Khushnood Naqshbandi, David N. Milne, Ben Davies, Sophie Potter, Rafael A. Calvo, and Simon Hoermann. Helping young people going through tough times: Perspectives for a peer-to-peer chat support system. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction*, OzCHI '16, pages 640–642. ACM, 2016. ISBN 978-1-4503-4618-4. doi: 10.1145/3010915.3011848. URL `http://doi.acm.org/10.1145/3010915.3011848`.

Safiya Umoja Noble. *Algorithms of oppression: How search engines reinforce racism*. nyu Press, 2018.

Mark L. Norris, Katherine M. Boydell, Leora Pinhas, and Debra K. Katzman. Ana and the internet: A review of pro-anorexia websites. *International Journal of Eating Disorders*, 39(6):443–447, 2006.

Ihudiya Finda Ogbonnaya-Ogburu, Angela D. R. Smith, Alexandra To, and Kentaro Toyama. Critical race theory for hci. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2020.

Jung Sun Oh, Daqing He, Wei Jeng, Eleanor Mattern, and Leanne Bowler. Linguistic characteristics of eating disorder questions on yahoo! answers-content, style, and emotion. In *Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries*, page 87. American Society for Information Science, 2013.

Kathleen O'Leary, Arpita Bhattacharya, Sean A. Munson, Jacob O. Wobbrock, and Wanda Pratt. Design opportunities for mental health peer support technologies. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 1470–1484. ACM, 2017. ISBN 978-1-4503-4335-0. doi: 10.1145/2998181.2998349. URL `http://doi.acm.org/10.1145/2998181.2998349`.

Alexandra Olteanu, Sarah Vieweg, and Carlos Castillo. What to expect when the unexpected happens: Social media communications across crises. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pages 994–1009. ACM, 2015.

Online Censorship. Offline-Online, 2019. URL `https://onlinecensorship.org/content/infographics`. Retrieved September 20, 2019.

Petar Opalić. Informal and formal social control of mentally ill persons. 55(3):78–92, 2007. ISSN 0003-2565, 2406-2693. URL `https://www.ceeol.com/search/article-detail?id=602416`.

John E. Pachankis. The psychological implications of concealing a stigma: A cognitive-affective-behavioral model. *Psychological bulletin*, 133(2):328, 2007.

Raul Pacheco-Vega and Kate Parizeau. Doubly engaged ethnography: Opportunities and challenges when working with vulnerable communities. *International Journal of Qualitative Methods*, 17(1):1609406918790653, 2018.

Lawrence A. Palinkas, Sarah M. Horwitz, Carla A. Green, Jennifer P. Wisdom, Naihua Duan, and Kimberly Hoagwood. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. 42(5):533–544, sept 2015. ISSN 1573-3289. doi: 10.1007/s10488-013-0528-y. URL `https://doi.org/10.1007/s10488-013-0528-y`.

Stephanie Pappas. Purging the myth of the vomitorium, August 2016. URL `https://www.scientificamerican.com/article/purging-the-myth-of-the-vomitorium/`. Accessed: 2020-08-18.

Minsu Park, David W. McDonald, and Meeyoung Cha. Perception differences between the depressed and non-depressed users in twitter. In *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.

Hester Parr. Mental health, the arts and belongings. *Transactions of the Institute of British Geographers*, 31(2):150–166, 2006.

Jessica Pater and Elizabeth Mynatt. Defining Digital Self-Harm. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 1501–1513, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4335-0. doi: 10.1145/2998181.2998224. URL `http://doi.acm.org/10.1145/2998181.2998224`.

Jessica A. Pater, Oliver L. Haimson, Nazanin Andalibi, and Elizabeth D. Mynatt. "Hunger Hurts but Starving Works": Characterizing the Presentation of Eating Disorders Online. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1185–1200, New York, NY, USA, 2016a. ACM. ISBN 978-1-4503-3592-8. doi: 10.1145/2818048.2820030. URL `http://doi.acm.org/10.1145/2818048.2820030`.

Jessica A. Pater, Moon K. Kim, Elizabeth D. Mynatt, and Casey Fiesler. Characterizations of online harassment: Comparing policies across social media platforms. In *Proceedings of the 19th International Conference on Supporting Group Work*, GROUP '16, pages 369–374. ACM, 2016b. ISBN 978-1-4503-4276-6. doi: 10.1145/2957276.2957297. URL `http://doi.acm.org/10.1145/2957276.2957297`.

Jessica A. Pater, Brooke Farrington, Alycia Brown, Lauren E. Reining, Tammy Toscos, and Elizabeth D. Mynatt. Exploring indicators of digital self-harm with eating disorder patients: A case study. *Proceedings of the ACM on Human-Computer Interaction*, 3 (CSCW):1–26, 2019a.

Jessica A. Pater, Lauren E. Reining, Andrew D. Miller, Tammy Toscos, and Elizabeth D. Mynatt. "Notjustgirls": Exploring Male-related Eating Disordered Content Across Social Media Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 651:1–651:13, New York, NY, USA, 2019b. ACM. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300881. URL `http://doi.acm.org/10.1145/3290605.3300881`. event-place: Glasgow, Scotland Uk.

Andrew R Payton. Mental health, mental illness, and psychological distress: same continuum or distinct phenomena? *Journal of health and Social Behavior*, 50(2):213–227, 2009.

Jon Penney. Internet surveillance, regulation, and chilling effects online: A comparative case study. *Internet Policy Review*, 2017.

Heather J. Peters, Haley N. Schwenk, Zachary R. Ahlstrom, and Lyndzie N. McIalwain. Microaggressions: The experience of individuals with mental illness. *Counselling Psychology Quarterly*, 30(1):86–112, 2017.

Shruti Phadke, Jonathan Lloyd, James Hawdon, Mattia Samory, and Tanushree Mitra. Framing hate with hate frames: Designing the codebook. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 201–204. ACM, 2018.

Cassandra Phoenix and Andrew C. Sparkes. Being fred: big stories, small stories and the accomplishment of a positive ageing identity. 9(2):219–236, apr 2009. ISSN 1468-7941. doi: 10.1177/1468794108099322. URL `https://doi.org/10.1177/1468794108099322`.

Leah Lakshmi Piepzna-Samarasinha. *Care work: Dreaming disability justice*. arsenal pulp press, 2018.

Sarah Pink, Heather A. Horst, John Postill, Larissa Hjorth, Tania Lewis, and Jo Tacchi. *Digital ethnography: Principles and Practice*. SAGE, 2016.

Radha Iyengar Plumb. An independent report on how we measure content moderation. *Facebook Newsroom*, May 2019. URL `https://newsroom.fb.com/news/2019/05/dtag-report/`.

Michael Pollan. Our national eating disorder. *The New York Times Magazine*, 27, 2004.

Jennifer M. Poole and Jennifer Ward. Breaking open the bone: Storying, sanism and mad grief. *Mad matters: A critical reader in Canadian Mad Studies*, pages 94–104, 2013.

Imogen Prism. We've been too patient: How to create mutual aid societies in the 21st century, June 2017. URL `https://thebodyisnotanapology.com/magazine/weve-been-to-patient-how-to-create-mutual-aid-relationships-in-the-21st-century-imogen-prism/`. Accessed: 2020-07-15.

Devin Proctor. Policing the fluff: The social construction of scientistic selves in otherkin facebook groups. *Engaging Science, Technology, and Society*, 4:485–514, 2018.

Devin Proctor. So you want to "do" digital ethnography, March 2020. URL `https://thegeekanthropologist.com/2020/03/25/so-you-want-to-do-digital-ethnography/`. Accessed: 2020-08-17.

Yada Pruksachatkun, Sachin R. Pendse, and Amit Sharma. Moments of change: Analyzing peer-based cognitive support in online mental health forums. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.

Rebecca Puhl and Young Suh. Stigma and eating and weight disorders. *Current psychiatry reports*, 17(3):10, 2015.

María Puig de la Bellacasa. 'nothing comes without its world': thinking with care. *The Sociological Review*, 60(2):197–216, 2012.

Julian Rappaport. Narrative studies, personal stories, and identity transformation in the mutual help context. 29(2):239–256, june 1993. ISSN 0021-8863. doi: 10.1177/0021886393292007. URL `https://doi.org/10.1177/0021886393292007`.

Reddit. r/proed has been banned from reddit. URL `https://www.reddit.com/r/ProED`. Accessed: 2020-07-15.

Bryce J. Renninger. "where i can be myself ... where i can speak my mind" : Networked counterpublics in a polymedia environment. 17(9):1513–1529, oct 2015. ISSN 1461-4448. doi: 10.1177/1461444814530095. URL `https://doi.org/10.1177/1461444814530095`.

Catherine Kohler Riessman. Strategic uses of narrative in the presentation of self and illness: A research note. *Social science & medicine*, 30(11):1195–1200, 1990.

Catherine Kohler Riessman. Analysis of personal narratives. In James Holstein and Jaber F. Gubrium, editors, *Inside Interviewing: New Lenses, New Concerns*, pages 331–346. SAGE, mar 2003. ISBN 978-0-7619-2851-5.

Kathryn E. Ringland. A place to play: the (dis) abled embodied experience for autistic children in online spaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

Kathryn E. Ringland, Christine T. Wolf, LouAnne E. Boyd, Mark S. Baldwin, and Gillian R. Hayes. Would you be mine: Appropriating minecraft as an assistive technology for youth with autism. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 33–41, 2016.

Sarah Roberts. Digital Refuse: Canadian Garbage, Commercial Content Moderation and the Global Circulation of Social Media's Waste. *Wi: Journal of Mobile Media*, January 2016. URL `https://ir.lib.uwo.ca/commpub/14`.

Sarah T. Roberts. *Behind the Screen: Content Moderation in the Shadows of Social Media*. Yale University Press, 2019.

Jennifer A. Rode. Reflexivity in digital anthropology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 123–132, 2011.

James P. Roehrig and Carmen P. McLean. A comparison of stigma toward eating disorders versus depression. *International Journal of Eating Disorders*, 43(7):671–674, 2010.

Emma J. Rose, Avery Edenfield, Rebecca Walton, Laura Gonzales, Ann Shivers McNair, Tetyana Zhvotovska, Natasha Jones, Genevieve I. Garcia de Mueller, and Kristen Moore. Social justice in ux: Centering marginalized users. In *Proceedings of the 36th ACM International Conference on the Design of Communication*, pages 1–2, 2018.

Gillian Rose. *Visual methodologies: An introduction to researching with visual materials*. sage, 2016.

Paul Rozin, Claude Fischler, Christy Shields, and Estelle Masson. Attitudes towards large numbers of choices in the food domain: A cross-cultural study of five countries in europe and the usa. *Appetite*, 46(3):304–308, 2006.

Nicolas Rüsch, Matthias C. Angermeyer, and Patrick W. Corrigan. Mental illness stigma: Concepts, consequences, and initiatives to reduce stigma. *European Psychiatry*, 20 (8):529–539, December 2005. ISSN 0924-9338. doi: 10.1016/j.eurpsy.2005.04.004. URL `http://www.sciencedirect.com/science/article/pii/S0924933805000908`.

Gerald F. M. Russell. The history of bulimia nervosa. *Handbook of treatment for eating disorders*, 2:11–24, 1997.

Tony Salvador, Genevieve Bell, and Ken Anderson. Design ethnography. *Design Management Journal (Former Series)*, 10(4):35–41, 1999.

Shruti Sannon, Elizabeth L. Murnane, Natalya N. Bazarova, and Geri Gay. ”I Was Really, Really Nervous Posting It”: Communicating About Invisible Chronic Illnesses Across Social Media Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI ’19, pages 353:1–353:13, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300583. URL `http://doi.acm.org/10.1145/3290605.3300583`. event-place: Glasgow, Scotland Uk.

Frank Schaap. The words that took us there. ethnography in a virtual reality. 2002.

Morgan Klaus Scheuerman, Stacy M. Branham, and Foad Hamidi. Safe Spaces and Safe Places: Unpacking Technology-Mediated Experiences of Safety and Harm with Transgender People. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW):155:1–155:27, November 2018. ISSN 2573-0142. doi: 10.1145/3274424. URL `http://doi.acm.org/10.1145/3274424`.

Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. Intersectional hci: Engaging identity through gender, race, and class. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 5412–5427, 2017.

Ari Schlesinger, Kenton P. O’Hara, and Alex S. Taylor. Let’s talk about race: Identity, chatbots, and ai. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 315. ACM, 2018.

Melissa Schrift. Angola prison art: Captivity, creativity, and consumerism. *Journal of American Folklore*, pages 257–274, 2006.

Lynne Schrum. Framing the debate: Ethical research in the information age. *Qualitative Inquiry*, 1(3):311–326, 1995.

James C. Scott. *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press, 1998.

Andrew Scull. Psychiatry and social control in the nineteenth and twentieth centuries. 2(6):149–169, june 1991. ISSN 0957-154X, 1740-2360. doi: 10.1177/0957154X9100200603. URL `http://journals.sagepub.com/doi/10.1177/0957154X9100200603`.

Andrew Scull. *Madness in Civilization: A Cultural History of Insanity, from the Bible to Freud, from the Madhouse to Modern Medicine*. Princeton University Press, 2015.

Joseph Seering, Robert Kraut, and Laura Dabbish. Shaping Pro and Anti-Social Behavior on Twitch Through Moderation and Example-Setting. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 111–125, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4335-0. doi: 10.1145/2998181.2998277. URL `http://doi.acm.org/10.1145/2998181.2998277`. event-place: Portland, Oregon, USA.

Joseph Seering, Tony Wang, Jina Yoon, and Geoff Kaufman. Moderator engagement and community development in the age of algorithms. *New Media & Society*, 21(7): 1417–1443, July 2019. ISSN 1461-4448. doi: 10.1177/1461444818821316. URL `https://doi.org/10.1177/1461444818821316`.

Yukari Seko. Picturesque wounds: A multimodal analysis of self-injury photographs on flickr. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, volume 14, 2013.

Yukari Seko and Stephen P. Lewis. The self—harmed, visualized, and reblogged: Remaking of self-injury narratives on tumblr. 20(1):180–198, jan 2018. ISSN 1461-4448. doi: 10.1177/1461444816660783. URL `https://doi.org/10.1177/1461444816660783`.

Leslie Regan Shade. Weborexics: The ethical issues surrounding pro-ana websites. 2003.

Siobhan Sharkey, Ray Jones, Janet Smithson, Elaine Hewis, Tobit Emmens, Tamsin Ford, and Christabel Owens. Ethical practice in internet research involving vulnerable people: lessons from a self-harm discussion forum study (sharptalk). *Journal of medical ethics*, 37(12):752–758, 2011.

Sins Invalid. Sins invalid website. URL `https://www.sinsinvalid.org`. Accessed: 2020-07-16.

Thomas Smyth and Jill Dimond. Anti-oppressive design. 21(6):68–71, oct 2014. ISSN 1072-5520. doi: 10.1145/2668969. URL `http://doi.acm.org/10.1145/2668969`.

Jaime Snyder, Elizabeth Murnane, Caitie Lustig, and Stephen Voida. Visually encoding the lived experience of bipolar disorder. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

Andreas Sonderegger, Klaus Heyden, Alain Chavaillaz, and Juergen Sauer. Anisam & aniavatar: Animated visualizations of affective states. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4828–4837, 2016.

spez. Update to our content policy, June 2020. URL `https://www.reddit.com/r/announcements/comments/hi3oht/update_to_our_content_policy/`. Accessed: 2020-07-15.

Katta Spiel, Os Keyes, Ashley Marie Walker, Michael Ann DeVito, Jeremy Birnholtz, Emeline Brulé, Ann Light, Pınar Barlas, Jean Hardy, and Alex Ahmed. Queer (ing) hci: Moving forward in theory and practice. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–4, 2019.

Katta Spiel, Kathrin Gerling, Cynthia L. Bennett, Emeline Brulé, Rua M. Williams, Jennifer Rode, and Jennifer Mankoff. Nothing about us without us: Investigating the role of critical disability studies in hci. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2020.

Gayatri Chakravorty Spivak. Can the subaltern speak? *Die Philosophin*, 14(27):42–58, 2003.

Staff. Follow-up: Tumblr's new policy against pro-self-harm blogs, March 2012. URL `https://staff.tumblr.com/post/18563255291/follow-up-tumblrs-new-policy-against`. Accessed: 2020-07-15.

Mel Starkman. The movement. In Brenda A LeFrançois, Robert Menzies, and Geoffrey Reaume, editors, *Mad matters: A critical reader in Canadian mad studies*, pages 27–37. Canadian Scholars' Press, 2013.

Kevin F. Steinmetz. Message received: Virtual ethnography in online message boards. *International Journal of Qualitative Methods*, 11(1):26–39, 2012.

Emma J. Stinson, Susanne B. Votruba, Colleen Venti, Marisol Perez, Jonathan Krakoff, and Marci E. Gluck. Food insecurity is associated with maladaptive eating behaviors and objectively measured overeating. *Obesity*, 26(12):1841–1848, 2018.

P. Stout, J. Villegas, and N. Jennings. Images of mental illness in the media: Directions for future research. *Schizophrenia Bulletin*, 30(3):543–561, 2004.

Strike Your Note. Selfies, hashtags, and the mechanisms of ideology | strike your note, 2014. URL `https://strikeyournote.wordpress.com/2014/05/21/selfies-hashtags-and-the-mechanisms-of-ideology/`.

Albert J. Stunkard. Eating patterns and obesity. *Psychiatric quarterly*, 33(2):284–295, 1959.

Jacqueline Anne Sullivan. Models of mental illness. In Harold Kincaid, Jeremy Simon, and Miriam Solomon, editors, *The Routledge Companion to the Philosophy of Medicine*, pages 455–464. Routledge, 2016.

Malin Sveningsson. Ethics in internet ethnography. In *Readings in virtual research ethics: Issues and controversies*, pages 45–61. IGI Global, 2004.

Heather Anne Swanson. Landscapes, by comparison. In Kei'ichi Omura, Grant Otsuki, Shiho Satsuka, and Atsuro Morita, editors, *The World Multiple: The Quotidian Politics of Knowing and Generating Entangled Worlds*, chapter 7. Routledge, 2018.

Thomas Szasz. *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. Harper, 1961.

T-Maps Community. The origins of this tool. URL `https://tmapscommunity.net/the-origins-of-this-tool/`. Accessed: 2020-07-15.

Erin Tatum. Soft grunge: Mental illness is not a style, December 2013. URL `https://everydayfeminism.com/2013/12/soft-grunge/`. Accessed: 2020-07-14.

Alex Taylor. Life less normal. *Interactions*, 27(4):79–82, 2020.

June M. Tester, Tess C. Lang, and Barbara A. Laraia. Disordered eating behaviours and food insecurity: A qualitative study about children with obesity in low-income households. *Obesity research & clinical practice*, 10(5):544–552, 2016.

Anja Thieme, Cynthia L. Bennett, Cecily Morrison, Edward Cutrell, and Alex S. Taylor. " i can do everything but see!"–how people with vision impairments negotiate their abilities in social contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2018.

Marika Tiggemann and Mia Zaccardo. 'strong is the new skinny': A content analysis of #fitspiration images on instagram. *Journal of health psychology*, 23(8):1003–1011, 2018.

Austin Toombs, Shad Gross, Shaowen Bardzell, and Jeffrey Bardzell. From empathy to care: a feminist care ethics perspective on long-term researcher–participant relations. *Interacting with Computers*, 29(1):45–57, 2017.

Edwin Fuller Torrey and Judy Miller. *The invisible plague: The rise of mental illness from 1750 to the present*. Rutgers University Press, 2001.

Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki. Recognizing depression from twitter activity. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pages 3187–3196, 2015.

Natalie M. Underberg and Elayne Zorn. *Digital Ethnography: Anthropology, Narrative, and New Media*. University of Texas Press, apr 2013. ISBN 978-0-292-74433-2. Google-Books-ID: NqmOeY04twYC.

Theo Van Leeuwen. Semiotics and iconography. *Handbook of visual analysis*, pages 92–118, 2001.

Theo Van Leeuwen and Carey Jewitt. *The handbook of visual analysis*. Sage, 2001.

Walter Vandereycken and Ina Van Humbeeck. Denial and concealment of eating disorders: a retrospective survey. *European Eating Disorders Review*, 16(2):109–114, 2008. ISSN 1099-0968. doi: 10.1002/erv.857. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/erv.857`.

Piia Varis. Digital ethnography. *The Routledge handbook of language and digital communication*, pages 55–68, 2016.

Aditya Vashistha, Abhinav Garg, Richard Anderson, and Agha Ali Raza. Threats, abuses, flirting, and blackmail: Gender inequity in social media voice forums. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 72. ACM, 2019.

Jessica Vitak, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab. Identifying women's experiences with and strategies for mitigating negative effects of online harassment. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 1231–1245. ACM, 2017.

Shirin Vossoughi and Meg Escudé. What does the camera communicate? an inquiry into the politics and possibilities of video research on learning. *Anthropology & Education Quarterly*, 47(1):42–58, 2016.

Otto F. Wahl. *Media madness: Public images of mental illness*. Rutgers University Press, 1995.

Ashley Marie Walker, Michael Ann DeVito, Juan Fernando Maestre, Katie A. Siek, Cassie Kresnye, Ben Jelen, Patrick C. Shih, Maria Wolters, and Mona Alqassim. Arc: Moving the method forward. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–4, 2019.

Evelyn Wan, Aviva de Groot, Shazade Jameson, Mara Păun, Phillip Lücking, Goda Klumbyte, and Danny Lämmerhirt. Lost in translation: an interactive workshop mapping interdisciplinary translations for epistemic justice. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 692–692, 2020.

Rui Wang, Min S. H. Aung, Saeed Abdullah, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Michael Merrill, and Emily A. Scherer. Crosscheck: toward passive sensing and detection of mental health changes in people with schizophrenia. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 886–897, 2016.

Rui Wang, Weichen Wang, Min S. H. Aung, Dror Ben-Zeev, Rachel Brian, Andrew T Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, and Emily A. Scherer. Predicting symptom trajectories of schizophrenia using mobile sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–24, 2017a.

Tao Wang, Markus Brede, Antonella Ianni, and Emmanouil Mentzakis. Detecting and characterizing eating-disorder communities on social media. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 91–100. ACM, 2017b.

Kate Werning. Practice: Mad maps with the icarus project's rhiana anthony, April 2019. URL `https://irresistible.org/podcast/35penglish`. Accessed: 2020-07-15.

Michael Westerlund. The production of pro-suicide content on the internet: a counter-discourse activity. *New Media & Society*, 14(5):764–780, 2012.

Rua M. Williams and LouAnne E. Boyd. Prefigurative politics and passionate witnessing. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, pages 262–266, 2019.

Sarah Wilson. From the asylum to the museum. marginal art in paris and new york, 1938-1968. In *Parallel Visions, Modern Artists and Outsider Art*, pages 120–149. Los Angeles County Museum of Art, 1992.

Langdon Winner. Do artifacts have politics? *Daedalus*, pages 121–136, 1980.

Donghee Yvette Wohn. Volunteer moderators in twitch micro communities: How they get involved, the roles they play, and the emotional labor they experience. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 160. ACM, 2019.

Donghee Yvette Wohn, Casey Fiesler, Libby Hemphill, Munmun De Choudhury, and J. Nathan Matias. How to handle online risks?: Discussing content curation and moderation in social media. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 1271–1276. ACM, 2017. ISBN 978-1-4503-4656-6. doi: 10.1145/3027063.3051141. URL `http://doi.acm.org/10.1145/3027063.3051141`.

Markus Wolf, Florian Theis, and Hans Kordy. Language use in eating disorder blogs: Psychological implications of social online activity. *Journal of Language and Social Psychology*, 32(2):212–226, 2013.

Maria K. Wolters, Zawadhafsa Mkulo, and Petra M. Boynton. The emotional work of doing ehealth research. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 816–826, 2017.

Alice Wong. Ep 48: Care work. URL `https://disabilityvisibilityproject.com/2019/04/07/ep-48-care-work/`. Accessed: 2020-07-16.

David Wood. Foucault and panopticism revisited. *Surveillance & Society*, 1(3):234–239, 2003.

Peter Wright and John McCarthy. *Experience-Centered Design: Designers, Users, and Communities in Dialogue*. Morgan and Claypool Publishers, 2010. ISBN 978-1-60845-044-2.

Diyi Yang, Zheng Yao, Joseph Seering, and Robert Kraut. The Channel Matters: Self-disclosure, Reciprocity and Social Support in Online Cancer Support Groups. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 31:1–31:15, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-5970-2. doi: 10.1145/3290605.3300261. URL `http://doi.acm.org/10.1145/3290605.3300261`. event-place: Glasgow, Scotland Uk.

Anon Ymous, Katta Spiel, Os Keyes, Rua M. Williams, Judith Good, Eva Hornecker, and Cynthia L. Bennett. " i am just terrified of my future"—epistemic violence in disability related technology research. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2020.

Elad Yom-Tov, Luis Fernandez-Luque, Ingmar Weber, and Steven P. Crain. Pro-anorexia and pro-recovery photo sharing: a tale of two warring tribes. *Journal of medical Internet research*, 14(6):e151, 2012.

Kimball Young. *Sociology: A Study of Society and Culture*. American Book Company, 1949.

Renwen Zhang, Jordan Eschler, and Madhu Reddy. Online support groups for depression in china: Culturally shaped interactions and motivations. 27(3):327–354, dec 2018. ISSN 1573-7551. doi: 10.1007/s10606-018-9322-4. URL `https://doi.org/10.1007/s10606-018-9322-4`.

APPENDIX A

Expressions of Mental Health Conditions Interview Guide

Thank you for participating in our interview today. We are interested in learning more about the ways you express and share your experiences with mental health. There are no right or wrong answers, and you may take a break or stop the interview at any time. Do you have any questions before we get started?

What are your experiences with mental health?

What types of creative processes do you engage with to express these experiences? [Probe: Individual or group activities, types of expression (e.g., writing, painting, sketching, sculpture, memes), time of process in relation to mental health experience or condition, length of process.]

Why do you engage with creative expression of your mental health experiences?

For this interview, we're focusing on expression of mental health experiences. Prior to this interview, you shared several examples of artifacts you produced that were related to your experiences with mental health. We'll talk about these. [For each artifact shared, ask the following questions.]

Describe the artifact you created. [Probe: Original, recycled, or remixed.]

Walk me through, step-by-step, how you decided to create this artifact, and the process that followed. [Probe: Tools, mediums, and technologies used; aesthetic and style; length of time; when the artifact was generated in relation to mental health experiences; involvement of others; feelings at beginning, end, and during the process.]

Did you experience any challenges to the creation of this artifact? Please explain.

What does this artifact mean to you or represent?

Why did you create this artifact? [Probe: Emotions tied to creative process and artifact (e.g., How do you feel about what you created?); reasons for artifact creation (e.g., expression, distraction, reflection).]

What did you do following the creation of this artifact? [Probe: sharing, archiving, photographing, destroying.]

Did you share this artifact with anyone? [Probe: Online and offline.]

Why did you share it?

Were you hoping to communicate or express your experience to others? Please explain.

Do you feel this artifact communicates your experience? [Probe: Why or why not? If "why not," ask why the artifact doesn't communicate experience and how it might be modified to do so. Also, ask why the artifact hasn't been modified to facilitate communication of experience.]

Did you have any expectations around how others would communicate or interact with you following viewing this visual artifact?

What was the response you received after sharing? [Probe: Who responded? How did they respond? How did their reaction make you feel? Did you respond to them? When and through which medium? Did anyone else respond to them?]

Have you ever commented or interacted with an artifact that someone else has created and shared (e.g., Instagram post, painting, writing)? Please explain.

Thank you for your responses.

Are there any other questions you think I should be asking you about how you express your experiences with mental health?

APPENDIX B

Content Moderation and Eating Disorders Interview Guide

Thank you for participating in our interview today. We are interested in learning more about the ways you have been personally impacted by online platform content moderation practices. These practices include the deletion and reporting of posts, as well as the banning of accounts and communities. There are no right or wrong answers. You may take a break or stop the interview at any time. Do you have any questions before we get started?

What is your experience with eating disorders? [Probe: diagnosis or self-applied label, length of time, current or past. Also explore this experience with respect to identity – both how they identify and how they perceive others as identifying them.]

Which online communities and social media do you belong to?

How long have you been a member of [Insert name of community/ies mentioned by participant]?

Do you currently post or read online social media or community content related to eating disorders? [Probe: Specific sites or forums, community belonging, individual accounts, lurking and posting practices, types of content, secret accounts.]

[If participant does not currently engage with this content, ask the previous question in the context of past practices.]

For this interview, we're focusing on moderation of content related to eating disorders. This means the deletion, ban, or reporting of posts, accounts, and online communities or forums.

Have you ever had personal content, including posts and accounts, reported or deleted? Please tell me about the content. [Probe: Specific sites, type of content, history and personal meaning of the content (i.e., why something was posted or why was account was started), when this happened, how long after the content was posted was it removed, who or what reported or removed the content, how they became aware the content was reported or removed, why they thought the content was removed.]

What was this like for you? [Probe: Type of moderation, initial thoughts and feelings, later reflection, feelings of being cut off from friends, if their contact information was lost, impact on identity – digital and otherwise.]

What did you do? [Probe: actions the participant took (e.g., reposting content, not posting that type of content anymore, making a new account, joining other sites, requesting for content to be restored, reaching out to others, disengaging with communities)]

How do you feel about this experience now? [Probe: Online identity, online community belonging, impact on experience with an eating disorder and general mental health]

Is there another way you would have like the content to be handled by the platform?

Have you ever self-moderated your own content?

Have you ever had experience as a member of a community, where a community was banned or reported? Please describe the community. [Probe: Specific sites, type of content on site, size of the site, how active the user was, when the community was banned or reported, who reported ore removed the content, how they became aware the content was reported or removed.]

What was this experience like for you? [Probe: Type of moderation, initial thoughts and feelings, later reflection, actions the participant took (e.g., joining other sites, creating new sites or ways of contacting prior members, requesting for content to be restored, reaching out to other community members (or had members reach out to them), disengaging with communities)]

How do you feel about this experience now? [Probe: Online identity, online community belonging, impact on experience with an eating disorder and general mental health] Is there another way you would have like the community to be handled by the platform?

Have you ever participated as a moderator on one of these online forums? [Probe: Moderator responsibilities and actions, such as deleting content and making communities private]

If you could say anything to people who report content, what would you say?

If you could say anything to the platforms moderating content, what would you tell them? [Probe for differences between platforms.]

What content would you like to moderate?

How would you moderate this content?

How would you deliver the news that content had been moderated? [Probe: transparency, platform accountability, and recourse – e.g., getting a post back.]

If you were forming an online community for people with eating disorders, what moderation rules would you create? [Also probe for: How would you organize and run the site, and why?]

Thank you for your responses.

Are there any other questions you think I should be asking you about your experience with content moderation?