"Dialing it Back:" Shadowbanning, Invisible Digital Labor, and how Marginalized Content Creators Attempt to Mitigate the Impacts of Opaque Platform Governance

SENA A. KOJAH, School of Information, University of Michigan, USA
BEN ZEFENG ZHANG, School of Information, University of Michigan, USA
CAROLINA ARE, Centre for Digital Citizens, Northumbria University, United Kingdom
DANIEL DELMONACO, Rutgers University Libraries, Rutgers, The State University of New Jersey, USA
OLIVER L. HAIMSON, University of Michigan, USA

Content creators with marginalized identities are disproportionately affected by shadowbanning on social media platforms, which impacts their economic prospects online. Through a diary study and interviews with eight marginalized content creators who are women, pole dancers, plus size, and/or LGBTQIA+, this paper examines how content creators with marginalized identities experience shadowbanning. We highlight the labor and economic inequalities of shadowbanning, and the resulting invisible online labor that marginalized creators often must perform. We identify three types of invisible labor that marginalized content creators engage in to mitigate shadowbanning and sustain their online presence: mental and emotional labor, misdirected labor, and community labor. We conclude that even though marginalized content creators engaged in cross-platform collaborative labor and personal mental/emotional labor to mitigate the impacts of shadowbanning, it was insufficient to prevent uncertainty and economic precarity created by algorithmic opacity and ambiguity.

CCS Concepts: • **Human-centered computing → Collaborative and social computing**; **Empirical studies in collaborative and social computing**; **Collaborative and social computing theory, concepts and paradigms**; **Social media**;

Additional Key Words and Phrases: Shadowbanning, invisible labor, content creator collaboration, marginalized identities, content moderation

ACM Reference Format:
Sena A. Kojah, Ben Zefeng Zhang, Carolina Are, Daniel Delmonaco, and Oliver L. Haimson. 2025. "Dialing it Back:" Shadowbanning, Invisible Digital Labor, and how Marginalized Content Creators Attempt to Mitigate the Impacts of Opaque Platform Governance. *Proc. ACM Hum.-Comput. Interact.* 9, 1, Article GROUP12 (January 2025), 22 pages. https://doi.org/10.1145/3701191

1 INTRODUCTION

Shadowbanning is a cross-platform light, insidious social media censorship technique whereby platforms avoid recommending or outright hide, content and profiles from their main discovery feeds, greatly affecting the visibility, earnings, mental health, and communications of the users

Authors' Contact Information: Sena A. Kojah, School of Information, University of Michigan, Ann Arbor, Michigan, USA, senami@umich.edu; Ben Zefeng Zhang, School of Information, University of Michigan, Ann Arbor, Michigan, USA, bzfzhang@umich.edu; Carolina Are, Centre for Digital Citizens, Northumbria University, Newcastle, Newcastle, United Kingdom, carolina.are@northumbria.ac.uk; Daniel Delmonaco, Rutgers University Libraries, Rutgers, The State University of New Jersey, Newark, New Jersey, USA, dan.delmonaco@rutgers.edu; Oliver L. Haimson, University of Michigan, Ann Arbor, Michigan, USA, haimson@umich.edu.

ACM 2573-0142/2025/1-ARTGROUP12
https://doi.org/10.1145/3701191

affected [7, 16, 30]. Known to particularly affect those who conduct work through social media - e.g. content creators, small businesses, and marginalized communities such as LGBTQIA+ and BIPOC accounts [25, 30] - shadowbanning is a controversial aspect of platform governance in continuous development.

Shadowbanning is a form of content moderation, but while content moderation typically is thought of as a removal of social media content or accounts, shadowbanning instead refers to content reduction or demotion, often accomplished by opaque algorithms [36]. Because it is difficult to detect or confirm, shadowbanning is less risky for platforms than removing content, yet it still enables platforms to maintain substantial power over users and content [36]. Shadowbanning is at heart a mechanism to adjust one's social media visibility in hidden and inscrutable ways [55], and although not as damaging as outright content removal, it can greatly affect those whose work depends on visibility, such as content creators [37].

To highlight the importance of users' experiences in interrogating opaque platform governance, and to examine the myriad types of hidden labor that shadowbanning may require for creators, we investigated the following research questions:

RQ1: How do marginalized content creators attempt to navigate posting content when trying to avoid shadowbanning?

RQ2: How might shadowbanning create different forms of hidden and unpaid labor affect marginalized content creators' wellbeing?

RQ3: How does shadowbanning affect marginalized content creators' earnings and wellbeing?

Through a diary study and interviews with eight marginalized content creators who are women, pole dancers, plus size, and/or LGBTQIA+, this paper examines how content producers from marginalized communities experience shadowbanning. Building on previous work on shadowbanning's impact on sex workers [7], pole dancers [30], influencers, and LGBTQIA+ individuals [16], this paper showcases how marginalized content creators navigate this opaque but often distressing censorship technique to highlight its economic and wellbeing impacts on those who need social media platforms not just to express themselves, but also to work. To this end, this paper is structured as follows: we begin with the latest research and developments related to shadowbanning across digital platforms, and then delve into the context of the type of employment we are observing - the creator economy. We draw from studies on creator precarity [6, 24, 37], research on the 'folk theories' [28] users develop to explain algorithms' decisions, and [19] idea of invisible labor - the unpaid and de-valued work women are often expected to perform to enable societal development. We then discuss our methods and analysis. Next, we present our data through a thematic analysis focusing on three primary themes prevalent in our data: the inequalities of shadowbanning, particularly when it comes to the moderation of content by marginalized creators and then we explore the hidden labor that that content creators find themselves performing to circumvent or tackle shadowbanning as well as circumvention and mitigation strategies they develop independently and through online communities.

We conclude that shadowbanning has substantial impacts on content creators at the margins, affecting their financial earnings and their well-being, leading them to have to perform different types of invisible labor to circumvent it. These types of labor include mental and emotional labor, misdirected labor, and community labor. Our findings expand on prior research into shadowbanning by highlighting labor that flies under the radar when creators rely on collaborative labor and revealed that even though marginalized content creators engaged in cross-platform personal and collaborative labor to varying degrees to mitigate the impact of shadowbanning, it was never sufficient to prevent economic precarity on social platforms created by algorithmic ambiguity.

2 BACKGROUND AND RELATED WORK

2.1 The History of Shadowbanning

Although platform governance techniques akin to shadowbanning have been used in Internet forums since the 1970s, the origin of the term seems to date back to 2001, when moderators on the Something Awful website began to limit the reach of those joining forums to harass others [14]. The hiding or lack of recommendation of content known as shadowbanning has been used by various social media platforms, including Twitter, which prevents shadowbanned accounts' usernames from appearing in search [67]; TikTok, which has been accused of hiding Black creators from its For You page [50]; and, notoriously, by Instagram, which has had to apologize to swathes of users for restricting the reach and views of their content [30]. A key characteristic of shadowbanning is that, often, platforms do not notify users it is happening (ibid) - a benefit and a hindrance. Avoiding notifying those spreading harmful views or content that their posts are being restricted is useful in running digital spaces. Yet the lack of communication is also frustrating and, often, damaging for users, who are made to think that no action is being taken on their content, and that their posts' lack of views is strictly due to their failure to create engaging posts - a practice dubbed 'blackbox gaslighting' [16].

Content moderation, including shadowbanning has been found to disproportionately affect marginalized communities [25, 41], and in particular those social media users who post nudity, who work as sex workers, or who are part of the LGBTQIA+ community [7, 30, 60]. Marginalized users' content suppression has often been linked to the approval of FOSTA/SESTA by the United States' Congress.

A 2018 exception to Section 230 of the US Telecommunications Act 1996, FOSTA/SESTA – the Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA) and the Stop Enabling Sex Traffickers Act (SESTA) - made social media platforms legally liable for facilitating sex trafficking (a crime) and sex work (a job) [9]. By lumping sex trafficking in with sex work, FOSTA/SESTA led platforms to protect themselves from liability by over-censoring swathes of content revolving around nudity and sexual expression, leading to censorship of content and profiles by sex workers, athletes, sexual health brands, sex educators and activists worldwide [12, 30, 42, 60]. Shadowbanning is one of those censorship techniques, often administered in relation to increasingly sex-averse community guidelines focused on nudity, sexual activity and solicitation [8, 9, 60].

The impact of FOSTA/SESTA has been found to be devastating, not just on sex workers (ibid), but on women's bodies and on LGBTQIA+ expression as a whole: after the law's approval, increasing academic work and journalistic reporting have highlighted how women's and LGBTQIA+ users' bodies are coded as inherently sexual by platforms [31, 58]. With this inherently sexual dimension comes an assumption that these users are dangerous, and do not deserve visibility [30, 60]. Another important consideration is that FOSTA/SESTA is primarily a US-based law, yet social media platforms' deployment and shadowbanning's application is usually global, thus impacting creators all over the world who are not subject to US laws, particularly in the Global South. Internet governance as it exists today is therefore undemocratic, reinforcing global asymmetries, and adding a new dimension to the North-South divide [75].

2.2 Shadowbanning: Issues, Developments, Questions

Research on shadowbanning takes place within the platform governance field, which involves research and activism comprising of important questions on the impact that a platform features, as well as regulatory policy, have on the freedoms of users and liability of platforms [73]. However, precisely because we are discussing features and issues pertaining to fast-moving digital platforms, we are faced with the challenge of rapid changes in platform policy and infrastructure beyond the

time of writing and publishing. As such, we must address some of the most relevant updates and controversies with regards to shadowbanning, which come not only from academic research, but from platforms' own communications as well.

One of the most notorious examples of shadowbanning was on Instagram, which previously had to apologize to pole dancers [30], admitting that a form of light suppression of hashtags and content was taking place [56], although Instagram's CEO subsequently stated that 'shadowbanning is not a thing' [16]. The platform then admitted that they do limit the visibility of specific content, such as content by Black creators and content showing people's skin and bodies [30]. In recent times Instagram's notorious censorship and de-prioritising of content depicting bodies has come under the scrutiny of Meta's own Oversight Board, which overturned the platform's ban of posts depicting transgender and non-binary nudity, criticizing its often confusing and discriminatory moderation of bodies and calling for clearer, fairer nudity governance that is consistent with international human rights and self-expression [10].

A Guardian investigation found that algorithms by major companies such as Google and Microsoft, which are often then utilized by mainstream social media platforms, code women and LGBTQIA+ users' bodies as inherently sexual, even in situations depicting exercise, healthcare or maternity [58], showing the strength of platform governance against content depicting skin.

Following repeated backlash against shadowbanning, in December 2022 Instagram announced they were going to notify users of this content moderation technique [34]. Following this move, Instagram users became able to see whether their content and/or profiles had become 'non-recommendable' within the settings and Account Status area of their profiles, where they could also appeal the decision [34].

It has been argued that Instagram's disclosure of their recommendations guidelines has mitigated one problem – the lack of transparency and gaslighting of users [16]. However, it has also effectively only rebranded shadowbanning (a user-generated term that Instagram did not use [30] into 'non-recommendable' [30]. Thus, while Instagram may now be disclosing their policies, the effect is the same: swathes of content, and particularly content featuring nudity and LGBTQIA+ expression, is often disproportionately 'non-recommendable' (ibid).

Aside from these recent events and changes, shadowbanning as we know it may change or decrease in some parts of the world after the approval of the European Union's Digital Services Act (DSA), which is the first major legislation to regulate the transparency of visibility remedies [55], making shadowbanning a legal matter. Leersen [55] argues that the DSA's due process requirements to notify users about what affects their content result in a general prohibition on shadowbanning, with very few exceptions for high-volume deceptive commercial content.

2.3 Shadowbanning as Invisible Labor in the Creator Economy

Shadowbanning greatly impacts the creator economy on social media; it directly affects creators' income, network, and visibility [30, 37]. Content creators and brands establishing relationships and promoting products through social media are part the 'gig' or ; 'platform economy', a form of work characterized by insecure, often short-term or piecemeal, employment, frequently facilitated by a platform or app. [27].

The creator economy includes influencer partnerships, social media self-promotion of services provided offline, and different forms of sex work [2, 30]. Just like other forms of gig work, content creation offers flexibility without the protection and benefits of traditional employee status ibid; [27]. Yet, despite providing different avenues for creativity and creating flexible, aspirational work opportunities for people previously excluded from cultural production and from a shot at fame and wealth [4, 25, 37], the creator economy has also resulted in increased precarity across the creative industries [37]. Since the digitization of labor has brought forth forms of 'ambiguous

work' that blur the boundaries between life and work [17], platform-mediated work calls for new conceptualizations of precarity that are not just industry-related, but also focused on the individual [35]. In this sense, individuals' identities, vulnerabilities, and backgrounds, which are often a key part of the carefully crafted online profiles they create, can deeply affect their experiences of digital precarity and censorship [2].

On social media, precarity is connected to a complex ecosystem of numbers: creators' success is often outside their control, and tied to an elusive visibility [24, 25, 37] measured through the accumulation of metrics such as likes, followers, and engagement [4] which Glatt calls "popularity metrics" [37]. Content views, channel subscriptions, and followers assign legitimacy to users and are therefore becoming increasingly sought-after since depending on the number of views and subscriptions that a channel achieves, it is accordingly deemed relevant and attention-worthy or not [33]. Yet, social media platforms do not guarantee such visibility, which is afforded sparsely and arbitrarily according to compliance with platforms' often broad, unclear, and unequally applied rules, as well as their processes and agendas [30, 60]. Views are governed by algorithms, codified step-by-step processes through which platforms afford or restrict visibility [6]. This opaque, inscrutable "algorithmic boss" is often viewed as the main entity responsible for the precarity of creator labor, an entity that can wreak havoc on creators' lives and livelihoods [24]. This job insecurity triggers financial anxiety, which negatively impacts well-being [23, 29] in a field that already provides inadequate healthcare and lacks most forms of organizational support [78], creating a notable power differential between platforms and their workers [2] [30].

Precarity of, at, and from work, and the opacity of social media governance, in turn, produce invisible labor, requiring users to perform a range of unpaid tasks to guess what may lead to success at gaining popularity metrics. The concept of invisible work was previously used by Arlene Daniels [20] to define women's work, which is often unpaid, invisible, and devalued. Through such invisible labor, Daniels [20] argued, women help construct daily life and maintain and/or develop institutions.

Human computation tied to economic activity relies on workers' invisibility [45]. Mainstream platforms do not recognize the value in the labor that users and stakeholders are performing to repair algorithmic failings even though these acts make the platform economy viable. [18]

Much of content creators' digital labor [49] can be viewed not only as invisible labor, but also as immaterial labor, which refers to affective labor that is undervalued and often not considered actual labor, yet nonetheless commodified [54] (in this case by social media platforms). That is, though creating digital content is not often viewed as work, it actually generates substantial financial benefit for platforms [49] - a form of online "free labor" [72]. Further, marginalized creators, who are often not recognized in the more mainstream influencer industry, must rely on their followers for income, which requires a substantial amount of relational labor [37]. When mitigating shadowbanning falls on top of this existing work, it thus requires an increase in immaterial, invisible, digital labor. Building on this line of scholarship, our study uses shadowbanning as a context to explore various forms of hidden and unrecognized labor that content creators engage in to make sense of, circumvent, and support each other, and to highlight the emerging harms that stem from shadowbanning in their daily content creation practices.

It can be argued that content creators working through social media have to perform a new kind of invisible labor, which is unpaid but strictly tied to their survival on platforms mitigating shadowbanning.

Platform users are not passive recipients of platform governance; they employ creative strategies to appropriate or work around issues [48]. On social media, this invisible labor takes the shape of attempting to reverse-engineer platform governance: inevitably, users in precarious working conditions seek to find out more about how algorithms, shadowbanning and content moderation

work to mitigate their powerlessness and lack of agency [28, 31]. To this end, social media users often find themselves developing and sharing 'folk theories' about how algorithms work, so they can plan their work around them [28], a behavior that Bishop called "algorithmic gossip," or a set of "communally and socially informed theories and strategies pertaining to recommender algorithms" [6]. Algorithmic systems are unstable due to the dynamic nature with which social media platforms change their values/morality; this shifting moral space means that users increasingly have trouble gaining control over systems that render them visible and that are tied to their identities [40, 65].

The unpaid cooperative labor that content creators perform to understand opaque platform governance policies or to circumvent them are similar to the pragmatic stance that people in use in other contexts when their labor is tied to economic activity [52]. For instance, in a study about co-operatives (member-owned organizations) and technology, Lampinen et al. [53] found that there is a reciprocal cost of participation in systems where people rely on cooperative labor, which manifests as an interactional cost. Similarly, gig workers often develop folk literacies by engaging, experimenting, and sharing ideas with each other or more experienced workers to learn the behavior of opaque algorithms and gain a competitive economic advantage by working with or around a platform to negotiate space for themselves [47, 59, 68] .

Folk theories and algorithmic gossip are a known form of protection and defense, particularly for women and marginalized communities [6] [22]. However, they are also a form of knowledge-generation which is not recognised, and sometimes even fought or undermined, by platforms themselves: as Cotter [16] has argued, the opacity of platforms' algorithms and their governance, and the technicalities behind technological infrastructure, enable platforms to mislead users about content moderation and suppression processes and their effects. Yet, as Savolainen [63] argues, this 'algorithmic folklore' does not always result in successfully gaming the algorithm and thriving on platforms, making invisible digital labor all the more disheartening for users.

3 METHODS

To understand how content creators with marginalized identities experience shadowbanning as a form of content moderation, we conducted a four-week diary study (n = 7) and then interviews (n = 8), with seven of the diary study participants and one additional participant) during two different stages - the first one before Instagram released its aforementioned Account Status feature [34] and the second one several months after users were able to evaluate their experience with it. We also conducted an additional round of follow-up interviews with seven participants to understand in greater depth the labor they perform to make sense of and circumvent shadowbanning, which they had discussed in the prior interviews and diary study. Dairy studies typically use a periodic data collection temporality that can provide unique insight into participants' day-to-day experiences, actions, patterns, and emotional state as they play out and develop across time [69]. The diary is a feminist practice and a way of discovering new combinations of meaning in an unfamiliar process that the research has no part in censoring [44]. Diary studies collect participants' descriptions of their feelings and behaviors in social contexts and real-life situations [46]. They are a way of documenting specific experiences in a way that is not possible using traditional study designs [11]. Diary studies are tailored to collect real-time data recorded by participants as subjective experiences; as they occur in natural and spontaneous contexts, they reveal insights into how people ascribe meanings to actions and social interactions [76].

This method of data collection allows the participant to communicate without a mediator in an accessible way that is non-hierarchical, accessible, and subversive to patriarchal social norms of literature and data that promotes a ranked system of what is acceptable data using arbitrary standards [44]. This reflective process gives participants agency in deciding the quantity and quality of information they wish to divulge, which is an ethical practice, especially for vulnerable and

marginalized populations [69]. In our study, we invited the participants to introspect and share their thoughts with us about the activities on their social media page and how the reality of content suppression and their attempts to circumvent it plays out in their minds on a day-to-day basis. Their responses, although brief, reveal the hope, frustration, and resignation that marginalized content creators experience when they are shadowbanned on social media platforms and face economic precarity they experience as a result of the bans. Our study was approved by our university's ethics review board.

3.1 Recruitment and Participants

We recruited participants for both the diary study and the interviews by posting our recruitment materials on social media (Instagram, TikTok, and Twitter). Specifically, we recruited from the third author's specialized community network on Instagram and TikTok as well as by posting on our personal Twitter accounts. Due to the niche nature of the participants we sought, we did not post in groups.

Our criteria required that participants believed they had experienced shadowbanning in the past year, belonged to a marginalized group (e.g., LGBTQIA+ users), and posted content on social media either as part of a brand partnership or to promote their own brand or small business. We used a screening survey to assess participant interest and gain basic demographic information, then chose a diverse sample of participants from those who expressed interest in participating. Participant demographics were as follows. Regarding gender, 50% of those who took part were women (three cisgender and one trans), 37.5% men (two cisgender and one questioning), and 12.5% nonbinary trans. In terms of race, 25% of participants were white, 25% Black, 25% Asian, and 25% mixed race. For sexual orientation, 25% of participants were lesbian, and 12.5% (one each) were queer, bisexual, straight, and allosexual, while two participants did not disclose. Participants' ages ranged from 23-42 (mean: 29.4, standard deviation: 7.5). Most participants (62.5%) were living in the US, with one each in Canada, Australia, and Greece.

The identities of the participants fall under three broad categories: racial/ethnic minorities, gender minorities, and minorities based on sexual orientation. We excluded people who were not from one or more marginalized groups (racial/ethnic/gender/sexual minorities).

The small sample size influenced our findings because we did not hear from a wide range of people, and so particular perspectives may have been overrepresented in our data, analysis, and results. Also, the results are not generalizable to content creators more broadly, or even marginalized content creators; but this was not the goal of our work. Rather, we set out to understand particular types of hidden and unpaid labor that marginalized content creators carry out when their content is suppressed using ethnographic approaches.

3.2 Diary Study Design and Data Collection

In our diary study, we asked participants to document their experiences with shadowbanning, their efforts to circumvent it, and how it impacted their socioeconomic life as creators. Our prompts asked participants about six main areas of their social media experiences throughout each of the four weeks of data collection: 1) a list of all social media platforms they used 2) questions about actions and strategies that they took weekly to increase their content's visibility to their followers and in attempts to reach new audiences 3) reflections on actions they took to avoid censorship or suppression of their social media posts 4) reflections on their post performance during the week to gauge engagement metrics 5) thoughts about why they received that level of engagement on their content, and 6) their experiences with engagement, content moderation, and content suppression, including emotional processes as they encountered it. We chose these prompts to

understand participant observation and behavior in situ as a way of minimizing researchers' effect on participants and how they reported their experiences [13]. In addition to answering our specific questions weekly, we also provided each participant with an optional editable document so that they could document supplementary thoughts and additional context related to their posting habits and content moderation and shadowbanning experiences.

3.3 Interviews

We conducted interviews in late 2022, 2023, and in 2024, through an ethnographic framework intended to complement data already collected through the diary study. Interviews were conducted in English, online via Zoom, and lasted between 44 to 77 minutes (*mean: 60.5*). All interviews were recorded and transcribed for data analysis. The questions were derived from the diary study data and our research questions. At the start of every interview, we informed participants about the study's objectives and asked for informed consent to record the conversation and use the data for the purpose of this research. We asked participants questions about their social media use across platforms, their typical engagement levels, their behavior to avoid censorship or suppression, their perceived relationship between their identity (e.g women, sex educators) and their engagement levels online, how they defined and experienced shadowbanning, how their experiences might compare to the experiences of people with more normative identities with similar content and their experiences with the financial implications of shadowbanning.

Following the first round of interviews and diary study, the research team observed and identified emerging themes related to content creators' necessary and hidden labor caused by shadowbanning. To learn more about the emergent themes as well as the participants' situated experiences, we contacted each of the initial participants for follow-up interviews and received responses from seven participants.

3.4 Data Analysis

Following data collection, we conducted qualitative data analysis. First, we used open coding, an interpretive process where data is broken down analytically with the aim of providing insights into ways of interpreting the phenomena generated from the data [15]. For both the interviews and the diary study text, we identified codes that embodied marginalized content creators' experiences with content moderation and shadowbanning when they used social media. The process of line-by-line open coding enabled us to closely examine each part and identify dimensions of the emerging phenomena [64] - in this case, participants' relationship with shadowbanning from a content creator standpoint. Next, we collaboratively conducted axial coding on a Miro[1] board.

During axial coding, the relationships of categories and subcategories generated during open coding are tested against data as a way of verifying the themes as plausible or not [15]. Here, we collaboratively examined the codes and grouped them into themes using parameters outlined by [15]. During open and axial coding, we carefully considered the conditions that gave rise to shadowbanning, the contexts where shadowbanning occurred, interactions (or lack thereof) that triggered content moderation, and the consequences and implications on content creators. After deliberation and discussion during the collaborative coding process, we agreed on six major themes covering the scope of our study. Of these, our paper focuses on two: the inequalities of content moderation on creators with marginalized identities, and the invisible labor that these creators are undertaking to mitigate and circumvent shadowbanning.

4 LIMITATIONS

Our study highlights the experiences of marginalized content creators and so our recruitment attracted a wide range of people with diverse ethnicities, genders, and sexual orientations. However,

we acknowledge that the majority of participants live in the Global North and primarily speak English and that moderation policies, especially those that target marginalized populations, are deployed globally. Our research is limited in understanding Global South experiences - an important area for future content moderation research. The limitation of the diaries in this study is that participants provided relatively short, punchy sentences to the prompts but did not expand on them, as compared to the more in-depth interviews. Yet the diary entries provided unique insights regarding marginalized social media users' everyday experiences with shadowbanning.

5 RESULTS

Content creators with marginalized identities experience disproportionate content moderation [41]and engage in different types of invisible labor to circumvent it. In this section, we explore the inequalities that this study reveals and the hidden labor of shadowbanning mitigation that the participants in this study engaged in to circumvent and manage their content online.

5.1 "Dialing it back": The Inequalities of Shadowbanning Experiences of Marginalised Populations with Shadowbanning

The content creators who participated in this study perceived that people with marginalized identities (e.g., women, pole dancers, and people whose bodies do not fit mainstream beauty standards) were more affected by shadowbanning and other forms of censorship than men who created similar content, and than people who fit conventional beauty standards.

The gender disparities in content moderation were particularly acute across various platforms, with women pole dancers in particular perceiving that men who posted pole dancing content were not as heavily shadow-banned or censored. For instance, P9, a woman from California, shared the following, comparing her content with that of similar male creators: *"I mean, I have also worked [with] and posted images of male models or male performers, but there's a huge discrepancy in the engagement"*. P6, a female pole dancer from Sydney, Australia, said male pole dancers get to express themselves fully – without fear of censorship – wearing different types of clothing that would be considered provocative when a woman pole dancer wears the same:

> *I also wear stilettos as well, but they [men pole dancers] would be wearing full fishnets and super raunchy, like, just a bikini brief, dancing… there was that challenge where I did a trick, and then I would have my crotch quite close to the camera. Theirs would probably be way more, and then I don't hear anyone get censored for that.*

In these instances, from participants' perspectives, gender seemed to substantially impact whether or not content was removed from a social media site, which they considered inequitable. This inequity translates to the labor that participants say they have to undertake, which they tie directly to their gender, body features and size.

In another example, P5 a trans woman living in Chicago, said that she spent extra time and consideration in curating the content that she posts to eliminate aspects of her sexuality that Instagram's moderation might deem problematic: "*On Instagram, I have to be extremely deliberate about posting something that is censored or cropped or whatever in a way where there is not even a hint of the sexuality that they seem to especially flag my account for"*. This additional time and effort that creators with marginalized identities say they spend self-censoring is perceived as inconsistent and unequal.

In weeks 1 and 2 of the diary study, P7, who is trans and non-binary, said they experienced less than average engagement on their post and had to use a custom hashtag as a way to confirm that they had been shadowbanned. P7 had to spend time devising a formula to test their suspicions and

then formulate a plan to circumvent shadowbanning, which involved withdrawing their visibility from the platform for a few days:

> *I didn't post much this week because I suspected my account was being shadowbanned. I use a custom hashtag to verify when I feel I am being shadowbanned. When my higher engagement posts are at the bottom of the hashtag search screen over reposts, it seems to confirm my suspicions that my visibility is being deprioritized by the platform. From past experience, taking 2-4 consecutive days off from posting has helped shorten the duration of the ban.*

Marginalized creators withdrawing from the platform in order to later become visible is an oxymoron; it is a complication of labor that marginalized creators must perform in the absence of clear communication from platforms about content violations and how to recover their engagement. Reflecting on his experience, P2, a bisexual male from Greece, said that he found himself being extra careful with the type of content he posts, in an attempt to avoid censorship and content suppression: ' *'...careful about language, sensitive images, clear messaging leaving no room for interpretation."* When he undertook these types of labor, he reported experiencing average to high performance on his posts.

P4, a Black man, similarly said he refrained from posting photos from his new photoshoot because of the ambiguity around platform policies on shadowbanning: *"I explicitly avoided posting anything from my new fetish shoots, even though they're excellent because I'm unclear on what constitutes "sexual activity" on Instagram"*. This uncertainty, a common theme throughout the data led to anxiety that the creators are typically unsure of how to navigate.

Further, creators argued that disparities in censorship often intersected with LGBTQIA+ and plus-size discrimination. For instance, P7, a plus size creator from Toronto, described their experience with being removed from Instagram, while *"for other people, thinner people, all that stuff stays up, but for me it tends to get taken down"*.

P1, a white lesbian, had a similar experience with lesbian-themed content on TikTok, which she perceived to get lower engagement than heterosexual creators' content. Participants suspected that the word "lesbian" was made less visible by platforms: for example, they described how when trying to search for lesbian content in TikTok's search bar, the platform did not show hashtags like it did for other words, leading them to believe 'lesbian' was shadowbanned. This would cause them to type it out with different spellings, i.e. "Le$bian", which creates additional labor in their everyday platform use. We further explore Algospeak, and its connection to the invisible labor that content creators must perform, in a later Results section.

Participants blamed changes in content moderation for these discrepancies in visibility, arguing that before platforms became overzealous in moderating, content like theirs would normally receive higher engagement. Consistent with previous research on content moderation disparities [30, 58, 60], participants in our study described feeling that their appearance, their profiles, and ultimately, their identity negatively affected how their social media content performed. In these instances, content moderation seems to be blending with offline inequalities, replicating the status quo in the creator and gig economy rather than subverting it [37]. As a result, participants who pole dance shared they felt they had to self-censor in order to continue to thrive under platforms' opaque and capricious management, triggering a chilling effect on art and expression. In this sense, they found that content moderation on platforms such as Instagram and TikTok stifled their creativity and personality, reducing their views, diluting their art, and even affecting their practice. P6 an Asian woman, described her self-censorship strategies on TikTok:

> *I'm not a super 'always wearing lingerie' type of person, but even then, due to TikTok's terms and conditions, you are not allowed to show your buttocks or something like that, so*

> *I had to sometimes just wear leggings instead of wearing pole wear, which would hinder my movement. It was quite tricky.*

P4 also described wearing clothing that they expected would not trigger TikTok's content moderation algorithms, which negatively impacted her pole dancing. P8 similarly reported that she "dialed down" aspects of their content that may lead to censorship - yet doing so made their content less compelling:

> *So when the shadowbanning came into effect, if I sort of dialed down the causes for censorship, then naturally I'll get less likes, because there'll be less people gaining their immediate attention, so to speak … with the shadowban in effect, if you continue to poke them, given their warnings, they say, if you've breached the whatever terms and conditions they have, if you continue to post this type of content, we will disable or block your account entirely*

The scaling back of content that P8 and other participants reported intersects with the identities of the creators and their ability to be creative and express themselves online. The ambiguity surrounding the reasons for shadowbanning directly correlates with when and how the content creators initiated shadowbanning.

5.2 Shadowbanning Mitigation as Mental and Emotional Labor

To mitigate the negative effects of the job insecurity and invisibility that shadowbanning costs them, most of the creators in our study shared they had to perform different types of invisible labor to manage their presence on social media platforms. Invisible labor included activities performed before posting, to try to avoid censorship and content suppression, and after posting, to mitigate the impacts of shadowbanning when it occurs. This mitigation labor shows up as mental/emotional labor, misdirected labor, and community labor, each of which we will focus on individually. Participants described these types of labor as often preceded by significant mental forethought and planning for unpredictable moderation decisions that they have no control over.

The attempt to reclaim control by the content creator often manifests as pre-emptive self-censorship. For example, P6 described having to make substantial readjustments to a pole dancing routine to avoid potential moderation:

> *There was a video transition between me wearing pajamas to lingerie, and then I wanted to do this move that is a bit more… it looks cool, but there's a moment where I sort of have to swing my body around the pole, and then my crotch is literally in the camera, and I couldn't do that because of the outfit I was wearing. I also had to angle myself better so that my crotch wasn't in the camera, if that makes sense. I had to kind of do the trick in the other direction so that it doesn't look like too revealing. It was little bits and pieces where I had to really think ahead and angle myself better so that I'm not doing things, like everything to the camera.*

The labor that P6 had to perform was extraordinary, involving difficult body positions and complicated choreography. These adjustments to what could have been a routine social media post take up valuable time that creators could instead use to create more content, which would in turn generate revenue for them. P6 added that other "regular" content creators do not have to engage in this time-wasting endeavor:

> *Managing shadowbanning means spending more time on the platform to ensure the content is 'Instagram-friendly' and keeping everything – from the thumbnails, content, and profile – more educational rather than sexy… Having to take precautions of reading*

community guidelines.... These activities take much longer to plan rather than posting anything I'd like to.

The examples in the sections that follow illustrate the hidden labor resulting from platform policy opacity that marginalized content creators must perform when they use social media and attempt to avoid and mitigate content suppression. Content creators showed deep awareness not just of the inequalities of shadowbanning, but also of content moderation's impact on the lives and livelihoods of creators and how these inequalities lead to more permanent consequences, such as de-platforming. Participants experimented with thriving under shadowbanning by self-censoring in ways that are common in their communities to chase after the elusive visibility that platforms award sparingly.

There are psychological impacts that result from the loss of income due to content moderation, such as the loss of a digital identity, isolation from established social communities, and the feeling of powerlessness [3]. These are further compounded for content creators with marginalized identities. These inequalities that are a feature of moderating the participants in this study such as shadowbanning, temporary suspensions, and in some cases de-platforming directly affected them financially and emotionally. On the financial side, participants found that the lack of visibility on platforms such as Instagram caused them to reach fewer customers than before, resulting in lost awareness and earnings for their businesses. P6, an Asian woman content creator who runs a pole dancing business, said other pole dancers were finding it difficult to reach their audiences, which was distressing and concerning for the prospect of her own content:

It obviously affects the business side of things as well. They [other pole dancers] can't reach a lot of people anymore, and that really scared me because I felt like if I get my content shadowbanned ever, then all the hard work that I put into it, especially Instagram... it would make a huge impact on, firstly, my reach, and also reaching my students or all prospects

Various participants described working through platforms as getting "no return on investment", be that through time or money. Some aptly described the challenges of beating or gaming the algorithm using gambling metaphors. US-based creator P1, for instance, said posting *"used to be kind of like playing a slot machine almost, where I would post a bunch of stuff and then just randomly one of them would blow up"*. Now, however, they find that posting on platforms is still like playing a slot machine - except that they are losing. Others added that content moderation made them angry and frustrated, particularly because they could not control who would see their content, whether it would become invisible and generate no income due to shadowbanning or attract online abuse (e.g. in the form of homophobia) when viewed by the wrong audience.

Economic uncertainty and feeling less seen on social media in turn affected creators' inspiration and desire to engage with platforms. Participants shared feelings of hopelessness and disenchantment with platforms. Shadowbanning and moderation as a whole discouraged them from posting which led them to worry about their finances, with some arguing social media did not provide a "return on investment" of time and resources, particularly if they find themselves posting to a large audience yet receiving few views. P9, a Black creator, elaborated:

I think at the beginning, it did worry me a bit more, like this is when I was a lot more active with Instagram. But once I crossed that line, soon after, I was like, 'Ugh.' I didn't give a damn anymore, if that makes sense. It was like, 'This is my work. If you like it, you like it. If you don't, that's fine either way.' Because I did not see any value in putting any more focused attention to it if it wasn't going to reciprocate that as a benefit towards me.

Next, we will further analyze shadowbanning's negative implications by examining the specific types of labor that creators undertake to try to avoid and circumvent shadowbanning.

The redirection of labor that content creators with marginalized identities would have otherwise invested in their creative process is invested in understanding vague policies on shadowbanning places a mental and emotional burden on creators, making content creation a stressful and burdensome exercise. P4 shared that the extra labor required to prevent content from being shadowbanned has become a hindrance to his consistency as a creator, negatively impacting his account's engagement and growth:

> *Diverting my attention from my creation to troubleshooting issues with aligning with the policies comes as a distraction, which hinders my consistency in the production of my content, disrupting engagement with my followers and my overall growth online. It's also stressful through my reevaluation of my strategies which is the cautious approach in my content creation.*

The labor they perform, which cannot be quantified, has an impact on their mental health and is linked to the shadowbanning that they are experiencing or have experienced in the past. The stress is further amplified when the participants feel that they cannot be themselves and show up authentically online because of the invisibilized work they must do to curate an arbitrary, non-offensive personality to meet a policy goalpost that always seems to shift. P7's frustration stems from the invisible labor around self-censorship. The boundaryless nature of emotional and cognitive labor means that aspects of personhood and sense of self are not spared from the impact of invisibilized labor and this is further heightened when the individual already has identities that are subjected to societal scrutiny and invalidation.

> *It has created an environment where I cannot share my images and words completely authentically. I have to censor myself according to other's ideas about what is acceptable for someone my size and perceived gender. It has frustrated me and succeeded in making me feel unwelcome and unworthy of being seen and heard.*

When P7 mentions that their shadowbanning experience "has frustrated" them, this directly points to the mental and emotional burdens they face. It also ties into self-presentation and how thorough the content creator has to be in order to avoid these mental and emotional burdens. That frustration hinders and impacts their sense of self-worth and their visibility online, to the extent that they say they feel "unworthy of being seen and heard", they try to shrink themselves to align with the algorithm. Shadowbanning impacted their emotions and feelings of frustration, unwelcomeness, and unworthiness.

When platforms do not provide clarity on their policies regarding shadowbanning as a form of content moderation, it places a labor demand on creators to figure it out. Importantly, it is a demand that places a significant mental burden on social media users, especially marginalized creators, who often have little knowledge of content moderation. For example, P1 lamented that she *"Just does not have time to do that [anymore],"* a signal to the time-consuming requirements of this type of labor and an indicator that her time is better spent doing other things. She added that it is "demoralizing" to do all that work and still have a "low reach", which means that even when creators undertake significant cognitive and emotional labor, it does not typically translate into tangible or profitable results for their content creation goals.

Despite cultivating thousands of followers on social media platforms such as TikTok, some participants described having to abandon their massive following that took substantial time and labor to cultivate because their content is not given visibility when they are shadowbanned. For P1, there is no emotional gratification from seeing low engagement numbers despite investing a significant amount of time into content creation, and this leads to frustration that results from a loss in value from invested labor:

> *I don't use TikTok anymore. I sometimes will cross-post reels every few months to see if my reach is better, but I've pretty much entirely left the platform because it's too frustrating to deal with a perpetual 200 views per video at over 10k followers*

Due to social media platforms' ambiguous content moderation and content suppression policies, it is clear that the stress and frustration that marginalized creators experience are byproducts of the emotional and mental labor they undertake because social media have vague policies on shadowbanning and do not communicate this effectively to content creators. This labor is not recognized or rewarded, and keeps creators with marginalized identities in a loop where they must perform labor that is itself invisible, and also often renders their content and online identity invisible – especially when they end up having to leave the platform.

5.3 Shadowbanning Mitigation as Misdirected Labor

Social media content creators are aware that they must perform some sort of labor to grow their platform and following or generate revenue from their content. They expect this labor to be linear – that is, to result in the creators achieving their goals of increased audience or revenue. Instead, participants in this study described experiencing misdirected labor [5]: a subversion of their labor in service of an opaque algorithm that does not fulfill their goals despite substantial effort.

The creators we spoke with sometimes seemed to curate a performance for the algorithm, with the hopes of bypassing automated moderation triggers. To fill the void created by lack of clear moderation policies, some participants engaged in this one-sided dance with algorithms in the hopes that their labor would produce some reward. P1 described one practice of misdirected labor to counteract perceived shadowbanning: "I've also seen creators post selfies "for the algorithm" after they've posted a lot of non-face content like art or high-suppression content like pro-Palestine activism. Personally, the only thing I really do on Instagram is post a reel the day before I post an image I really want people to see. Instagram loves to boost reels and in my experience will boost everything a little bit after you post one.

The act of learning what an algorithm loves becomes misdirected labor because what the algorithm loves often changes based on arbitrary measures. One example of how this dissipation occurs is when creators meticulously modify each aspect of their content as a preemptive account preservation strategy against shadowbanning. The problem with this tactic lies in the fact that they do not know how effective their approach will be, yet they dedicate time and effort to it because of its unproven potential to sleight the algorithm. As P1 said:

> *Do these tactics work or are they just superstitious? I don't really know, but I still use them because there's a perception that it makes your account less vulnerable to random deletion. I try to avoid posting images with even the implication of nudity, but it feels completely arbitrary*

Similar to previous findings [6, 33, 43, 63], participants like P1 attempted to tackle algorithmic insecurity [43] with algorithmic gossip and folk theories [22, 28] about recommender algorithms, sharing these amongst each other to both make sense of platform governance. The folk theorization and sensemaking that the creators in this study engage with point to the currents of labor that are diverted from creative processes in service of vacuums that are created by lapses in moderation policy by social media companies.

In addition to labor performed on behalf of "the algorithm" and folk theorization, other forms of misdirected labor that participants performed self-censorship via practices like Algospeak – that is, intentionally misspelling words or blocking images to trick the algorithm and bypass content moderation and suppression – a practice that requires significant time and effort. P5 listed several self-censorship techniques she has observed: *"placing an emoji or graphic to show less skin"*,

censoring words by spelling them phonetically, and logging off for two or three days and then upon return reposting popular content to reverse the effects of shadowbanning.

For P1, incorporating Algospeak is one strategy that content creators like herself are using to self-censor in other to bypass the moderation algorithm: *"It means that everything I post needs to be specially modified and curated for the platform, down to weird censoring like writing 'pxxn' or 'prrn' instead of 'porn,' etc. Like how people have adopted 'unalive' to mean 'suicide'."*

Despite giving into to this type of labor by many of the participants, there are some visible currents of resistance to the digital irritation that they arrive at after folk theorization [80]. For example, P1 chooses not to obfuscate important content using Algospeak, despite the apparent costs: *"I refuse to censor words like 'lesbian' or 'autism' or 'disability', which is probably why I will never recover on TikTok. Other creators definitely do that with some success".*

In this section, we drew connections between participants' work to counteract shadowbanning and capitalistic notions of misdirected labor – labor that diverts individuals' productive labor into serving needs that do not align with their visions and goals. We showed how misdirected labor typically ends up filling a labor gap that social media companies refuse to provide by having clear policies on shadowbanning. This burden then offloads on marginalized users and impacts their options and ability for creativity and self-determination.

5.4 Shadowbanning Mitigation as Community Labor

In response to algorithmic-mediated governance of the platforms where they are creatively expressing themselves, participants engaged with forms of labor tied to community engagement as ways to resist and bypass what they suspected to be shadowbanning on social media platforms. Most participants shared how they learned different "tips" and "tricks" from other creators or from engagement pods (online groups created to inorganically drive engagement to the posts of other members of the group through actions such as liking, sharing, and reposting) to either drive organic or inorganic engagement to content after suspected shadowbanning. This type of labor was typically a departure from what they originally would have engaged in while creating content on social media if they were not experiencing shadowbanning.

P5 told us about a trick she learned from a creator community on Instagram. Since losing her original account to shadowbanning, she has not been linking her work to her new main account. Instead, she created a dedicated account where she puts the link for her work *'so that if Insta does a sweep on anyone with AllMyLinks containing OnlyFans or whatever, it'll delete the placeholder account and not my main. That's a trick I picked up from some of the bigger porn people."*

The knowledge of creating an alternative account to circumvent shadowbanning which in itself is additional labor was also derived from the labor of bigger creators. The reliance on information sourced through community labor in the engagement pods they create or are a part of by some creators seems to stem from an understanding that social media companies have an entirely different moderation strategy for celebrity accounts. In this way, they rationalize how risque posts made by celebrities like Kim Kardashian stay up, while a similar post by a regular model would get deleted. In the same vein, P4 said he frequently checks other accounts, "especially influencers", on how they "play their cards right".

Another way that participants benefit from community labor to mitigate shadowbanning is by relying on information that other creators share about their experiences with content moderation. This insight can be helpful in understanding their own experiences. For example, P1 told us: *"I don't think I'm shadowbanned on Instagram, but I only know that because of other creators sharing their engagement stats and showing this is just a general platform trend for image-based artists".* P1 said she relies on "screenshots of Instagram notices" posted by other creators, which contain warnings that their post will no longer have a wider audience because of the type of content they are posting:

"I do like that I can at least check and confirm on Instagram that I haven't missed a helpful step, unlike with all other platforms where it's purely guesswork". Platforms' lack of transparency about their moderation trends forces creators to rely on each other's labor as they theorize on reasons for moderation by sharing their private engagement metrics with each other.

To build solidarity and assist each other in coping with the precarity of the platform economy [81], creators often rely on the labor of other creators through community labor from engagement pods, which inorganically drives up engagement on each other's posts using mutual aid. P1 admits that she has not had much luck with engagement pods in the past, while P7 had some degree of success: *"I have worked with other creators to help boost engagement during a shadowban. We will promote and engage each other's content during a suspected ban and report missing content for each other".* Despite the different needs of the creators based on their marginalized identities, we found that participants connected with the community they found online directly or indirectly to overcome perceived obstacles to their shared algorithmic and moderation frustrations.

6 DISCUSSION

This study focused on marginalized content creators' experiences of shadowbanning, observing how they navigated this insidious but upsetting censorship technique, with a specific emphasis on how creators attempted to mitigate it. We highlight shadowbanning's impact on marginalized content creators' finances [57] and well-being. We argue that the invisible labor they perform to avoid it replicates offline gender and sexuality inequalities, which actively worsens their experiences on platforms. In addition, we pay special analytical attention to participants' everyday invisible labor that shadowbanning sets into motion.

We extend prior research that underscores critical themes such as marginality, invisibility, and precarity in the creator economy [62] by crystalizing three distinct types of labor marginalized creators perform to make sense of and circumvent shadowbanning. Many participants in our study, who depend on social media platforms as a source of livelihood, faced substantial work, including mental and emotional labor, misdirected labor, and community labor, to cope with the perceived impacts and uncertainties that shadowbanning posed to their source of livelihood, which in turn further exacerbated their financial precarity and negatively impacted their wellbeing.

The data in this study shows that the invisible labor often takes place both before and after a post is made to avoid content suppression or to mitigate the impact of shadowbanning. In some instances, participants shared that because they were confused by the opacity of the moderation algorithm, and in an attempt to reduce the likelihood of getting shadowbanned, they deleted posts they had made in the past, erasing their own invested labor. For instance, the shadowbanning directly impacted participant P6's reaching out to her audience and made her concerned about the prospects of her own content. Therefore, while the creator economy is often praised due to its entrepreneurial prospects, the possibility to "be your own boss" and to reach inordinate amounts of new customers [37], this promise is precarious and requires more labor even for users who do not have marginalized identities [26] which means that marginalized content creators are subjected to more entry-level barriers that further marginalizes them. Instead of being their own bosses, they are subject to the opaque and capricious management of the "algorithmic boss"[24], leading them to face economic uncertainty.

In line with studies on offline work precarity, what [43] defines as "algorithmic insecurity", we also found that many participants, including P6 and P9, felt powerless over their work and its success, triggering negative psychological effects due to financial insecurity [23, 29] and due to the feeling of posting into a void [12, 30]. Additionally, our findings show that the layered invisible labor of making sense of and circumventing shadowbanning is not just a tiring, frustrating, unpaid attempt at survival by reverse-engineering platform governance: it is also an exercise in mitigating

one's presence, in dimming one's light, and in "dialing back" one's identity. Dialing it back is often an attempt to appeal to the "algorithmic boss" [24] by relying on gossip and folk theories that are not foolproof, but that are worth a shot [6, 63].

Informed by scholarship on platform work and labor [25, 49, 54, 72] and drawing from our empirical data, we extend prior work and contribute to understanding invisible labor in the creator economy by identifying three unique categories of labor that marginalized creators perform to mitigate the perceived impact of shadowbanning. When [21] defined cognitive labor, the intention was to describe the often invisible work of thinking, planning, scheduling, and organizing that accompanies family life and the emotional labor that comes with it. Mental and emotional labor is often boundaryless and without limit manifests internally in spatial and temporal ways that render it invisible [21]. For the purpose of this study, we extrapolate this definition to contextualize the invisible labor that content creators with marginalized identities are engaging with. Additioanlly, informed by participants' experiences, these types of labor are 1) *mental and emotional labor*, which involves the meticulous but unpaid cognitive and psychological efforts to create content that does not trigger shadowbanning algorithms; 2) *misdirected labor*, which involves the hidden labor that does not directly contribute to creators' content production but is done in the hopes of avoiding shadowbanning; and 3) *community labor*, which we argue is a form of collaborative survival [74] and involves sharing knowledge and strategies within creator communities to tackle and mitigate the precarity and uncertainties related to opaque algorithmic content suppression [24]. Each of these types of labor is often unnoticed by others and rendered invisible.

The precarity that marginalized content creators experience, particularly when they are women, LGBTQIA+, disabled, or plus size, is indeed more pronounced than their less marginalized counterparts [30], and it is directly linked with the opacity of social media governance. As a result, the incentive for these users to produce invisible labor in order to circumvent shadowbanning is strong, requiring various tasks such as guessing what may guarantee success at gaining popularity metrics, erasing their previous labor involved in content production, and initiating collective actions to resist content suppression from social media. Importantly, similar to the dominant literature on invisible labor [20], the invisible labor involved in shadowbanning mitigation is also often performed by racial and gender minorities. Additionally, this additional labor is essentially devalued, as it is not typically rewarded with additional visibility.

Previous research in the context of gig economy, micro work, and crowdsourcing [38, 61, 71, 77, 78] has demonstrated that digital labor communities frequently engage in mutual support and collective action to confront with challenges and exploitations from online platforms. Our findings reaffirm this collective labor, showing how creators, in attempts to mitigate the impact of shadowbanning, often rely on communal efforts to understand how to circumvent it. We argue that such community labor should be recognized as a form of invisible labor, as it constitutes additional work that does not directly translate into financial gain for marginalized creators. However, this community labor, though practiced in response to the opacity and ambiguity of algorithmic systems, could also offer potential benefits for marginalized content creators by fostering online community building and mutual support beyond the context of shadowbanning. For instance, our findings illuminate how such community efforts can lead to forming engagement pods and strengthening solidarity in response to platform-induced frustrations and precarity.

We advocate for worker-centered technology development and adoption in shadowbanning, informed by worker-centered design approach [32, 66]. A shift towards a worker-centered approach requires moving from design goals that prioritize platforms' profit and efficiency, to design approaches that include marginalized content creators and address their daily challenges with these systems. Additioanlly, in line with the worker-centered technology development scholarship [39, 51, 70], which encourages both academics and practitioners in fields such as CSCW and GROUP

to play a "strategic role in supporting labor organizations," our study not only documents the tools participants use (e.g., appeal systems) and the challenges they face within the broader economic context, but also proposes alternative design considerations to better recognize and address participants' invisible labor. Specifically, beyond existing content moderation appeal systems, there is a need to develop alternative mechanisms or tools to better document and address the often invisible labor associated with shadowbanning. Such systems or tools would not only restore invisible labor and provide evidence of the labor wasted, which contributes to the precariousness of creators' livelihoods, but also foster solidarity among them. While it is true that social media platforms usually must compete with incentives and stakeholders in designing moderation policy[1], social media companies currently design moderation policies to appease regulatory agencies by cracking down on accounts they consider borderline in regard to content policies [3] while keeping those accounts invisible to maintain user numbers, yet this approach is counterproductive and alienates marginalized content creators. We suggest a design approach that relies on collaboration with marginalized content creators and takes into account cultural and cross-cultural complexities in demystifying what types of content are likely to result in shadowbans.

Additionally, as both Greenbaum [39] and Tang and colleagues [70] have advocated, there is a need to reframe "labor" over "work" and return to examining the labor process in the research and design of sociotechnical systems. This shift in perspective can enable scholars to more effectively analyze the "consequences of information system design." Building on this line of research, we strategically use "labor" instead of "work" when describing participants' experiences interacting with algorithmic systems. This approach allows us to critically engage with the consequences of these interactions from both the users' and platforms' perspectives. We argue that the various forms of invisible labor associated with shadowbanning are detrimental not only to marginalized creators but also to technology companies. The process of navigating shadowbanning becomes a form of misdirected labor for content creators, as it forces them to continuously adjust to elusive and changing algorithms. This effort diverts their attention from improving their craft and expressing their creativity. Our research further reveals that the consequences resulting from participants' shadowbanning experiences have a material impact that renders creators' content and online identities invisible, leading some to disengage from or leave these platforms entirely. As platforms profit from creator-generated content, this not only harms the creators but could also represent significant losses for the platforms.

Taken together, adopting a worker-centered design approach in shadowbanning to develop alternative systems and tools would benefit social media companies by helping retain content creators and further unleashing their creativity, but also making the content creation ecosystem more equitable and sustainable. This work also presents opportunities for future research into how other groups are experiencing shadowbanning in order to design more inclusive content moderation policies. For example, while we focused on people with marginalized identities in the Global North for this study, it is unclear what the experiences of other marginalized populations might be in the rest of the world, and the collaborative processes they are practicing to circumvent shadowbanning. Other groups with high visibility on social media, such as journalists, politicians, and activists, are also critical pieces in understanding the arbitrariness and opacity surrounding social media shadowbanning, and how platforms may adjust content moderation design and policy to better support marginalized users. While design implications could be more powerful when aligned with appropriate policy implications [79], we acknowledge that most sociotechnical systems, including content moderation systems, still rely heavily on ambiguity. Consequently, one-size-fits-all policies are less likely to yield satisfactory results [1]. While advocating for more content-sensitive policy

[1]Shedding Light on Shadowbanning: https://cdt.org/insights/shedding-light-on-shadowbanning/

implications, it is crucial to leverage community members' situated understanding of the content, while also recognizing their labor and expertise.

7 CONCLUSION

This study examines how content creators with marginalized identities experience shadowbanning, a controversial aspect of platform governance characterized by the use of opaque algorithms to reduce or demote content. Through interviews and a diary study with content creators from marginalized communities (e.g., women, plus-size individuals, and/or LGBTQIA+ members), we found that participants are disproportionately affected by shadowbanning. This content suppression not only negatively impacts their financial earnings and well-being but also necessitates additional invisible labor to circumvent the effects. Specifically, we found that the layered invisible labor involved in understanding and circumventing shadowbanning often occurs both before and after posts are made, as efforts to mitigate its impact. Drawing on participants' experiences and informed by scholarship on platform work and labor, we identify three unique categories of invisible labor – mental and emotional labor, misdirected labor, and community labor – that are integral to participants' work practices around shadowbanning but often go unnoticed and are rendered invisible by platforms. We argue that such invisible labor, essential for mitigating shadowbanning, is both racialized and gendered and harmful not only to marginalized creators but also to technology companies. Lastly, informed by worker-centered design approaches, we propose that alternative mechanisms and tools must better document and address the often invisible labor associated with shadowbanning.

ACKNOWLEDGMENTS

We thank our study participants for sharing their valuable insights and experiences. We are also grateful to the members of the Community Research on Identity and Technology (CRIT) Lab at the University of Michigan, Ann Arbor, School of Information (UMSI) for their insightful feedback on our work. Special thanks to Tianling Yang and Renkai Ma for their additional feedback. We appreciate the constructive comments our anonymous reviewers provided that have improved this work. This research was supported by National Science Foundation grant #1942125.

REFERENCES

[1] Mike Ananny. 2018. *Networked press freedom: Creating infrastructures for a public right to hear.* MIT press.
[2] Carolina Are and Pam Briggs. 2023. The emotional and financial impact of de-platforming on creators at the margins. *Social media+ society* 9, 1 (2023), 20563051231155103.
[3] C. Are and P. Briggs. 2023. The Emotional and Financial Impact of De-Platforming on Creators at the Margins. *Social Media + Society* 9, 1 (2023). https://doi.org/10.1177/20563051231155103
[4] Sarah Banet-Weiser. 2018. *Empowered: Popular feminism and popular misogyny.* Duke University Press.
[5] E. B. Bigelow. 1878. The Relations of Labor and Capital. *The Atlantic* (October 1 1878). https://www.theatlantic.com/magazine/archive/1878/10/the-relations-of-labor-and-capital/521106/
[6] Sophie Bishop. 2019. Managing visibility on YouTube through algorithmic gossip. *New media & society* 21, 11-12 (2019), 2589–2606.
[7] D. Blunt, E. Coombes, S. Mullin, and A. Wolf. 2020. *Posting Into the Void.* Hacking/Hustling. https://hackinghustling.org/posting-into-the-void-content-moderation/
[8] D. Blunt and Z. Stardust. 2021. Automating Whorephobia: Sex, Technology and the Violence of Deplatforming. *Porn Studies* 8, 4 (2021), 350–366.
[9] D. Blunt and A. Wolf. 2020. Erased: The impact of FOSTA-SESTA and the removal of Backpage on sex workers. *Antitraffickingreview.org* 14 (2020), 117–121.
[10] Oversight Board. 2023. *Gender identity and nudity.* Decisions. https://oversightboard.com/decision/BUN-IH313ZHJ/
[11] N. Bolger, A. Davis, and E. Rafaeli. 2003. Diary methods: capturing life as it is lived. *Annual Review of Psychology* 54 (2003), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

[12] Carolyn Bronstein. 2021. Deplatforming sexual speech in the age of FOSTA/SESTA. *Porn Studies* 8, 4 (2021), 367–380. https://doi.org/10.1080/23268743.2021.1993972

[13] Scott Carter and Jennifer Mankoff. 2005. When participants do the capturing: The role of media in diary studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 899–908.

[14] S. Cole. 2018. *Where Did the Concept of 'Shadow Banning' Come From?* Vice. https://www.vice.com/en/article/a3q744/where-did-shadow-banning-come-from-trump-republicans-shadowbanned

[15] Juliet M Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1 (1990), 3–21.

[16] K. Cotter. 2021. 'Shadowbanning is not a thing': black box gaslighting and the power to independently know and credibly critique algorithms. *Information, Communication Society* 26, 6 (2021), 1226–1243. https://doi.org/10.1080/1369118X.2021.1994624

[17] Mary Crain, Walter Poster, and Michael Cherry. 2016. *Invisible labor: Hidden work in the contemporary world*. University of California Press, Berkeley.

[18] Samantha Dalal, Ngan Chiem, Nikoo Karbassi, Yuhan Liu, and Andrés Monroy-Hernández. 2023. Understanding Human Intervention in the Platform Economy: A case study of an indie food delivery service. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[19] A. K. Daniels. 1987. Invisible work. *Social Problems* 34, 5 (1987), 403–415. https://doi.org/10.2307/800538

[20] Arlene Kaplan Daniels. 1987. Invisible work. *Social problems* 34, 5 (1987), 403–415.

[21] Liz Dean, Brendan Churchill, and Leah Ruppanner. 2022. The mental load: building a deeper theoretical understanding of how cognitive and emotional labor over load women and mothers. *Community, work & family* 25, 1 (2022), 13–29.

[22] Michael A. DeVito. 2021. Adaptive Folk Theorization as a Path to Algorithmic Literacy on Changing Platforms. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–38. https://doi.org/10.1145/3476080

[23] Sandra MA Dijkstra-Kersten, Karolien EM Biesheuvel-Leliefeld, Johannes C van der Wouden, Brenda WJH Penninx, and Harm WJ van Marwijk. 2015. Associations of financial strain and income with depressive and anxiety disorders. *J Epidemiol Community Health* 69, 7 (2015), 660–665.

[24] Brooke Erin Duffy. 2020. Algorithmic precarity in cultural work. *Communication and the Public* 5, 3-4 (2020), 103–107.

[25] Brooke Erin Duffy and Colten Meisner. 2022. Platform governance at the margins: Social media creators' experiences with algorithmic (in) visibility. *Media, Culture & Society* 45, 2 (2022), 285–304.

[26] Brooke Erin Duffy, Annika Pinch, Shruti Sannon, and Megan Sawey. 2021. The nested precarities of creative labor on social media. *Social media+ society* 7, 2 (2021), 20563051211021368.

[27] Gwyn Easterbrook-Smith. 2023. OnlyFans as gig-economy work: A nexus of precarity and stigma. *Porn studies* 10, 3 (2023), 252–267.

[28] Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I "like" it, then I hide it: Folk Theories of Social Feeds. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2858036.2858494

[29] Lisa Fiksenbaum, Zdravko Marjanovic, and Esther Greenglass. 2017. Financial threat and individuals' willingness to change financial behavior. *Review of Behavioral Finance* 9, 2 (2017), 128–147.

[30] Anonymized for Review. 2021. Anonymized title. (2021).

[31] Anonymized for Review. 2022. Anonymized title. (2022).

[32] Sarah E Fox, Vera Khovanskaya, Clara Crivellaro, Niloufar Salehi, Lynn Dombrowski, Chinmay Kulkarni, Lilly Irani, and Jodi Forlizzi. 2020. Worker-centered design: Expanding HCI methods for supporting labor. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 1–8.

[33] Florencia García-Rapp. 2017. Popularity markers on YouTube's attention economy: the case of Bubzbeauty. *Celebrity Studies* 8, 2 (2017), 228–245.

[34] Tom Gerken. 2022. How to check if your Instagram posts are being hidden. *BBC News* (2022). https://www.bbc.co.uk/news/technology-63907699

[35] Emily Gilbert. 2023. Beyond the usual suspects: Invisible labor(ers) in futures of work. *Geography Compass* 17 (2023), e12675, 1–12. https://doi.org/10.1111/gec3.12675

[36] T. Gillespie. 2022. Do Not Recommend? Reduction as a Form of Content Moderation. *Social Media + Society* 8, 3 (2022), 20563051221117550. https://doi.org/10.1177/20563051221117552

[37] Z. Glatt. 2022. 'We're all told not to put our eggs in one basket': Uncertainty, precarity and cross-platform labor in the online video influencer industry. *International Journal of Communication* 16 (2022), 1–19. https://ijoc.org/index.php/ijoc/article/view/15761

[38] Mary L Gray, Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni. 2016. The crowd is a collaborative network. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 134–147.

[39] Joan Greenbaum. 1996. Back to labor: returning to labor process discussions in the study of work. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work* (Boston, Massachusetts, USA) *(CSCW '96)*. Association

for Computing Machinery, New York, NY, USA, 229–237. https://doi.org/10.1145/240080.240259
[40] Cheng Guo. 2020. Identity and user behavior in online communities. In *Companion Proceedings of the 2020 ACM International Conference on Supporting Group Work*. 35–38.
[41] Oliver L. Haimson, Daniel Delmonaco, Peipei Nie, and Andrea Wegner. 2021. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 466 (oct 2021), 35 pages. https://doi.org/10.1145/3479610
[42] Oliver L. Haimson, David Delmonaco, Peng Nie, and Arman Wegner. 2021. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–35. https://doi.org/10.1145/3479610
[43] K. Hill. 2021. Algorithmic insecurity, schedule non standardness, and gig worker wellbeing. (2021). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3773164 Population Association of America Poster Presentation. Working Paper.
[44] Cynthia Huff. 1989. " That Profoundly Female, and Feminist Genre": The Diary as Feminist Practice. *Women's Studies Quarterly* 17, 3/4 (1989), 6–14.
[45] Lilly C Irani and M Six Silberman. 2013. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 611–620.
[46] Karin AM Janssens, Elisabeth H Bos, Judith GM Rosmalen, Marieke C Wichers, and Harriëtte Riese. 2018. A qualitative approach to guide choices for designing a diary study. *BMC medical research methodology* 18 (2018), 1–12.
[47] Mohammad Hossein Jarrahi and Will Sutherland. 2019. Algorithmic management and algorithmic competencies: Understanding and appropriating algorithms in gig work. In *Information in Contemporary Society: 14th International Conference, iConference 2019, Washington, DC, USA, March 31–April 3, 2019, Proceedings 14*. Springer, 578–589.
[48] Mohammad Hossein Jarrahi, Will Sutherland, Sarah Beth Nelson, and Steve Sawyer. 2020. Platformic management, boundary resources for gig work, and worker autonomy. *Computer supported cooperative work (CSCW)* 29 (2020), 153–189.
[49] Kylie Jarrett. 2022. *Digital labor*. John Wiley & Sons.
[50] Trinady Joslin. 2020. Black Creators Protest TikTok's Algorithm with# imblackmovement. *Daily Dot* (2020).
[51] Vera Khovanskaya. 2023. The Cyberunion 20 Years Later. *Interactions* 30, 2 (feb 2023), 48–51. https://doi.org/10.1145/3583079
[52] Kalle Kusk and Claus Bossen. 2022. Working with wolt: an ethnographic study of lenient algorithmic management on a food delivery platform. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–22.
[53] Airi Lampinen, Moira McGregor, Rob Comber, and Barry Brown. 2018. Member-Owned Alternatives: Exploring Participatory Forms of Organising with Cooperatives. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–19.
[54] Maurizio Lazzarato. 2006. The Concepts of Life and the Living in the Societies of Control. *Deleuze and the Social* (2006), 171–190.
[55] P. Leerssen. 2023. An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law & Security Review* 48 (2023). https://doi.org/10.1016/j.clsr.2023.105790
[56] M. Leybold and M. Nadegger. 2023. Overcoming communicative separation for stigma reconstruction: How pole dancers fight content moderation on Instagram. *Organization* 0, 0 (2023). https://doi.org/10.1177/13505084221145635
[57] Renkai Ma and Yubo Kou. 2021. " How advertiser-friendly is my video?": YouTuber's Socioeconomic Interactions with Algorithmic Content Moderation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.
[58] G. Mauro and H. Schellman. 2023. 'There is no standard': investigation finds AI algorithms objectify women's bodies. *The Guardian* (2023). https://www.theguardian.com/technology/2023/feb/08/biased-ai-algorithms-racy-women-bodies
[59] Isabel Munoz, Michael Dunn, Steve Sawyer, and Emily Michaels. 2022. Platform-mediated Markets, Online Freelance Workers and Deconstructed Identities. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–24.
[60] Susanna Paasonen. 2021. *Dependent, distracted, bored: Affective formations in networked media*. MIT Press.
[61] Julian Posada. 2022. Embedded reproduction in platform data work. *Information, Communication & Society* 25, 6 (2022), 816–834.
[62] Bernhard Rieder, Erik Borra, Òscar Coromina, and Ariadna Matamoros-Fernández. 2023. Making a living in the creator economy: A large-scale study of linking on YouTube. *Social Media+ Society* 9, 2 (2023), 20563051231180628.
[63] Laura Savolainen. 2022. The shadow banning controversy: perceived governance and algorithmic folklore. *Media, culture & society* 44, 6 (2022), 1091–1109.
[64] Catherine Scott and Megan Medaugh. 2017. *Axial Coding*. https://doi.org/10.1002/9781118901731.iecrm0012
[65] Ellen Simpson, Andrew Hamann, and Bryan Semaan. 2022. How to tame" your" algorithm: LGBTQ+ users' domestication of TikTok. *Proceedings of the ACM on Human-computer Interaction* 6, GROUP (2022), 1–27.

[66] Franchesca Spektor, Sarah E Fox, Ezra Awumey, Ben Begleiter, Chinmay Kulkarni, Betsy Stringam, Christine A Riordan, Hye Jin Rho, Hunter Akridge, and Jodi Forlizzi. 2023. Charting the Automation of Hospitality: An Interdisciplinary Literature Review Examining the Evolution of Frontline Service work in the Face of Algorithmic Management. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–20.

[67] L. Stack. 2018. What Is a 'Shadow Ban,' and Is Twitter Doing It to Republican Accounts? *The New York Times* (July 26 2018). https://www.nytimes.com/2018/07/26/us/politics/twitter-shadowbanning.html

[68] Will Sutherland, Mohammad Hossein Jarrahi, Michael Dunn, and Sarah Beth Nelson. 2020. Work precarity and gig literacies in online freelancing. *Work, Employment and Society* 34, 3 (2020), 457–475.

[69] Karoliina Talvitie-Lamberg, Vilma Lehtinen, and Sanna Valtonen. 2022. Tactics of invisibility: How people in vulnerable positions make datafied everyday life livable. *new media & society* (2022), 14614448221136077.

[70] Joice Tang, McKane Andrus, Samuel So, Udayan Tandon, Andrés Monroy-Hernández, Vera Khovanskaya, Sean A. Munson, Mark Zachry, and Sucheta Ghoshal. 2023. Back to " Back to Labor": Revisiting Political Economies of Computer-Supported Cooperative Work. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing* (Minneapolis, MN, USA) *(CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 522–526. https://doi.org/10.1145/3584931.3611285

[71] Arianna Tassinari and Vincenzo Maccarrone. 2020. Riders on the storm: Workplace solidarity among gig economy couriers in Italy and the UK. *Work, employment and society* 34, 1 (2020), 35–54.

[72] Tiziana Terranova. 2012. Free labor. In *Digital Labor*. Routledge, 33–57.

[73] Katrin Tiidenberg. 2021. Sex, power and platform governance. *Porn Studies* 8, 4 (2021), 381–393.

[74] Anna Lowenhaupt Tsing. 2015. *The mushroom at the end of the world: On the possibility of life in capitalist ruins*. Princeton University Press.

[75] C. Vark. 2021. Sex, Money, & Politics: Big Tech Made FOSTA-SESTA Global for Sex Workers. *American University Business Law Review* (2021). https://aublr.org/2021/10/sex-money-politics-big-tech-made-fosta-sesta-global-for-sex-workers/#_ftn22

[76] Katie Waddington. 2013. Using qualitative diary research to understand emotion at work. In *A Day in the Life of a Happy Worker*, Arnold B. Bakker and Kevin Daniels (Eds.). Psychology Press, Hove, 132–149.

[77] Mark E Whiting, Dilrukshi Gamage, Snehalkumar (Neil) S Gaikwad, Aaron Gilbee, Shirish Goyal, Alipta Ballav, Dinesh Majeti, Nalin Chhibber, Angela Richmond-Fuller, Freddie Vargus, et al. 2017. Crowd guilds: Worker-led reputation and feedback on crowdsourcing platforms. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing*. 1902–1913.

[78] Alex J Wood, Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019. Networked but commodified: The (dis) embeddedness of digital labour in the gig economy. *Sociology* 53, 5 (2019), 931–950.

[79] Qian Yang, Richmond Y Wong, Thomas Gilbert, Margaret D Hagan, Steven Jackson, Sabine Junginger, and John Zimmerman. 2023. Designing Technology and Policy Simultaneously: Towards A Research Agenda and New Practice. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–6.

[80] Brita Ytre-Arne, Trine Syvertsen, Hallvard Moe, and Faltin Karlsen. 2020. Temporal ambivalences in smartphone use: Conflicting flows, conflicting responsibilities. *new media & society* 22, 9 (2020), 1715–1732.

[81] Zizheng Yu, Emiliano Treré, and Tiziano Bonini. 2022. The emergence of algorithmic solidarity: unveiling mutual aid practices and resistance among Chinese delivery workers. *Media International Australia* 183, 1 (2022), 107–123.

Received May 2024; revised August 2024; accepted October 2024