



Flagging as a silencing tool: Exploring the relationship between de-platforming of sex and online abuse on Instagram and TikTok

new media & society

1–19

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14614448241228544

journals.sagepub.com/home/nms**Carolina Are** 

Centre for Digital Citizens, Northumbria University.

Abstract

This article investigates Instagram and TikTok's approach to malicious flagging through users' experience. Similar to liking, commenting and sharing, flagging is a reaction social media platforms allow users to highlight content that potentially violates community guidelines. However, flagging's influence on moderation remains opaque: users who flag are largely unaware about the success of their reports; those who are de-platformed cannot be sure if or why their content has been reported, making them feel targeted not just by platforms' processes, but by the retaliation of audiences themselves. Since the impact of de-platforming on users, and particularly on content creators who work through platforms, can be huge, this study provides scope to investigate flagging as an online abuse technique.

Keywords

Censorship, content moderation, flagging, Instagram, platform governance, TikTok

Introduction

The moderation—for example, the deletion or censorship—of online content is a key aspect of platform governance, without which digital spaces would be unusable (Diaz and Hecht-Felella, 2021). Platforms decide what types of content are allowed and visible in their spaces through their governance infrastructure's affordances, or the possibilities

Corresponding author:

Carolina Are, Centre for Digital Citizens, Northumbria University, College Ln, Newcastle upon Tyne NE1 8SG, UK.

Email: carolina.are@northumbria.ac.uk

objects, and therefore technologies, offer for action (Graves, 2007). Flagging is one of these affordances, allowing users to report content to social media platforms and express their concerns about their governance (Crawford and Gillespie, 2016). However, flagging can also be weaponised against accounts other users disagree with, through ‘user-generated warfare’ (UGW; Fiore-Silfvast, 2012). This can become an effective, crippling online abuse technique, aiding malicious actors in banishing users from online spaces and adopting silencing strategies against those who have been already disproportionately targeted by content moderation, which has so far over-focused on nudity and sexuality instead of on violence (Are, 2023b; Duffy and Meisner, 2023; Lumsden and Morgan, 2017). To understand the relationship between malicious flagging and de-platforming, this study focuses on two mainstream social media platforms used by professional content creators: the Meta-owned photo and video sharing platform Instagram, and its fast-growing Chinese rival, the video sharing platform TikTok.

Social media platforms became liable for content posted in their spaces in 2018, after the United States government approved the Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA) and the Stop Enabling Sex Traffickers Act (SESTA; 115th Congress, 2017–2018). Intended to fight sex trafficking, FOSTA/SESTA caused platforms to over-censor posts worldwide to be seen to be complying with the new law, applying US legislation to global content (Are, 2020; Bronstein, 2021; Haimson et al., 2021 etc.). Since then, sex workers’ accounts have been deleted without warning; athletes, lingerie, sexual health brands, sex educators and activists have had their content deleted or shadowbanned—whereby ‘vaguely inappropriate content’ is hidden from the platform’s main pages without the user knowing (Are, 2022; Cotter, 2023). Creators at the margins, and particularly those sharing nudity and sex-related content, are therefore often excluded from the visibility and work opportunities that platforms offer because their presence is viewed as inherently dangerous (Are and Paasonen, 2021).

Given the importance of social media as work and civic spaces (Are, 2020; Duffy and Meisner, 2023), platforms’ over-moderation of nudity has been affecting the livelihoods, lives and visibility of many users, who find themselves powerless in trying to understand or appeal these decisions, and in attempting to re-gain control over their profiles, content and social media history (Are and Paasonen, 2021; Blunt and Wolf, 2020). Therefore, this article emphasises the effects legislation has had on platform governance, arguing that the study of flagging and content governance cannot be separated from its initial targets: sex work and nudity. However, while nudity and sexuality have been bearing the brunt of platform censorship post-FOSTA/SESTA, malicious flagging by opposing factions in conflicts in Palestine and Ukraine have also triggered content removals, targeting activists and journalists alike (Stokel-Walker, 2022).

Platform governance research has so far explored shadowbanning (Are, 2022; Are and Paasonen, 2021; Cotter, 2023), flagging (Crawford and Gillespie, 2016; Fiore-Silfvast, 2012; Peterson, 2013) and the challenges of automated moderation (Gillespie, 2010; 2018; Kaye, 2019; etc.). Research also examined the censorship of sex or sexual content (Tiideberg and Van der Nagel, 2020, etc.) and its offline effects on sex workers (Blunt and Wolf, 2020). However, studies have yet to examine the challenges that flagging poses to marginalised users, particularly once they are de-platformed and directly faced with said lack of transparency, communications and accountability.

Platform governance in the risk society

Platform governance is ‘a set of actions with wide-ranging socio-political effects on transnational communication and workers’ rights’ (Bronstein, 2021: 370). As part of this, content moderation, or ‘removing the content, making it inaccessible, and/or suspending or terminating the accounts’ of those who posted it is an integral part of the running of Internet platforms (Goanta and Ortolani, 2021: 2). Its rules are set by platforms’ ‘community guidelines’ or ‘standards’, which are often enforced by a blend of human and automated action (Gillespie et al., 2020; Kaye, 2019, etc.).

TikTok and Instagram delete accounts that go against their community guidelines, often without warning (Instagram, n.d.; TikTok, n.d.). ‘Instagram is not a place to support or praise terrorism, organised crime or hate groups’, the Meta-owned platform’s guidelines state, adding they prohibit offering and selling of sexual services, buying or selling firearms, alcohol, drugs and tobacco (Instagram, n.d.). Nudity and sexuality are also banned on TikTok, which prohibits content that is ‘overtly revealing of breasts, genitals, anus, or buttocks, or behaviors that mimic, imply, or display sex acts’ (TikTok, n.d.).

Community guidelines are enforced through algorithms, ‘codified step-by-step processes’ platforms use ‘to afford or restrict visibility’ (Bishop, 2019: 2590). Algorithmic governance has, however, been anything but transparent, creating a power and information asymmetry between social media companies and users ‘as platforms withhold, obscure, and strategically disclose details about their algorithms’ (Cotter, 2023: 1). A lack of investment in human moderation and the over-reliance on opaque algorithms has resulted in flawed platform governance (Tedroff, 2022).

Algorithmically enforced community guidelines are similar across different platforms, covering content already prohibited by most countries’ criminal law (Gillespie, 2010; 2018; Goanta and Ortolani, 2021; Kaye, 2019 etc.). Yet, while their juxtaposing of something as humanly necessary as sex with violence may seem excessive, it reveals Instagram and TikTok’s management of risk. Connected to platforms’ commercial interests (Goanta and Spanakis, 2020) and to their over-zealousness in the face of FOSTA/SESTA (Blunt et al., 2020; Blunt and Wolf, 2020; Nolan-Brown, 2022; Tiideberg and Van der Nagel, 2020; etc.), platform governance can be conceptualised through Beck’s (1992, 2006) and Giddens’ (1998) World Risk Society theory, arguing that businesses and institutions address uncertainty by micromanaging risks produced by late modernity. However, the risks they identify are socially constructed, and the focus on preventing them means identifying certain populations as ‘high-risk’, further marginalising society’s ‘others’ (Beck, 2006). Already used to conceptualise platforms’ use of shadowbanning of nude content (Are, 2022), World Risk Society theory can help analyse loopholes surrounding malicious flagging and censorship, particularly in the aftermath of FOSTA/SESTA.

Through the joint bill, the US government applied a World Risk Society logic, forcing platforms to adopt its approach and transfer it not only to what was covered by FOSTA/SESTA, but also to adjacent content posted globally (Beebe, 2022; Blunt and Wolf, 2020; Nolan-Brown, 2022 etc.). Big Tech have therefore fully embraced World Risk Society logic, expanding it further; while FOSTA/SESTA has not been found to reduce sex trafficking and has only resulted in one charge so far (Blunt et al., 2020), its

risk-focused, carceral governance has resulted in a significant chilling of sexual speech online, with policies ‘more restrictive than necessary to avoid liability’, and platforms going ‘into overdrive in their efforts to create distance from sexual speech and sex work to avoid potential legal problems and to attract investors and advertisers’ (Bronstein, 2021: 368).

If in Beck’s World Risk Society businesses’ and governments’ increasing preoccupation with risks largely benefitted private insurance firms (Hudson, 2003), who defined and managed them (Beck, 2006; Giddens, 1998), platforms’ governance of risk post-FOSTA/SESTA is inevitably connected to their payment providers, their advertisers, their public image and their self-protection from legal charges. The risk platforms attempt to manage therefore is not only reputational: it is financial, tying governance of most online sex work and sexual expression to the United States. Indeed, while laws governing sex work and sexual expression differ around the world, financial industry regulations are mostly concentrated in the United States, meaning most nations have to connect with US-based credit card companies and comply with their terms of service (Beebe, 2022). Therefore, platforms—already largely born and based in the United States and showing a puritanical approach to governance—must abide by the whims of the United States’ financial system. This hits users with a ‘double whammy’ of puritanical cultural and financial governance, based on the potential risks of platforming sex, sex work and sexual expression (Nolan-Brown, 2022).

This risk-averse and puritanical financial and platform governance of sex shows financial discrimination is applied even to services that are legal in the United States, such as the multi-million-dollar pornography industry or stripping, extending even to countries where governance of sex work is less strict (Beebe, 2022).

The affordances of flagging

Legislation such as FOSTA/SESTA brought platforms to add a series of affordances to manage content visibility in their spaces (Bucher and Helmond, 2018). Originally developed in psychology (Gibson, 1977; Norman, 1988) and, later, used in communication studies (Bucher and Helmond, 2018; Graves, 2007), affordances can be viewed as a set of functionalities that allow both users *and* platforms to perform different actions.

Platforms’ moderation affordances include shadowbanning, account takedowns, content removals and temporary posting bans (Cotter, 2023; Duffy, 2020; Gillespie et al., 2020; etc.). In this article, I focus on a ‘double’ moderation Instagram and TikTok affordance: flagging, which is afforded to both audiences *and* platforms themselves, and its relationship with content and account takedowns. Similar to liking, commenting and sharing, flagging is a reaction platform that allows users to highlight content that potentially violates community guidelines (Crawford and Gillespie, 2016). It is both a mechanism for platforms to learn from users, distributing labour to volunteers (Crawford and Gillespie, 2016) and a form of crowdsourcing algorithmic progress ‘to facilitate the evasion of liability by platforms as intermediaries’ (Goanta and Ortolani, 2021: 13), but its influence on moderation remains opaque. Users have so far been largely unaware as to why or if their reports have been taken up, or indeed as to why their content has been reported (Crawford and Gillespie, 2016).

With social media being cultural, work and public discourse spaces, flags can become an easily gamed way to police content for users who may view a gay kiss or an artistic nude as too inappropriate for social media (Crawford and Gillespie, 2016). Flags are also extremely convenient for platforms: tech companies do not have to honour them, and can use them to justify the removal of contentious content (Crawford and Gillespie, 2016). In this sense, companies' over-zealousness in adhering to legislation and common decency and its related incentivisation of quick responses to risks has facilitated the proliferation of malicious flagging (Griffin, 2022).

Examples of malicious flagging, also known as organised flagging (Crawford and Gillespie, 2016), user-generated censorship (Peterson, 2013) or UGW (Fiore-Silfvast, 2012), are many: conservative group Truth4Time's coordinated effort to flag pro-LGBT+ Facebook groups; countering online Cyber-Jihad YouTube videos; a coalition of incels joining forces to 'purge' adult performers from Instagram (Clark-Flory, 2019; Crawford and Gillespie, 2016; Fiore-Silfvast, 2012). Flagging can therefore be a form of 'digilantism', or 'politically motivated extrajudicial practices in online domains that are intended to punish or bring others to account in response to a perceived or actual lack of institutional remedies' (Jane, 2017: 461), escalating different social beliefs into 'something resembling mass vigilantism' (Schoenebeck and Blackwell, 2021: 7). This way, platforms like Instagram and TikTok can simultaneously *facilitate* online harms through their affordances—for example, flagging—and *perpetrate* them through their design and governance (Schoenebeck and Blackwell, 2021).

Flagging has left users feeling targeted not just by platforms' processes, but by the retaliation of audiences themselves (Duffy and Meisner, 2023; Myers West, 2018). Cox (2019, 2021) found that many de-platformed users paid hackers *working for* platforms to reinstate their accounts, and that malicious actors also pay hackers to *trigger* de-platforming, meaning malicious flagging can be considered a silencing strategy to drive users, and particularly women, Lesbian, Gay, Bisexual, Transgender, Queer, Intersex and Asexual (LGBTQIA+) and marginalised users, off digital spaces. The impact of de-platforming on users, and particularly on professional content creators, can be huge, since their career success (e.g. gaining brand partners or monetising their audience) 'is directly related to platformed indices of visibility (i.e. views, likes, favorites, shares)' (Duffy and Meisner, 2023: 2). There is thus scope to investigate it as a form of online abuse, which includes actions such as harassment, flaming, doxing and image-based abuse (Lumsden and Morgan, 2017).

In this increasingly uncertain governance scenario where particularly users who work, express themselves and communicate through their bodies find themselves under threat of de-platforming at the hands of both users and platforms, their own explanations and interpretations of the governance they are targeted with become all we have to understand social media companies' opaque governance.

Resisting opaque platform governance with gossip

The opacity of social media infrastructure means users are not always privy to what exactly triggered their account's deletion (Schoenebeck and Blackwell, 2021): although both Instagram and TikTok show users when their accounts accumulate

multiple violations and may be at risk of deletion (Are, 2023a), users do not always have access to specific information, for example, whether deletions were caused by a single post, a succession of posts, one or a series of reports by other users. What is known following a decision to be reviewed by Meta's independent oversight body, the Oversight Board, is that in the case of partial nudity picturing trans and non-binary bodies, as little as three reports were enough for content to be taken down by Instagram (Marks, 2022).

Faced with this lack of essential information, users have taken matters in their own hands to be able to continue working, networking and expressing themselves through platforms. Bishop, for example, found that beauty bloggers on YouTube shared their knowledge about the platform's moderation through 'algorithmic gossip', or 'communally and socially informed theories and strategies pertaining to recommender algorithms, shared and implemented to engender financial consistency and visibility on algorithmically structured social media platforms' (Bishop, 2019: 2602). While gossip is often dismissed as biased or frivolous, it becomes a knowledge resource for marginalised groups, a tool to fight power and facilitate resistance (Bishop, 2019). Indeed, the use of gossip, for Bishop, reflects the power inequalities between creators and platforms. Similarly, Savolainen found that users 'fight' platforms' algorithms through 'algorithmic folklore', or 'beliefs and narratives about moderation algorithms that are passed on informally and can exist in tension with official accounts' (Savolainen, 2022: 2). Combining personal experience with platforms, media reports and friends' tales, folklore can provide insights into the opaque functioning of algorithms (Bishop, 2019; Duffy and Meisner, 2023; Savolainen, 2022). Yet, while users' theories represent a form of resistance to platform governance, they do not always result in successfully managing the algorithm—and, in fact, they are often denied by platforms themselves (Bishop, 2019; Duffy and Meisner, 2023; Savolainen, 2022).

Platforms' outright denials of moderation targeted against specific communities—even in the face of their own apologies (Are, 2022)—has been likened to gaslighting, a strategy used to both 'neutralize particular criticisms that such individuals might lodge' and 'to neutralize the very possibility of criticism by undermining the victim's conception of herself as an autonomous locus of thought, judgement, and action' (Cotter, 2023: 4). This gaslighting, which Blunt et al. mention in the case of shadow-banning, can be applied to other mysterious moderation techniques such as flagging, when targeted users are 'made to feel crazy, as their reality is being denied publicly and repetitively by the platform' (Blunt et al., 2020: 79). This way, users are made to look 'incapable of assessing algorithms independently of what platforms say about them' (Cotter, 2023: 14). Thus, in denying users' experiences and knowledge, platforms are minimising their expertise which, given that a variety of women and LGBTQIA+ users have been affected by censorship *and* use gossip, shows a patriarchal, patronising approach to criticism. Instead, approaching the study of algorithms through a feminist lens means broadening the sources we learn from and reframing who we think is an 'algorithmic professional' (Bishop, 2019: 2602). By centring the experiences of de-platformed users in the study of potential links between malicious flagging and censorship, this article plans to do just that.

Methods

This study responded to the following research questions:

RQ1. What is the relationship between negative comments and content and/or profile removal on Instagram and TikTok?

RQ2. Which techniques do users utilise to recover their accounts?

Instagram and TikTok were chosen because they are crucial for maintaining networks, promoting work, expressing oneself and organising and, crucially, they are largely free of charge for users, who often utilise them to gain visibility for their work (Are, 2022; Duffy and Meisner, 2023 etc.).

Data were gathered in a 3-month period through an anonymous qualitative survey circulated through my own social media profiles: Facebook, Twitter, Instagram and TikTok. Conversely to quantitative methods, which are used to conduct a statistical analysis of large samples, qualitative studies examine smaller data samples to understand specific contexts, such as subcultures, using specific information or data to make inferences about case studies (e.g. Goertz and Mahoney, 2012).

Qualitative surveys feature open-ended questions centred on a particular topic and presented in a fixed order to all participants (Braun et al., 2021). They can produce rich accounts of participants' experiences, allowing them to respond in their own words instead of having to select from multiple choice answers (Braun et al., 2021). Furthermore, given that the nature of this study meant engaging with participants from marginalised communities who had to share personal, often traumatic experiences of de-platforming, qualitative surveys allowed them to present their own narrative, in their own safe space, without me interfering with their narration.

The inclusion criteria for participants were intentionally narrow: this survey only collected data from users over of above 18 years of age, and who had experienced *both* negative comments *and* account and/or content deletion. While this has significantly narrowed the number of participants involved, it has provided a clearer snapshot of the potential links between malicious flagging and de-platforming, returning more specific experiences of platform governance, particularly in the case of sex workers. As such, qualitative surveys are also relevant from 'a social justice and inclusion point of view', as they 'offer an accessible method to research beyond the "usual suspects"' (Braun et al., 2021: 643, 644).

I circulated this survey through my social media networks because, as a known researcher with experiences of de-platforming and of organising anti-censorship activism campaigns, I have a sizable following of over 400,000, comprising artists, athletes, sex workers, activists, researchers and journalists with whom I have built relationships and who helped me share the call for participants. By circulating the survey among communities that knew me, I wanted to make sure users, and particularly sex workers, felt comfortable when sharing their experiences of de-platforming. In addition, XBIZ, a leading news website for the adult industry, wrote an article about this study, helping me circulate information about it among their audiences, which explains why sex workers are over-represented in this article (Parkman, 2022).

Survey questions were modelled on personal and documented experiences of de-platforming (Are, 2023a;b; Stokel-Walker, 2021, 2022), asking participants 10 questions in a blend of open-ended ones and multiple choice, demographic-related questions for screening. The former featured questions about participants' experiences of de-platforming, their understanding of the reasons behind their deletion and the steps they took towards account recovery. Multiple choice questions were informed by previous research findings highlighting platforms' disproportionate censorship of Black, Indigenous and people of color (BIPOC), LGBTQIA+, plus-sized and disabled users (Haimson et al., 2021), and asked participants about their profiles' aims, their age, gender identity, sexual orientation, ethnic background and whether they identified as plus size or as having disability.

Responses are organised through thematic analysis (TA) in the form of answer excerpts. This method allows researchers to identify, analyse and report themes or patterns within data (Braun and Clarke, 2006). Characterised by minimal data organisation, TA describes data sets in rich detail and can be a 'realist method, which reports experiences, meanings and the reality of participants' (Braun and Clarke, 2006: 81), providing insights into relevant realities, particularly when researching on stigmatised and excluded communities.

Themes are 'creative and interpretive stories about the data, produced at the intersection of the researcher's theoretical assumptions, their analytic resources and skill, and the data themselves' (Braun and Clarke, 2019: 594). They are developed through the creative labour of researchers' coding, and often reflect data collection questions (Braun and Clarke, 2019). Yet, since TA is a qualitative method, the importance of a theme does not necessarily depend on quantifiable measures or replicability, but rather on whether it captures relevant details about the research questions (Braun and Clarke, 2019).

This study presents a set of limitations which have affected the data gathered. First, conducting research on platforms *through* platforms was challenging. With this study focusing on the Meta-owned and the Chinese tech giants, where most of my following lies, I experienced different levels of invisibility and, in some cases, an outright chilling effect on my ability to share research and collect data. While the call for participants received an exponentially reduced audience on my TikTok profile compared to my usual posts, Instagram flagged my survey—hosted on the conventional and university-approved survey platform Qualtrics—as a 'dangerous link', potentially affecting the number of users who, already under threat of de-platforming and hacking, may have decided not to complete it. Due to the high shares my call for participants received among sex working communities, Instagram may have classed my link as spam or 'sexual solicitation'—bitterly ironic, considering that platform governance is the topic of this study.

Second, the narrow inclusion criteria set during data collection meant that participants were limited to a specific experience, potentially excluding those who *may have been* maliciously flagged but who did not receive negative comments on their posts. However, given users' previous experiences of malicious flagging (Clark-Flory, 2019; Fiore-Silfvast, 2012; Stokel-Walker, 2021, 2022), the inclusion criteria I set were the most tangible factor to continue investigating malicious reporting.

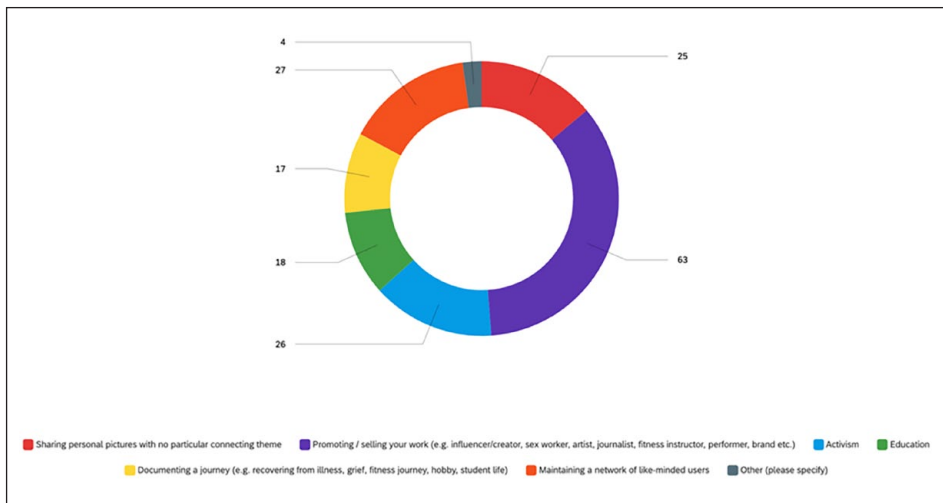
Third, having received media coverage about my call for participants from XBIZ meant that the data collected is skewed towards sex workers' experiences, although a

variety of users from different backgrounds also took part in the survey. Still, given that sex workers are over-represented among de-platforming targets (e.g. Blunt and Wolf, 2020; Stardust et al., 2020), the responses to this survey represent an accurate picture of the current platform governance landscape.

Finally, given the experiences reported in this article, it may appear that this study is nothing more than a set of conjectures. Yet, while it is not possible to fully prove that users have been targeted with malicious flagging, utilising their experiences and conjectures about content moderation is precisely my intention: with little to no platform communications and transparency about governance processes, this study sets out to recognise the content moderation and algorithmic expertise of those affected by it (Bishop, 2019).

User demographics and result themes

A total of 123 participants took part in this study. In total, 98 of them reported having been affected by both censorship *and* negative comments on Instagram and TikTok. Out of 123, 41 participants provided full, detailed responses about their experiences. Most users reported using Instagram and TikTok to promote and/or sell their work, be it journalism, sex work, content creation, fitness and so on (63). Participants also utilised platforms to keep in touch with like-minded users (27), to engage in activism (26), to share pictures without a connecting theme (25), to share and receive education (18) and to document a journey (17).



Most of the users surveyed were aged 25–34 (18), with the second biggest group being 35–44 (13) and the third 18–24 (5). Respondents were largely White (34), with only a few coming from mixed race (3), Black (2), Latine (2) or Indigenous (1) backgrounds. Most of those who took part in the survey were cisgender women (28), followed by cisgender men (5), non-binary (3) and gender fluid (2) users, with only one transgender woman taking part. Most respondents were bisexual (13), followed by

pansexual (12), heterosexual (7), queer (6) and gay (2) users. Only eight users identified as plus size, and 16 identified as having a disability.

Theme 1. Flagging as online abuse

Respondents' experiences highlight significant exploitable loopholes in platform governance, leaving particularly feminist, sex working and activist users vulnerable to silencing by those who disagree with them. Users highlight receiving negative comments on their posts to then be de-platformed shortly after, consistently with Are (2023b), as well as being flagged due to their political views or their references to being sex workers in their social media posts or bios. Worryingly, users reported campaigns orchestrated by single individuals egging their followers on, even targeting sexual assault survival content:

- 'I think that one person or a small group of people added my account with the sole purpose to do targeted harassment, post negative comments and flag my photos. [. . .] At some point a hacker gained control of my account. They posted child pornography which flagged Instagram to block and delete both my Instagram account and my Facebook'
- 'My thoughts are that I was maliciously reported by a fan that I had an online altercation with on my OnlyFans platform. He was an incel who had taken offence to me being married and my Instagram was taken down on my wedding anniversary, after I posted about it, which I don't think was a coincidence'
- 'A person with a much larger following exploited their fans to abuse a system that is already broken. Once it hits a certain amount of reports it bans the content without any manual need or reason for appeal or approval'
- 'I truly think people reported my sexual assault story post and after that since it got so many reports Instagram just deleted me. They said it was for sexual solicitation which I was not doing. I think they are a right-wing company that wants to silence feminist voices'.

Fully ascertaining that one or more reports caused deletions may be challenging without platforms' confirmation, but these users' experiences highlight that flagging can be an effective silencing strategy and online abuse technique by those who wish to cause harm to those running accounts that have already been disproportionately targeted by platform censorship such as sex working accounts, erotic art and body-positive accounts (Are, 2022; Blunt and Wolf, 2020 etc.). As such, malicious flagging can be likened to other deliberately harmful online behaviours such as flaming or doxing (Lumsden and Morgan, 2017). Exploiting Instagram and TikTok's track record of de-platforming, malicious flaggers can stage full-blown attacks (Fiore-Silfvast, 2012) against specific users, relying on trigger-friendly moderation processes and on time-poor and under-paid moderators who have to make conservative decisions to deliver on their targets, leaving very little room for context (Suri and Grey, 2018). Therefore, online abusers are motivated to give flagging a shot, in the case that the very exploitable loophole of an under-funded, context-lacking platform governance can help them achieve their aims.

Theme 2. Making sense of opaque platform governance

Participants reported an array of content and profile deletions surrounding topics from sex work to activism, from art to mental health advocacy, from memory-making to self-expression:

- ‘I’m banned, bankrupted and demonized for what? Linework illustrations of breasts?’
- ‘My paintings have been removed several times without warning, [. . .] flagged for nude content and sometimes sexual solicitation. Paintings that include LGBTQ+ themes or POC tend to be removed faster and without hesitancy compared to my other pieces that depict white figures. Example: 2 paintings that had the same angle and pose were flagged, but the white model was kept on my page while the black model was removed.’
- ‘I was sharing LEGAL SAFE abortion pill information and it was deleted and my account is in jeopardy of being removed from Instagram. No option to appeal either. TikTok removed a pro-abortion meme of a historical painting of a woman bare breasted with a bunch of babies saying “me meeting all of my abortions when I get to heaven.”’
- ‘I have a business about destigmatizing STIs and fighting slut shaming—essentially it’s a place for empowering women. I’ve had a few posts taken down on Instagram but for reasons that didn’t make sense. Like one time I shared info about abortion pills and it got taken down because it said I was selling illegal goods. I shared a personal story about sexual assault and my account was deleted within 24 hours.’
- ‘All images of any flesh, i.e. showing off hands with painted nails, removed due to “sexual content”’.

Consistently with research on the aftermath of FOSTA/SESTA, participants’ experiences show that censorship of sex work is trickling down to other groups (Are, 2022; Are and Paasonen, 2021). Sexual health and activism information such as posts related to abortion or to surviving sexual assault, as well as artful displays of bodies, were caught in the net of platform censorship. Participants’ experiences show that bodily displays or women’s health are a shorthand for sex on platforms, replicating previous research findings on content moderation sexualising bodies without users’ consent (Are, 2023a;b).

While some users claimed to have been deleted due to ‘plus size stigma’, ‘misogyny’ and due to ‘being a woman’ however, sex workers reported that merely associating with adult content platforms has been enough for them to be potentially flagged before being deleted by Instagram and TikTok. Although most users claimed to be following and creating content in accordance with community guidelines, the mere mention to sex work or OnlyFans in their biographies or link sections made them more vulnerable to deletion. Users recounted having had as many as six accounts deleted by Instagram and TikTok, despite showing no nudity:

- ‘My content was taken down off of Instagram for “nudity” when there was no nudity whatsoever. I showed too much of my shoulder. Other content was deleted for “solicitation of sexual services” because I simply mentioned OnlyFans.’

- ‘My very first account was removed after I used the word “Goddess” in my profile bio—nothing specific to fetish or anything of that nature just put “redhead goddess” in my bio and within 10 minutes my account was gone permanently with no warning’
- ‘My 2nd account got suspended after I had my 5th “photo violation” the photo in question that was the 5th photo was a headshot of me in a gold metallic tank top which they saw as nudity—it was not even sheer and picture was from the chest up with my arm across my chest- nothing about the photo was lewd or in violation of their policies. The other photos were basic bikini or lingerie photos that were not as risqué as what many other models post on IG. One of them was a pin up photo even—tight clothing but nothing showing.’
- ‘I have had videos removed for adult nudity when I’m in a hoodie and sweatpants. [. . .] Swimming with my kids has been nudity, skateboarding with my daughter was also somehow nudity. I’ve had my livestreams removed for solicitation when I’m just talking about being a single parent.’

In this sense, platforms seem to be profiling users, and particularly sex workers, assuming that even when they are posting personal or non-sex work related content, they must be ‘soliciting’ regardless. This assumption is reflected in the variety of users who claim to not even post nudity, but to only feature an OnlyFans link. In doing so, platforms greatly underestimate their roles in their users’ lives: far from being just a creative outlet or a tool to promote work, TikTok and Instagram are also a portfolio, a way to access vital information, to network with one’s community and to make memories with friends and family. In assuming that users may be ‘soliciting’ 24/7, Instagram and TikTok are profiling sex workers as dangerous and dehumanising them, stigmatising their work and denying them any other sort of human interaction.

In the absence of detailed information from platforms about the reasons behind users’ deletion, this study’s participants have resorted to finding their own explanations, consistently with research on algorithmic gossip and folklore (Bishop, 2019; Savolainen, 2022). Participants communicated a blend of confusion surrounding platform processes—for example, confusing flagging by other users with platforms’ deletion of content—and explanations relying on the discrimination of user groups arising from gossip, folklore and media reports.

Users’ reliance on gossip, conjectures and folklore may appear as engaging with conspiracy theories. Yet, in similar scenarios such as in instances of shadowbanning (Are, 2022), users have instead become experts in their own workspaces, reacting to a power dynamic that does not help them (Bishop, 2019). Following reports that Instagram and TikTok have previously discriminated against women, plus size users, people of colour, LGBTQIA+ creators and, largely, sex workers (Haimson et al., 2021), and given platforms’ often dry, public relations-heavy responses to critiques of their governance, users have fully embraced negative narratives about the way the apps are run.

Theme 3. ‘I don’t think I ever spoke to a human ever’. Battling faceless governance

A striking majority of respondents were not successful at recovering their Instagram and/or TikTok accounts through platforms’ official appeal systems. This highlights significant

failures within Instagram and TikTok's moderation infrastructure and, in particular, an over-reliance on automated moderation that leaves users feel lost and disheartened.

Participants reported a lack of clarity, accessibility and transparency by platforms when dealing with account deletions. Often, despite self-censoring and following community guidelines, they experienced deletions regardless, and were not able to speak to platform employees to ask for further clarifications about this censorship:

- 'I have had to use strategically sportier looking items to promote my products. I have also sent messages when appealing the decision to delete products with very short 'this is a crop top'. With not much luck.'
- 'I have emailed tiktok and IG and done support requests to get to the bottom of things and see what I could do differently but no response ever.'

Responses to my questions surrounding the steps users had taken to recover their accounts are a succession of 'I tried contacting Instagram/TikTok and never heard back'. When dealing with the distressing aftermath of account deletions that led to loss of income and of network, most users reported not being able to reach human moderators or platform workers and having to deal with slow or ineffective appeals systems—that is, even receiving wrong appeal links from platforms, or facing deletion despite having been a victim of hacking:

- 'I've had countless days of appealing to the faceless and nameless masters of Instagram. [. . .] I scrounged for any direct connection; emailing people who never answered me, trying any way of reaching to someone who'd understand that this is my livelihood.'
- 'I tried for over two months to get my page back only to hit a brick wall. I don't think I ever spoke to a human ever.'
- 'I've appealed every single case and once again it feels like the lottery. Sometimes it's put back up, other times it's removed and the penalty isn't fully processed. I have tickets that have been 'in process' for over a year now.'
- 'Even though I told Instagram that my account has been hacked and they had seen signs of this and even sent me emails reporting strange activity they did absolutely nothing. I don't think messages even get seen by a person. They told me they had a received my appeal and that a decision would be made in one to two days. At 30 days they decided to block and remove my accounts permanently.'
- 'They gave me a link that doesn't work and I couldn't retrieve it. I tried to email and call them. nothing since March 2022.'

Without being able to reach the 'nameless masters' of platforms or 'to speak to a human ever', most responders could not understand *which* specific element of their content triggered their deletions, meaning that, were they to create other accounts, they may once again be facing censorship due to the inability to learn from previous experiences. Furthermore, with respondents finding their appeals were not processed or reviewed, they were left in limbo: without hearing back from platforms, they had to take matters into their own hands. In these cases, users have two possibilities: striking a deal with

hackers to attempt to retrieve their accounts or resign themselves to the fact that their profiles will not be coming back.

With most respondents reporting it is shared knowledge that platforms do not engage with deleted users, many have given up on recovering their banned accounts entirely:

- ‘Of all of the times I was deleted, they gave my account back only once. So, I have completely started over several times.’
- ‘I tried to recover my Instagram account, but gave up after weeks. I had to create a new account.’
- ‘It’s easier to create a new account.’

Users therefore had to start over, creating new accounts that may still be under threat of de-platforming, while also having to engage in the taxing digital labour of re-building a following. In line with Cox’s (2019, 2021) findings, respondents’ experiences seem to show that, when faced with platforms’ lack of communication, the inability to recover their accounts through conventional means and the possibility to lose one’s work and network, users are prepared to spend money and run risks in the hope to recover their profiles:

- ‘I have emailed them every day for weeks with no reply. I even paid a guy to get my Instagram account back and he eventually gave up and just took my money and I never got my account back. I have had friends make reports for me and they never got replies back.’
- ‘[My deletion] also lent more legitimacy to scammers who use my content to create catfish profiles. If those stay up while my genuine profile is deleted then people are unlikely to see my real profiles with lower follow numbers as genuine.’

Social media platforms’ approach to moderation and their opaque granting of verification badges has therefore inadvertently created a market for scammers, from hackers promising to recover accounts (Cox, 2019) and to delete accounts (Cox, 2021), all the way to those posing as a deleted profile to scam others out of nude images or money (Are, 2018).

Participants also reported that their communities have had to step in and fill the gaps left by platforms’ inaction through awareness-raising, pulling contacts’ strings and gathering support from freedom of expression focused non-profits. Users’ followers reported them as ‘missing’, submitting Help Centre tickets and emailing platforms:

- ‘When my first account was banned, I was able to receive assistance from dozens of friends, and hundreds of followers, who all messaged/emailed tiktok to request I get my account back. Sadly they ignored all of these and never restored my account, which is why I created my new one.’
- ‘I appeal each removal (when it is available), and I’ve written to the Oversight Board whenever it is offered. When my @nipeople account was removed twice last year, I received support from many influential users and accounts including @feminist and @bloggeronpole, as well as the National Coalition Against Censorship and Freemuse.’

- ‘I have asked people that follow me to help with complaints and messages. I also contacted people inside the platform.’

Consistently with my own experiences of de-platforming (Author, 2023a; Stokel-Walker, 2021), the users surveyed have felt more supported by their own communities and often recovered their accounts only when someone in their network was able to directly speak to platform insiders. This highlights a shared understanding among communities that platforms’ infrastructure is at best inefficient and inaccessible, and at worst, directly biased against them.

Conclusion

This article highlighted the exploitable loopholes Instagram and TikTok’s flagging affordances provide to malicious actors, increasing the vulnerabilities of creators who are already under threat of de-platforming due to their work, their characteristics or the content they post. The experiences shared seem to show that Instagram and TikTok allowed personal or moral crusades against them, affording power over content and profiles to anyone but the users who create them.

The joint power of malicious flagging and de-platforming resulted in the following isolating experiences: trying to recover one’s profile, re-building a network and a workplace from scratch, having to rely only on their network, on scammers, on individuals with contacts within Instagram and TikTok to still be able to work or, otherwise, having to resign themselves to the loss of their profile and subsequent loss of livelihoods, memories and networks.

Participants felt they had been targeted by both de-platforming *and* negative comments, which may have triggered flagging. While their beliefs about the moderation they received might not coincide with Instagram and TikTok’s agendas, in line with Bishop (2019), I take their experiences seriously. As experts in their fields, about their communities and their subcultures, these users are highlighting failures within platform governance—which, according to both Instagram and TikTok, is allegedly aimed at users’ safety (Instagram, n.d.; TikTok, n.d.)—that are leaving them vulnerable instead. Users’ experiences show that platforms are both *facilitating* and *perpetrating* harms against them (Schoenebeck and Blackwell, 2021), allowing others to report them and triggering their de-platforming. Their experiences are both valid and concerning. To this end, more research is needed: on the moderator and platform end, there is a need to access workers and uncover the factors triggering companies’ decisions; on the user end, broadening the scope of research even outside already marginalised users will be beneficial towards clearer understandings of platform governance.

My findings show that, similarly to other online abuse techniques such as flaming or doxing (Lumsden and Morgan, 2017), malicious flagging is a digital silencing strategy driving users offline, a strategy particularly effective against users and topics that are stigmatised or have been the target of platform governance: sex work, art, activism, queer expression and sexual health education. Intentionally or not, those who engage in flagging are playing a part in users’ loss of network, livelihood and education, exploiting a platform governance that is already unequal (see Are, 2022; Blunt and Wolf, 2020; Duffy and Meisner, 2023; Haimson et al., 2021 etc.) and skewed against marginalised communities, and that thrives off of opacity (Crawford and Gillespie, 2016; Kaye, 2019 etc.).

What is clear from this article's findings is that sex workers are particularly vulnerable to this type of behaviour post-FOSTA/SESTA, and that platform governance seems to be blending with malicious flaggers' moral code, bringing offline whorephobia—the hatred and disgust towards sex workers—into an already whorephobic, risk-focused platform governance (Beebe, 2022; Blunt et al., 2020).

Malicious flagging's success rests on the fact that platforms do very little to restore accounts they disable, showing a carceral, punitive form of governance that does not allow for rehabilitation (Schoenebeck and Blackwell, 2021). Indeed users—from sex workers to artists, from activists to educators—reported distressing experiences such as being targeted by rogue fans, by those who disagree with their political views or their approach to morality, without being protected, leading to the loss of accounts that were so crucial to their livelihoods.

A further aggravating factor of malicious flagging is the targeting of users with disabilities, low incomes and/or childcare needs who had finally found types of online work and expression that fit with their living situation and needs. Therefore, the supposedly safety-focused governance of content and profiles on Instagram and TikTok begs the question: *safety for whom?*

The experiences shared by participants in this study therefore show a trigger-heavy approach to governance, a facilitation of flagging loopholes, an inadequate appeals system and an approach to user safety which is cavalier at best, and careless at worst. Through withholding information about their practices and processes and through over-policing sexual content, social media are effectively handing the reins of their most conservative tools to misogynist, abusive accounts that benefit from the de-platforming of sex education, sexual information, feminist education and sex work.

Acknowledgements

The author would like to thank everyone who shared and took part in this survey, particularly those who shared personal, traumatic experiences of de-platforming towards advancing knowledge in this field.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Research Councils UK > Engineering and Physical Sciences Research Council EP/T022582/1. <https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/T022582/1>.

ORCID iD

Carolina Are  <https://orcid.org/0000-0003-1110-3155>

References

- Are C (2018) Does Instagram's verification discriminate against nudity? *Blogger On Pole*. Available at: <https://bloggeronpole.com/2020/08/does-instagrams-verification-process-discriminate-against-nudity/>
- Are C (2020) A corpo-civic space: a notion to address social media's corporate/civic hybridity. *First Monday* 25(6). DOI: 10.5210/fm.v25i6.10603.

- Are C (2022) The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram. *Feminist Media Studies* 22: 2002–2019.
- Are C (2023a) An autoethnography of automated powerlessness: lacking platform affordances in Instagram and TikTok account deletions. *Media, Culture & Society* 45(4): 822–840.
- Are C (2023b) The assemblages of flagging and de-platforming against marginalised content creators. *Convergence*: 0(0). <https://doi.org/10.1177/13548565231218629>
- Are C and Paasonen S (2021) Sex in the shadows of celebrity. *Porn Studies Forum* 8(4): 411–419.
- Beck U (1992) *Risk Society: Towards a New Modernity* (trans M. Ritter). London: Sage.
- Beck U (2006) Living in the world risk society. *Economy and Society* 35(3): 329–345.
- Beebe B (2022) ‘Shut up and take my money!’ Revenue chokepoints, platform governance, and sex workers’ financial exclusion. *International Journal of Gender, Sexuality and Law* 2: 140–170.
- Bishop S (2019) Managing visibility on YouTube through algorithmic gossip. *New Media & Society* 21(11–12): 2589–2606.
- Blunt D and Wolf A (2020) Erased: the impact of FOSTA-SESTA and the removal of backpage on sex workers. *Anti-Trafficking Review* 14: 117–121.
- Blunt D, Coombes E, Mullin S, et al. (2020) Posting into the void. *Hacking/Hustling*. Available at: <https://hackinghustling.org/posting-into-the-void-content-moderation/>
- Braun V and Clarke V (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology* 3: 77–101.
- Braun V and Clarke V (2019) Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise & Health* 11(4): 589–597.
- Braun V, Clarke V, Boulton E, et al. (2021) The online survey as a qualitative research tool. *International Journal of Social Research Methodology* 24(6): 641–654.
- Bronstein C (2021) Deplatforming sexual speech in the age of FOSTA/SESTA. *Porn Studies* 8(4): 367–380.
- Bucher T and Helmond A (2018) The affordances of social media platforms. In: Burgess J, Poell T and Marwick A (eds) *The SAGE Handbook of Social Media*. London and New York: Sage, pp. 233–253.
- Clark-Flory T (2019) A troll’s alleged attempt to purge porn performers from Instagram. *Jezebel*. Available at: <https://jezebel.com/a-trolls-alleged-attempt-to-purge-porn-performers-from-1833940198>
- Cotter K (2023) ‘Shadowbanning is not a thing’: black box gaslighting and the power to independently know and credibly critique algorithms. *Information, Communication & Society* 26: 1226–1243.
- Cox J (2019) Hacked Instagram influencers rely on white-hat hackers to get their accounts back. *Vice*. Available at: <https://www.vice.com/en/article/59vnyk/hacked-instagram-influencers-get-accounts-back-white-hat-hackers>
- Cox J (2021) Scammer service will ban anyone from Instagram for \$60. *Vice*. Available at: <https://www.vice.com/en/article/k78kmv/instagram-ban-restore-service-scam>
- Crawford K and Gillespie T (2016) What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society* 18(3): 410–428.
- Diaz A and Hecht-Felella L (2021) *Double Standards in Social Media Content Moderation*. New York: Brennan Centre for Justice.
- Duffy BE (2020) Algorithmic precarity in cultural work. *Communication and the Public* 5(3–4): 103–107. <https://doi.org/10.1177/2057047320959855>
- Duffy BE and Meisner C (2023) Platform governance at the margins: social media creators’ experiences with algorithmic (in)visibility. *Media, Culture & Society* 45(2): 285–304.

- Fiore-Silfvast B (2012) User-generated warfare: a case of converging wartime information networks and coproductive regulation on YouTube. *International Journal of Communication* 6: 1–24.
- Gibson JJ (1977) The theory of affordances. In: Shaw R and Bransford J (eds) *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, NJ: Erlbaum, pp. 67–82.
- Giddens A (1998) *The Third Way: The Renewal of Social Democracy*. Cambridge: Polity Press.
- Gillespie T (2010) The Politics of Platforms. *New Media & Society* 12(3): 347–364.
- Gillespie T (2018) *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. London: Yale University Press.
- Gillespie T, Aufderheide P, Carmi E, et al. (2020) Expanding the debate about content moderation: scholarly research agendas for the coming policy debates. *Internet Policy Review* 9(4). DOI: 10.14763/2020.4.1512.
- Goanta C and Ortolani P (2021) Unpacking content moderation: the rise of social media platforms as online civil courts. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3969360
- Goanta C and Spanakis J (2020) Influencers and social media recommender systems: unfair commercial practices in EU and US Law. *TTLF Working Papers* 54: 1–27.
- Goertz G and Mahoney J (2012) *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton, NJ: Princeton University Press.
- Graves L (2007) The affordances of blogging, a case study in culture and technological effects. *Journal of Communication Inquiry* 31(4): 331–346.
- Griffin R (2022) The sanitised platform. *JIPITEC* 13(36): 36–52.
- Haimson O, Delmonaco D, Nie P, et al. (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: 1–35.
- Hudson B (2003) *Justice in the Risk Society: Challenging and Re-Affirming Justice in Late Modernity*. London: Sage.
- Instagram (n.d). Available at: <https://help.instagram.com/477434105621119>
- Jane EA (2017) ‘Dude . . . stop the spread’: antagonism, agonism, and #manspreading on social media. *International Journal of Cultural Studies* 20(5): 459–475.
- Kaye D (2019) *Speech Police: The Global Struggle to Govern the Internet*. New York: Columbia Global Reports.
- Lumsden K and Morgan H (2017) Media framing of trolling and online abuse, silencing strategies, symbolic violence and victim blaming. *Feminist Media Studies* 17(6): 926–940.
- Marks A (2022) How are trans bodies monitored on Instagram? Meta’s oversight board takes up its first gender identity case. *Rolling Stone*. Available at: <https://www.rollingstone.com/culture/culture-news/meta-oversight-board-trans-body-gender-identity-case-1388117/>
- Myers West S (2018) Censored, suspended, shadowbanned: user interpretations of content moderation on social media platforms. *New Media & Society* 20(11): 4366–4383.
- Nolan-Brown E (2022) The new campaign for a sex-free internet. *Reason*. Available at: <https://reason.com/2022/04/09/the-new-campaign-for-a-sex-free-internet/>
- Norman DA (1988) *The Psychology of Everyday Things*. New York: Basic Books.
- Parkman D (2022) Sex workers, performers invited to participate in platform discrimination study. *XBIZ*. Available at: <https://www.xbiz.com/news/266649/sex-workers-performers-invited-to-participate-in-platform-discrimination-study>
- Peterson C (2013) User-generated censorship: manipulating the maps of social media. *MIT Graduate Program in Comparative Media Studies*. Available at: <https://cms.mit.edu/user-generated-censorship/>

- Savolainen L (2022) The shadow banning controversy: perceived governance and algorithmic folklore. *Media, Culture & Society* 44(6): 1091–1109.
- Schoenebeck S and Blackwell L (2021) Reimagining social media governance: harm, accountability, and repair. *Yale Journal of Law and Technology* 23: 113–152.
- Stardust Z, Garcia G and Egwatu C (2020) What can tech learn from sex workers? Sexual ethics, tech design & decoding stigma. *Berkman Klein Center Collection*. Available at: <https://medium.com/berkman-klein-center/what-can-tech-learn-from-sex-workers-8e0100f0b4b9>
- Stokel-Walker C (2021) TikTok censored a pole-dancing PhD who studies how social media silences women. *Input Mag*. Available at: <https://www.inverse.com/input/culture/tiktok-censored-banned-pole-dancer-phd-carolina-are>
- Stokel-Walker C (2022) TikTok was designed for war. *WIRED*. Available at: <https://www.wired.com/story/ukraine-russia-war-tiktok/>
- Suri S and Gray ML (2019) *Ghost Work: How to Stop Silicon Valley From Building a New Global Underclass*. Boston, MA: Houghton Mifflin Harcourt.
- Tedroff M (2022) Content moderation at TikTok: b*obs, or no b*obs? *Zoning Out*. Available at: <https://zoningout.substack.com/p/content-moderation-at-tiktok-bobs>
- Tiideberg K and Van der Nagel E (2020) *Sex and Social Media*. Melbourne, VIC, Australia: Emerald Publishing.
- TikTok (n.d) Adult nudity and sexual activities. *Community Guidelines*. Available at: <https://www.tiktok.com/community-guidelines?lang=en#30>

Author biography

Carolina Are is a platform governance researcher with a PhD in online abuse and conspiracy theories, currently working as Innovation Fellow at Northumbria University's Centre for Digital Citizens. Following her experiences of online censorship, she has been researching on algorithmic bias against nudity and sexuality on social media, and has published the first study on the shadow-banning of pole dancing in *Feminist Media Studies*. Her work has been published in *Social Media + Society*, *Journalism*, *First Monday* and *Porn Studies*, and it has appeared in *The New York Times*, *The Atlantic*, *The Conversation*, the BBC, *Wired*, the *MIT Technology Review*. She is also a content creator and blogger, as well as an activist and a pole dance instructor, on social media at @bloggeronpole.