
“Our Content Should Be Treated Differently”: Race, Gender, Language, and the Experiences of Quechua Women on Social Media

Dhanaraj Thakur

I examine online gender-based violence (GBV) targeted at Quechua women social media users who post in Quechua and how trust and safety systems address this problem. This builds on previous research on how such systems operate in low-resource languages (Elsawah, Bhatia, and Thakur 2025) and the literature on online GBV in the Majority World. I use an intersectional framework to examine how the overlapping identities of race, gender, and language shape who has power when posting in a low-resource language (i.e., Quechua). Using a 2024/25 survey of Quechua social media users; interviews with content moderators, linguistic activists, and others (Thakur 2025b); and social media comments about Quechua women influencers on TikTok and YouTube during 2025, I find that Quechua women are targeted with online GBV because of their race, gender, and language. The analysis also suggests that trust and safety systems fail to provide protection for these women, in part because moderation and reporting policies are ineffective and because the AI-based tools used by social media companies perform worse in Quechua. I argue that social media platforms should treat the content of historically oppressed groups differently—with greater protection, community engagement, and investment.

1 Introduction

The problem of online gender-based violence (GBV) is well known, particularly in terms of scale and impacts (Dunn 2020). For the purposes of this paper, I use the term online GBV to cover a range of actions that ultimately are designed to control how women, transgender people, and others express themselves online. These actions include violent threats, harassment, stalking, theft of private data, sextortion, the non-consensual distribution of intimate images (NDII), and doxing (Dunn 2020; Kee 2005). Researchers

also examine how online GBV presents unique and unresolved challenges for trust and safety systems within social media. These systems aim to protect users from harmful content and behaviors while promoting trust in the company's products and services (TSPA, [n.d.](#)). They include content moderation, which often combines machine learning (ML) systems with human review and decision-making. Some of the challenges of addressing online GBV include the limitations of automated content moderation (Shenkman, Thakur, and Llansó [2021](#)), and the problems of reporting mechanisms (Lo and Vilk [2023](#)).

A related issue is how social media platforms moderate content in low-resource languages, or languages for which there is little data online to train automated tools such as large language models (LLMs). This contrasts with high-resource languages (e.g., many European languages) for which there is sufficient data online to train models. Research in this area points to the limitations of LLMs for analyzing content in low-resource languages (Nicholas and Bhatia [2023a](#)), and more broadly how trust and safety systems designed in one cultural and linguistic context (high-resource languages) perform worse for content in low-resource languages (Elsawah, Bhatia, and Thakur [2025](#)).

The problem I am concerned with in this paper is related to both of these issues. More specifically, I examine online GBV targeted at Quechua women social media users and how trust and safety systems attempt to address this problem in a low-resource language post-colonial context (i.e., Quechua). While there is an important body of research on online GBV in the Majority World (APC [2017](#)), few have focused on Quechua communities, and fewer still have examined how language “resourcedness” impacts how trust and safety systems address online GBV.

To interrogate this problem, I use an intersectional framework to assess how power is distributed, and how oppression is felt specifically for those holding identities related to race, gender, and language. Researchers have pointed out that by using an intersectional lens we can better understand how multiple sources of oppression and violence are meted out to specific groups of women (Crenshaw [1991](#)). Here, I apply this framework to a specific colonial and linguistic context of social media users in a low-resource language. Thus, I look at the problem of online GBV in terms of race, gender, and language by examining the experiences of Quechua women who post in their language online. I argue that as language is an important means of developing and expressing individual and cultural identity (Bucholtz and Hall [2005](#)), it is also an important variable in intersectional analysis, particularly for communities in the Majority World who have been subject to colonialism and linguistic oppression such as the Quechua (Rivera Cusicanqui [2010](#)).

More specifically, I attempt to understand how language combines with categories of race and gender when it comes to online GBV meted out to Quechua social media users and the extent to which trust and safety systems minimize or enable that abuse. To do that, I employ a mixed-method approach with new analysis of data from a

2024/25 survey of Quechua social media users and interviews with influencers, content moderators, and linguistic activists focused on Quechua (Thakur 2025b). I also conducted an analysis of comments about Quechua women social media influencers on TikTok and YouTube.

I find that Quechua social media users have general concerns about online GBV and that women are more likely to report harassment and sexualized content to the platforms. More specifically, Quechua women social media influencers who post in Quechua or use Quechua in their videos are targeted with distinct types of online GBV from different groups. The analysis also suggests that the trust and safety systems (both the human and automated elements) that underpin these social media platforms fail to provide equitable protections for these women online. The work contributes to the literature on online GBV in the Majority World and how trust and safety systems operate in low-resource language contexts. The findings are also relevant for trust and safety teams and others who want to understand how their human and machine-based processes can better support users in Indigenous and Majority World contexts from an intersectional perspective.

2 Related Literature and Background

I focus on research on content moderation in low-resource languages and research on online GBV, particularly in the Majority World. A significant amount of work on online GBV has examined the prevalence of the problem (Dunn 2020; Hicks 2021; Plan International 2020; Vogels 2021). Others have looked more specifically at the problem from a Majority World perspective (APC 2017; Bansal et al. 2024). In terms of research examining how trust and safety systems address online GBV, scholars have generally found them to be ineffective (Suzor et al. 2019) and even harmful to the humans tasked with addressing the problem (Olsen and Shulruff 2026). Several have critiqued reporting mechanisms which fail to provide survivors access to accountability and justice (Cover et al. 2026; Lo and Vilk 2023; Qiwei et al. 2026) or looked to solutions aimed at perpetrators including for example, deterrence messaging (Rao and Thakur 2025). Others have focused on the problems of using ML tools to address online GBV including classifiers that miscategorize LGBTQ+ creators as adult content creators (Southerton et al. 2021), or systems that unfairly remove content and censor accounts (Are 2020).

The more recent manifestation of AI within trust and safety is the use of large language models (Nicholas and Bhatia 2023b). A major critique in this area is that the trust and safety systems of global social media platforms are built with a Western user base in mind (Shahid and Vashistha 2023), and therefore skew the global distribution of content moderation resources to favor that base (De Gregorio and Stremlau 2023). This includes using models that undervalue morphologically complex languages of the Majority World or common practices such as code-mixing (Shahid, Elswah, and Vashistha 2025). The overall result is less effective moderation outcomes for speakers of low-resource languages on

popular social media platforms (Elsawah, Bhatia, and Thakur 2025).

This can have significant consequences for speakers of low-resource languages, such as increased hate speech and incitement to violence (Tadesse and Bradshaw 2025), vulnerabilities that allow bad actors to target speakers of those languages (Nigatu and Raji 2024), over-moderation and restrictions on free speech (Balendra 2025). However, there is very little work that looks specifically at consequences in terms of online GBV. Indeed, when we consider either the literature on content moderation in low-resource languages or the more mature work on online GBV and content moderation, neither sufficiently examines how trust and safety systems address online GBV in low-resource language contexts. This is the gap that I focus on in this paper. More specifically, I examine online GBV targeted at Quechua women social media users and how trust and safety systems attempt to address this problem in the low-resource language context of Quechua.

To examine this issue, I employ an intersectional analytical framework. Researchers have argued for an intersectional approach to understanding and addressing online GBV given the reality that individuals traverse multiple identities all the time and that online violence targets different identities simultaneously. The oppression people face based on a combination of different identities can be unique and point to disproportionate harms compared to other groups (Collins and Bilge 2020; Crenshaw 1991), suggesting that intersectionality as an approach is critical to understanding online behaviors as well (Bailey 2021; Madden et al. 2018; Finkel et al. 2025; Thakur et al. 2022).

When it comes to trust and safety and specifically content moderation, researchers have suggested existing models of content moderation often replicate societal biases and harms on oppressed groups and that it's therefore important to account for distributions of power (Gilbert 2023). This includes understanding how moderation reinforces existing structures of power across groups through policies and platform affordances (Register et al. 2024).

One aspect of intersectionality that is perhaps less examined within research on online behaviors is language. Although language is a central concept in understanding the impacts of trust and safety systems on users (Elsawah, Bhatia, and Thakur 2025), language is often not part of an intersectional framework used to examine such systems. At the same time, linguists and others have incorporated language into intersectional approaches to their work (Levon 2015). Indeed, language is an important means of identity development and expression, particularly for communities who have faced oppression through colonization or other forms of subjugation such as Indigenous communities. Some researchers have argued that intersectionality is therefore important to examining the context of such communities (Olsen 2018).

Using an intersectional lens (e.g., through language, gender, and race) can help us understand how trust and safety systems address online GBV in low-resource language contexts. I pose three main research questions in this paper:

- **RQ1:** What are attitudes toward online GBV among Quechua social media users?
- **RQ2:** What are the experiences of online GBV among Quechua women social media users?
- **RQ3:** How do trust and safety systems address online GBV targeted at Quechua women social media users who post in Quechua?

2.1 The Quechua Context

Estimates suggest that there are 9–10 million Quechua speakers across South America and elsewhere (Galiano Román 2024). Starting with the Spanish invasion of the region in the sixteenth century, Quechua speakers (of which there are numerous languages and dialects) have been subject to systematic and implicit attempts at erasing their language and culture (Rivera Cusicanqui 2010; Quijano 2019). Discrimination is also evident in various forms of broadcast media, where Quechua voices were historically suppressed (Ojeda Ludena 2025). In practice, because of the long history of linguistic repression, many Quechua speakers today are actually bilingual and use forms of code-switching (alternating between Quechua and Spanish) in offline and online communications (Limachi Pérez 2020; Sánchez 2012).

Social media is viewed as a particularly important means of breaking from power structures that minimized Quechua voices in traditional media while offering more control to Quechua content creators and communities (Coronel-Molina 2005). This includes opportunities to teach Quechua languages (Herrera Corahua 2022), and for Quechua content creators to share their culture (e.g., language, music, and food) with their own communities and others (Mendoza-Mori and Sanchez 2023).

Linguistic activists and others use social media platforms to promote tools and resources to support the language (Galiano Román 2024). Researchers point to other forms of online activism including those that complement movements against femicide and intimate partner violence in Peru (Cavagnaro and Shenton 2019). Given that these are major problems in the region, we would expect there to also be concerns about online GBV, but there is only limited research on this in the Quechua context (Hiperderecho 2018).

3 Methods and Approach

I use a mixed-method approach drawing on three sources of data: (1) an online survey of Quechua social media users, (2) interviews with linguistic activists, social media influencers, and content moderators, and (3) comments to posts from Quechua women social media influencers. The first two are from a previous research project that examined how social media content moderation systems operate in Quechua (Thakur 2025b).

For the first research question (RQ1: What are attitudes toward online GBV among Quechua social media users?), I draw on the survey data. It was conducted by a Peruvian NGO (Hiperderecho) and a U.S.-based non-profit advocacy organization (the Center for Democracy & Technology). Using a convenience sample, the survey included 161 Quechua-speaking persons who used social media frequently (at least several times a week) from Peru, Bolivia, and Ecuador, all countries with significant Quechua-speaking populations. The two organizations administered the online survey between December 2024 and March 2025, and it focused on users' experiences with and perceptions of content moderation, and trust in social media platforms.

To address RQ1, I conducted a statistical analysis of the survey data (using R) to look specifically at perceptions and experiences of online GBV on social media by gender. In addition, I conducted a keyword analysis of one open-ended survey question following Seale and Charteris-Black (2010). This question was a follow-up to whether the respondent had ever reported content to the platform and asked them to describe that content. Here, I focus primarily on the identity of gender within Quechua communities. A summary of the survey respondents is presented in Table 1.

For RQ2 (What are the experiences of online GBV among Quechua women social media users?), I examined comments to posts by Quechua women social media influencers. In all, the research partner identified six influencers in the previous study as part of our effort to understand the Quechua social media ecosystem. Five of the accounts were on TikTok and one on YouTube as summarized in Table 2.¹

While not representative of all Quechua social media influencers, these are some of the well-known women influencers and their posts focusing on Quechua language, music, food, and culture provide a good overview of this type of content. In this analysis, I am concerned with identities of race (Quechua), gender (primarily the women influencers), and language.

For each account I pulled comments for videos posted only in 2025. These were the most viewed videos for the year (capped at 500) and the most recent comments for each video (capped at 1000 comments per video) with the assumption that these limits were reasonable based on review of the number of videos and comments per account. The dataset only includes comments and not the videos themselves. Also, I only included the text of each comment and so emojis, gifs, etc. were not captured.

I then used a fine-tuned XLM-RoBERTa² transformer model to classify comments based on distinct categories of online GBV.³ This model is trained on a corpus of 100 languages and can be used for classification tasks. I fine-tuned the model on two datasets covering online GBV in English and Spanish based on content from Twitter—HatEval (Basile et

1. To maintain user privacy, I do not refer to the account names, nor the comment poster names in this paper. I also do not include the full comment as quotes nor in the original Spanish or Quechua. Interested researchers can contact the author for the complete dataset.

2. <https://huggingface.co/FacebookAI/xlm-roberta-base>

3. Note that I used Claude (Sonnet 4.6) to help develop the Python script for fine-tuning the model.

Table 1: Summary characteristics of survey respondents. $N = 161$.

Characteristic	N (%)
Age group (years)	
18–24	10 (6.2%)
25–34	51 (32%)
35–44	45 (28%)
45–54	25 (16%)
55–64	18 (11%)
65–74	12 (7.5%)
Gender	
Female	73 (45%)
Male	86 (54%)
Other	2 (1%)
Education level	
Secondary or less	7 (4.3%)
Some university	6 (3.7%)
University degree	101 (63%)
Post-graduate	36 (22%)
Refused/unknown	11 (6.8%)
Country of residence	
Bolivia	18 (11%)
Ecuador	53 (33%)
Other/unknown	16 (9.9%)
Peru	73 (45%)
USA	1 (0.6%)

Table 2: Social media accounts used for the comment analysis.

Account	Followers	Platform
1	≈700K	TikTok
2	≈40K	TikTok
3	≈80K	TikTok
4	≈30K	TikTok
5	≈30K	TikTok
6	>1M	YouTube

al. 2019) and from Twitter and TikTok—EXIST 2025 (Plaza et al. 2025).

As a way of operationalizing online GBV, I build on categories developed in the underlying papers for these datasets and employed the following six categories for the classification of comments:

- Hate speech: disparaging or dehumanizing comments on the basis of gender, race, sexual orientation, etc. (Basile et al. 2019).
- Misogynistic threats: comments that target women combined with aggressive language, including direct or indirect threats of physical or psychological (and non-sexual) violence (Plaza et al. 2025).
- Sexual harassment: comments consisting of harassment of a sexual nature including threats of sexual violence (Plaza et al. 2025).
- Stereotyping and dominance: comments that falsely suggest women are not good at certain tasks/roles or that men are superior to women in those roles or generally

Table 3: Model performance summary. Validation classification report; overall accuracy 0.69 on 1,233 validation comments.

Category	Precision	Recall	F1-score	Support
Neutral	0.86	0.72	0.78	712
Hate speech	0.43	0.43	0.43	21
Sexual harassment	0.29	0.38	0.33	32
Stereotyping and dominance	0.43	0.37	0.40	86
Ideological inequality	0.31	0.69	0.43	65
Misogynistic threats	0.71	0.81	0.76	275
Objectification	0.23	0.21	0.22	42
Macro average	0.47	0.52	0.48	1,233
Weighted average	0.72	0.69	0.70	1,233

(Plaza et al. 2025).

- Objectification: includes comments where women are described as objects, with an emphasis on physical appearance, and/or suggestions that women must maintain a certain beauty standard (Plaza et al. 2025).
- Ideological inequality: comments that “reject inequality between men and women, or present men as victims of gender-based oppression” (4).
- Neutral: comments with none of the above characteristics.

Table 3 presents a summary of the model performance across categories. The strongest performance metrics are Misogynistic threats and Neutral.

To address RQ2 and RQ3 (How do trust and safety systems address online GBV targeted at Quechua women social media users who post in Quechua?) I draw on the interview data mentioned earlier. The interviews were conducted between November 2024 and March 2025. There were 15 in all of which I conducted 14 in Spanish and 1 in English; they included 6 women and 9 men. The Peruvian NGO Hiperderecho identified interviewees using a snowball sampling approach. They included Quechua content creators, trust and safety professionals at social media companies, content moderators contracting with those companies, linguistic activists, and academics. Using these transcripts, I conducted a thematic analysis to help understand how trust and safety systems address online GBV for Quechua-related content and how identities of race, gender, and language were treated. More specifically, I used an inductive approach starting with the transcript text, identifying keywords, generating codes, and finally generating overall themes following Naeem et al. (2023). I used NVivo as the main tool to support this analysis.

4 Findings

4.1 RQ1: What are attitudes toward online GBV among Quechua social media users?

The survey data suggests that there were differences in what users were concerned about on social media. For example, a Wilcoxon signed-rank test indicated a significant difference in concern about hate speech on social media ($W = 3603.5$, $p = 0.0172$) between women (median on a 5-point Likert scale was 5) and men (median = 4).

There were also significant results in terms of what people were more likely to report to social media as problematic or harmful content. Although the incidence of reporting between men and women was not significantly different, when asked about the kinds of content they reported on, women were more likely to report about harmful posts including sexual content or sexual harassment. Using a keyword analysis of an open-ended question asking what kinds of content respondents reported on, I grouped responses ($n = 86$) into three categories: (1) sexual content and harassment, (2) misinformation and fraud, and (3) language and culture (about attacks on Quechua language/culture). While there were no significant gender differences in reporting the latter two categories, women (27.5%) were more likely to report posts about sexual content and harassment than men (4.5%) (with a Fisher's exact test $p = 0.0053$).

4.2 RQ2: What are the experiences of online GBV among Quechua women social media users?

Content advisory: This section includes examples of racist, sexist, threatening, and other abusive language.

Several interviewees noted that posting content in Quechua or about Quechua could initiate online GBV. They suggested this attracts criticism and abuse given the history of colonization and current racist attitudes toward Quechua communities. Perhaps unsurprisingly, as one interviewee noted, this abuse is often in Spanish:

“Well, the comments are not in Quechua, they are in Spanish, and they are very aggressive. It is difficult at that point because they verbalize what they feel and they put it there.”

—Content creator (man), November 2024

In most cases, interviewees noted that online GBV referred to women content creators:

“I know that because I’ve talked at length about this with (name), with (name), and she is very discouraged by this, by racist people saying things like Quechua is a dying language. Why do you waste time? You should be learning Chinese or teach English. Things like those are, of course, very hurtful.”

—Linguistic activist (man), January 2025

Another aspect of online GBV is objectification or the idea of controlling how women look or act (i.e., the logic of patriarchy). For example:

“Now you have visible Quechua speakers with hundreds of thousands of followers even, they are going to be targeted and they are going to be insulted, because you don’t represent Quechua women. You have short hair and you are showing too much skin, you’re wearing a miniskirt. When that happened, there was a backlash saying that, well, that skirt, that pollera is too short. This is for, I don’t know, for like the gringo or for the male gays, blah, blah, blah.”

—Linguistic activist (man), January 2025

The harm from these comments can also be severe, particularly when dealing with threats and sexual harassment. At least one interviewee noted the distinction in the language of the comment as well:

“There are comments in Quechua that are sexualized comments and I’m going to be a little bit open with that. They are quite, as I mentioned, they are bold comments or suggestive. What I do is try to filter in the comments, try to filter the specific words that I don’t want to be shown, I don’t want them to appear.”

—Content creator (woman), December 2024

Several interviewees suggested that these problems stemmed from biased content moderation practices among the platforms. Some suggested that Quechua was less likely to be moderated: “If we comment in Spanish or in English, those words will be already regulated but not in Quechua” (Content creator (woman), December 2024). Making a similar point, another interviewee noted that regardless of language, the platforms are generally not good at moderating online GBV:

“Just the day before yesterday, I read such a bold comment. It was very naughty and sassy. I felt bad asking who is seeing my content. Those are comments that are not being moderated on TikTok or Facebook.”

—Content creator (woman), December 2024

Another way to assess experiences of online GBV is to examine social media comments specifically to posts from Quechua women who post in Quechua or use Quechua. Here I focused on social media influencers who primarily post videos on TikTok and YouTube teaching their audiences how to speak Quechua, or content about Quechua food, music, fashion, and culture.

Table 4 below summarizes the classification results from our model. The results include separate classifications by platform as well. We can see that TikTok has proportionally more flagged comments than YouTube. The differences across platforms suggest that

Table 4: Model classification results for all user comments. Online GBV flagged is the sum of the six category rows; percentages are of total comments on each platform.

	All platforms	TikTok	YouTube
Total comments	28,625	24,114	4,511
Online GBV flagged	4,366 (15.3%)	3,895 (16.2%)	471 (10.4%)
Neutral	24,259 (84.7%)	20,219 (83.8%)	4,040 (89.6%)
Misogynistic threats	3,591 (12.5%)	3,261 (13.5%)	330 (7.3%)
Objectification	357 (1.2%)	297 (1.2%)	60 (1.3%)
Stereotyping and dominance	244 (0.9%)	186 (0.8%)	58 (1.3%)
Hate speech	96 (0.3%)	86 (0.4%)	10 (0.2%)
Ideological inequality	66 (0.2%)	58 (0.2%)	8 (0.2%)
Sexual harassment	12 (0.0%)	7 (0.0%)	5 (0.1%)

platform dynamics matter. However, it is hard to draw definite conclusions without a larger sample, and is not the aim of this study.

Of the 28,625 comments in the dataset, 15.3% were flagged as containing some form of online GBV as defined earlier. The most significant category was misogynistic threats followed by objectification. However, the results also suggest that a significant portion of the comments, 30.5% or 8,746 are in Quechua or are mostly Quechua. It is unlikely that the model can perform well on low-resource languages. Further, the model calculates the probabilities for a given comment being classified into each of the six categories for online GBV and then uses the category with the highest probability (the function used in this case will ensure that the sum of the probabilities is 1). To take a more conservative approach I restrict the model outputs to those classifications that have a probability of 70% or more and for comments that are in high-resource languages only (e.g., Spanish and English). Table 5 summarizes those results.

Here only 4 categories are present with misogynistic threats being the most prominent among the GBV categories. However, the overall proportion of comments with some type of online GBV has dropped to 5.9% with TikTok videos again much more likely to contain such comments than YouTube.

The categorization of comments suggests that social media users react to these women and their use of language in particular ways. The two most prominent (following our conservative model approach—Table 5) are misogynistic threats and stereotypes.

Misogynistic threats in this case are those that target women including direct or indirect threats of physical or psychological (and non-sexual) violence. Some of these comments⁴ were direct: “Shut up!” or “ARE YOU A LESBIAN?” (the original was in all caps). However, most of these comments were questions asking the content creator where she lives or is now, ostensibly so that the commenter (a stranger) could meet them: “Hello dear (name),

4. The comments used as examples in this section were originally in Spanish except for those that I note were in Quechua.

Table 5: Model classification results for high-resource-language comments. Comments in Spanish and English with high-confidence results (0.7 or higher) only. Online GBV flagged is the sum of the category rows; percentages are of total comments on each platform.

	All platforms	TikTok	YouTube
Total comments	19,879	15,679	4,200
Online GBV flagged	1,169 (5.9%)	1,059 (6.8%)	110 (2.6%)
Neutral	18,710 (94.1%)	14,620 (93.2%)	4,090 (97.4%)
Misogynistic threats	1,119 (5.6%)	1,017 (6.5%)	102 (2.4%)
Stereotyping and dominance	34 (0.2%)	26 (0.2%)	8 (0.2%)
Ideological inequality	16 (0.1%)	16 (0.1%)	0 (0.0%)

where do you live now?” and “Where will you be so I can see you?”

In some cases, the comment was in Quechua and was particularly violent: “Suck this and choke on it.”

Stereotyping and dominance refer to comments that falsely suggest women are not good at certain tasks/roles, and/or that men are superior to women in those roles or generally. Some of the comments here focused on the women’s role as producer of Quechua-related content. For example: “It seems like she’s not putting much effort into her videos anymore, this girl is getting boring” and “You seem sad and lacking energy. Take a break from videos and watch shorts.”

Some of the comments and insults in Quechua were unique to the language: “You are speaking pure nonsense in Quechua with a dog’s mouth,” where reference to the “dog’s mouth” suggests that they are not credible. While these attacks about language were in response to posts by the content creators, some were explicitly focused on the fact that they were women: “Is this little Indian woman crazy? Why does she act like that, I saw this video by mistake and it’s horrible, this person speaks very poorly.”

Ideological inequality was the final category flagged in our conservative model (Table 5). This refers to comments that reject notions of inequality between men and women. While seldom, one example of that is the comment: “‘How difficult it is to be a woman’... hahaha Good one.”

Objectification was not part of the conservative model, but it was part of the classification of comments in the full model (see Table 4) and worth noting since this was mentioned in the interview analysis earlier. This includes comments on physical appearance, and/or suggestions that women must maintain a certain beauty standard. Indeed, many of the comments in this category were just that and often disconnected from the content of the videos which were mostly about Quechua language, music, and food: “That old woman really got ugly” and “Mamacity, rich, beautiful, gorgeous, hot.”

Hate speech includes disparaging or dehumanizing comments on the basis of gender,

race, sexual orientation, etc. Again, while not frequent, one illustration of this was: “You look all ordinary and poor, and from the mountain it looks like you’re like Tarzan.”

Finally, **sexual harassment**, including threats of sexual violence, seldom occurred in the comments (it is unclear why based on these data). However, some remained and often consisted of people asking the content creator to be their girlfriend or partner. Here commenters often used the content creator’s name to make it more personal: “Name, will you be my girlfriend? You’ve been my crush since I met you” and “Name, my love, I want you to be my wife.”

The same pattern occurred in comments in Quechua: “Name, I am going to gaze at your little eyes.”

4.3 RQ3: How do trust and safety systems address online GBV targeted at Quechua women social media users who post in Quechua?

The findings so far suggest that Quechua social media users, and women in particular, are concerned about and report on harmful and abusive content on social media. In addition, Quechua women social media influencers who post in Quechua are subject to various forms of online GBV. It’s important to understand the way trust and safety systems on these platforms are designed and managed when it comes to addressing comments that relate to the language, gender, and race of the content creator. To do that, I draw on the interview data described earlier. The analysis points to several relevant themes.

Several interviewees noted the ways in which the trust and safety systems of the platforms they are active on negatively impact their experiences. In most cases, the multinational companies behind popular platforms outsourced content moderation functions to regional firms. These firms typically did not hire sufficient staff fluent in Quechua or other Indigenous languages, which limited how Quechua content was moderated, something several interviewees pointed to, as noted earlier.

The ways in which content moderation policies were operationalized for Quechua content had direct impacts on the moderation of online GBV. For example, these policies, which typically outlined which types of content including specific phrases, images, etc. were permissible on the platform and which were not, lacked guidance for phrases and terms in Quechua. This was partly because the policies were originally written in English and then translated to Spanish and then there were attempts to translate that to Quechua as described by one interviewee:

“The company would ask us to make a specific translation. In many cases, you cannot translate the same word with the term. There is no word for that same term. In those cases, it was very complicated. We had to submit a whole report to explain why that violation is similar to one in Spanish and that it had to be treated in the same way.”

—Content moderator (woman), February 2025

To compound matters there were many different policies in Spanish for various content categories. One interviewee suggested as many as 18, which they were ultimately required to work with and adapt to Quechua as best as they could. By assuming that all languages have equivalent words to those in Spanish or English the companies create a system that privileges those European languages over others such as Quechua. This in turn had implications for the moderation of online GBV:

“However, with the original [Indigenous] languages, a term of hate is very difficult because it is always a phrase or a concept that has all of the meaning of the hate. It cannot be enclosed within a word for that matter, which is how the organization approached it”

—Content moderator (woman), February 2025

Following this pattern of privileging Spanish over Quechua, it is also worth noting that the classifiers and other ML-based tools used by platforms for moderation such as large language models will also be more effective in Spanish than in Quechua (Thakur 2025b). Indeed, as noted earlier, one interviewee reported that abusive comments in Quechua were often left online while the Spanish ones were not.

5 Discussion

“The fact of being a native, a woman, an Indigenous woman, also has its own ways.”

—Quechua content creator (woman), December 2024

The findings offer insight into how language interacts with perceptions of gender and race to influence the comments that Quechua women content creators receive on social media. Among Quechua social media users, both men and women are concerned about abuse meted out to them when they or others use Quechua online. Women are also more concerned about hate speech and sexual harassment than men.

These concerns align with the actual experiences of Quechua women content creators who post videos in Quechua or about Quechua language, food, music, fashion, etc. All the women were subject to abuse in their comments. Most of this abuse revolves around misogynistic threats, stereotyping and dominance, and ideological inequality (using our conservative classification model).

These forms of online GBV were potentially enabled by trust and safety systems of the social media platforms that either did not sufficiently address the online abuse or privileged Spanish over Quechua, leaving abuse in the latter online. This was most evident in how the content moderation policies, originally designed in English, were badly operationalized for Quechua content.

Indeed, throughout I found that language identity compounded targeted abuse based on related identities of race and gender. Interestingly, the explicit use of Quechua (often as means of promoting the Quechua identity and culture) also made the women content creators targets of abuse from both within and outside the community. These experiences of Quechua women on social media resemble those of Black women in the U.S. who are subject to hypervisibility (Newton 2023). That is, their race and gender together make them more visible in communities where men dominate (e.g., within Quechua communities) or where Spanish dominates (i.e., more broadly). Taken together, Quechua women who post in Quechua become more visible targets for discrimination online.

The focus on the intersection of race, gender, and language is therefore important to understand the experiences of these women. It also points to ways in which trust and safety systems shift power away from these women, particularly their ability to express themselves equally on social media, and toward men and Spanish speakers.

These patterns are not ad hoc but instead a function of the asymmetrical power relationship between Spanish and Quechua, which was built through centuries of colonialism (Rivera Cusicanqui 2010). The historical and current oppression of Quechua leads content creators to post about their language and culture in the first place (Mendoza-Mori and Sanchez 2023). Yet that historical relationship, including expressions of misogyny and hate also manifests itself in abusive comments levied at the women content creators. Indeed, that same asymmetry also informs how trust and safety systems and content policies are built and operated often privileging Spanish over Quechua and enabling some of the observed online GBV.

The cumulative effect of all this is to undermine the power of Quechua women social media users to use the platforms in ways that they wish. In this way, the platforms through their policies replicate colonial and patriarchal systems of power (Register et al. 2024) that subjugate women and specifically Quechua women. These findings may point to similar experiences for women who post in other Indigenous and low-resource languages around the world.

6 Conclusion

The findings in this paper contribute to the literature on how trust and safety systems operate in low-resource languages and in addressing online GBV in the Majority World. They also have implications for trust and safety teams. Intersectionality is not just an analytical tool but also one that can shift power through intentional outcomes (Collins and Bilge 2020). Thus, recognizing that an intersectional approach can help trust and safety systems to be more equitable could be a first step for the social media platforms. These companies should, however, resist the urge to focus simply on intersecting identities (e.g., developing more precise labeling techniques) for automated analysis (Thakur 2025a).

An intersectional approach must consider different and unique forms of oppression and how unequal distributions of power shape that. Thus, a second step might be to prioritize the needs and safety of those who have been and continue to face discrimination and abuse because of their intersecting identities—e.g., the trust and safety teams translated policies from Spanish to Quechua but never started with Quechua to understand how online GBV manifests in that culture.

As one interviewee noted:

“I believe that the contribution that we as cultural ambassadors make and as people who struggle for our rights, should be recognized. We think that this is a very important service we are doing to know ourselves and recognize ourselves and contribute to a culture of harmony between human beings. I think that these companies should prioritize the distinct contribution that we make. Our content should be treated differently.”

—Content creator (woman), February 2025

In other words, by treating everyone the same, social media companies risk ignoring the fact that many people are treated unfairly in society. Real equity will require steps to address centuries of oppression and that might mean treating them differently and with priority. This can include investing more in protecting Quechua speakers and women content creators, engaging with Quechua communities to better understand how platforms can avoid replicating systems of oppression, and investing in research and community-led efforts to address the low-resource language challenges facing Quechua.

6.1 Limitations

The dataset of social media comments consists of those on the platform at the time of data collection and not those removed by the platforms for violating policies. The comments themselves sometimes include Quechua or a combination of Quechua and Spanish. Multilingual models such as XLM-RoBERTa may be trained on many languages but are not as effective on low-resource languages such as Quechua and thus I used a higher confidence level and focused on Spanish comments as noted earlier. My analysis did not include gifs or emojis, a common element in user comments and something future research can address.

A different approach to the interviews could have focused on the problem of online GBV directly with women social media influencers, although that would require a more supportive approach for respondents engaging in potentially traumatic discussions (which was not covered in the IRB protocol for this research). These are gaps that future research can address.

References

- Are, Carolina. 2020. "How Instagram's Algorithm Is Censoring Women and Vulnerable Users but Helping Online Abusers." *Feminist Media Studies* 20 (5): 741–44. <https://doi.org/10.1080/14680777.2020.1783805>.
- Association for Progressive Communications. 2017. *Online Gender-Based Violence: A Submission from the Association for Progressive Communications to the United Nations Special Rapporteur on Violence against Women, Its Causes and Consequences*. Association for Progressive Communications' Women's Rights Programme. <https://www.apc.org/en/pubs/online-gender-based-violence-submission-association-progressive-communications-united-nations>.
- Bailey, Moya. 2021. *Misogynoir Transformed: Black Women's Digital Resistance*. New York: New York University Press.
- Balendra, Soorya. 2025. "Meta's AI Moderation and Free Speech: Ongoing Challenges in the Global South." *Cambridge Forum on AI: Law and Governance* 1:e21. <https://doi.org/10.1017/cfl.2025.5>.
- Bansal, Vaiddehi, Mayumi Rezwan, Mithila Iyer, Erin Leasure, Connor Roth, Poulomi Pal, and Laura Hinson. 2024. "A Scoping Review of Technology-Facilitated Gender-Based Violence in Low- and Middle-Income Countries Across Asia." *Trauma, Violence, & Abuse* 25 (1): 463–75. <https://doi.org/10.1177/15248380231154614>.
- Basile, Valerio, Cristina Bosco, Elisabetta Fersini, Debora Nozza, Viviana Patti, Francisco Manuel Rangel Pardo, Paolo Rosso, and Manuela Sanguinetti. 2019. "SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter." In *Proceedings of the 13th International Workshop on Semantic Evaluation*, edited by Jonathan May, Ekaterina Shutova, Aurelie Herbelot, Xiaodan Zhu, Marianna Apidianaki, and Saif M. Mohammad, 54–63. <https://doi.org/10.18653/v1/S19-2007>.
- Bucholtz, Mary, and Kira Hall. 2005. "Language and Identity." In *A Companion to Linguistic Anthropology*, edited by Alessandro Duranti, 369–94. John Wiley & Sons. <https://doi.org/10.1002/9780470996522.ch16>.
- Cavagnaro, Kellie, and Jamie E. Shenton. 2019. "Territories of Contested Womanhood: Pussyhats, the Pachamama, and Embodying Resistance in the Era of Hashtag Feminism." *Journal of Cultural Analysis and Social Change* 4 (1). <https://doi.org/10.20897/jcasc/5847>.
- Collins, Patricia Hill, and Sirma Bilge. 2020. *Intersectionality*. Cambridge, UK: Polity Press.
- Coronel-Molina, Serafín M. 2005. "Lenguas originarias cruzando el puente de la brecha digital: Nuevas formas de revitalización del quechua y el aimara." In *Lenguas e identidades en los Andes: perspectivas ideológicas y culturales*, edited by Linda L. Grabner-Coronel and Serafín M. Coronel-Molina, 31–82. Quito: Abya-Yala.

- Cover, Rob, Jennifer Beckett, Benedetta Brevini, Catharine Lumby, Rhyle Simcock, and Jay Daniel Thompson. 2026. "Reporting Online Abuse to Platforms: Factors, Interfaces and the Potential for Care." *Convergence* 32 (1): 142–58. <https://doi.org/10.1177/13548565251324508>.
- Crenshaw, Kimberlé. 1991. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43 (6): 1241–99. <https://doi.org/10.2307/1229039>.
- De Gregorio, Giovanni, and Nicole Stremlau. 2023. "Inequalities and Content Moderation." *Global Policy* 14 (5): 870–79. <https://doi.org/10.1111/1758-5899.13243>.
- Dunn, Suzie. 2020. *Technology-Facilitated Gender-Based Violence: An Overview*, Supporting a Safer Internet Paper 1. Centre for International Governance Innovation. <https://www.cigionline.org/publications/technology-facilitated-gender-based-violence-overview/>.
- Elsawah, Mona, Aliya Bhatia, and Dhanaraj Thakur. 2025. *Content Moderation in the Global South: A Comparative Study of Four Low-Resource Languages*. Center for Democracy & Technology. <https://cdt.org/insights/content-moderation-in-the-global-south-a-comparative-study-of-four-low-resource-languages/>.
- Finkel, Müge, Dhanaraj Thakur, Steven Finkel, Firat Duruşan, Erdem Yörük, Işık Topçu, Melih Can Yardi, Amanda Zaner, and Jungmin Han. 2025. "Intersectional Hatred - An Application of Large Language Models to Detect Hate and Offensive Speech Targeted at Congressional Candidates in the 2024 U.S. Election." In *Companion Proceedings of the ACM on Web Conference 2025*, 2776–81. <https://doi.org/10.1145/3701716.3716880>.
- Galiano Román, Juan. 2024. "Activismo digital en redes sociales y la promoción de los derechos lingüísticos de los quechuas de Cusco." *Syntagmas* 3 (2): 171–99. <https://doi.org/10.51343/syntagmas.v3i3.1316>.
- Gilbert, Sarah. 2023. "Towards Intersectional Moderation: An Alternative Model of Moderation Built on Care and Power." *Proceedings of the ACM on Human-Computer Interaction* 7 (CSCW2). <https://doi.org/10.1145/3610047>.
- Herrera Corahua, Juan Bautista. 2022. "El uso de las TIC en la EIB como herramienta para la revitalización del quechua." *Lengua y Sociedad* 21 (1): 463–79. <https://doi.org/10.15381/lengsoc.v21i1.22494>.
- Hicks, Jacqueline. 2021. *Global Evidence on the Prevalence and Impact of Online Gender-Based Violence (OGBV)*. The Institute of Development Studies and Partner Organisations. <https://doi.org/10.19088/K4D.2021.140>.
- Hiperderecho. 2018. *Conocer para resistir: Violencia de género en línea en Perú*. Hiperderecho. https://hiperderecho.org/tecnorendencias/wp-content/uploads/2019/01/violencia_genero_linea_peru_2018.pdf.

- Kee, Jac S. M. 2005. *Cultivating Violence through Technology?* Association for Progressive Communications Women's Networking Support Programme. <https://www.apc.org/en/pubs/cultivating-violence-through-technology>.
- Levon, Erez. 2015. "Integrating Intersectionality in Language, Gender, and Sexuality Research." *Language and Linguistics Compass* 9 (7): 295–308. <https://doi.org/10.1111/lnc3.12147>.
- Limachi Pérez, Vicente. 2020. "El translenguaje digital, estrategia discursiva ecológica de jóvenes bilingües quechua–Castellano en Facebook y Whatsapp." *Ecolinguística: Revista brasileira de ecologia e linguagem (ECO-REBEL)* 6 (1): 83–103. <https://periodicos.unb.br/index.php/erbel/article/view/29898>.
- Lo, Kat, and Viktorya Vilks. 2023. *Shouting into the Void: Why Reporting Abuse to Social Media Platforms Is So Hard and How to Fix It*. PEN America. <https://pen.org/report/shouting-into-the-void/>.
- Madden, Stephanie, Melissa Janoske, Rowena Briones Winkler, and Amanda Nell Edgar. 2018. "Mediated Misogynoir: Intersecting Race and Gender in Online Harassment." In *Mediating Misogyny: Gender, Technology, and Harassment*, edited by Jacqueline Ryan Vickery and Tracy Everbach, 71–90. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-72917-6_4.
- Mendoza-Mori, Américo, and Mathias A. Becerra Sanchez. 2023. "Quechuactivism in Social Media: Digital Content and Indigenous Language Awareness." *ReVista*, April 25, 2023. Accessed April 8, 2025. <https://revista.drclas.harvard.edu/quechuactivism-in-social-media-digital-content-and-indigenous-language-awareness/>.
- Naeem, Muhammad, Wilson Ozuem, Kerry Howell, and Silvia Ranfagni. 2023. "A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research." *International Journal of Qualitative Methods* 22. <https://doi.org/10.1177/16094069231205789>.
- Newton, Veronica A. 2023. "Hypervisibility and Invisibility: Black Women's Experiences with Gendered Racial Microaggressions on a White Campus." *Sociology of Race and Ethnicity* 9 (2): 164–78. <https://doi.org/10.1177/23326492221138222>.
- Nicholas, Gabriel, and Aliya Bhatia. 2023a. "Lost in Translation: Large Language Models in Non-English Content Analysis." Center for Democracy and Technology, May 23, 2023. Accessed April 28, 2025. <https://cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis/>.
- . 2023b. "Toward Better Automated Content Moderation in Low-Resource Languages." *Journal of Online Trust and Safety* 2 (1). <https://doi.org/10.54501/jots.v2i1.150>.

- Nigatu, Hellina Hailu, and Inioluwa Deborah Raji. 2024. ““I Searched for a Religious Song in Amharic and Got Sexual Content Instead”: Investigating Online Harm in Low-Resourced Languages on YouTube.” In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 141–60. <https://doi.org/10.1145/3630106.3658546>.
- Ojeda Ludena, Jermani. 2025. “Indigenous Voices of the Andes: Quechua People and Their Contestation to Misti Settler-Colonial Broadcasting.” *Latin American and Caribbean Ethnic Studies* 20 (2): 201–19. <https://doi.org/10.1080/17442222.2025.2468554>.
- Olsen, Erica, and Toby Shulruff. 2026. “Trust and Safety: An Approach to Countering Gender-Based Violence.” In *Trust, Safety, and the Internet We Share*, edited by Maia Levy Daniel, Amanda Menking, Marlyn Savio, and Jean Claffey, 235–44. New York, NY: Routledge. <https://doi.org/10.1201/9781003621072-27>.
- Olsen, Torjer A. 2018. “This Word Is (Not?) Very Exciting: Considering Intersectionality in Indigenous Studies.” *NORA - Nordic Journal of Feminist and Gender Research* 26 (3): 182–96. <https://doi.org/10.1080/08038740.2018.1493534>.
- Plan International. 2020. *Free to Be Online? A Report on Girls’ and Young Women’s Experiences of Online Harassment*. Plan International. <https://plan-international.org/publications/freetobeonline>.
- Plaza, Laura, Jorge Carrillo-de-Albornoz, Iván Arcos, Paolo Rosso, Damiano Spina, Enrique Amigó, Julio Gonzalo, and Roser Morante. 2025. “EXIST 2025: Learning with Disagreement for Sexism Identification and Characterization in Tweets, Memes, and TikTok Videos.” In *Advances in Information Retrieval*, edited by Claudia Hauff, Craig Macdonald, Dietmar Jannach, Gabriella Kazai, Franco Maria Nardini, Fabio Pinelli, Fabrizio Silvestri, and Nicola Tonellotto, 442–49. https://doi.org/10.1007/978-3-031-88720-8_65.
- Qiwei, Li, Katelyn Kennon, Nicole Bedera, Asia A. Eaton, Eric Gilbert, and Sarita Schoenebeck. 2026. “Platforms as Crime Scene, Judge, and Jury: How Victim-Survivors of Non-Consensual Intimate Imagery Report Abuse Online.” In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3772318.3791115>.
- Quijano, Aníbal. 2019. “Colonialidad del poder, eurocentrismo y América Latina.” *Espacio Abierto* 28 (1): 255–301. <https://www.redalyc.org/journal/122/12262976015/>.
- Rao, Varun, and Dhanaraj Thakur. 2025. “Think Twice Before You Search: Deterrence Messaging Designs to Prevent Searches for Non-Consensual Intimate Images.” Center for Democracy and Technology, November 20, 2025. Accessed June 29, 2026. <https://cdt.org/insights/think-twice-before-you-search-deterrence-messaging-designs-to-prevent-searches-for-non-consensual-intimate-images/>.

- Register, Yim, Izzi Grasso, Lauren N. Weingarten, Lilith Fury, Constanza Eliana Chinae, Tuck J. Malloy, and Emma S. Spiro. 2024. "Beyond Initial Removal: Lasting Impacts of Discriminatory Content Moderation to Marginalized Creators on Instagram." *Proceedings of the ACM on Human-Computer Interaction* 8 (CSCW1). <https://doi.org/10.1145/3637300>.
- Rivera Cusicanqui, Silvia. 2010. *Ch'ixinakax utxiwa: una reflexión sobre prácticas y discursos descolonizadores*. Buenos Aires: Tinta Limón.
- Sánchez, Liliana. 2012. "Convergence in Syntax/Morphology Mapping Strategies: Evidence from Quechua–Spanish Code Mixing." *Lingua* 122 (5): 511–28. <https://doi.org/10.1016/j.lingua.2011.10.004>.
- Seale, Clive, and Jonathan Charteris-Black. 2010. "Keyword Analysis: A New Tool for Qualitative Research." In *The SAGE Handbook of Qualitative Methods in Health Research*, edited by Ivy Bourgeault, Robert Dingwall, and Raymond De Vries, 536–56. SAGE Publications. <https://doi.org/10.4135/9781446268247.n28>.
- Shahid, Farhana, Mona Elswah, and Aditya Vashistha. 2025. "Think Outside the Data: Colonial Biases and Systemic Issues in Automated Moderation Pipelines for Low-Resource Languages." *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 8 (3): 2331–44. <https://doi.org/10.1609/aies.v8i3.36719>.
- Shahid, Farhana, and Aditya Vashistha. 2023. "Decolonizing Content Moderation: Does Uniform Global Community Standard Resemble Utopian Equality or Western Power Hegemony?" In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. <https://doi.org/10.1145/3544548.3581538>.
- Shenkman, Carey, Dhanaraj Thakur, and Emma Llansó. 2021. *Do You See What I See? Capabilities and Limits of Automated Multimedia Content Analysis*. Center for Democracy & Technology. <https://cdt.org/insights/do-you-see-what-i-see-capabilities-and-limits-of-automated-multimedia-content-analysis/>.
- Southerton, Clare, Daniel Marshall, Peter Aggleton, Mary Lou Rasmussen, and Rob Cover. 2021. "Restricted Modes: Social Media, Content Classification and LGBTQ Sexual Citizenship." *New Media & Society* 23 (5): 920–38. <https://doi.org/10.1177/1461444820904362>.
- Suzor, Nicolas, Molly Dragiewicz, Bridget Harris, Rosalie Gillett, Jean Burgess, and Tess Van Geelen. 2019. "Human Rights by Design: The Responsibilities of Social Media Platforms to Address Gender-Based Violence Online." *Policy & Internet* 11 (1): 84–103. <https://doi.org/10.1002/poi3.185>.
- Tadesse, Kirubel, and Samantha Bradshaw. 2025. "Ethiopia's Civil War and Patterns of Exclusion in Global Content Moderation Regimes." In *The Routledge Companion to Digital Media and Democracy*, edited by Zizi Papacharissi, 349–61. New York, NY: Routledge. <https://doi.org/10.4324/9781003477570-35>.

- Thakur, Dhanaraj. 2025a. "Intersectionality and the Problems of Using Artificial Intelligence to Address Online Gender-Based Violence and Gendered Disinformation." In *Routledge Handbook of Social Media, Law and Society*, edited by Kim Barker and Olga Jurasz, 19–32. Routledge. <https://doi.org/10.4324/9781003252740-7>.
- . 2025b. *Moderating Quechua Content on Social Media*. Center for Democracy & Technology. <https://cdt.org/insights/moderating-quechua-content-on-social-media/>.
- Thakur, Dhanaraj, DeVan L. Hankerson, Michal Luria, Saiph Savage, Maria Rodriguez, and Miriam G. Valdovinos. 2022. *An Unrepresentative Democracy: How Disinformation and Online Abuse Hinder Women of Color Political Candidates in the United States*. Center for Democracy & Technology. <https://cdt.org/insights/an-unrepresentative-democracy-how-disinformation-and-online-abuse-hinder-women-of-color-political-candidates-in-the-united-states/>.
- Trust & Safety Professional Association. n.d. "Trust & Safety Curriculum." <https://www.tspa.org/curriculum/ts-curriculum/>.
- Vogels, Emily A. 2021. "The State of Online Harassment." Pew Research Center, January 13, 2021. Accessed March 10, 2026. <https://www.pewresearch.org/internet/internet/2021/01/13/the-state-of-online-harassment/>.

Authors

Dhanaraj Thakur is the Director of the Fair Technology Initiative, Multiracial Democracy Project, George Washington University Law School. Email: [dhanaraj.thakur@gwu.edu](mailto:ghanaraj.thakur@gwu.edu).

Positionality Statement

I identify as a man from Latin America and the Caribbean and as a social scientist and advocate. I recognize my privilege in navigating issues of online GBV while not typically being subject to it myself. I am also not Quechua and acknowledge the limitations that brings to this work.

Acknowledgments

Thanks to the reviewers and editors for their valuable feedback. Also, thanks to the Center for Democracy & Technology for access to the survey and interview data used in this research. I am very grateful to the respondents and interviewees who participated in the surveys and interviews.

Data Availability Statement

Please email the author for access to the dataset of social media comments used in this paper.

Funding Statement

No external funding supported the work for this paper.

Ethical Standards

The study received an exemption after review from a U.S. based independent institutional review board, Sterling IRB (protocol 11812), in March 2024. I did not present the comments to the posts of the Quechua women social media users in their original Spanish or Quechua. Instead, I only presented the English translations to protect the privacy of the commenters and women. The original text is available in the social media comments dataset available from the author.

Keywords

Online GBV; trust and safety; low-resource languages; Quechua; intersectionality.