



DigitalRightsFoundation
"KNOW YOUR RIGHTS"

CASE STUDY

VIRAL MISOGYNY AND THE KILLING OF SANA YOUSAF



About

Digital Rights Foundation (DRF) is a women-led, not-for-profit organization based in Pakistan, working since 2013 to advance digital rights, freedoms, and online safety for all. While our work centers on the lived experiences of women and gender minorities, we actively support and collaborate with religious minorities, human rights defenders, journalists, and civil society organizations across the region. We believe in the power of a free and secure internet, one that enables expression without fear, protects personal privacy, and resists censorship. At DRF, digital rights are not abstract ideals; they are the foundation for dignity, equality, and justice in the digital age. Through direct interventions, policy advocacy, and strategic partnerships, DRF delivers digital security trainings, provides emergency assistance, and builds resilience among those most at risk. We work on the frontlines to combat technology-facilitated gender-based violence, promote access to information, and strengthen protections for those under threat, especially in marginalized communities. As surveillance intensifies and digital spaces become more hostile, DRF confronts both state and corporate actors through research, monitoring, and rights-based advocacy. We challenge repressive policies, expose violations, and propose practical, rights-affirming solutions. Our goal is clear: to create digital environments where safety, agency, and freedom are not privileges but guarantees for everyone.

Contact information:

info@digitalrightsfoundation.pk
www.digitalrightsfoundation.pk

Acknowledgements

This case study examines the brutal murder of Sana Yousaf, a popular online content creator and influencer, and the subsequent wave of online misogyny, victim-blaming, gendered disinformation, and public glorification of her killing across digital platforms. These patterns collectively constitute technology-facilitated gender-based violence (TFGBV). Through a review of the hateful and violent commentary directed at Sana even after her death this report exposes how digital spaces in Pakistan are increasingly weaponized to dehumanize women and incite harm. Sana's murder is not an isolated incident; it mirrors a broader pattern of femicide in the country, most notably the 2016 killing of Qandeel Baloch, another woman targeted for her digital visibility.

This case study underscores the urgent need for social media platforms to:

Strengthen content moderation in regional languages,

Recognize and act upon cultural nuances in online abuse, and

Implement comprehensive protections for women and marginalized communities.

Authors: Sara Imran, Research Associate
Seerat Khan, Senior Research Associate
Maria Nazar, Research Associate
Nighat Dad, Executive Director

Introduction

On June 2, 2025, 17-year-old [Sana Yousaf](#) was murdered in her home in Islamabad after repeatedly rejecting the advances of a 22-year-old man from Faisalabad. A prominent social media influencer, Sana had nearly 800,000 followers on TikTok and close to 500,000 on Instagram. Her murder is the latest in a disturbing pattern of violence targeting women content creators in Pakistan. In the aftermath, social media platforms were flooded with graphic content and disturbing praise for her killer, exposing deeply entrenched misogyny, gendered hate speech, and technology-facilitated gender-based violence (TFGBV).

Earlier this year in Quetta, another shocking case emerged. A man who had recently returned from the United States admitted to killing his [15-year-old daughter over her TikTok videos](#). Initially claiming she had been shot by unidentified attackers, he later confessed. Police said the family objected to her digital presence and lifestyle.

These are not isolated incidents. In 2021, TikTok star [Muskan Sheikh](#) and her three male colleagues, Amir, Rehan, and Sajjad, were gunned down in Karachi. While police cited a personal dispute, public discourse frequently framed their content as justification for violence, labeling it “obscene.”

The same year, content creator [Ayesha Akram was assaulted](#) by a mob of over 400 men while filming at Lahore’s Greater Iqbal Park on Independence Day. Despite the severity of the assault, which included public groping, harassment, and viral footage, many blamed her rather than the perpetrators. Though she later pardoned the suspects, the case remains emblematic of the public hostility women face for simply existing online.

These incidents echo the tragic legacy of [Qandeel Baloch](#), a digital influencer who in 2016 was murdered by her brother for “bringing dishonor” to the family. Although he was sentenced to life in prison, he was acquitted in 2022 after their parents pardoned him, highlighting systemic flaws in laws surrounding so-called “honour” crimes.

Despite TikTok’s widespread popularity in Pakistan, with over [54 million users](#), the app has faced [repeated bans](#) for allegedly “spreading

obscenity.” In 2021 alone, it was [blocked four times](#). Though currently accessible, TikTok now operates under heavy [content moderation policies](#) and complies with frequent state takedown requests. The state’s narrative portraying the platform as immoral, obscene and vulgar has fueled a culture of public hostility and surveillance, particularly targeting women creators.

It is pertinent to mention that Pakistan remains one of the most unequal countries in terms of gender parity, ranking 145th out of 146 in the [World Economic Forum’s Global Gender Gap Report 2024](#). Women’s participation in the digital economy remains “alarmingly low,” due to structural barriers like digital illiteracy, affordability, lack of safety, and restrictive gender norms. According to the [OECD’s Bridging the Digital Gender Divide report](#), 327 million fewer women than men globally have access to smartphones and mobile internet. These disparities are further intensified in Pakistan by cultural and systemic inequities that continue to sideline women from digital spaces.

DRF’s own [Digital Security Helpline recorded](#) 3,171 new cases in 2024 alone, a majority involving women. Over the past eight years, the helpline has documented over 20,000 cases of TFGBV and online threats, numbers that have only grown.

DRF worked on this case study by analysing **180 comments across TikTok, Instagram, Facebook, and X** after Sana’s murder. In addition, DRF also received comments from the public on its Digital Security Helpline around escalating online hateful comments against Sana Yousaf. The comments revealed propagated false or exaggerated claims rooted in misogynistic stereotypes for example, implying that Yousuf’s social media presence itself was an immoral act and “justifying” the attack under conservative patriarchal norms. This narrative plays into Pakistan’s deep-rooted history of *honour culture*, where women’s behavior is heavily policed both in online and offline spaces. On various platforms, users glorified the killer and painted the murder as an almost inevitable consequence. Disinformation specifically targeted Sana’s gender and public presence, because she was a young woman in the public eye, some argued her fate was self-inflicted due to the choice she made of making her life public over social media. This victim-blaming through false moral framing is a form of gendered disinformation and gendered hate speech which has shifted the blame from the perpetrator to the victim’s character or actions.

In the wake of Sana Yousuf's murder, DRF observes that online platforms must take urgent and meaningful action to protect women content creators, especially those from the Global South. Current content moderation systems are failing to adequately detect or address hate speech, particularly when it appears in regional languages or different dialects like Roman Urdu. Without culturally informed safeguards and stricter enforcement of policies, digital spaces will remain dangerous and exclusionary for women.

Platforms Obligations

[Online gender-based violence \(GBV\)](#) is widespread, with **38%** of women globally experiencing abuse and **85%** witnessing it. This includes misogynistic slurs, harassment, doxxing, and non-consensual imagery, especially targeting women in public roles such as journalists and content creators. In the Global South, the impact is intensified by language gaps and weak moderation. Social media platforms have a responsibility to curb TFGBV.

The [UN Guiding Principles on Business and Human Rights \(UNGPs\)](#) outline that companies must respect human rights by preventing and addressing harms like online GBV. This includes assessing features for misuse, monitoring gendered impacts, and conducting due diligence. In the past, independent bodies like the Oversight Board have given verdicts around cases of visuals relating to gender based violence and gendered content appeals that count as best practices when looking at content that is fueling online hate against women and gender minorities. In August 2023, the board [overturned Meta's decision](#) to leave up a Facebook post mocking a woman targeted by violence (an Eritrean activist in Iraq). The Board found a "gap in Meta's rules" because content praising, justifying, celebrating or mocking gender-based violence (GBV) was not explicitly banned. The Board recommended that Meta "undertake a policy development process" to prohibit content that normalizes GBV in any form. The Board also noted that enforcement may have failed here due to "inadequate sensitivity to different dialects of Arabic" and a lack of transparency on the classifiers in that region. This is particularly important seeing that social media platforms lack recognizing different dialects of Urdu which in turn in Sana's case, leads to weaker enforcement when comments are fueled with online hate and disinformation.

UNGPs also call for grievance mechanisms, easy abuse reporting, timely takedowns, appeals, and support referrals. Platforms must ensure internal accountability and transparency. UN bodies have emphasized that online gender based violence silences women's voices and must be treated as a human rights violation.

Analysis

When news of Sana Yousuf's murder first surfaced, it was widely reported that a 17-year-old TikTokker had been killed in her home by an intruder. With limited verified information available, social media was quickly flooded with speculation and misinformation about the case. Many users, seeing Sana's public profile as an influencer, assumed the killing was an "honour" crime, blaming her family for disapproving of her online presence. Later, it was confirmed that the perpetrator was a man who had been repeatedly rejected by Sana, making it a case of gendered violence rooted in rejection and consent, not family-driven honour killing.

In the immediate aftermath, while some users expressed condolences under her TikTok videos, a much larger wave of online content turned hateful. Across platforms, users spread disinformation, glorified the murderer, and vilified Sana for her visibility and lifestyle as an influencer. Many invoked religious and cultural values to justify the violence, accusing her of dishonoring her family or modesty. Comments in local languages described her as a "bad influence" and used slurs to dehumanize her. This digital backlash reflects a broader, toxic narrative that women who are publicly visible or active online deserve punishment. It normalizes technology-facilitated gender-based violence and reinforces societal misogyny.

Cross-Platform Themes in Comments

DRF observed cross-platform themes that emerged from the surge of comments particularly with some users glorifying and celebrating the murder under posts related to the deceased.

Glorification and Celebration of the Murder

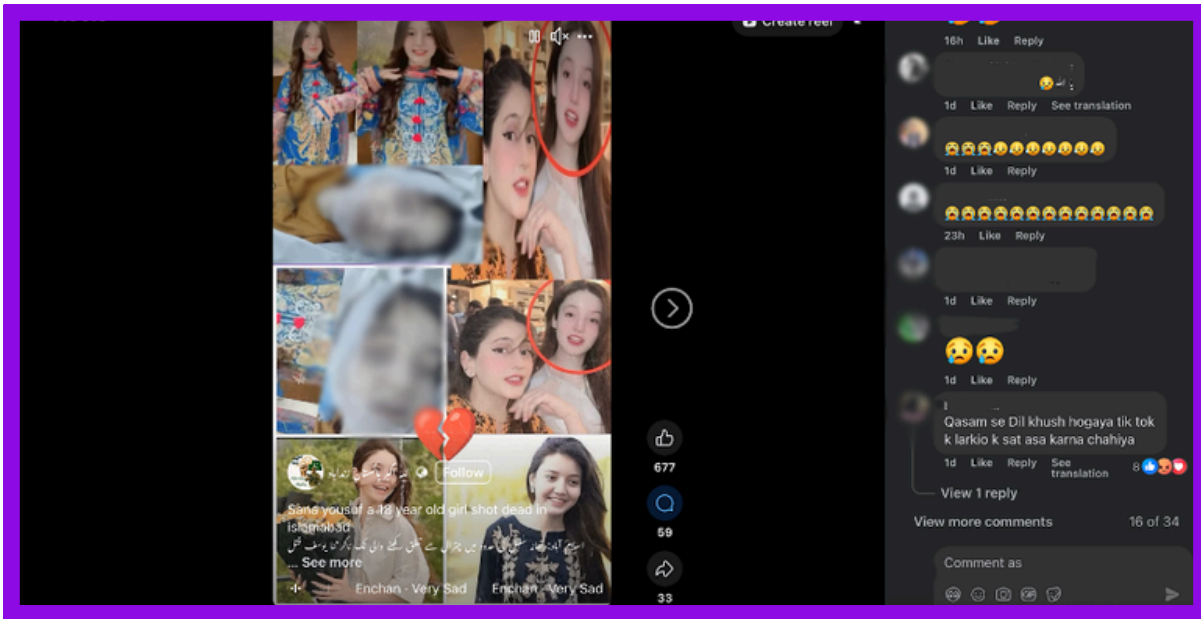
Hundreds of users went beyond blaming Sana and glorified her murder outright. DRF observed posts and comments that celebrated the killing, hailing the murderer, or suggesting this violence was praiseworthy. On Facebook and TikTok, especially, several comments contained phrases like “*Alhamdulillah*” and “*MashaAllah*” (“Praise be to God”), “*Jazakallah*” in direct reference to Sana’s death, effectively treating the murder as a blessed or justified outcome. These celebratory posts were sometimes accompanied by emoji reactions commonly used for positive events, for instance, the “👏” (clapping) emoji or the “🎉” (party popper) emoji (grotesquely repurposed to cheer a killing). In one screenshot shared widely, multiple Facebook users commented “*MashAllah*” with a ¹⁰⁰ (hundred points) emoji under a news post about Sana’s death. The use of such emojis and religious praises in this context indicates an extremist and misogynist mindset that views violence against women as virtuous or necessary, cloaking hatred in cultural or religious approval.



Victim-Blaming Narratives (Subtle and Overt)

A major trend in the content was victim-blaming, ranging from subtle insinuations to overt declarations that Sana “brought it on herself.” Subtle victim-blaming often took the form of backhanded condolences or moralizing statements. While some commenters expressed sorrow for Sana’s death in the same breath they advised that [“girls should be careful on social media”](#) or *“this is why modesty is important,”* [one TikTok user with 183k](#) followers suggested that “every girl who has access to mobile phone” and “brothers (who do) not keep an eye on their sisters” will eventually face such consequences, he also suggested that universities and colleges are immoral places to send our sisters to. This type of content stops short of explicitly praising the crime, but implicitly faults the victim for transgressing social norms. It reinforces the idea that women’s public visibility or personal choices (like using TikTok) are to blame for violence against them. Such coded blame was prevalent, reflecting deep-rooted societal attitudes. An analysis of posts on TikTok, Instagram, Facebook and X found many users referencing “honour” and “decency,” indirectly suggesting Yousuf’s fate, though tragic, was a lesson for other girls to stay within traditional Islamic boundaries. These dog-whistle criticisms resonate with conservative audiences while providing plausible deniability to the posters, making moderation challenging.

On the other hand, overt victim-blaming was alarmingly common as well. Many comments explicitly stated that Sana Yousuf *deserved* or *provoked* her own murder. For instance, one widely cited comment under a news post read: *“You reap what you sow”*. This stark statement directly suggests Sana’s killing was a consequence of her actions (i.e. her social media presence), blatantly shifting culpability onto the 17-year-old victim. Others echoed similar sentiments, openly agreeing that because she was a TikTok influencer, a fact repeatedly emphasized in media headlines, her murder was somehow warranted. The label “TikToker” was used pejoratively by detractors to dehumanize her, as if being a young woman on TikTok made her a valid target. This reflects a larger misogynistic trope of demonizing women for using digital platforms and blaming them when they face violence. By treating a girl’s online visibility as a justification for murder, these posts normalize femicide as an almost acceptable response to female agency, which is deeply troubling. The breadth of victim-blaming, from insinuations to outright endorsement, shows how ingrained these attitudes are in online discourse.



Sexualizing the Victim

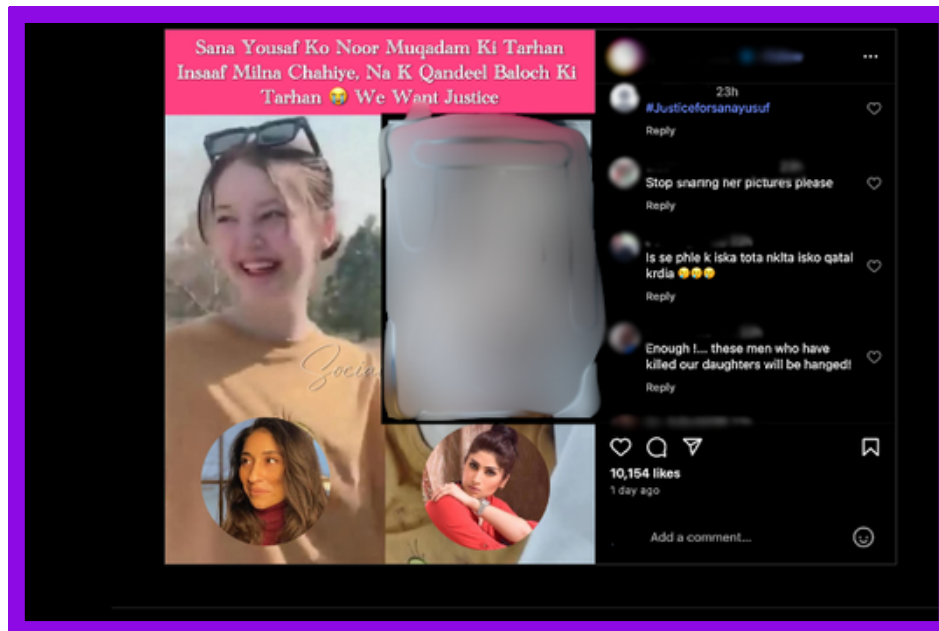
Following Sana Yousuf's murder, social media was flooded with deeply disturbing and sexualized commentary about the 17-year-old. Despite her being a minor, many users reduced her to derogatory labels such as "taxi," "nachne wali," and "kanjari." Some even falsely speculated about an impending leak of explicit content, with one user writing, "Is se pehle ke iska tota nikalta, isko qatal kar diya." ("Before her private video came out, she was killed.")

This wave of gendered disinformation not only dehumanized Sana but also further incited degrading and misogynistic language. Many users justified her murder by questioning her moral character, citing her friendships with boys or assumptions about her relationships. [One comment](#) stated:

ناجائز کاموں کا عموماً یہی انجام ہوتا ہے۔ قرآن کریم کے واضح احکامات کو جوتے کی نوک پر رکھنے والوں کا انجام بد ہوتا ہی ہوتا ہے۔ فرق صرف اتنا ہے کہ کسی کا جلدی، کسی کا دیر سے، کسی کا دنیا میں اور کسی کا آخرت میں۔ مرحومہ کو اللہ معاف فرمائے، مسلمان تھی، دعا گو ہیں۔ لیکن ایسی عورتوں کی حمایت نہ کریں، یہ کھلم کھلا قرآن و حدیث کی منکر ہوتی ہیں۔

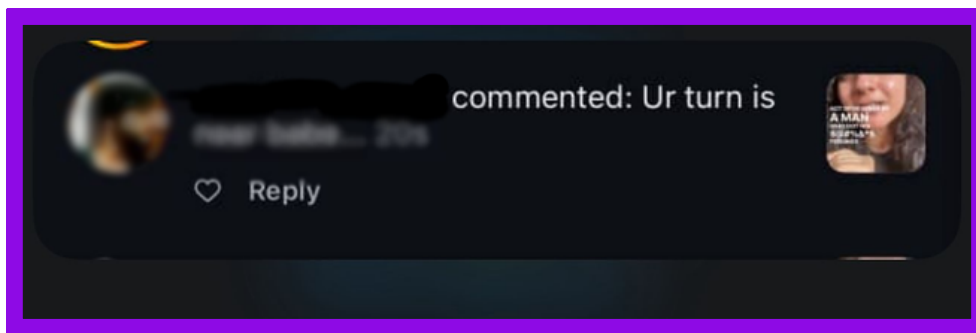
("This is usually the outcome of immoral acts. Those who disregard the clear commands of the Quran face bad ends, some sooner, some later, some in this life, some in the next. May Allah forgive the deceased, she was a Muslim, we pray for her, but do not support such women; they openly reject the teachings of the Quran and Hadith.")

Such rhetoric not only perpetuates the sexual objectification of women and girls but also dangerously links a woman's online presence or perceived morality to justifications for violence. This form of digital vigilantism contributes to a broader culture of victim-blaming, where abuse is normalized and accountability is shifted away from the perpetrator.

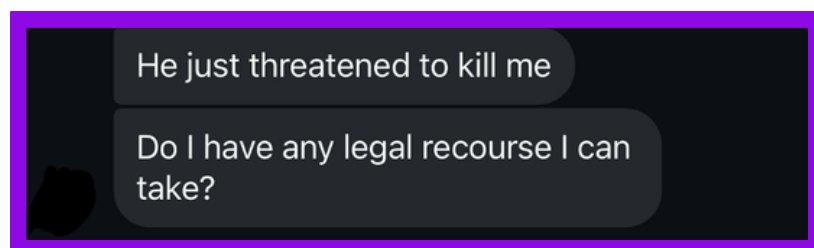
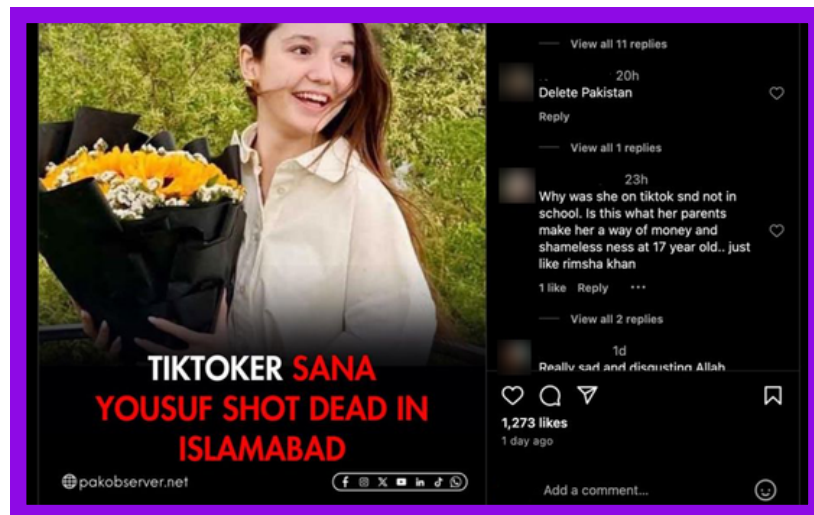


Incitement to Violence and Threats Toward Others

Beyond commenting on Sana's case, some posts escalated into incitement of further violence, including rape and death threats directed at other women influencers. DRF's analysis found that in the comment threads and discussion groups across platforms following the murder, other female content creators were explicitly named and threatened. In a few chilling instances, users tagged or referred to popular Pakistani women on TikTok/Instagram, suggesting they would meet the same fate as Sana Yousuf. For example, one commenter warned that a certain outspoken female TikToker "*will be next*," effectively turning Sana's murder into a threat held over the heads of other women. Such copycat threats create a climate of terror for women online, the message is that any woman who gains visibility could be killed for it.



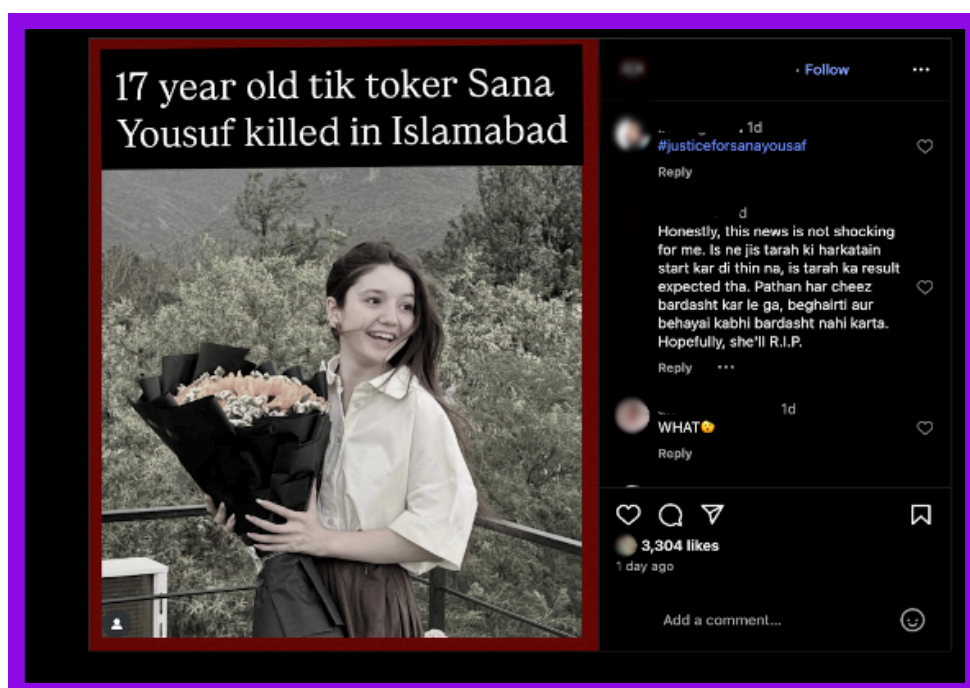
Notably, other women influencers and activists reacted by speaking about how unsafe this made them feel. The account of one female rights activist highlighted that female social media influencers [routinely face](#) threats, blackmail, and intimidation, and Sana's case tragically underscores their vulnerability.



Ethnic Hate Speech and Harmful Comparisons in Online Commentary

Another troubling pattern that emerged was of ethnic hate speech surfacing in online comments following Sana Yousaf's murder. Many users focused on her Chitrali background, suggesting that women from conservative areas like Chitral "shouldn't be on TikTok" and implying that her murder was somehow justified. One comment read, 'sy area se tw bnai thi chitral ke دشمنی belong krne wali asy tiktok pe viral ho gi tw .krti ha پرده khawateen

Additionally, several users racialized the incident by referring to Sana as Pakhtun and weaponizing ethnic stereotypes. One particularly disturbing comment stated, "Pathan har cheez bardasht kar le ga, beghairti aur behayai kabhi bardasht nahi karta. Hopefully, she'll R.I.P." Such remarks not only reinforce misogyny but also use ethnic identity to rationalize violence against women, perpetuating harmful narratives about entire communities.

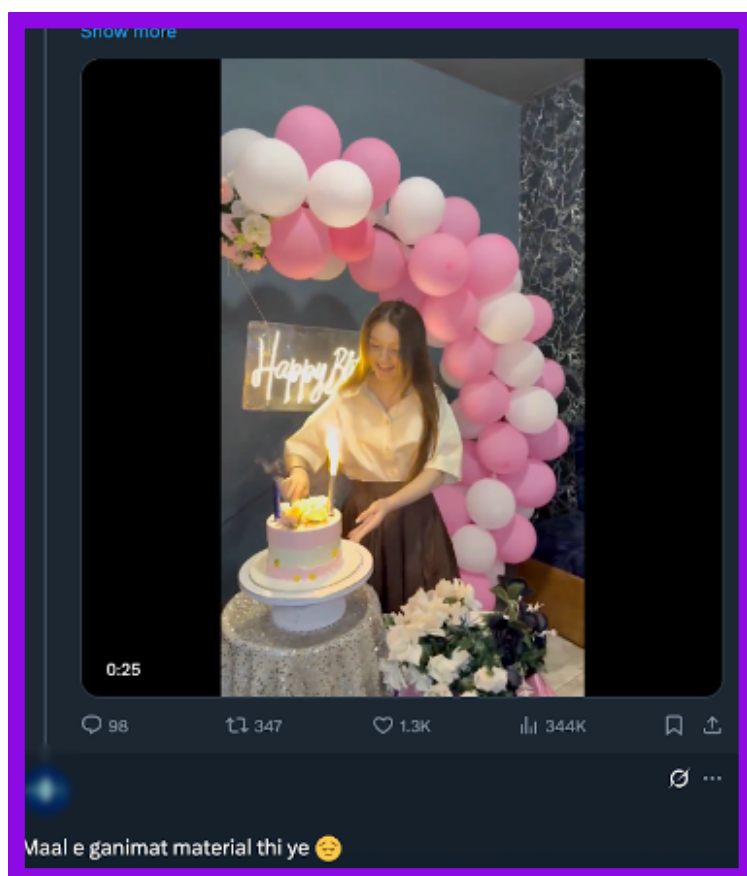


Another trend that emerged was when some users downplayed the murder by comparing it to violence in conflict zones like Waziristan and Palestine, suggesting that coverage of Sana's death was disproportionate. These comparisons are deeply problematic; they erase the specific context of gender-based violence and diminish the severity of a targeted killing by framing it as less newsworthy or justified because other global crises exist.

Cross-Border Hate Speech

Sana Yousaf's murder also sparked disturbing cross-border commentary, with users from India contributing to the online abuse. Amid ongoing India-Pakistan tensions, many Indian commenters used the tragedy to promote hateful narratives, referring to Pakistan as "Porkistan" and mocking Islamic beliefs with comments like, "Karma returns 🌀 A new hoor added in Jannat."

Some users went further, dehumanizing Sana by calling her "maal e ghanimat", a term historically used to describe the spoils of war, while others spread misinformation, suggesting, "Maybe because she knew India won the war 😂."

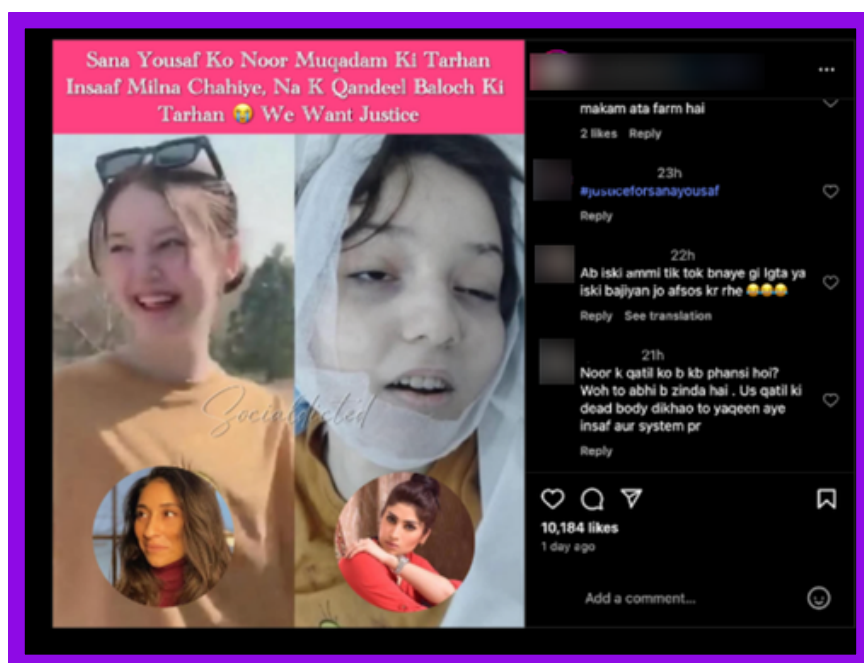


These remarks mirror the misogynistic and violent rhetoric seen from within Pakistan and show how geopolitical animosity can fuel gendered disinformation and online hate. These cross-border interactions turned the brutal murder of a teenage girl into mockery, hate speech, and digital violence. This particular instance highlights the urgent need for platforms to better moderate transnational abuse that intersects gender, religion, and nationalism, which has been highlighted in DRF's [previous report](#) on India-Pakistan escalations and TFGBV.

Parental Shaming

In the wake of Sana Yousaf's murder, much of the online discourse turned toward blaming her parents, especially her mother, for the tragedy. Many commenters suggested that stricter parental control could have prevented her death, shifting responsibility away from the perpetrator. Some even mocked her grieving family, with one user writing, "Ab iski ammi TikTok banaye gi lagta hai, ya uski bajiyan jo afsos kar rahi hain 😂😂" ("Looks like now her mother or sisters will be making TikToks").

Others went further, condemning the family's supposed failure to provide "proper upbringing," with one comment reading: "Bahut bura hua lekin in influencers ko promote mat karein! Yeh deen nahi phaila rahe thay balkay beghairti. Kal inko dekh ke hamare bachay bhi aisa karne lagte. Walidain ka kaam banta hai ke apne bachon ki achi tarbiyat karein, na ke social media se paisa kamayein. Phir logon ki nazar lagti hai, aur anjaam aap ke samne hai." ("What happened is terrible, but we shouldn't promote influencers. They weren't spreading religion, but shamelessness. Our children might follow them. Parents should raise their children properly, not earn off their social media. Then people take notice and this is the result.")



These narratives are deeply problematic since they place the burden of violence on the victim and her family, absolving the actual perpetrator. By framing women's public presence or self-expression as a provocation, such commentary reinforces a culture where gender-based violence is normalized and excused.

Platform-Specific Analysis

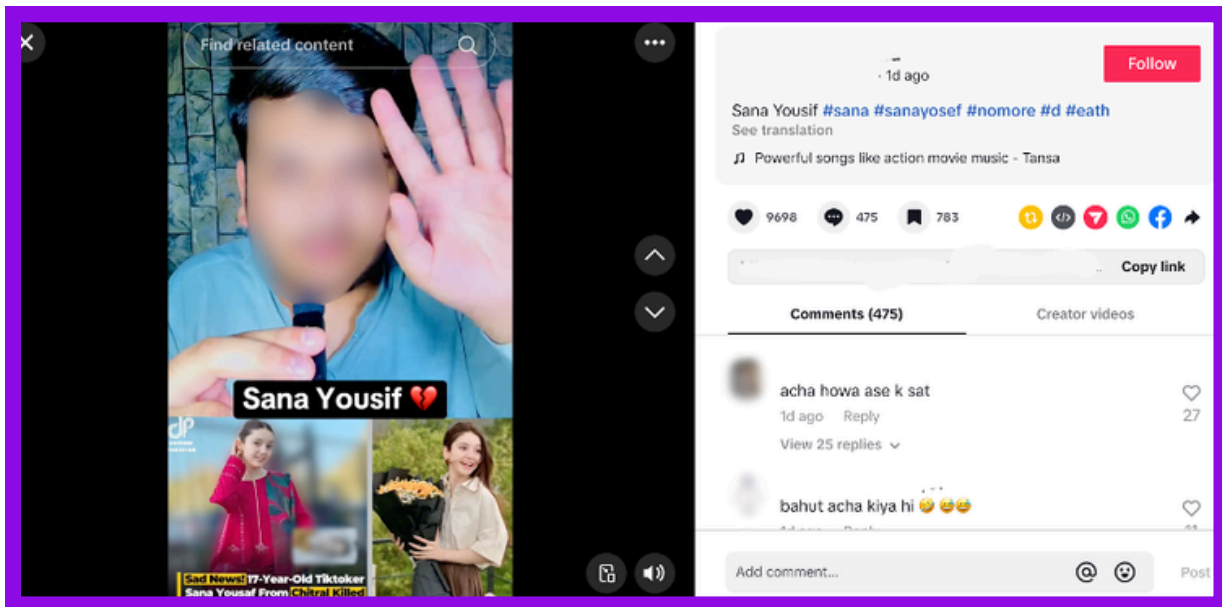
DRF was able to analyze content from four major platforms during the time, which were TikTok, X, Facebook and Instagram. DRF's collection of 180 videos, pictures and comments reflects the opinions of users after hearing news about Sana's murder and the perpetrator being caught.

TikTok

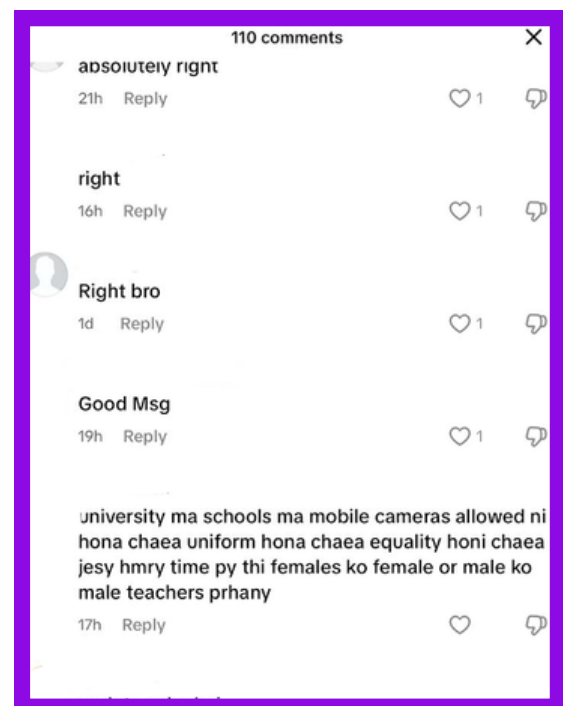
Multiple self-proclaimed journalists on TikTok who, along with sharing news updates about the killing, were passing judgments on Sana's character and proclaiming her death to be a direct consequence of her actions (i.e., her content and online presence). [Such videos](#), which over the course of just a day have received thousands of views, likes, and comments, which pose as news sharing accounts, victim blame Sana as well as other girls who have over the course of this year (and as recent as last week) been murdered in connection with them being prominent personalities on social media.



Equally disturbing is the consistent use of this incident in particular, as an example of what happens when women are allowed to be on TikTok (and other social media platforms). DRF's analysis on TikTok found countless videos that begin by detailing facts and updates around Sana's case but soon turn into misogynist rants (usually by men) around how girls and their honors are 'ruined' because they are allowed to maintain public social media accounts and have access to the internet. [One video](#) goes as far as to caution men into keeping the women in their households 'under control' so that they don't fare the same fate as Sana.



While these videos in and of themselves are sexist and a form of gendered hate speech, they also inadvertently end up creating a space for fellow sexist, misogynistic men to spew hate, blame victims of gender-based violence both online and offline, and use this as an opportunity to push patriarchal agendas.



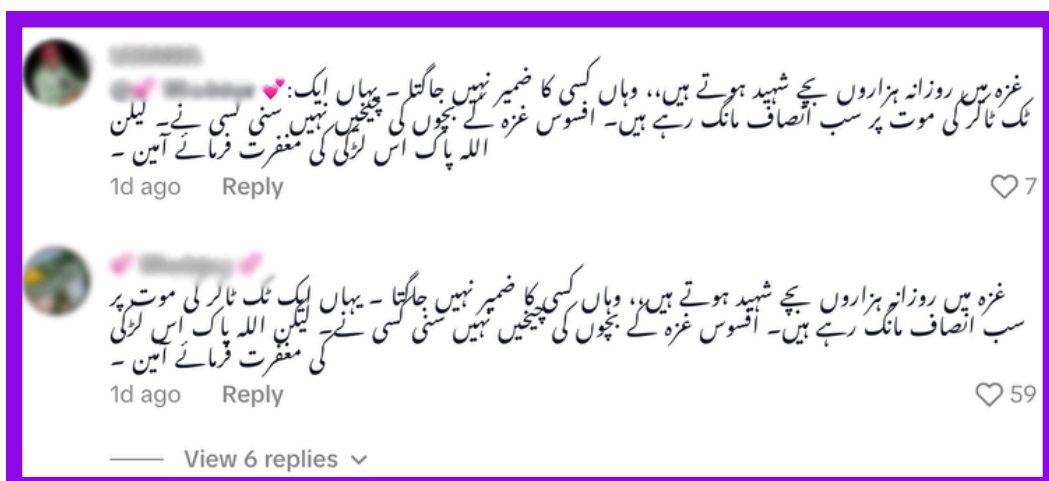
Additionally, many comments under videos of support for Sana and holding her perpetrator accountable invoked religion and implied that violence against women in this context (and many others) is a direct consequence of women not covering themselves and spreading behayai.



Given the fact that Sana's killing is a reality that women in Pakistani society are made to face on a daily basis in one way or another, many took to TikTok to express their fear, concern and, simultaneously, sexist remarks including comments that mentioned the only way for women to protect themselves against something like this was by not posting pictures/videos and/or staying off social media. These comments are similar to the ones made by Hasan Iqbal Chishti in his song "Othy Dance Kardi Payi Ae," with derogatory remarks around women dancing in educational institutions and undermining girls' education. Chishti was later [arrested](#) for his remarks, however, his comments and the hateful rhetoric he spread online remains the same albeit through other social media users.

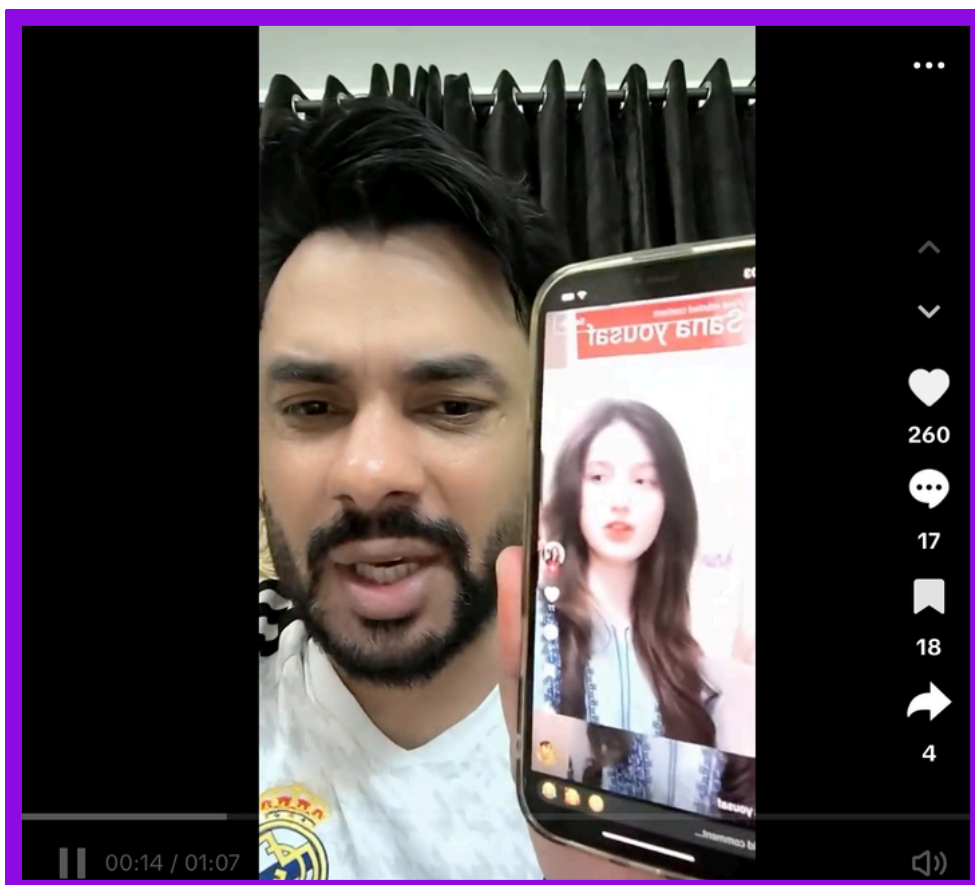


Another trend found under emerging news reports was how many users compared the coverage Sana's death was getting to other unrelated national and international events. Multiple TikTok users were seen saying how 'unfair' it is for a TikToker's death to gain such media outrage and coverage while the ongoing genocide of Palestinians in Gaza and murder of Pakistanis in Balochistan and Khyber Pakhtunkhwa was not given the same importance and attention. Interestingly, users were copy-pasting the same messages and posting them across different videos on TikTok about Sana's death.



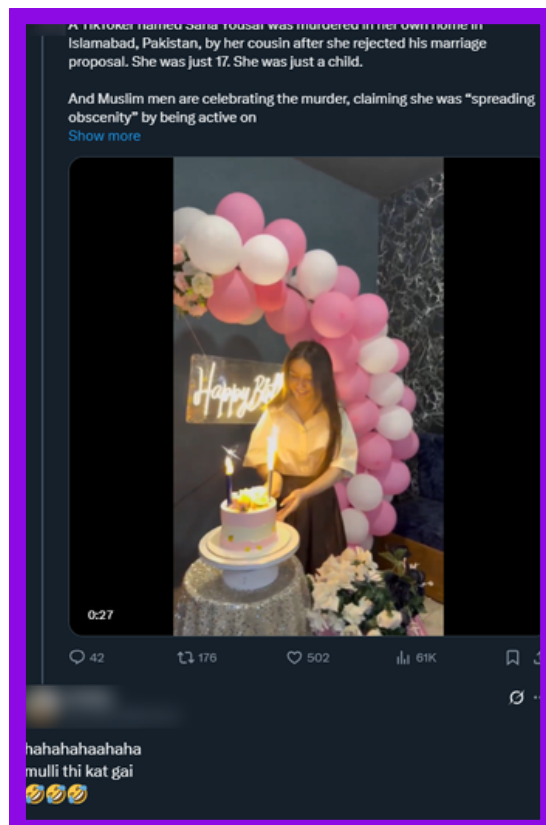
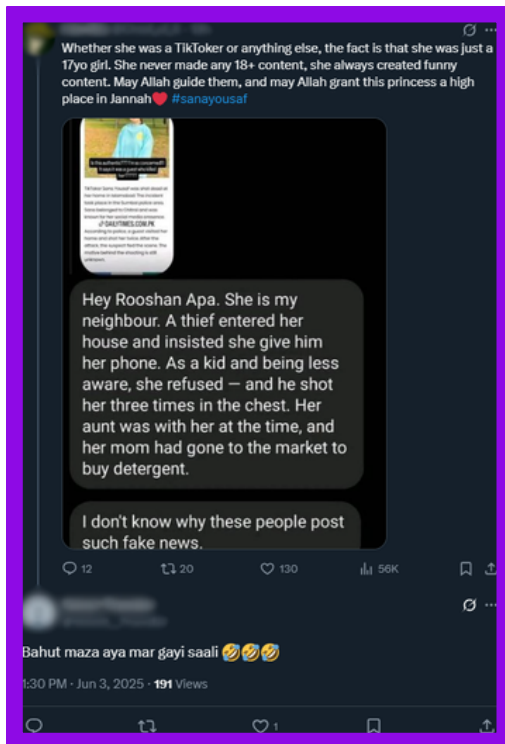
While the unfortunate deaths of Pakistanis across the country and the millions of Palestinians' lives lost are concerning and are all stories that deserve to be covered by media outlets and the public, the implication that the murder of a young girl who was a public personality on TikTok is any less important is deeply disturbing. Women and girls are exposed to different forms of violence every day, with numerous cases that often than not never make the headlines, let alone garner social media attention. Placing the death of human beings on a pedestal, be it those who died under colonial oppression or those silenced and killed under the patriarchy, wrongfully pits people against each other and causes a societal divide that has the potential to unite against a common enemy.

In line with the sentiments found on Facebook, Instagram, and X (which are explained later in this report), TikTok had its fair share of videos and comments that blamed Sana and the content she made for her own death. One [prominent TikTok user](#), for example, shared the news of Sana's killing while proclaiming that her TikTok content, which is a cause of her family's badmashi, was the reason why she was killed out of honor. Not only is this statement untrue, but it further pushes a patriarchal rhetoric that is all too common in Pakistan's society; namely that a family, specifically a man's, honour is tied to the actions of the women.

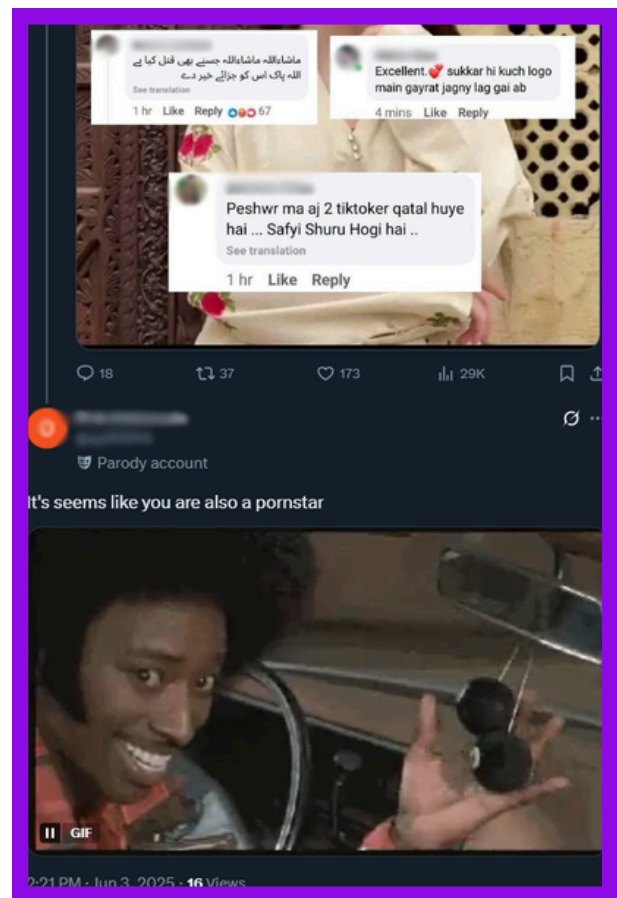


X (formerly Twitter)

DRF observed multiple X (formerly Twitter) users responding to news of Sana's cold-blooded murder with [celebrations](#), [jubilation](#) and [congratulatory messaging](#) on one hand, and casting aspersions on the nature of her 'relationship' with the killer on the other.



Several instances of users victim-blaming Sana by using derogatory terms for her such as “[pornstar](#)”, “[OnlyFans](#)”, and “[dancer](#)”, were recorded.

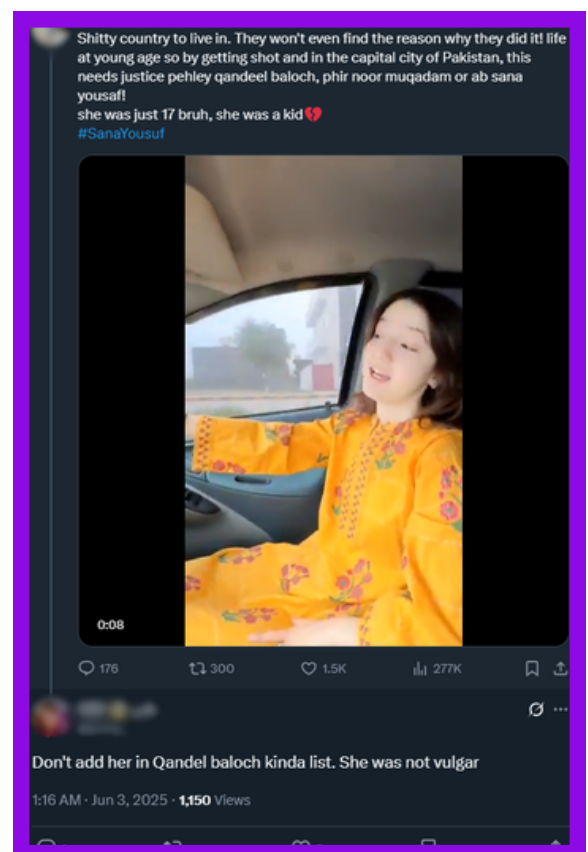
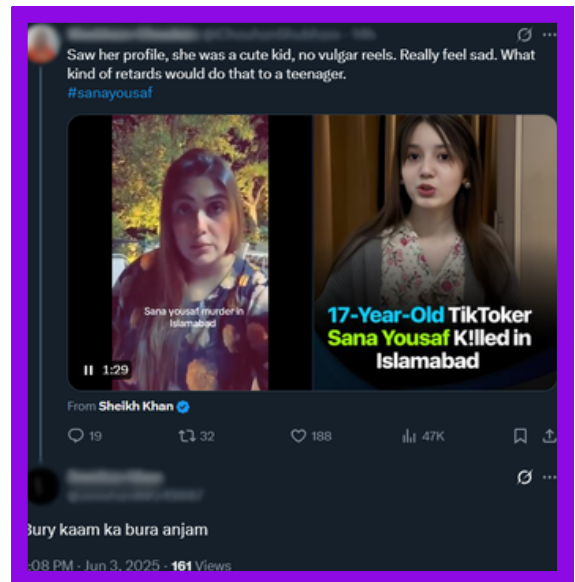


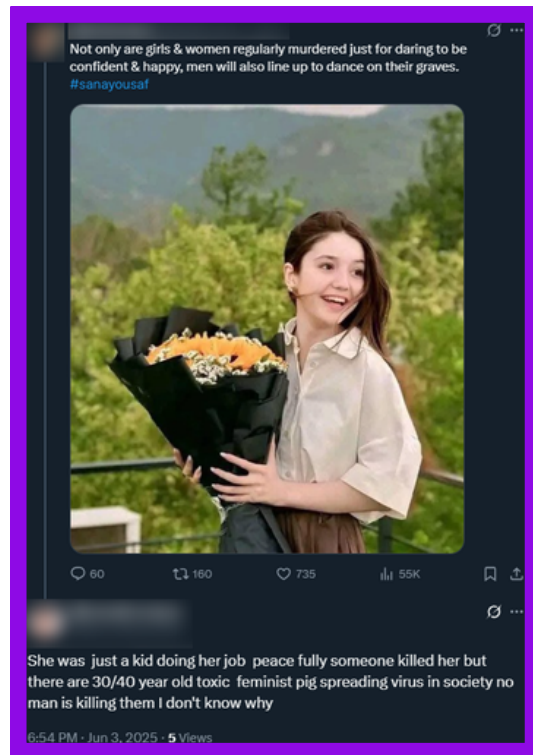
Another user [termed](#) the murder “bury kaam ka bura anjam” (bad punishment for a bad deed), implying that Sana’s content was immoral just by virtue of its being, and that this ‘immorality’ legitimised her murder.

This nature of content not only slandered Sana with baseless allegations, but also revealed the disturbing mindset of some Pakistani social media users wherein ending the life of a woman who is creating content that is unacceptable to their personal moralities is justified and considered “the needful”.

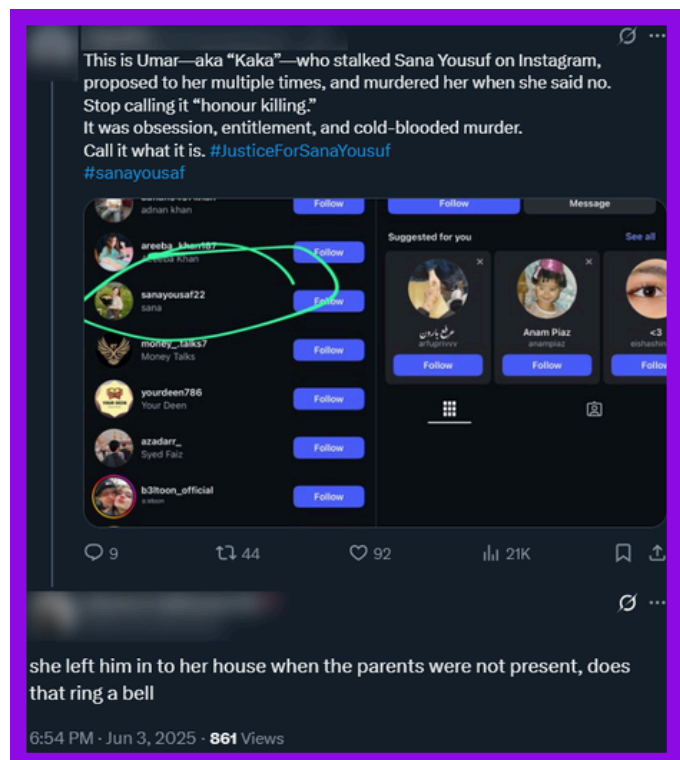
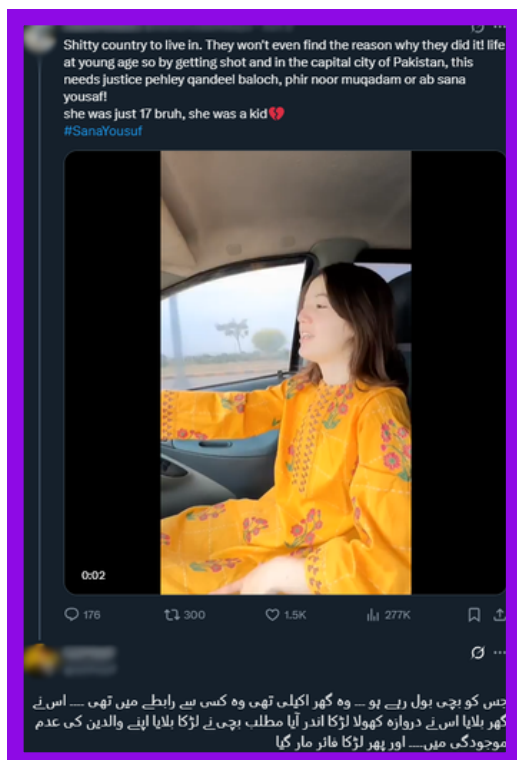
This is why, even some of the accounts who ostensibly opposed the killing, appeared to do so at the expense of another woman who was killed in Pakistan in 2016–Qandeel Baloch—whose name started cropping up in conversations again because of the similarities of her case to Sana’s heinous murder. Users attempted to make a distinction between the “[not vulgar](#)” Sana—whom they opposed the murder of—and the ‘vulgar’ Qandeel—whom it was ‘justified’ to murder, solely based on which woman’s content was deemed ‘acceptable’ to users.

The false dichotomy of the ‘good’ vs. ‘bad’ woman made a return in Sana’s case, with users contemplating why a “[good](#)” woman like her got killed, instead of the “[toxic feminist pig\[s\]](#)” who should have actually been targets of violence.

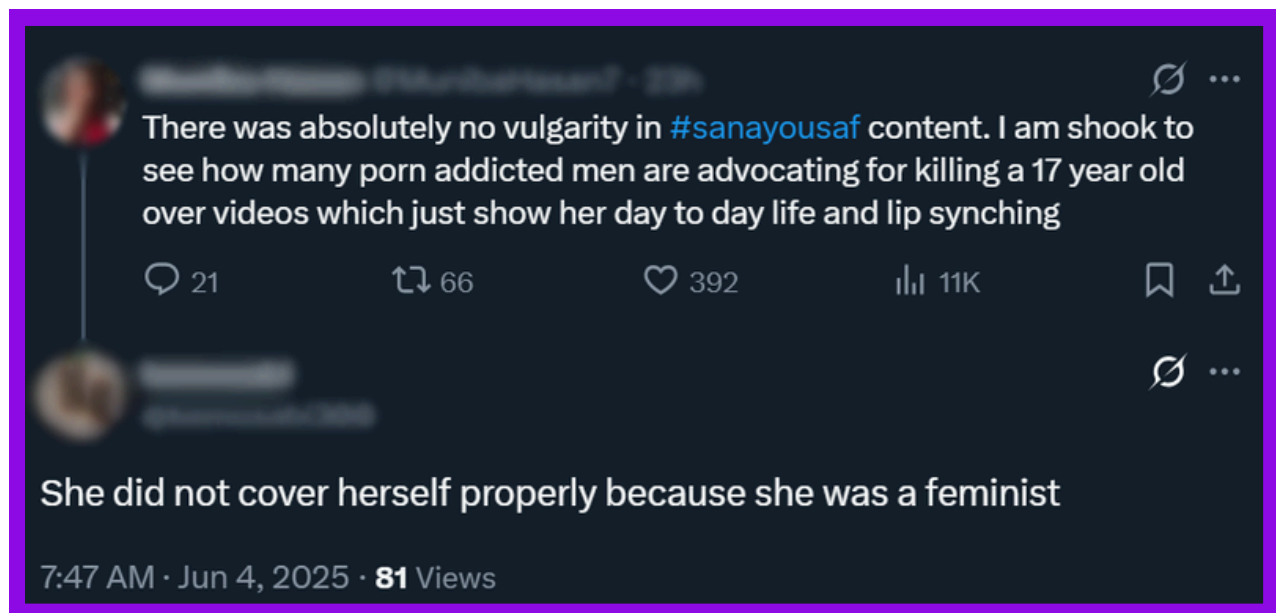




Yet, again, in some replies it was noticed that Sana became the target of victim-blaming despite the more ‘acceptable’ nature of her content, when users started questioning the nature of her relationship to her killer. Some [claimed](#) she invited him into the house, while others [claimed](#) she was already in contact with him, shifting the blame for her own murder onto her.



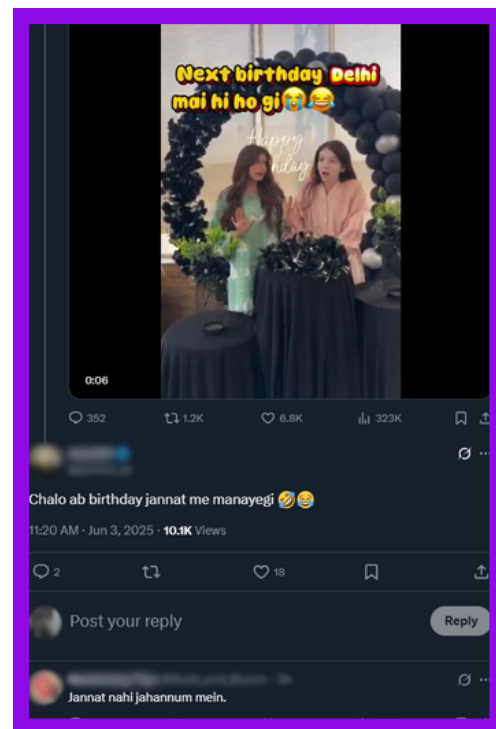
Others still found ways to police Sana's clothing, [claiming](#) that she "did not cover herself properly because she was a feminist".



The fact that as far as Pakistani social media users are concerned, the jury remains out as to whether the cold-blooded murder of a minor was justified according to her content, her dressing, and the people she was or was not in contact with, goes to show just how dangerous it is to be a woman creating content online in Pakistan, and how there is no acceptable mold a woman can take which will guarantee her safety, or the condemnation of violence committed against her.

A separate trend noticed was a higher than usual influx of Indian users commenting on the killing. This seems to be linked to a video Sana had posted during the early-May Indo-Pak escalations, which depicted her 'joking' with her friends about how she would celebrate her next birthday in Delhi (in reference to a meme trend about Pakistan taking over India). Indian users [posted](#) that video, making [statements](#) such as "only a de[a]d Pakistani is a useful Pakistani for India".

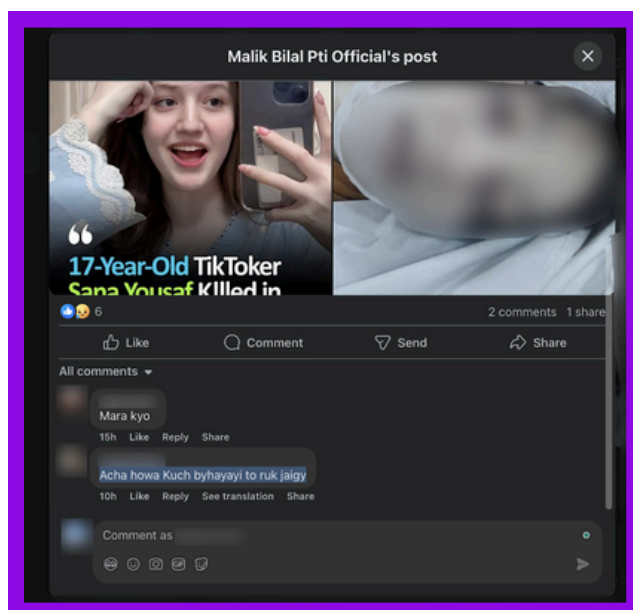


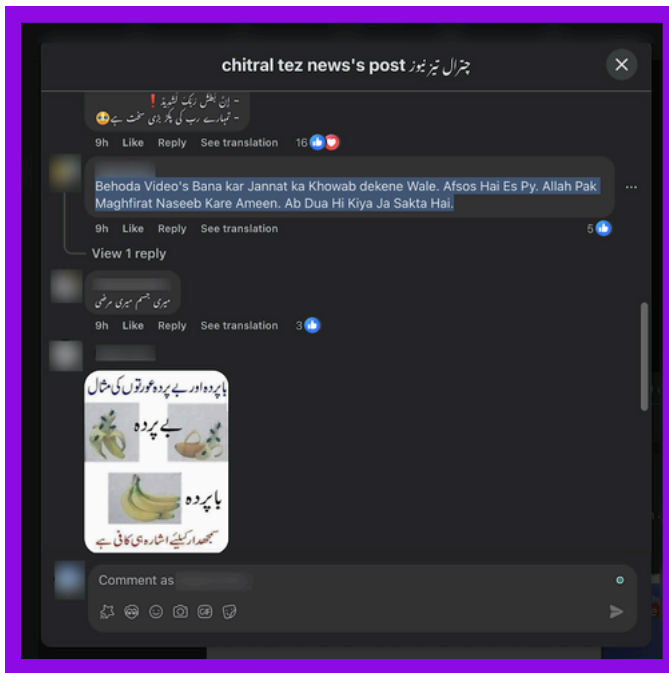


The influx of Indian user [commentary](#) on a Pakistani case not only demonstrates the global shockwaves that this case caused, but in this particular thread also demonstrates how women's bodies are often made the vehicle of cross-border animosity.

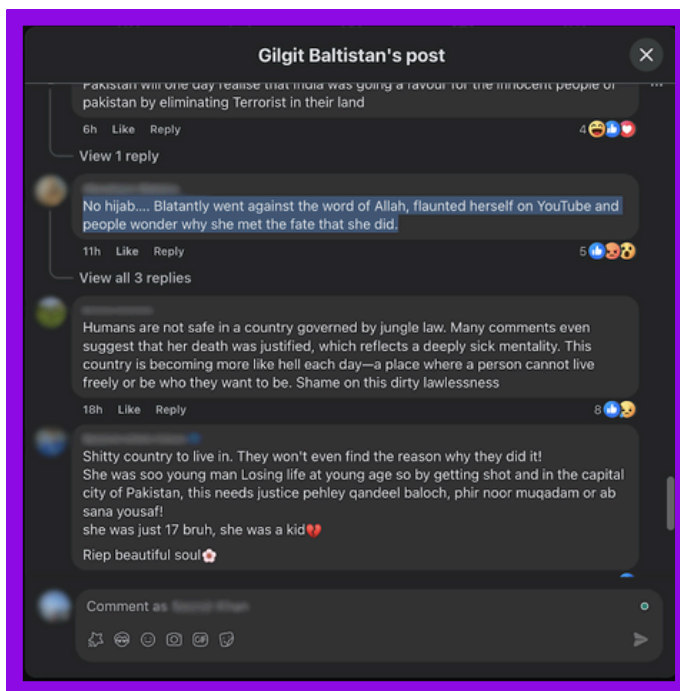
Meta (Facebook and Instagram)

Comments by users on both of Meta's platforms, Facebook and Instagram, were rife with disturbing hate-filled comments openly celebrating and justifying Sana's murder, linking her career as a TikToker to ideas of "[beghairati](#)" (shamelessness) and "[byhayayi](#)" (immodesty).

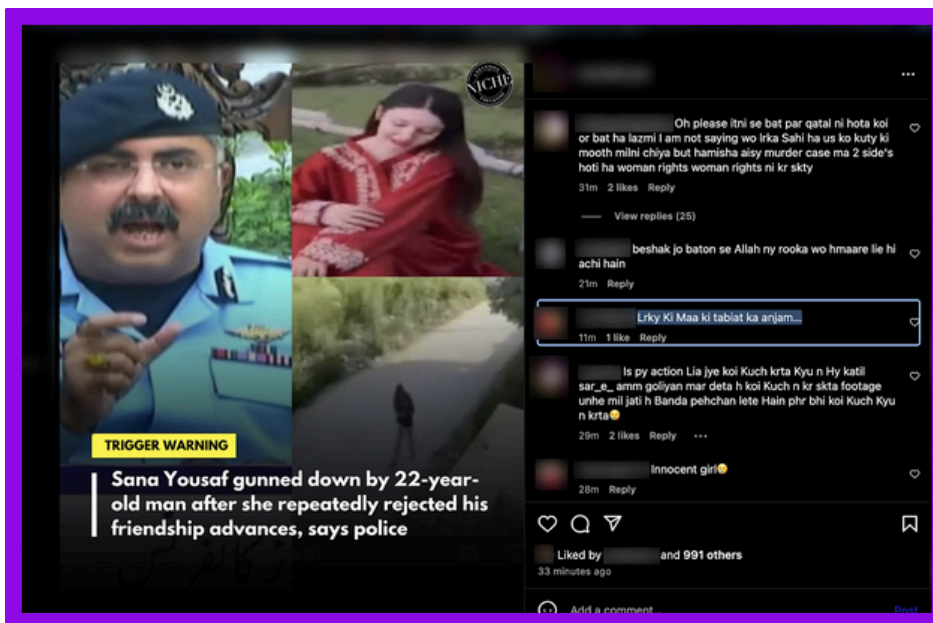




Others still went further, [commenting](#) on Sana’s dressing which lacked a hijab, and her going “against the word of Allah” by “flaunting herself on YouTube”, stating that it was no wonder that she “met the fate she did”.

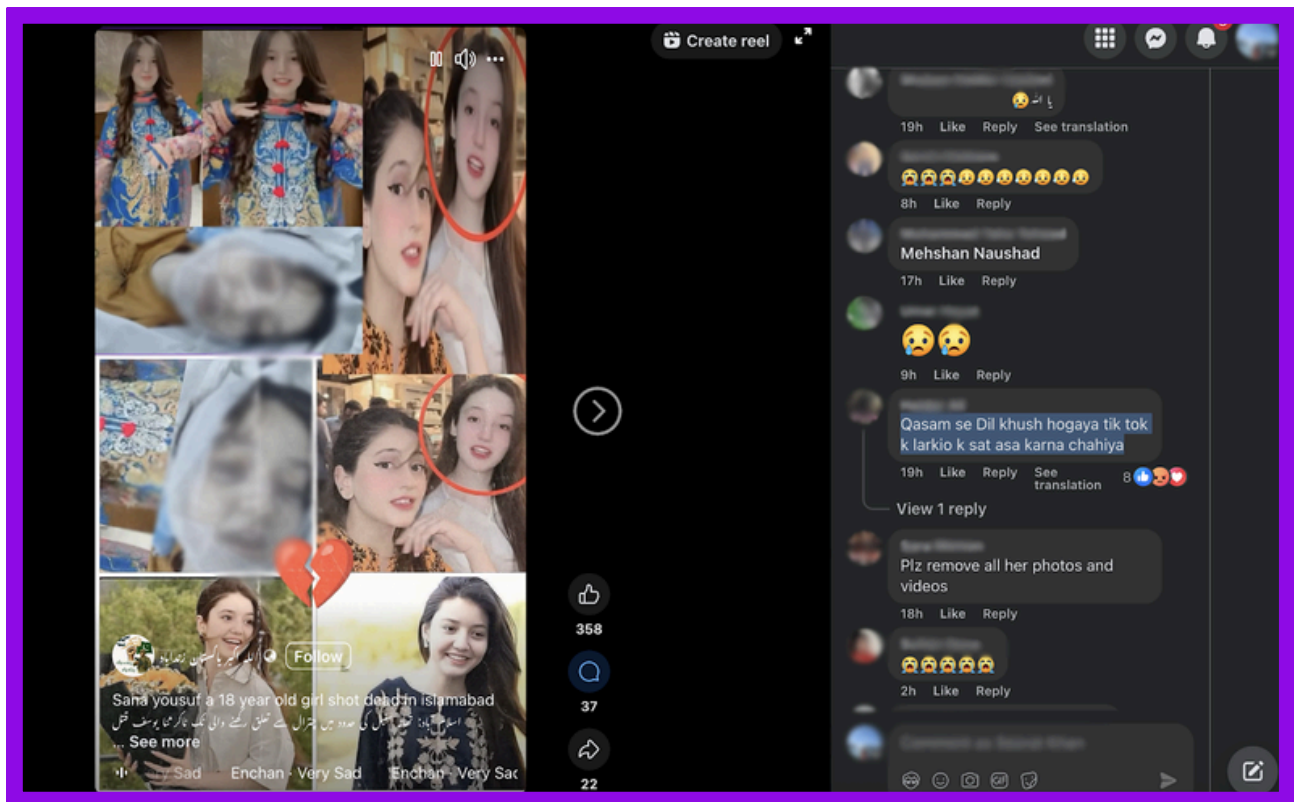


Some Facebook and Instagram users attempted to [blame Sana’s parents](#) for the way they raised her and gave her “the [freedom](#) to do whatever she wants”, which led to her becoming a TikToker and going viral.



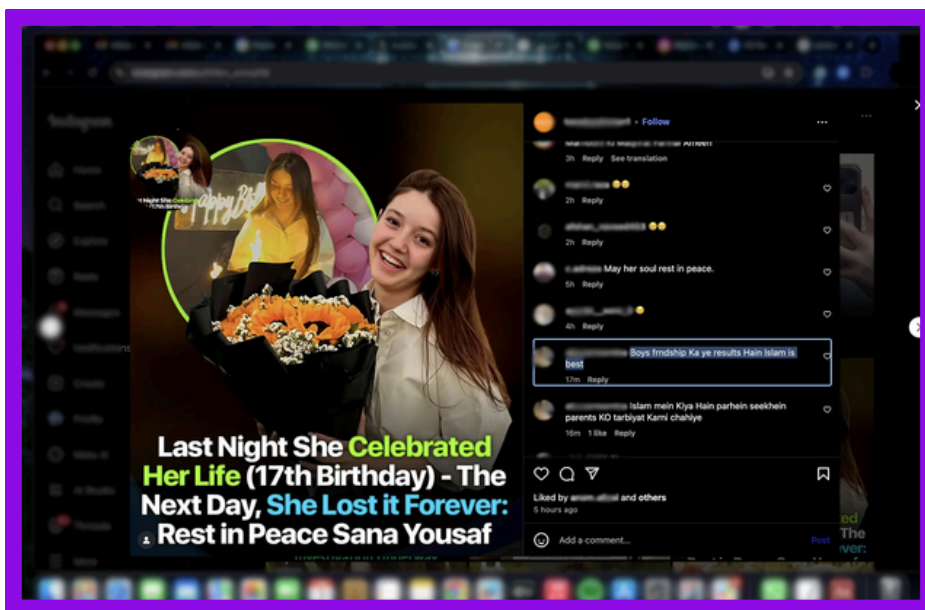
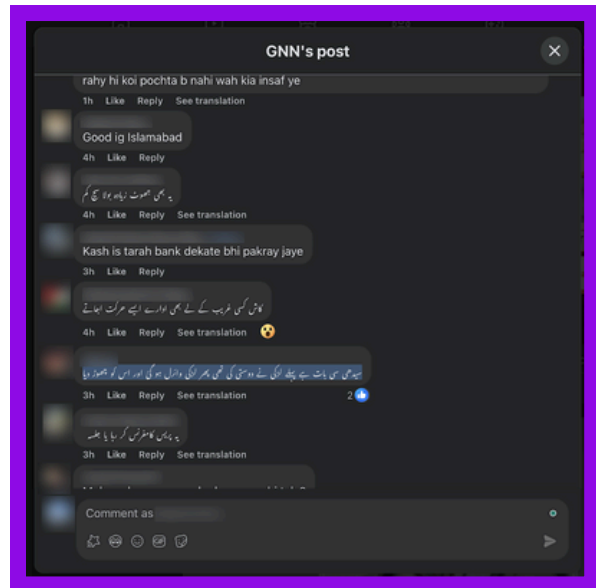
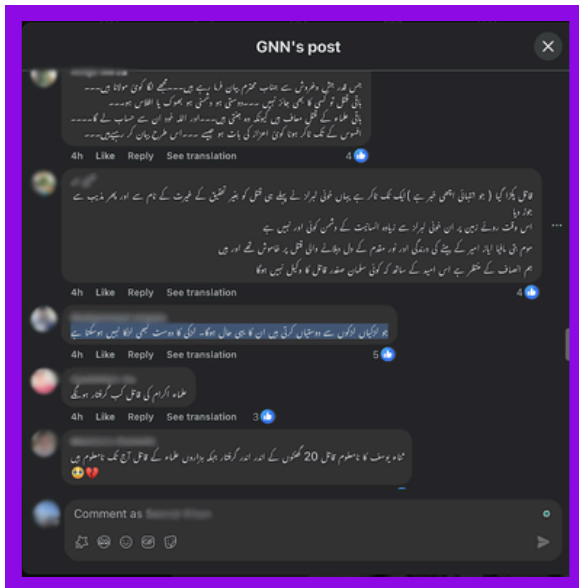
The correlation of the mere fact of Sana's career as a TikTokker with her being a shameless and immodest or immoral figure who deserved to die is reflective of a deeply disturbing mindset prevalent within Pakistani social media users (and Pakistani society in general) wherein women appearing in public, be it online or offline, is considered transgressive to social norms.

An unsettling number of comments displayed outright [joy](#) and [celebration](#) in reaction to the murder news, claiming that this is the treatment "[TikTok girls](#)" deserve.



Jubilant reactions to a murder not only show the cold-blooded nature of a society where such killers are bred, but the framing of Sana and others like her as “TikTok girls” dehumanises and demonises content creators who are living breathing human beings with rights, thus delegitimising sympathetic reactions to TFGBV.

Users on both Facebook and Instagram also brought into question Sana’s relationship with the killer, making baseless claims about [her friendship](#) with him, and the [consequences](#) of women [befriending](#) men.

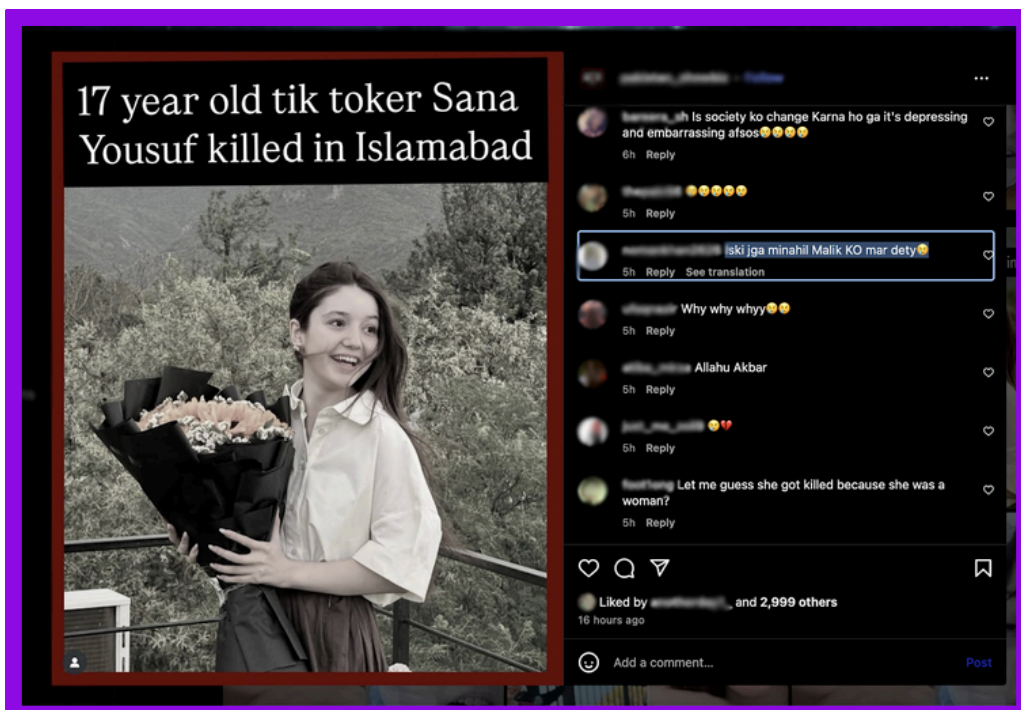
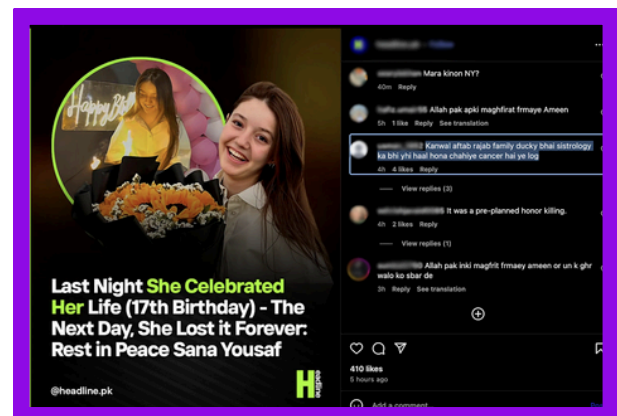
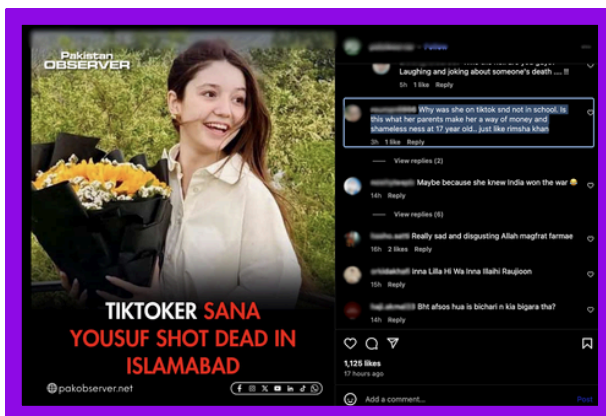


Even the condemnations for the murder had an undertone of moral judgement, with comments [stating](#) the murder was reprehensible, but people shouldn't promote social media influencers, as they are promoting "beghairati" instead of "deen" (religion).



The abandonment of nuance and painting the entirety of social media influencers, and content creation as a whole, as a morally corrupt practice only serves to make women more unsafe online, and to shift the blame away from the perpetrators to the victims themselves.

A particularly dangerous trend was seen on Instagram wherein users were using the news of Sana's murder to [incite violence](#) against [other TikTokers](#) and [content creators](#) whose views they ostensibly disagreed with.



These incitements to violence in comments underneath popular Instagram posts remain unremoved or censored in any way, which raises the question of platforms failing to tackle violent content that can translate into real-world harm and even loss of life if left unaddressed.

Where Platforms Come In

Viral Spread and Lingering Content: Platform Response Failures

A striking aspect of the digital aftermath was how quickly the offensive content spread and how slowly platforms reacted. Within hours of the murder news, hateful posts and comments began accumulating thousands of views and interactions. Some particularly egregious content, for example hate-filled video reels blaming Sana or [celebrating her death, went viral, amassing views in the millions within 48 hours](#), according to DRF's review. Alarmingly, even as users reported such content, much of it remained live at least two days after the incident, continuing to rack up views and shares. For instance, a [TikTok reel](#) that subtly cheered the murder by overlaying news footage with approving captions was still circulating freely two days later, with over 1 million plays and a flurry of supportive comments.

Platform failure

All four platforms, TikTok, Instagram, Facebook, and X, appeared to struggle with the surge of violent content against Sana Yousaf. TikTok's algorithm, known for boosting popular clips, likely accelerated the spread of some hate content before moderators could intervene; that is why the platform saw amplified hate against the influencer.

In Instagram and Facebook's (both under Meta) case, users saw that certain content was removed after public outcry, with some particularly high-profile abusive comments being deleted. Still, the reactive approach meant harmful content had already reached a wide audience, with many individuals taking screenshots of harmful content on platforms.

In the case of X (Twitter), which in recent times has reduced its moderation workforce, there were exacerbated amounts of unfiltered hate until [users themselves shamed the offenders](#) or appealed to authorities to intervene. An example of this is when one user posted about another social media user who was inciting hate towards Sana. The slow takedown response allowed a 'pile-on' effect, with the more

people who saw hateful or celebratory posts unpunished, the more emboldened others became to chime in with the same.



These instances point towards a systemic moderation failure where platforms lacked an effective crisis protocol to rapidly detect and curb gendered disinformation and violence-inciting content in the immediate aftermath of a gender-based killing. The persistence and virality of these posts not only deepened the victim's family's trauma and spread fear among women but also risked amplifying the dangerous mindset that such violence is acceptable. These findings highlight an urgent need for better platform policies and enforcement.

Recommendations

In light of the issues identified in this case study, UN Guiding Principles on Business and Human Rights (UNGPs), decisions made by Meta's independent Oversight Board, and digital rights civil society groups, the following are actionable steps platforms must adopt to curb TFGBV, gendered disinformation, and gendered hate speech in Pakistani and similar contexts:

TikTok

1. **Improve Local Moderation:** Expand investment in AI and human moderators trained in Urdu and regional languages (Punjabi, Pashto, Sindhi) to better detect misogynistic language, coded hate, and religious justifications used to glorify violence.
2. **Crisis Response Protocol:** Implement a real-time escalation process during high-profile incidents. This should include temporary prioritization of reports under relevant hashtags and increased manual oversight to remove hateful or violent content before it gains traction.
3. **Gender-Sensitive AI:** Enhance AI systems to recognize gendered hate using emojis (e.g., party, clapping, or prayer emojis in response to a woman's murder) and phrases often used in coded incitement or victim-blaming narratives.
4. **Policy Update:** Amend community guidelines to clearly prohibit content that praises, celebrates, or justifies gender-based violence. Enforce immediate takedowns and apply account suspensions or bans for repeat offenders.

Instagram

1. **Trending Topic Moderation:** Proactively moderate content under trending tags related to femicide or abuse cases. Use keyword triggers to surface harmful reels, comments, and posts for faster review.
2. **Fast Takedown Process:** Ensure that all reported rape threats, murder threats, and celebration of violence are reviewed and removed within a 24-hour window, with priority response protocols for high-risk cases.
3. **Context-Aware Filters:** Update automatic filters and AI moderation tools to recognize Roman Urdu, slang, and cultural references that may slip through existing filters but contribute to victim-blaming or

hate

4. **User Safety Tools:** Broaden availability and visibility of safety features like “Hidden Words” and bulk comment controls. Introduce a “Crisis Mode” allowing users to restrict interactions during harassment spikes.

Facebook

1. **Enhanced Human Review:** Trigger additional manual review teams to moderate viral content involving women’s murders or gender-based violence. News post comment sections should receive extra scrutiny.
2. **Policy Gap Fix:** Close loopholes in content policies by banning glorification, mockery, or justification of gender-based violence even if the content targets no specific named individual.
3. **Cultural AI Updates:** Improve AI systems to understand the cultural and linguistic context of abuse, including celebratory or sarcastic religious phrases and emoji usage that implicitly condone violence.
4. **Women’s Support Features:** Expand access to resources in Urdu and other regional languages, including safety toolkits and emergency response features. Partner with local groups for survivor-centered tools.

Instagram and Facebook Policy must implement recommendations by oversight board in multiple cases relating to TFGBV

- Glorifying gender-based violence, even [without naming the victim](#), should be banned.
- Content that [condemns abuse](#) must be protected and not misclassified as hate speech.
- Platforms should [treat veiled or indirect threats](#) as violations when contextual harm is evident.
- Stakeholder engagement and localized input are essential to enforce gender-sensitive moderation effectively.

X (Twitter)

1. **Restore Moderation Teams:** Re-establish trust and safety teams with regional language experts to moderate hate during crises. Build capacity to respond rapidly to gendered abuse.
2. **Zero-Tolerance for Threats:** Adopt a strict enforcement protocol for accounts that glorify murder, or issue rape and death threats ensuring swift takedowns and permanent suspensions.
3. **Veiled Threat Detection:** Train moderation AI to identify veiled or coded language used to threaten women (e.g., “someone should

teach her a lesson”) in local languages and dialects.

4. **Protected Profile Mode:** Introduce an opt-in feature for vulnerable users, especially women influencers, to restrict replies and mentions, filter threats, and alert moderation teams to targeted campaigns.

Cross-Platform Recommendations

- **Civil Society Collaboration:** Engage local digital rights groups to inform content policy, moderation context, and cultural nuances.
- **Reporting and Transparency:** Simplify reporting tools for gender-based abuse and publish detailed enforcement data in transparency reports.
- **Protect Counter-Speech:** Ensure content that condemns violence or calls for justice is not mistakenly removed by automated systems.
- **Hire Local Content Moderators:** Employ more moderators from the Global South familiar with local dialects, slang, and context
- **Develop Language-Specific AI:** Invest in machine learning models trained on local languages and scripts to better understand hate speech
- **Conduct Human Rights Impact Assessment:** Regularly test moderation systems for bias and ensure consistent protection across languages
- **Enhance Reporting Tools:** Create user-friendly reporting interfaces in native languages and collaborate with regional experts, like trusted partners, to keep detection tools up-to-date
- **Expand Language Support:** Improve AI and moderation capabilities for non-English languages like Urdu and Roman Urdu to detect misogynistic content effectively
- **Addressing GBV Content:** An example of this is the [oversight board](#) recommendation to Meta to fill policy gaps around gendered violence. In this particular instance, the board explicitly asked Meta to prohibit any normalization of GBV (including mockery).



DigitalRightsFoundation
"KNOW YOUR RIGHTS"



@DigitalRightsFoundation



@digitalrightsfoundation



@digitalrightsfoundation



@digitalrightsfoundation



Digital Rights Foundation



@digitalrightspk.bsky.social



@DigitalRightsPK



@DigitalRightsPK

digitalrightsfoundation.pk