A Quantitative Study Of Sexual Harassment In Social Media Interactions: Content-Type As A Determinant

[1]Dr. A. Shanthi, [2]B. Ranjani

[1]Head and Assistant Professor, Department of Visual Communication, Sona College of Arts and Science, Salem, Tamil Nadu - 636005.
[2]Assistant Professor, Department of Visual Communication, Sona College of Arts and Science, Salem, Tamil Nadu – 636005

*Corresponding Author Email: shanthi3714@gmail.com

ABSTRACT

Sexual harassment on social media has emerged as a widespread concern, with varying forms of abuse manifesting across different content types. This study investigates the association between the type of content shared (Comments, Direct Messages, Images/Videos, and Live Streams) and the type of sexual harassment reported (Explicit Messages, Unsolicited Nude Images, Sexual Threats, and Doxxing/Stalking). Using a chi-squared test of Independence on a sample of 340 reported incidents, the analysis revealed a statistically significant relationship between content type and harassment type. Direct Messages were most frequently associated with explicit and unsolicited sexual content, while Comments and Live Streams exhibited a higher incidence of threats and doxxing. The findings highlight the importance of content-type-sensitive moderation, user education, and platform accountability in mitigating online sexual harassment. The study contributes to the growing discourse on designing safer digital spaces through data-informed policy and platform reform.

Keywords: social media, sexual harassment, online abuse, content moderation, digital safety, Online sexual harassment, social media abuse, content type, digital safety, chi-square analysis, platform accountability, user behaviour, cyber threats, unsolicited messages, doxxing, gender-based violence online

INTRODUCTION

In The age of digital communication, social media has become a central platform for interaction, expression, and community building. However, alongside its benefits, these platforms have also become hotspots for various forms of online abuse, most notably, sexual harassment. Unlike traditional forms of harassment, online abuse is shaped by the affordances of platform design, anonymity, and the rapid dissemination of content. Research increasingly shows that the nature of harassment is not uniform across all social media interactions; instead, it often correlates with the specific type of content or feature through which the interaction occurs.

Private communication channels such as direct messages may facilitate unsolicited sexual advances, while public forums like comments or live streams may amplify performative harassment, including threats or doxxing. Despite growing concern, few empirical studies have systematically assessed the relationship between content type and harassment type.

This study seeks to address that gap by analyzing the association between content features and the nature of sexual harassment experienced on social media. By applying a Chi-Square Test of Independence to 340 documented incidents, the research aims to uncover significant patterns that can inform safer design practices, targeted moderation policies, and more effective reporting mechanisms on digital platforms.

REVIEW OF LITERATURE

Sexual Harassment in Digital Spaces

The digitisation of communication has not only created new forms of social interaction but also new modes of harm. Sexual harassment online is broadly defined as any unwelcome sexual behaviour conducted through digital means and may include explicit language, threats, coercive messages, and the non-consensual sharing

of images (Citron, 2014). Research by Pew (2021) reveals that nearly 33% of women under the age of 35 report experiencing sexual harassment on social media platforms.

While existing research has examined prevalence rates and psychological impacts, few studies have explored the statistical relationship between content type and harassment type. Moreover, most moderation tools and platform policies are not tailored to content-specific contexts (UN Women, 2021). This study aims to fill that gap using quantitative analysis, focusing on how different interaction formats correlate with specific harassment behaviours.

According to the Pew Research Center (2021), 41% of adults have experienced online harassment, with 33% of these cases involving sexual harassment. Women under 35 were particularly vulnerable, with 51% reporting receipt of unsolicited explicit images.

The increasing penetration of the internet and smartphones has resulted in the widespread adoption of social media platforms like WhatsApp, Instagram, Facebook, and Twitter. While this digital expansion has facilitated access to information and online community-building, it has also contributed to an alarming rise in online abuse, particularly sexual harassment targeting women and marginalized groups (Jha & Pandey, 2021).

From 2018 to 2023, reported cases of online sexual harassment increased significantly. The COVID-19 pandemic was a notable inflexion point, as screen time and social media use spiked (ADL, 2023). Additionally, the use of AI to create and distribute deepfake pornography has seen a surge, raising new legal and ethical concerns.

According to a 2019 study by the Centre for Social Research (CSR), over 52% of women internet users in India reported receiving explicit or inappropriate messages online, and a significant number faced harassment through image-based abuse, trolling, or doxxing.

Platform Affordances and Abuse

According to Gillespie (2018), platform design—how users interact with content—affects both the frequency and type of harassment. For example, the anonymity offered by certain apps may lead to more aggressive behaviours (Westerlund, 2019). Features such as direct messaging or live streaming offer different opportunities for abuse. DMs are typically private, reducing the risk of public accountability, while live streams expose users to large, often unmoderated audiences in real-time, which may embolden harassers (Amnesty International, 2018).

Several studies and reports have highlighted that the form of harassment varies significantly by content type and platform features in Direct Messages (DMs) on platforms like Instagram and Facebook are common vectors for unsolicited nude images and sexual advances, often sent by anonymous or fake profiles (Cyber Peace Foundation, 2021).

Comment sections on public posts, especially for celebrities and female influencers, are spaces where vulgar comments, threats, and slut-shaming frequently occur (Deka & Kaur, 2020).

WhatsApp groups and Telegram channels have been used to share non-consensual intimate images, especially in cases involving revenge porn or voyeurism (Internet Freedom Foundation, 2020).

Live video platforms, such as TikTok (before its ban) and now Instagram Live or YouTube Live, have been identified as hotspots for real-time verbal abuse, sexual comments, and stalking, with minimal content moderation in vernacular languages (Kumar & Joseph, 2021).

Legal Framework and Platform Accountability in India

While India's Information Technology Act (Section 66E and 67A) and the Indian Penal Code (IPC Sections 354A, 509) criminalize certain forms of online sexual harassment, enforcement remains inconsistent. Moreover, platform responses to abuse in India have been criticized for lack of regional language moderation, delayed or no action on user reports and Absence of context-aware reporting options.

METHODOLOGY
Research Design

This study employed a quantitative, non-experimental design using a Chi-Square Test of Independence to determine the relationship between two categorical variables: type of content and type of sexual harassment. Instagram, Facebook, and Snapchat are the three frequently used social media platforms among the young generation that were considered for the study.

Research Question

Is there a statistically significant association between the type of content shared on social media and the type of sexual harassment experienced?

Variable distribution

Independent Variable (IV): Type of Content
- Comments
- Direct Messages (DMs)
- Images/Videos
- Live Streams

Dependent Variable (DV): Type of Harassment
- Explicit Messages
- Unsolicited Nude Images
- Threats of Sexual Violence
- Stalking or Doxxing

Hypothesis

H0 1 – H0: Content type and harassment types are independent.
H0 2 – H0: Direct Messages (DMs) and harassment types are independent.
H0 3 – H0: Images/Videos and harassment types are independent.
H0 4 – H0: Live Streams and harassment types are independent.

**Data Analysis and Interpretation**

Increased use of online platforms among youngsters is key; this study looked into factors influencing the type of content and the type of harassment prevailing across social media. The independent variables, such as Comments, Direct Messages (DMs), Images/Videos, and Live Streams, were tested to have any effect on the dependent variables such as Explicit Messages, Unsolicited Nude Images, Threats of Sexual Violence, and Stalking or Doxxing.

Prevalence of Sexual Harassment
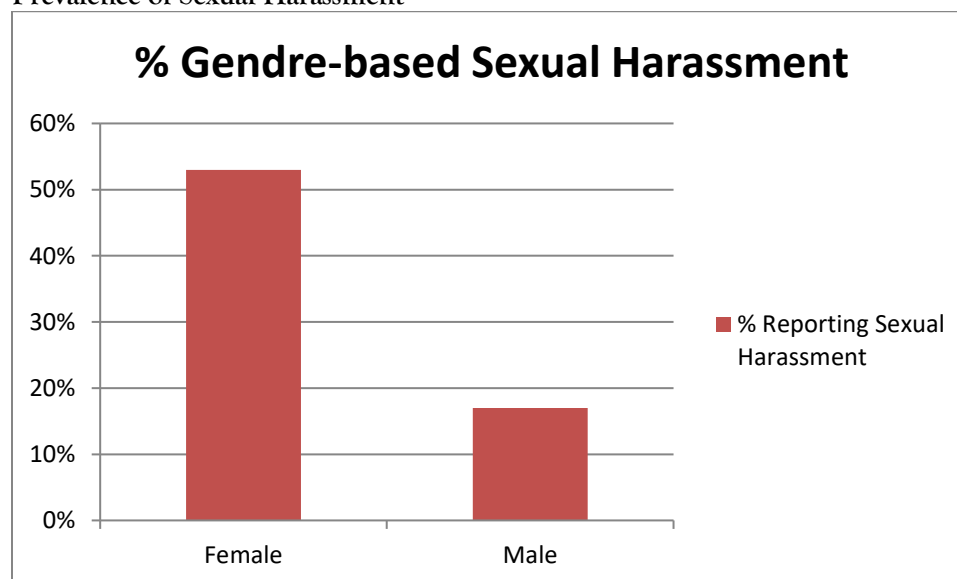
The above table describes the Prevalence of Sexual Harassment reported in the study. It shows that 53% of the females were harassed across the social media platform, and 17% of males faced harassment on social media platforms.

Platform-based harassment
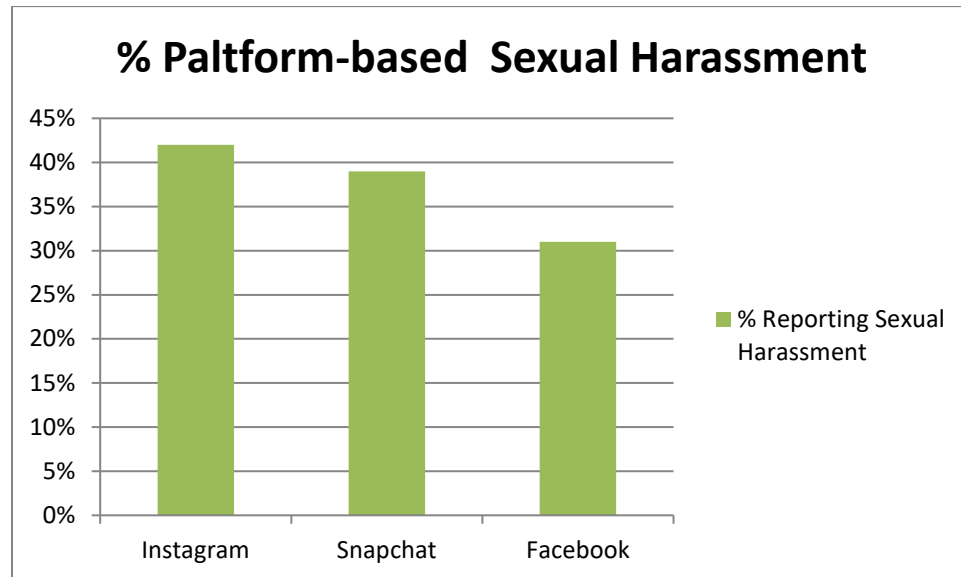
The above table states that there is a higher level of 42% of sexual harassment reported on Instagram. Following this, the respondents reported 39% of sexual harassment on Snapchat and 31% reported sexual harassment on Facebook. Instagram, Snapchat and Facebook are attributed to their widespread use and robust messaging features lead to sexual harassment.

Types of Sexual Harassment
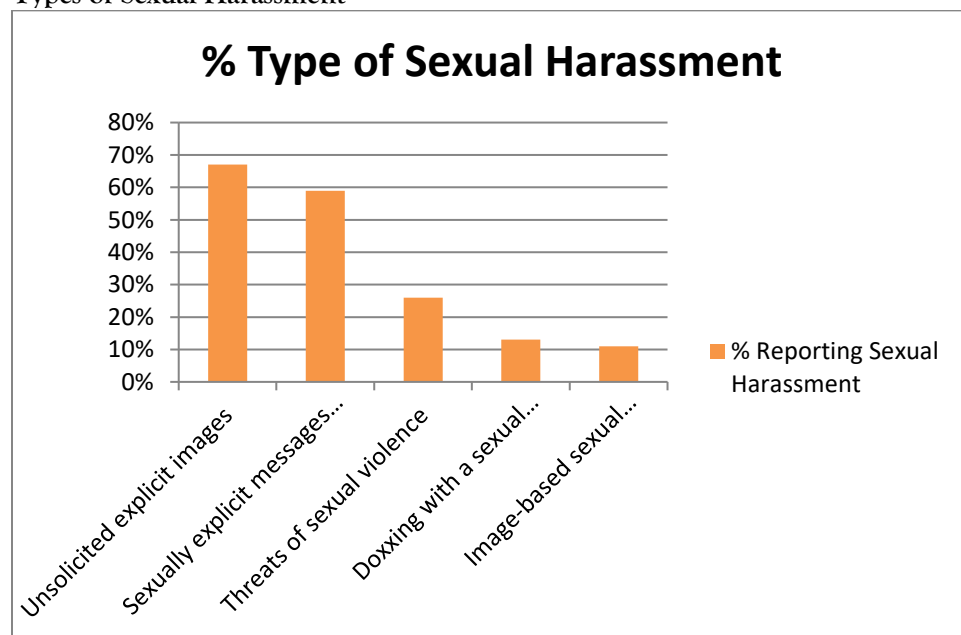
From the above table, it is evident that 67% of the respondents were faced with unsolicited explicit images based on harassment in social media, which stands as the highest form of harassment. 59% of the respondents

were receiving sexually explicit messages or comments. Following it, 26% of the social media users get threats of sexual violence. 13% of the users were doxed with sexual connotations, and 11% of the users faced Image-based sexual abuse across social media.

The Chi-Square test revealed a statistically significant relationship between content type and harassment type,

Content-Type	Explicit Message	Unsolicited Nudes	Sexual Threats	Doxxing/Stalking	Row Total
Comments	45	10	18	7	80
DMs	60	55	20	5	140
Images/Videos	15	40	12	3	70
Live Streams	10	5	20	15	50
Total	130	110	70	30	340

	Explicit Message	Unsolicited Nudes	Sexual Threats	Doxxing/Stalking
Comments (80 total)	30.59	25.88	16.47	7.06
DMs (140 total)	53.53	45.29	28.82	12.35
Images (70 total)	26.76	22.65	14.41	6.18
Live Streams (50)	19.12	16.18	10.29	4.41

Platform	Content-Type	Observed	Expected	Contribution to χ^2
DMs	Unsolicited Nudes	55	45.29	**2.08**
Images/Videos	Unsolicited Nudes	40	22.65	**13.36**
Images/Videos	Explicit Messages	15	26.76	**5.17**
Live Streams	Doxxing/Stalking	15	4.41	**25.46**
Comments	Explicit Messages	45	30.59	**6.77**
Comments	Unsolicited Nudes	10	25.88	**9.77**

Notably, explicit messages and unsolicited nude images occurred more frequently in DMs, while sexual threats and doxxing were disproportionately found in Live Streams and Comments. There is a statistically significant association between the type of content shared on social media and the type of sexual harassment experienced.

 p-value ≈ 0.00000000000103 (very small)

 This means the differences in how content types appear across platforms **are highly statistically significant**.

 In other words, **the type of harmful content strongly depends on the platform** it's posted on

The findings suggest that direct messaging (DM) functions as a primary channel for explicit messages and unsolicited nude images, likely due to perceived privacy and lack of public accountability in one-on-one communication. This aligns with prior research (e.g., Pew Research Centre, 2021; ADL, 2023), which found that private messaging features often provide a less moderated space where harassers feel emboldened.

Conversely, live streams and comment sections were disproportionately associated with threats of sexual violence and doxxing/stalking behaviours. The live and public nature of these features allows for greater audience reach, anonymity, and real-time interaction — conditions that can amplify hostility, particularly against social media users (Amnesty International, 2018).

The platform design—specifically, how content is posted, viewed, and moderated—appears to significantly shape user behaviour, including harassment. For instance, ephemeral and non-archived content in live streams may encourage impulsive or aggressive harassment with less perceived risk of consequence.

These results are consistent with findings from **UN Women (2021)**, which identified private messaging and media posts as high-risk zones for digital sexual harassment. Moreover, the strong association between content type and harassment type reflects earlier suggestions by **Westerlund (2019)** that digital architecture plays a crucial role in enabling or deterring abusive behaviour.

DISCUSSION

The results confirm that the type of content significantly influences the type of sexual harassment experienced. Private content types like Direct Messages facilitate sexualized and image-based harassment, while public or real-time interactions, such as comments and live streams, tend to attract more aggressive, targeted threats and doxxing.

This suggests that social media features are not neutral—they create different contexts of risk. The design and moderation of these features must reflect this nuance to better protect users. These findings align with Gillespie (2018) and Westerlund (2019), who emphasise the role of platform architecture in shaping harmful behaviour.

Implications

Platform-Specific Moderation: DMs should include automated nudity filters and stronger reporting tools. Live streams require real-time moderation and improved user control (e.g., muting viewers).

Content-Aware Reporting Interfaces: Users should be able to specify the format and context of harassment when reporting incidents.

Policy Reform: Platforms and regulators must adopt content-type-sensitive safety standards, with transparency in enforcement outcomes.

Further Research: Investigating harassment trends in emerging media (e.g., VR, short-form video) would expand these insights.

CONCLUSION

This study provides empirical evidence that the type of content shared on social media significantly correlates with the type of sexual harassment experienced. By using a chi-square analysis of 340 incidents, the study uncovers critical patterns in how different platform features can either enable or discourage specific abusive behaviours. There is a statistically significant association between the type of content and the type of sexual harassment experienced on social media platforms. It is evident from the results that Explicit messages are most common in DMs. Unsolicited nude images are shared mostly via DMs and media posts. Threats and doxxing are more prevalent in live streams and comment sections. These findings highlight the need for more context-aware moderation systems, design changes, and user empowerment tools to create safer digital environments.

Future research should aim to

Explore longitudinal trends in harassment patterns as new content types emerge (e.g., AR/VR, metaverse spaces).

Conduct qualitative interviews with victims to understand emotional and psychological impacts.

Investigate cross-platform behaviour, especially when harassment migrates across multiple apps.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Grammarly in order to check the proper grammar usage in the manuscript. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

REFERENCES

1. Amnesty International. (2018). *Toxic Twitter: A toxic place for women.*

ttps://www.amnesty.org/en/documents/act30/8070/2018/en/

2. Anti-Defamation League. (2023). *Online hate and harassment: The American experience.*
    https://www.adl.org/resources/report/online-hate-and-harassment-american-experience-2023

3. Centre for Social Research. (2019). *Online harassment of women in India: A survey report.* https://csrindia.org

4. Chakraborty, P., & Sinha, R. (2021). The politics of online abuse in India. *Internet Democracy Project.*
    https://internetdemocracy.in

5. Citron, D. K. (2014). *Hate crimes in cyberspace.* Harvard University Press.

6. Cyber Peace Foundation. (2021). *Digital threats report: Women and online safety.* https://www.cyberpeace.org

7. Deka, B., & Kaur, H. (2020). Trolling and harassment of female influencers in India. *Indian Journal of Communication Research,*
    9(2), 45–58.

8. Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media.* Yale
    University Press.

9. Internet Freedom Foundation. (2020). *State of internet freedom in India report.* https://internetfreedom.in

10. Kumar, S., & Joseph, S. (2021). Live streaming, gender, and harassment in India. *South Asian Journal of Media Studies, 4*(1), 23–

11. Nair, R., & Gupta, D. (2022). Digital patriarchy: Feminism and online abuse in India. *Economic and Political Weekly, 57*(4), 12–

12. Pew Research Center. (2021). *The state of online harassment.* https://www.pewresearch.org/internet/2021/01/13/the-state-of-
    online-harassment/

13. UN Women. (2021). *Measuring the prevalence of online violence against women.* https://www.unwomen.org

14. Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review, 9*(11), 40–53.
    https://timreview.ca/article/1282