
What are Effective Strategies of Handling Harassment on Twitch? Users' Perspectives

Jie Cai
New Jersey Institute of Technology
Newark, New Jersey, USA
jc926@njit.edu

Donghee Yvette Wohn
New Jersey Institute of Technology
Newark, New Jersey, USA
wohn@njit.edu

ABSTRACT

Harassment is an issue in online communities with the live streaming platform Twitch being no exception. In this study, we surveyed 375 Twitch users in person at TwitchCon, asking them about who should be responsible for deciding what should be allowed and what strategies they perceived to be effective in handling harassment. We found that users thought that streamers should be most responsible for enforcing rules and that either blocking bad actors, ignoring them, or trying to educate them were the most effective strategies.

INTRODUCTION

Online harassment is widespread in online communities and includes cyberbullying [1], hate speech [5, 7] and many other repeated annoyances or attacks. Racism, sexism, and many other prejudices flourish online [3] while trolling, flaming, spamming, and flooding messages disrupt users' online experience [6, 9]. Negative behavior online is a concern because it has contagious effects, as users are likely to imitate each other [4]. Furthermore, reducing unwanted behaviors is important because

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CSCW '19 Companion, November 9–13, 2019, Austin, TX, USA

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6692-2/19/11.

<https://doi.org/10.1145/3311957.3359478>

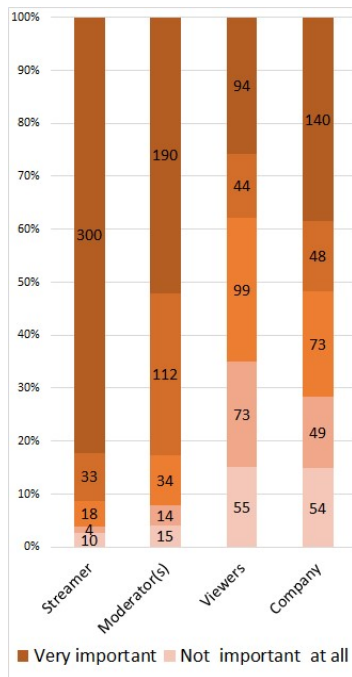


Figure 1: Who should be responsible for deciding how to enforce what is appropriate?

it may cause various adverse health outcomes, such as depressive symptoms, anxiety, loneliness, somatic complaints, or suicide [2, 8].

The live streaming platform Twitch applies both technical intervention and human moderators [10, 11], but is unique in that there are more opportunities to self-govern compared to social media such as Twitter or Facebook. Moreover, the communities are centered around the streamer, who has some control over what people are permitted to say. In this study, we ask users of Twitch who should take responsibility for handling harassment and which strategies they think are effective:

- RQ1: Who should be responsible for deciding how to enforce what is appropriate?
- RQ2: What are effective strategies in getting rid of harassment behavior?

METHODS

The survey data was collected during TwitchCon, an annual convention for Twitch enthusiasts that is hosted by Twitch. Six researchers walked around the convention and asked attendees (mainly people standing in line for something) to fill out a paper survey. Participants were given a small, custom pin that we designed for completing the survey. The survey included questions about their favorite streamer (not a part of this study) and about content moderation on Twitch (items development based on the pilot interview and brainstorm). Results from the paper surveys were then put into Survey Gizmo for digital archiving and subsequent analysis.

RESULTS

The sample (N=375) was mostly male (64.2%), 23.3% female, and two people who identified as non-binary. Age ($M=26.05$, $SD=6.56$) was between 12 and 52 years. Of the 80% of participants who reported race, most were White (44.4%), followed by Latino/Hispanic (13.1%), Asian (12.72%), Black (4%), and Pacific Islander (3.7%). 59% said that they were a streamer.

To answer the first research question, we asked in the survey, “How important are the roles of the following entities in terms of deciding how to enforce what is appropriate to say in chat? Please rate from 1 (not important at all) to 5 (very important).” The frequency table (see figure 1) displays users’ responses. Participants thought that the streamer should be the most responsible with the highest average score ($M=4.67$, $SD=.85$), followed by the moderator(s) ($M=4.23$, $SD=1.04$), the company (Twitch) ($M=3.47$, $SD=1.48$), and the viewers ($M=3.13$, $SD=1.39$). The differences between the results were statistically significant (Table 2). Also, independent t-tests showed no difference in results between the streamers and non-streamers.

To answer the second research question, we asked “How effective do you think are the following strategies in terms of getting rid of toxicity? Please rate from 1 (not effective at all) to 5 (very effective)” in the survey gave participants a list of strategies based on our earlier qualitative work. We conducted

Table 1: Exploratory Factor Analysis of Effective Strategies

Themes	Items	Loadings				
<i>Educating</i>	Explaining to the toxic person how to act properly	.88	.16	.08	.02	.15
	Educating the toxic person on the rules of the stream	.84	.10	-.08	.00	.17
	Telling the toxic person what they are doing is wrong	.83	.11	.06	.05	.10
	Asking the toxic person if they are feeling okay	.52	.44	.06	.18	.00
<i>Sympathizing</i>	Trying to have a discussion with the toxic person	.14	.78	-.07	.06	.01
	Sympathizing with the toxic person	.08	.77	.08	.02	-.04
	Asking the toxic person why they are toxic	.14	.76	.20	.14	.07
	Extending pity to the toxic pity	.14	.58	.38	.14	.01
<i>Shaming</i>	Saying rude things to the toxic person	.10	.08	.87	.04	-.04
	Shaming the toxic person	.03	.05	.72	.15	.14
	Being toxic back to them	-.08	.16	.71	.18	.01
<i>Humor</i>	Responding to toxicity with humor	.08	.10	.10	.90	.01
	Treating toxic statements as a joke	.04	.16	.29	.82	.01
<i>Blocking</i>	Banning the toxic person from the stream	.11	-.06	.10	-.04	.86
	Timing out the toxic person so they can't chat for a certain period of time	.22	.08	.01	.05	.80

a Principal Components Analysis with Varimax rotation method and eigenvalue greater than one. The exploratory factor analysis revealed five factors with a total explained variance of 68% (see Table 1). According to the description of items, We named these five variables: Educating ($M=3.10, SD=1.15, \alpha=.82$), Sympathizing ($M=2.02, SD=.94, \alpha=.76$), Shaming ($M=1.68, SD=.91, \alpha=.67$), Humor ($M=2.62, SD=1.26, \alpha=.74$), and Blocking ($M=4.01, SD=1.03, \alpha=.62$). Educating refers to telling or explaining to the violator how to act appropriately. Sympathizing refers to caring about the violator and trying to help. Shaming refers to responding to the violator with the same toxicity. Humor refers to laughing off the toxic comment. Blocking refer to banning the toxic person from speaking either temporarily or permanently.

We also asked participants to write in any strategies that were not listed above. The open-ended question revealed several themes to supplement the factor analysis results (Table 3). Many participants suggested to "simply ignore them"(M,30), and this strategy is effective because toxic people just want attention. A participant explained: "Ignoring even if they are not banned or timed out. If they do not get a reaction, they will go somewhere where they will." Not only would they ignore the toxicity, but also would be "telling the viewers to ignore them"(F,23).

Table 2: Difference Test

Pairs	T-value	P-value
Streamer vs Moderator(s)	7.11	.00
Moderators vs Company	8.51	.00
Company vs Viewers	3.46	.00

Table 3: Other Strategies Users Suggested to Combat Toxicity

Category	Code Count
Ignore	50
Encouraging positivity	13
Tolerance before ban	11
Making rules clear	7
Having good mods	5
A combo of options listed in Table 2	3
Bot intervention	2
Asking the community to help curb it	1

Table 4: Quotes of Other Strategies

<p><i>Making rules clear</i> Persistent and consistent applying the rules (M,37). Making sure your community is all on the same page of what is acceptable in your chat so they can help set the correct tone and support the chat while you are streaming (F,32).</p>
<p><i>Having good mods</i> Having good moderators that understand your wants in getting rid of toxicity in chat along with a supportive community (F,24).</p>
<p><i>A combo of options listed in Table 2</i> We usually time them out for 10 mins, tell the person what they did wrong then give them a chance to come back and stay (M,23).</p>
<p><i>Bot intervention</i> Posting help links with bot commands (M,39).</p>

Participants also suggested to “promote positivity”(F,24) such as “teaching the toxic person how to be positive”(M,28) because “positivity breeds positivity”(F,37). The streamer or mods should encourage “positive conversation and foster a healthy community”(F,28) and “have everyone involved in the community engage in a positive and friendly way”(F,20) when things do not go their way.

Many people were willing to give people opportunities to first-time violators. For example, “Just be kind, give them a chance, continue with a ban if it continues”(F,29), “Extend a second chance to first time offenders, but after that, a ban is in order”(F,25), and “Once they have been reported three times, impose a 30-day ban” (F,35). Similarly, “Track who bans by profile, not just in the channel, after three bans on different channels, either ban the profile or make a toxic emotes, although that is a form of shame (sad face)” (M,50). Other quotes are displayed in Table 4.

DISCUSSION

Twitch users thought that the streamer should be the most responsible entity to enforce the rule in the chat instead of the company; it would be interesting to see how this compares to users of social media like Facebook and Twitter. One possible explanation is that the live streaming community has a decentralized governing structure, and the users generate and moderate content autonomously.

Among the five strategies identified in the factor analysis, blocking and educating were the most effective strategies, and the other three (humor, sympathizing, and shaming) were perceived as less effective with the average score under three. Interestingly, we found ignoring was a popular strategy that was unprompted but mentioned by many users. It might be caused by the attribute of real-time interaction in the live streaming community and the fact that conversations are somewhat ephemeral. Without any action, the toxic messages in the chat will soon disappear as more comments emerge. Moderation on Facebook and Twitter often happened behind the scenes so that it is easy to block but difficult to educate the problematic viewers, In the live streaming community, the live interaction in the chat allowed moderators to block while educating at the same time. Design to facilitate educating and blocking or to help moderators to balance ignoring and actual educating and blocking should be considered. The attendees from TwitchCon were experienced users with an in-depth understanding of moderation, gaining insights into answering our research questions, but the limitation was that they would not represent the average Twitch users.

CONCLUSION

In this study, we asked users about who should be responsible for deciding how to enforce rules on Twitch and found that they held the streamer to be most responsible. We also conducted a factor analysis to identify five strategies (educating, sympathizing, shaming, humor, and blocking) and the open-ended questions revealed several more strategies (ignoring, encouraging positivity, tolerance before ban, etc.).

ACKNOWLEDGEMENT

This project was funded by the National Science Foundation [1841354]

REFERENCES

- [1] Zahra Ashktorab and Jessica Vitak. 2016. Designing Cyberbullying Mitigation and Prevention Solutions through Participatory Design With Teenagers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3895–3905. <https://doi.org/10.1145/2858036.2858548>
- [2] T. Beran and Q. Li. 2005. Cyberharrassment: A study of new method for an old behavior. *Journal of Educational Computing Research* 32, 3 (2005), 265–277.
- [3] Andrew Chadwick. 2006. *Internet politics: States, citizens, and new communication technologies*. 400 pages. <https://doi.org/10.1089/153312902753300042>
- [4] Stevie Chancellor, Jessica Annette Pater, Trustin A Clear, Eric Gilbert, and Munmun De Choudhury. 2016. #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*. ACM Press, New York, New York, USA, 1199–1211. <https://doi.org/10.1145/2818048.2819963>
- [5] Nemanja Djuric, Jing Zhou, Robin Morris, Mihajlo Grbovic, Vladan Radosavljevic, and Narayan Bhamidipati. 2015. Hate Speech Detection with Comment Embeddings. In *Proceedings of the 24th International Conference on World Wide Web - WWW '15 Companion*. <https://doi.org/10.1145/2740908.2742760>
- [6] Cliff Lampe, Paul Zube, Jusil Lee, Chul Hyun Park, and Erik Johnston. 2014. Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly* 31, 2 (2014), 317–326. <https://doi.org/10.1016/j.giq.2013.11.005>
- [7] Rijul Magu and Jiebo Luo. 2018. Determining Code Words in Euphemistic Hate Speech Using Word Embedding Networks. In *Proceedings of the Second Workshop on Abusive Language Online*.
- [8] C. L. Nixon. 2014. Current perspectives: The impact of cyberbullying on adolescent health. *Adolescent Health, Medicine and Therapeutics* 5 (2014), 143–158.
- [9] Bryan Pfaffenberger. 2011. "A standing wave in the web of our communication": Usenet and the socio-technical construction of cyberspace values. In *From Usenet to CoWebs: interacting with social information spaces*, Christopher Lueg and Danyel Fisher (Eds.). Springer, London, Chapter 2, 20–43. https://doi.org/10.1007/978-1-4471-0057-7_2
- [10] Joseph Seering, Tony Wang, Jina Yoon, and Geoff Kaufman. 2019. Moderator engagement and community development in the age of algorithms. *New Media and Society* (2019), 1–28. <https://doi.org/10.1177/ToBeAssigned>
- [11] Donghee Yvette Wohn. 2019. Volunteer moderators in Twitch micro communities: How they get involved, the roles they play, and the emotional labor they experience. In *In Proceedings of 2019 ACM Conference on Human Factors in Computing Systems*.