

Towards an Experience-Centric Paradigm of Online Harassment: Responding to Calling out and Networked Harassment

Haesoo Kim

haesookim@kaist.ac.kr
School of Computing, KAIST
Daejeon, Republic of Korea

Juho Kim

juhokim@kaist.ac.kr
School of Computing, KAIST
Daejeon, Republic of Korea

Juhoon Lee

juhoonlee@kaist.ac.kr
School of Computing, KAIST
Daejeon, Republic of Korea

Jeong-woo Jang

jwjang29@kaist.ac.kr
School of Digital Humanities and Computational Social
Sciences, KAIST
Daejeon, Republic of Korea

ACM Reference Format:

Haesoo Kim, Juhoon Lee, Juho Kim, and Jeong-woo Jang. 2023. Towards an Experience-Centric Paradigm of Online Harassment: Responding to Calling out and Networked Harassment. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (CHI '23 Workshop)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/XXXXXXXX.XXXXXXX>

1 INTRODUCTION

In recent years, recognizing and moderating harassment has become an increasingly essential task for social media platforms. Online harassment is damaging to the community in that it harms the individuals involved, but also due to its ability to negatively influence the overall social media landscape. Users often reduce their use of social media or change their behavior as a result of online harassment [6, 11, 13], resulting in a loss of users and discourse quality for platforms [2, 15, 16]. Furthermore, the prevalence of harassment causes psychological unrest and anxiety to users – leading to users preemptively change their behaviors or social media use habits to avoid being harassed even when they have not experienced it firsthand [9].

As online communities mature, harassment tactics also evolve with it. Malicious users often attempt to bypass moderation policies of the platform by introducing more subtle and complex ways of abusing others, to the level where it becomes hard to distinguish what constitutes harassment. One such type of harassment is ‘networked harassment’, where harassment is encouraged or instigated by members of an online network at a large scale [12]. In other cases, the harassment could be hidden behind ‘benign’ intent (or be truly unintentional), such as to criticize morally wrong behavior or prevent the spread of false information [7]. These forms of harassment are not clearly distinguishable by form, and there is a relative lack of scholarship that discuss the definition and detection of such

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '23 Workshop, April 22, 2023, Hamburg, Germany

© 2023 Association for Computing Machinery.
ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00
<https://doi.org/XXXXXXXX.XXXXXXX>

subtle forms of harassment. In this position paper, we introduce characteristics of a more expansive category of online harassment, while suggesting research implications for future explorations in online networked harassment.

2 CALLING OUT AND NETWORKED CRITICISM: EXPANDING THE BOUNDARIES OF HARASSMENT

The authors of this position paper have previously explored the phenomenon of ‘calling out’ as a predecessor to online harassment. Calling out refers to the public identification and criticism of individuals online, often due to a perceived transgression. Through qualitative interviews of various stakeholders in calling out, we observed how the justification of intent desensitizes users to the potential of harassment, and which elements determine users’ perception of calling outs as harassment. More detailed findings could be found in [7].

2.1 Harassment is Context-Dependent

Users’ perceptions of calling out as harassment are often defined by complex factors. Previous work on the retributive harassment has discussed the impact of prior transgressions on how acceptable the harassment is considered to be [1, 3, 4]. Policies and social norms surrounding harassment and antisocial behavior also influence users’ actions – and the users’ perceptions of what actions are and aren’t acceptable can also depend on these factors [5]. Our work found that overall attitude or prior experience surrounding a social group can influence hostile behavior, especially with subjective cases of moral judgment. Therefore, calling out and networked harassment could also be seen as an artifact of inter-group conflict and hostility, and not necessarily determined only by the individual’s behavior. Based on this, we emphasize the importance of viewing online harassment in relation to its situational contexts, which could be a significant factor in understanding calling out and networked harassment.

2.2 Harassment could be Unintentional

We also note that harassment can happen unintentionally, especially due to the networked characteristics of online social media environments. In the traditional, dyadic notion of harassment, harassment

is defined as a situation where one harasser repeatedly engages with the victim in an unwanted fashion [10, 14]. In networked harassment situations, the contribution of each individual can be minimal, and as the situation may evolve quickly, the perception of whether or not an incident is harassment could also differ. Even so, the subjective experience of harassment of the victim could be much worse as they are subject to more overall negative comments. Much previous work in the field of defining harassment considers malicious intent as a key factor [8, 17]. However, users may be unaware of the consequences of their behaviors, or the networked progression of the events may instigate harassment after the fact. Thus, we claim that harassment could happen as an unexpected consequence of criticism or benign social discourse, and that they must be taken into account when analyzing harassment at scale.

3 DESIGNING TO REDUCE THE NEGATIVE EFFECTS OF ONLINE HARASSMENT

Based on this extended perspective of online harassment, we argue that interventions for online harassment should also focus on understanding the experiences of victims rather than punishing or detecting offenders. Instead, we argue that more resources should be allocated to protecting the targeted user and giving them the sense of power to respond. For example, if a user claims that they are being harassed, platforms may employ methods to distance them from the harassers, reducing their psychological burden. This doesn't necessarily have to involve clear-cut definitions of harassment, but could be based on the subjective perception of each user.

In addition, we argue that increasing the perceived agency of victims and potential victims of online harassment could be beneficial to reducing the negative effects of online harassment. Perceived self-efficacy based on the existence of effective response measures benefit victims of harassment by encouraging them to take action [11] and providing a sense of safety. The reassurance that one will be able to respond to harassment even if it happens can make them feel safer, freer, and less anxious about potential negative responses to their posts. This paradigm could focus on allowing harassed users to apply interventions by and to themselves, providing them with immediate power to control the situation. We also propose that this could be a sustainable way to counteract the evolving tactics of online harassment by accommodating for the subjective experiences and individual situations of each harassment incident.

REFERENCES

- [1] Lindsay Blackwell, Tianying Chen, Sarita Schoenebeck, and Cliff Lampe. 2018. When Online Harassment Is Perceived as Justified. *Proceedings of the International AAAI Conference on Web and Social Media* 12, 1 (2018), 10.
- [2] Amanda Burgess-Proctor, Justin W Patchin, and Sameer Hinduja. 2009. Cyberbullying and online harassment: Reconceptualizing the victimization of adolescent girls. *Female crime victims: Reality reconsidered* (2009), 153–175.
- [3] Ann DeSmet, Sara Bastiaensens, Katrien Van Cleemput, Karolien Poels, Heidi Vandebosch, Greet Cardon, and Ilse De Bourdeaudhuij. 2016. Deciding whether to look after them, to like it, or leave it: A multidimensional analysis of predictors of positive and negative bystander behavior in cyberbullying among adolescents. *Computers in Human Behavior* 57 (2016), 398–415.
- [4] Ann DeSmet, Sara Bastiaensens, Katrien Van Cleemput, Karolien Poels, Heidi Vandebosch, and Ilse De Bourdeaudhuij. 2012. Mobilizing bystanders of cyberbullying: an exploratory study into behavioural determinants of defending the victim. *Annual review of cybertherapy and telemedicine* 10 (2012), 58–63.
- [5] Casey Fiesler, Cliff Lampe, and Amy S Bruckman. 2016. Reality and perception of copyright terms of service for online content creation. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. Association for Computing Machinery, New York, NY, USA, 1450–1461.
- [6] Jesse Fox and Wai Yen Tang. 2017. Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. *New media & society* 19, 8 (2017), 1290–1307.
- [7] Haesoo Kim, HaeEun Kim, Juho Kim, and Jeong-woo Jang. 2022. When Does it Become Harassment? An Investigation of Online Criticism and Calling Out in Twitter. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–32.
- [8] Colette Langos. 2012. Cyberbullying: The Challenge to Define. *Cyberpsychology, Behavior, and Social Networking* 15, 6 (June 2012), 285–289. <https://doi.org/10.1089/cyber.2011.0588>
- [9] Amanda Lenhart, Michele Ybarra, Kathryn Zickuhr, and Myeshia Price-Feeny. 2016. *Online harassment, digital abuse, and cyberstalking in America*. Data and Society Research Institute.
- [10] Rebecca Lewis, Alice E Marwick, and William Clyde Partin. 2021. "We Dissect Stupidity and Respond to It": Response Videos and Networked Harassment on YouTube. *American Behavioral Scientist* 65, 5 (2021), 735–756.
- [11] May O Lwin, Benjamin Li, and Rebecca P Ang. 2012. Stop bugging me: An examination of adolescents' protection behavior against online harassment. *Journal of adolescence* 35, 1 (2012), 31–41.
- [12] Alice E Marwick. 2021. Morally motivated networked harassment as normative reinforcement. *Social Media+ Society* 7, 2 (2021), 20563051211021378.
- [13] Fayika Farhat Nova, Michael Ann DeVito, Pratyasha Saha, Kazi Shohanur Rashid, Shashwata Roy Turzo, Sadia Afrin, and Shion Guha. 2021. "Facebook Promotes More Harassment" Social Media Ecosystem, Skill and Marginalized Hijra Identity in Bangladesh. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–35.
- [14] Jessica A Pater, Moon K Kim, Elizabeth D Mynatt, and Casey Fiesler. 2016. Characterizations of online harassment: Comparing policies across social media platforms. In *Proceedings of the 19th international conference on supporting group work*. 369–374.
- [15] Kusyanto Prihadi, Yen Ling Hui, Melissa Chua, and Calvin KW Chang. 2019. Cyber-Victimization and Perceived Depression: Serial Mediation of Self-Esteem and Learned-Helplessness. *International Journal of Evaluation and Research in Education* 8, 4 (2019), 563–574.
- [16] Martin EP Seligman. 1972. Learned helplessness. *Annual review of medicine* 23, 1 (1972), 407–412.
- [17] Peter K Smith, Cristina Del Barrio, and Robert S Tokunaga. 2012. Definitions of Bullying and Cyberbullying: How Useful Are the Terms? In *Principles of cyberbullying research*. Routledge, 54–68.