

Personal attacks decrease user activity in social networking platforms

Abstract. We conduct a large scale data-driven analysis of the effects of online personal attacks on social media user activity. First, we perform a thorough overview of the literature on the influence of social media on user behavior, especially on the impact that negative and aggressive behaviors, such as harassment and cyberbullying, have on users' engagement in online media platforms. The majority of previous research were small-scale self-reported studies, which is their limitation. This motivates our data-driven study. We perform a large-scale analysis of messages from Reddit, a discussion website, for a period of two weeks, involving 182,528 posts or comments to posts by 148,317 users. To efficiently collect and analyze the data we apply a high-precision personal attack detection technology. We analyze the obtained data from three perspectives: (i) classical statistical methods, (ii) Bayesian estimation, and (iii) model-theoretic analysis. The three perspectives agree: personal attacks decrease the victims' activity. The results can be interpreted as an important signal to social media platforms and policy makers that leaving personal attacks unmoderated is quite likely to disengage the users and in effect depopulate the platform. On the other hand, application of cyberviolence detection technology in combination with various mitigation techniques could improve and strengthen the user community. As more of our lives is taking place online, keeping the virtual space inclusive for all users becomes an important problem which online media platforms need to face.

Keywords: verbal aggression online, personal attacks, social media, artificial intelligence, online engagement

Significance. Despite many efforts put into researching and preventing online aggression, it is still commonplace in cyber-encounters. Experiencing cyberbullying and harassment has been shown to lead to depression, feeling of hopelessness, and social media fatigue. People who experienced online harassment also reported disengagement and stopping using the services in which undesirable behavior occurred. Unfortunately, the existing literature and nearly whole body of research mostly studies the effects of verbal aggression on well-being relying only on self-reported data. Therefore it is often unclear if being subject to verbal aggression extends its effects on the behavioral level beyond the self-reported effects. We observe, using a large-scale data-driven analysis, that experiencing verbal aggression online in the form of personal attacks indeed substantially decreases the victims' activity.

Disclaimer. During the course of the study, we have utilized content that is publicly available on Reddit.com and can be accessed via the Reddit API or other similar technologies. This study was not interventional research. Moreover, although Reddit usernames are anonymous and usually do not display any personal information, we have additionally anonymized each one of them. For these reasons, no informed consent was required (following point 8.05 of the *Ethical Principles of Psychologists and Code of Conduct* of the American Psychological Association).

1 Introduction

Recent decade brought an exponential growth of social networking services (SNS) and associated with them social media platforms (Ortiz-Ospina, 2019; Sheth, 2020), with Facebook¹ closing in on 2.5 billion users (one third of world population), YouTube² on close to 2 billion, with smaller scale platforms such as Twitter³ and Reddit⁴ converging on around 350 million (equivalent to the entire population of the United States of America).

At the same time social media have witnessed an over ten percentage-point drop in growth rate between 2012 and 2018,⁵ which suggests that the market is reaching over-satiation (Andersen, 2001) and most social media platforms will struggle for sustainable growth. This leads to a shift from quantity-based to quality-based growth policy, meaning that most social networking platforms will need to focus not on how many new users to invite, but rather on how to keep the existing users from leaving the platform and at the same time keep them actively engaged.

One problem growing within the SNS that seriously hinders user engagement is cyberbullying and various forms of Internet harassment (Ptaszynski & Masui, 2018). The global increase of the numbers of Internet harassment cases has been related to the perceived anonymity and thus the sense of impunity in users (Barlett, 2015; Barlett, Gentile, & Chew, 2016).

One of the most crude and common forms of cyberbullying are directed personal attacks. In a typical personal attack user-attacker directly verbally attacks user-victim with a harmful comment, or a set of harmful comments, often based on a made-up excuse (e.g. the victim expressed an unpopular opinion). Often the attack not only remains unrestrained, but also it escalates when other users join on either or both sides, causing a wider distress not only to the victim, but also to a larger group of users.

¹<https://www.facebook.com/>

²<https://www.youtube.com/>

³<https://www.twitter.com/>

⁴<https://www.reddit.com/>

⁵<https://www.searchenginejournal.com/growth-social-media-v-3-0-infographic/>

References

- Andersen, E. S. (2001). Satiation in an evolutionary model of structural economic dynamics. In *Escaping satiation* (pp. 165–186). Springer.
- Barlett, C. P. (2015). Anonymously hurting others online: The effect of anonymity on cyberbullying frequency. *Psychology of Popular Media Culture*, 4(2), 70. Educational Publishing Foundation.
- Barlett, C. P., Gentile, D. A., & Chew, C. (2016). Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture*, 5(2), 171. Educational Publishing Foundation.
- Ortiz-Ospina, E. (2019). The rise of social media. *Our World in Data*, 18. Retrieved from [url{https://ourworldindata.org/rise-of-social-media}](https://ourworldindata.org/rise-of-social-media)
- Ptaszynski, M. E., & Masui, F. (2018). *Automatic cyberbullying detection: Emerging research and opportunities: Emerging research and opportunities*. IGI Global.
- Sheth, J. N. (2020). Borderless media: Rethinking international marketing. *Journal of International Marketing*, 28(1), 3–12. SAGE Publications Sage CA: Los Angeles, CA.