|  |  |
| --- | --- |
|  | **Document CWG-COP-21/INF/13** |
| **18 September 2024** |
| **English only** |
|  |  |
| Contribution by the United States of America |
| CONSIDERATIONS WHEN DEVELOPING REPORTING MECHANISMS TO HELP PREVENT, MINIMIZE, AND MITIGATE BULLYING AND OTHER FORMS OF ONLINE HARASSMENT AND ABUSE |
| **Purpose**The purpose of this contribution is to share U.S. recommendations for child online protection reporting mechanisms to address online harassment and abuse of children and youth. **Action required**This report is transmitted to the Council Working Group on child online protection **for information**.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**References***Online Health and Safety for Children and Youth: Best Practices for Families and Guidance for Industry*[*https://www.ntia.gov/category/kids-online-health-and-safety/online-health-and-safety-for-children-and-youth*](https://www.ntia.gov/category/kids-online-health-and-safety/online-health-and-safety-for-children-and-youth) |

Access to digital media can be beneficial to young people, allowing them to build community and find learning opportunities, providing a benefit to their well-being. For many youths, social media and digital technology facilitate self-directed education, help them make connections, and reduce isolation, particularly for those who are marginalized.

At the same time, young people can be negatively impacted by various online harms, including harassment, cyberbullying, and sexual exploitation and abuse, threatening their safety and well-being. A 2022 survey of U.S. teens found that about half experienced some form of cyberbullying, including being harassed and being sent explicit images that they did not request.[[1]](#footnote-1)

Cyberbullying is a major concern for young people online. According to the Centers for Disease Control and Prevention (CDC), bullying among children and young people is defined as “any unwanted aggressive behavior(s) by another youth or group of youths who are not siblings or current dating partners that involves an observed or perceived power imbalance and is repeated multiple times or is highly likely to be repeated. Bullying may inflict harm or distress on the targeted youth including physical, psychological, social, or educational harm.”[[2]](#footnote-2) Cyberbullying – occurring in online spaces such as social media, online games, websites, instant messaging, chat rooms, text messages, and online forums – can take the form of social isolation, rumor spreading, cyberstalking, and public sharing of private sexual images (both real and AI-generated), and can result in depression, anxiety, and substance abuse that continue into adulthood. Cyberbullying is one of the most prevalent preceding risk factors for youth suicide-related behaviors.[[3]](#footnote-3)

In 2021, nearly 16 per cent of U.S. high school students reported being cyberbullied.[[4]](#footnote-4) A Pew Research Center survey revealed that nearly half (49 per cent) of youth aged 15-17, and 42 per cent of those aged 13-14, reported being threatened or harassed, or receiving explicit images they did not request.[[5]](#footnote-5) Further, LGBTQI+ youth, youth from racial and ethnic minority groups, and youth with disabilities are more likely to experience cyberbullying than their peers.

At the same time, there is limited understanding about the effectiveness of focused bullying prevention efforts, including cyberbullying prevention strategies, outside the school environment. Rigorous and timely evaluations of strategies to promote safe and healthy online environments for youth are required to prevent cyberbullying across different online platforms (e.g., social media, online gaming, and forums) and populations (e.g., across age groups and youth with developmental/physical disabilities or from different racial/ethnic backgrounds). Young people report being skeptical of social media companies’ willingness and ability to effectively respond to bullying online via their content moderation systems and other interventions. In consultations conducted by the Biden-Harris Administration’s Kids Online Health and Safety Task Force, children and teens also shared that the process of using safety tools when they encounter toxic online behaviors is overwhelming. Finally, protective factors —such as positive, prosocial interpersonal relationships (both in offline and virtual communities) with parents/caregivers, other adults, and peers— are important to protect against bullying and other forms of youth violence, suicidality, and self-harm.

Online service providers have a significant role to play in helping to mitigate cyberbullying and other forms of online harassment and abuse. There is a huge diversity in the online platforms and services that young people use, and there is no one-size-fits-all approach to making platforms safer for kids. Thus, it is pivotal that platform operators design their services based on research and performance assessments (e.g., user experience testing)

focused on youth’s experiences with cyberbullying, harassment, and other forms of online abuse. Providers can engage youth, parents and caregivers, relevant stakeholders, and experts from the field throughout the entire product development lifecycle —including young people from different backgrounds and age groups.

Online service providers can implement designs that help prevent, minimize, and mitigate bullying and other forms of online harassment and abuse. This could include, for example, embedding online civility norms across online spaces (e.g., onboarding, policies, design, monitoring, and resources). Providers can evaluate technical interventions aimed at minimizing online bullying, such as reporting and bystander support tools, for their effects on both youth who are being bullied and those who are bullying others. Providers can seek to use diverse, tailored approaches to help protect users from being bullied, prevent users from bullying others, and empower bystanders to stand up to bullying (e.g., by providing tools that allow users to reach out to trusted friends when they are experiencing bullying). Online service providers can also provide tools and adopt policies that give young people more control over their experiences online, such as allowing muting or blocking of problematic users, even if the behavior does not rise to the level of violating platform policies. It is also vital that online services understand the barriers that youth face to employing the reporting mechanisms that many services already provide, and take steps to remedy those issues, including improving response times and efficacy of interventions.

In addition to user reports, various data-driven methods and models are used across online spaces and platforms to detect cyberbullying. While different methods and models are best suited for different types of content and platforms, challenges persist in the consistent identification of cyberbullying across social media spaces. The accuracy and efficacy of both human moderators and cyberbullying detection models can be further affected by cultural differences among youth. Additional contextual information from users’ posts on a service can help improve the accuracy of online bullying detection. Online service providers can develop a classification framework for incorporating cultural differences to identify indicators of online bullying victimization. They can also ensure equitable access to online safety resources and mechanisms for diverse user audiences (e.g., consider literacy levels, accessibility across different devices, and languages). Providers can also adopt a risk-based approach to prioritizing interventions and identify specific groups that have experienced more bullying on their services, in order to develop tailored interventions to address cyberbullying.

It is important for online service providers and experts to share evidence-based bullying prevention resources with parents/caregivers and youth. Examples of those resources could include cyberbullying tactics; how to prevent cyberbullying; social media, apps, and sites commonly used by children and teens and the reporting mechanisms they provide; and information on cyberbullying and online gaming. [StopBullying.gov](https://StopBullying.gov) includes a comprehensive set of resources. Online service providers can also direct young people expressing self-harm or suicide-related behaviors to relevant national helplines, such as the 988 Suicide and Crisis Lifeline in the United States.

Finally, it is important that safety efforts focused on online harassment and abuse take into account that sextortion and non-consensual sharing or threatening to disseminate sexual images are also perpetrated by peers, including by current or former dating partners. Other forms of technology-facilitated abuse, such as cyberstalking, can take place in the context of dating violence, with different considerations and implications for youth safety. Online service providers can consider how their bullying and harassment reporting flows are triaged with respect to their reporting mechanisms for child sexual abuse. Online services can respond promptly to user-submitted reports of online child sexual exploitation and abuse, as well as bullying, harassment, and other abuse. It is vital for online services to address youth users’ skepticism of the utility of reporting mechanisms by providing prompt responses that effectively resolve the challenges the young person is facing.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. E. Vogels, Pew Research Center, “Teens and Cyberbullying 2022” (December 2022), <https://www.pewresearch.org/internet/2022/12/15/teens-and-cyberbullying-2022/>. [↑](#footnote-ref-1)
2. R. Gladden, A. Vivolo-Kantor, M. Hamburger, and C. Lumpkin, "Bullying Surveillance Among Youths: Uniform Definitions for Public Health and Recommended Data Elements, Version 1.0," Atlanta, GA: National Center for Injury Prevention and Control, CDC; (2014). [↑](#footnote-ref-2)
3. S. Sumner, B. Ferguson, B. Bason, J. Dink, E. Yard, M. Hertz, B. Hilkert, K. Holland, M. Mercado-Crespo, S. Tang, and C. Jones, "Association of Online Risk Factors with Subsequent Youth Suicide-Related Behaviors in the US," JAMA Network Open, 4(9):e2125860. doi: 10.1001/jamanetworkopen.2021.25860. [↑](#footnote-ref-3)
4. H. Clayton, G. Kilmer, S. DeGue, L. Estefan, V. Le, N. Suarez, B. Lyons, and J. Thornton, "Dating Violence, Sexual Violence, and Bullying Victimization among High School Students – Youth Risk Behavior Survey," United States, (2021): MMWR supp. 2023:72(1). [↑](#footnote-ref-4)
5. *Supra* n. 1. [↑](#footnote-ref-5)