Helping Teens Make Their Online Spaces Safer

Perspectives From the Online Use and Reporting Survey (OURS)

Child Mind Institute

Helping Teens Make Their Online Spaces Safer

Perspectives From the Online Use and Reporting Survey (OURS)

AUTHORS

Harry Kimball, Child Mind Institute

Aki Nikolaidis, PhD, Child Mind Institute

Mirelle Kass, Child Mind Institute

Sara Bagheri-Hamaneh, Child Mind Institute

Tatum Connell, Child Mind Institute

Lindsay Alexander, MPH, Child Mind Institute

Michael Milham, MD, PhD, Child Mind Institute

CONTRIBUTORS

Fu Wang, Child Mind Institute

Carolyn Chadwick, Child Mind Institute

Ramon Diah, Child Mind Institute

Francesca Fernandez, Child Mind Institute

Katarina Hajder, Child Mind Institute

Katie Moskowitz, Child Mind Institute

Jerry Johnson, Cascade Strategies

Executive Summary

Identifying and preventing online harassment and abuse are key concerns of parents, child safety advocates, and industry. However, youth who are targets of online harm are not always reliable reporting partners. Youth at heightened risk of harm because of emotional, behavioral, and learning challenges also face unique barriers to reporting.

The Online Use and Reporting Survey (OURS) study used a quantitative survey (n = 1,059) and qualitative interviews (n = 109) to explore how minors (age 9 to 16 years) experience online abuse, their willingness to report that abuse, and what individual-level factors (mental health, social aptitude, parenting environment) act as barriers or motivating factors.

Taken together, the quantitative and qualitative arms of this research suggest that youth in general, and youth with mental health symptoms in particular, are far more likely to have negative online experiences than they are to report them. **Two factors fuel this trend**

1. Personal barriers, strongly influenced by diagnostic status, socialization, and home environment, decrease the likelihood of reporting. These fall into four categories: a) embarrassment; and uncertainty about reporting b) process, c) policy, and d) outcome.

2. A consistent hierarchy of the seriousness of negative online experiences, openness to the myriad intentions of others, and flexible reasoning about which online behaviors deserve to be reported, mean that youth err on the side of not reporting.

The survey data includes some telling statistics:

o Respondents spend 4 hours online daily outside of school, mostly messaging and watching video.

o More than a quarter said that in the past year they had a negative online experience, defined as "a situation online that made you uncomfortable or scared."

o Only 20% said they reported the problem online when asked what they did "the last time you had problems with something or someone online that bothered or upset you in some way."

o 64% agreed or strongly agreed that "I am confident in managing these situations myself".

DATA SNAPSHOT

"THINKING ABOUT THE PLATFORMS YOU USE, HOW IMPORTANT IS IT TO YOU THAT YOU ARE PROVIDED WITH THE FOLLOWING INFORMATION WHEN YOU USE THESE PLATFORMS?"

95% felt it was important to have information on how to protect yourself online from uncomfortable or risky sexual experiences

94% felt it was important to have information on how to block people

92% felt it was important to have information on how to report people

91% felt it was important to have information on what situations are reportable

88% felt it was important to have tutorial videos on how to report or block someone

81% felt it was important to have support or counseling if you've had a bad experience

79% felt it was important to have pop-up messages that check in on how safe you feel

Executive Summary

Analysis of **quantitative data** found significant correlations with reporting and mental health symptoms:

- Youth with anxiety disorders and less positive parenting perceive significantly more barriers to reporting disturbing online experiences than their typical peers.

- Older participants, those with higher scores on social aptitude assessments, and those whose parents score high on positive parenting measures report fewer barriers.

The **qualitative interviews** reveal a consistent hierarchy of online abuse and reporting:

- Interviewees broke down online abuse into three categories: first-tier acts of malice, second-tier acts of malice, and slights.

- When the intent is clear, youths are well-versed in differentiating slights (like teasing or unsolicited attention) from first-tier acts of malice (demands for nude pictures or exposure of personal information).

- When the intent is unclear (second-tier acts of malice, like faked identities), youths are more likely to ask questions to discern the best course of action.

This research leads us to six design **recommendations** for improving reporting acceptability:

1. Provide accessible information about reporting policy, process, and outcomes

2. Increase discoverability of resources

3. Increase ease of reporting

4. Provide more examples of reportable behaviors

5. Make all reporting anonymous

6. Use age-appropriate language

In general, responses to cyberbullying and online abuse should focus on increasing understanding; destigmatizing reporting to overcome embarrassment; and putting special emphasis on solutions for kids most at risk.

Table of Contents

Introduction

There is relatively little academic research on how minors experience abuse online (e.g., harassment, cyberbullying, sexual grooming), and their mental models around how to deal with these kinds of experiences. An exception is a report published by a nonprofit organization (Thorn, 2021), which specifically surveyed 1,000 minors to better understand impediments to reporting and flagging bad behavior on online platforms. This study looked at harmful online experiences and responses to online harms, finding that a majority of minors who underwent a negative or harmful online experience choose to not seek help. Alternatively, minors who decided to take action were found to be more open to using online safety tools than turning to off-line support networks, such as a friend or caregiver. This underscores the critical role online platforms have to mitigate experiences of abuse online.

In the past decade, internet use has increased significantly in the lives of young people. In particular, technology has transformed the landscape of social interactions, creating new risks and opportunities to experience harm online (Williams & Guerra, 2007). Cyberbullying has been defined as an "aggressive, intentional act carried out by a group or individual, using electronic forms of contact, repeatedly and over time against a victim who cannot easily defend him or herself" (Smith et al., 2008). A study of more than 900 households with a child between the age of 12 and 17 years old found that almost 30% reported they had experienced at least one behavior indicative of online harassment, with adolescent females being more likely to report being a victim.

Online grooming is another opportunity for internet users to experience harm online. Along with the increased internet use, experiences of online grooming and sexual exploitation amongst youth has expanded tremendously. Online grooming is defined as "the process by which an adult builds an online relationship of trust and influence over a minor in order to obtain some type of sexual interaction" (Gámez-Guadix, M et al., 2021). Wolak, Mitchell, and Finkelhor (2006) interviewed 1,500 children ages 10–17 and found that 13% had been a victim of online sexual exploitation and 34% were unintentionally exposed to online sexual content. Online grooming has been linked to negative mental health outcomes for minors, including depression, lower self-esteem, substance abuse, and self-harm (Wachs et al., 2016; Wachs et al., 2018; Whittle et al., 2013).

The impact of cyberbullying, including adverse psychological and somatic health outcomes, has been well established (Nixon, 2014). A longitudinal study found that both traditional and cyber victimization are associated with higher levels of depressive symptoms — and that cyber victimization also predicted increases in depression symptoms severity overtime (Machmutow et al., 2012).

In the face of these new challenges, researchers are exploring how youth seek to protect themselves. A self-report online survey administered to a large group of adolescents found that a significant number felt that they would do nothing if cyberbullied themselves, fearing that "the cyberbully could get back and escalate the problem" (Li, 2010). Studies have shown with some consistency that a majority of cyberbullying

Introduction

victims choose not to report these incidents to adults (Juvonen & Gross, 2008; Li, 2007). Other research suggests that parental regulation factors (e.g., presence monitor program, parental check after use, etc) do not impact online victimization (Moore et al., 2010). In addition to barriers to reporting, a very large sample of 1,852 bullied children and adolescents explored positive motivations to report, which include a desire to exert agency and strong emotional responses to bullying (Craig et al., 2007).

Despite emerging evidence of the psychological impact of online harm, there is a limited body of research examining the distinct pathways between experiencing online abuse and mental health for children and adolescents. There is also a great opportunity to expand the limited research on reporting behaviors of online harassment, which reporting options children find most useful, and what support options or information they would like to see. Understanding how to increase motivation to report may help create a safer online environment for young people. Therefore, the purpose of this research is to expand our knowledge about how minors experience online abuse, their willingness to report that abuse, and what individual-level factors (mental health, social aptitude, parenting environment) act as barriers or motivating factors. We utilize the Child Mind Institute's comprehensive Healthy Brain Network database with brain imaging, genetics, psychiatric, behavioral, and cognitive information (Alexander et al., 2017).

This research is being funded by the Google Trust and Safety team and the Google Kids and Family team. These teams focus on creating safe and high quality product experiences at Google, this includes better understanding how adolescents interact with products and building protections for them. While the research teams at Google provided feedback and approved the final quantitative survey described below, they did not have a direct role in designing the survey, nor were they directly involved in data collection, data analysis, or interpretation of results.

Methods

Sampling and Procedures

HBN Overview

The Child Mind Institute Healthy Brain Network (HBN) is a large-scale community referred sample of 10,000 children (current N=6,820 in September 2023) and adolescents (ages 5–21) residing in the New York City area. The HBN Biobank includes behavioral and cognitive phenotyping, as well as multimodal brain imaging, electroencephalography (EEG), eye tracking, genetics, digital voice and video samples, and actigraphy (Alexander et al., 2017). The HBN Biobank has an extensive phenotyping protocol that includes comprehensive psychiatric and learning assessments, as well as instruments probing a range of familial, environmental and lifestyle variables (e.g., physical activity, nutrition).

A primary goal for the HBN is to generate a dataset that captures the broad range of heterogeneity and impairment that exists in developmental psychopathology. Accordingly, we adopted a community- referred recruitment model. We use advertisements to encourage participation of families who have concerns about psychiatric symptoms in their child. The 'announcements' are distributed to community members, educators and local care providers, as well as directly to parents via email lists and events. The advertisements highlight the potential value of participation for children who may require school-based accommodations. In particular, the comprehensive diagnostic evaluation reports provided by HBN include clinical impressions and actionable treatment recommendations; when appropriate, the reports can be used to acquire an Individualized Education Program (IEP)—a prerequisite for obtaining school accommodations, services, and specialized classroom placements. Based on the Consensus Diagnosis provided through the study, 22.5% of

participants have one mental health or learning disorder diagnosis, 68.7% have two or more diagnoses, and 8.8% have no diagnosis. Upon completion of the study, we offer participants referral information and up to three in-person feedback sessions. Modest monetary compensation for their time and expenses incurred are also provided.

Eligibility

Past and current HBN participants ages 9-15 were eligible to participate in the Online Use and Reporting (OURS) study.

Recruitment methods

The study recruited 1,059 participants from the HBN sample. 59% were male, 35% female, and 7% identified as other; and 3% were Asian, 10% Black, 6% Hispanic, 60% White, and 18% of two or more races. Past participants were recruited between January and June 2023, through emails and phone calls from research staff. Parents of past participants were sent recruitment emails, and received up to three phone calls from a research staff requesting their participation. Participants were offered a $75 Amazon gift card for completing the quantitative survey.

Current HBN participants were asked to participate at a scheduled visit between January and July 2023.

Families were invited to attend a town hall meeting at the conclusion of the study to learn about the results, and get practical advice on how to keep their kids safe online.

In addition, 109 9- to 16-year-old youth participated in an online bulletin board where they were asked about

Methods

their knowledge and understanding of how to report negative online interactions, and what would make that process easier for them. The online bulletin board was hosted by an outside contractor, Cascade Strategies (https://cascadestrategies.com/), with whom the Child Mind Institute had a BAA and service agreement, ensuring participant confidentiality. Participants in the bulletin board were asked to answer a combination of open- and close-ended questions for 30 minutes per day for 3 consecutive days. Participants were compensated $30 for each day of the bulletin board they participated in and a $20 bonus if they participated in all three days, receiving up to $110.

Study Procedures

Parents were required to provide consent for their child to participate. This consent was collected electronically prior to the study visit. At the study visit, research staff completed an assent process with the participant. In the event that a parent provided consent, but a child did not provide assent, the child was excluded from participation. Subjects were able to decide at any time to withdraw consent and quit the study. If a subject withdrew from the study data collected up until the time of withdrawal was included in the study dataset unless otherwise specified by the participant. In total, 2,745 were contacted for the study, and 1,129 agreed to participate. Of those, 27 declined to assent and 43 did not show up to their visit. Of the 1,059 who participated, 9 participants (<1%) did not complete the survey.

Once a participant completed assent, they were given the survey link to complete on their own. If a participant needed additional support, a research assistant was available to answer questions and read the questions. The research assistants checked in on participants throughout their remote visit.

This study was approved by the Advarra Institutional Review Board.

Quantitative Measures and Questionnaire Development

Online Use and Reporting Survey

A 135-item survey was developed to assess prevalence of negative online experiences, explore reporting knowledge and behaviors, and investigate how barriers to reporting negative experiences online could be associated with individual differences in children's social, psychiatric, and familial characteristics. Due to legal and ethical considerations while working with minors, participants were not asked if they personally sent or received any explicit messages, photos, or videos, or interacted with anyone online that they suspected to be an adult. As a proxy, participants were asked if they or their friends have had these experiences, and how often. This allows comparison to a peer group, without directly disclosing information that could have significant negative implications for participants. The survey was developed in consultation with a range of stakeholders including external experts comprising psychiatrists, psychologists and epidemiologists, and includes measures derived from a series of previously validated questionnaires:

Alabama Parenting Questionnaire (APQ). APQ - 9 is the abbreviated version we used that contains subscales used to derive measures of positive parenting, inconsistent discipline, and poor supervision (Elgar et al., 2007).

Social Aptitude Survey (SAS). The SAS is an assessment tool used to measure an individual's social skills, attitudes, and ability to navigate various social situations effectively (Liddle et al., 2009).

Internet Addiction Test (IAT). The IAT is a self-report assessment tool that evaluates the severity of self-reported compulsive use of the internet, assessing aspects like the extent of internet use interfering with daily life, work, social relationships, and emotional wellbeing (Faraci et al., 2013).

Methods

EU Kids Online. Sections from the EU Kids Online survey were used to assess participant's online experiences and exposure to harmful content. (Zlamal, 2020)

Thorn Barriers to Reporting Survey. The survey serves to shed light on factors that may deter individuals from reporting, thereby guiding efforts to improve reporting mechanisms and support systems (Thorn, 2021).

Healthy Brain Network Data

Data provided by participants during their participation in the Healthy Brain Network were used to provide a more detailed profile and mental health status. These measures include:

Wechsler Individual Achievement Test (WIAT). The WIAT is a comprehensive assessment tool of achievement skills, learning disability diagnosis, special education placement, and clinical appraisal. It is individually-administered and norms allow for assessment of individuals between the ages of 4 and 85 (Wechsler, 2005).

Barratt Simplified Measure of Social Status (BSMSS). The BSMSS is a measure of social status, which is a proxy for socio-economic status and is built on the work of Hollingshead (1957, 1975). It is a self-report measure completed by parents and produces a total score based on educational attainment and occupational prestige (Barratt, 2006).

Consensus Diagnosis. Mental health and learning disorder diagnoses made by clinicians after completing the full HBN evaluation. The 10 Consensus Diagnoses were defined as follows: ADHD-Combined; ADHD-Inattentive; ADHD-Hyperactive/Impulsive; Autism Spectrum Disorder; Anxiety: Agoraphobia, Generalized Anxiety Disorder, Selective Mutism, Separation Anxiety, Social Anxiety (Social Phobia), Specific Phobia, Unspecified Anxiety Disorder, Other Specified Anxiety Disorder; Depression: Major Depressive Disorder; Persistent Depressive Disorder (Dysthymia), Disruptive Mood Dysregulation Disorder, Other Specified Depressive Disorder; Behavior: Intermittent Explosive Disorder, Oppositional Defiant Disorder, Conduct Disorder - Childhood-onset type; Learning (Writing): Specific Learning Disorder with Impairment in Written Expression; Learning (Math): Specific Learning Disorder with Impairment in Mathematics, WIAT-Numerical Operations<85; Learning (Reading): Specific Learning Disorder with Impairment in Reading, WIAT-Word Reading <85.

Quantitative Data Analysis

We provide descriptive statistics for the population, and then stratified on disorder status. Rao-Scott chi-square tests were used to determine whether participants answered the reporting questions differently based on disorder status, and whether these differences were statistically significant.

We performed an Exploratory Factor Analysis of the 12-Item Barriers to Reporting Survey to identify four major domains of variance between individuals on their reporting behavior.

Factor 1 contained variables that related to individuals' uncertainty about the reporting process, including:

- o I am unsure how to report
- o The platform does not provide enough information to report
- o I do not know how to make an official report

Factor 2 contained variables that related to individuals' embarrassment, including:

- o I am too embarrassed to talk about these situations

Factor 3 contained variables related to individuals' uncertainty about platform policy, including:

- o I do not know of a policy that protects students against these situations
- o I do not know where to locate policy information
- o It is important to report these situations to help other kids like me

Factor 4 contained variables related to individuals' lack of confidence in positive outcomes, including:

- o I am not confident in support options offered by the platform
- o I would not expect a favorable outcome to occur if I reported

We performed linear regression models predicting each factor with the following measures: Age, gender identity, sexual identity, BSMSS, IQ, SAS, IAT, APQ Inconsistent Discipline, APQ Positive Parenting, APQ Poor Supervision, and the Consensus Diagnosis.

Qualitative Bulletin Board Guide Development

The focus of the online bulletin board was to provide more in-depth, qualitative knowledge of how children and adolescents perceive different forms of online harassment and what they would do or not do in different situations (e.g., reporting the situation to a platform). The online bulletin board modality was selected due to its satisfaction of our purpose; in addition to offering participant anonymity, each participant could share their own thoughts whilst still being able to engage in discussions with their fellow participants. Additionally, this platform mimics traditional online forums, which participants were likely familiar with. The content for the online bulletin board was derived from a variety of resources, such as EU Kids Focus Group Guide and literature about focus groups about online harassment (e.g., Jacobs, et al., 2015). The bulletin boards were structured to guide participants through five separate stages of inquiry. They were asked to respond to 1) scenarios featuring fictional adolescents facing common online challenges with privacy, disturbing behavior, and reporting and 2) hypothetical personal situations that they might find themselves in. The participants also answered specific questions about their own attitudes towards 3) reporting and 4) cyberbullying. Finally, they were asked to provide feedback on 5) how platforms could adapt to better support safety and reporting for youth.

Qualitative Data Analysis

Researchers from Cascade Strategies performed qualitative text analysis and prepared a summary of key points and trends from the four separate bulletin board groups.

Results/Key Findings

Overview

Respondents to the quantitative survey spent most of their time online messaging, watching videos, or playing games. More than a quarter (27%) of respondents to the quantitative survey said that they had a negative online experience in the past year. Of that group, 69% had negative experiences more than once.

Analysis of the quantitative portion of our mixed methods research found that youth with a consensus diagnosis of anxiety, lower social aptitude, and whose parents score lower on the APQ positive parenting measures perceive more barriers to reporting negative online experiences than their typical peers, and these differences are statistically significant. Older participants, those with higher scores on social aptitude assessments, and those with higher IQ are associated with having fewer barriers to reporting, with these differences also being significant.

The quantitative survey also revealed that the greatest "barrier" to reporting is the young people's own assessment that a disturbing online event was "not a big deal" enough to report. This trend is fully explored in the qualitative interviews, where the interviewees broke down online abuse into three categories: first-tier acts of malice, second-tier acts of malice, and slights. The respondents are remarkably coherent in their feeling that slights can be handled without appeal to authorities, e.g., "the situation could be resolved without the platform's help." More malicious

behavior (like demands for nude pictures or exposure of personal information) seems to trigger outrage in young people, who then act by telling an adult or making a report.

Most of the time, respondents use flexible reasoning to evaluate nuances in online behavior, grant others the benefit of the doubt, and prefer to let individuals work out their differences before appealing to platforms whose ability to effect change is not at all clear to them.

DATA SNAPSHOT

PARTICIPANTS WERE GIVEN A LIST OF REASONS THEY MAY NOT HAVE REPORTED AN INCIDENT THEY EXPERIENCED ON A PLATFORM, AND ASKED TO SELECT THE TOP 3 REASONS THAT APPLY TO THEM.

The top 3 reasons participants cited for not reporting were:

1. Felt this was not a big deal **(49%)**

2. Worried their report would not be anonymous **(18%)**

3. Felt embarrassed and worried about being judged **(20%)**

Results/Key Findings

Quantitative Survey

The data presented here include 1,059 participants with a mean age of 12 years. Nearly half of participants were 11 or younger (45%). The majority identified as male (590; 59%) %), with 349 (35%) females, 52 (5%) who identified as "other", and 17 (2%) who preferred not to share their gender identity. The majority reported their sexual identity as "straight" (585; 56%), with 137 (13%) reporting queer or other sexual identities and 331 chose not to share (31%). The sample was enriched in terms of mental health and learning disorder status, with 936 (91%) participants having at least one Consensus Diagnosis. The most prevalent Consensus Diagnoses were Anxiety (403; 38%), ADHD-Combined Type (305; 29%), and ADHD-Inattentive (291; 27%). Smaller numbers of participants had Consensus Diagnoses of Behavior (183; 17%), Learning-Math (168; 16%), Autism (112; 11%), Learning-Writing (94; 9%), Depression (59; 6%), and ADHD-Hyperactive/Impulsive (59; 6%).

DATA SNAPSHOT

Age	N	%
9	135	13%
10	154	15%
11	188	18%
12	189	18%
13	199	19%
14	151	14%
15	42	4%
16	1	0.1%

On average, respondents said they spent about 2 hours online each day for school and 4 hours online for personal use or socializing. Most of that time is spent messaging, watching videos, or playing games.

54% said they "never" shared photos online, 61% never comment, and 73% never post. 43% said they sent direct messages "several times a day" or "almost constantly." 45% spent 1-3 hours watching video each day, and 17% for 4 to 6 hours. Just 4% said they don't watch video on a daily basis. Of video watchers, 97% used YouTube and 41% used TikTok.

More than a quarter (27%) said that in the past year they had a negative online experience, defined as "a situation online that made you uncomfortable or scared." Of those, 69% reported having more than one upsetting experience in the past year. 18% said they or their friends had received sexual messages in the past year. 67% agreed or strongly agreed with the statement "it is important to report these situations to help other kids like me." 64% also agreed or strongly agreed that "I am confident in managing these situations myself. Finally, more than 70% disagreed or strongly disagreed that "I am unsure how to report."

And yet, only 20% said they reported the problem online when asked what they did "the last time you had problems with something or someone online that bothered or upset you in some way." More than 60% said they would be more likely to report if they knew a real human being that worked at the platform would evaluate their report, compared to an independent organization (33%) or an AI bot (16%). Nearly 50% would be more likely to report if they felt the perpetrator was lying about their identity.

We performed a linear regression analysis to predict the main factors summarizing barriers to reporting.

o Predictors of higher barriers to reporting due to uncertainties in the reporting process (Factor 1) included lower social aptitude, less positive parenting as measured by the APQ, younger age, and lower IQ.
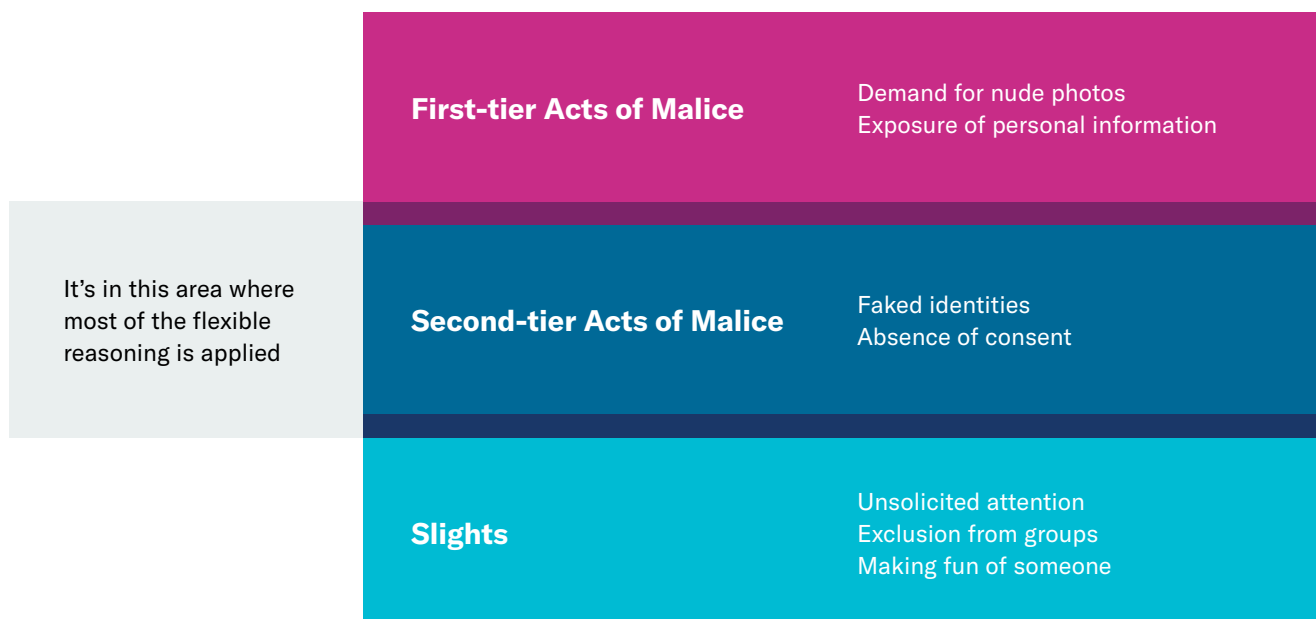
Results/Key Findings

- Predictors of higher barriers to reporting for kids due to embarrassment (Factor 2) include lower socioeconomic status, lower social aptitude, and less positive parenting as measured by the APQ.

- Barriers to reporting due to uncertainty about reporting policies (Factor 3) were lower in those with a behavioral diagnosis and older age.

- Predictors of higher barriers to reporting for kids due to uncertainty about the outcome of reporting (Factor 4) included the presence of an anxiety diagnosis. For Factor 4, higher social aptitude was associated with increased likelihood of reporting.

The Adjusted R-squared ranged from 4.2% to 11.8% across the four factors, indicating that these variables have sufficient power to predict individual differences in reporting. This also suggests that other variables may predict reporting tendencies with even higher accuracy and are worth investigating.

Qualitative Bulletin Board

The qualitative portion of this research sought to dig deeper into the nuanced attitudes and behaviors of the participants. Specifically, the research was designed to elicit responses around these questions:

- How do children and adolescents perceive common online harms?

- How do children and adolescents behave when faced with common online harms?

- How do different reporting policies influence children and adolescents' comfort levels with reporting?

First-tier Acts of Malice	Demand for nude photos Exposure of personal information
Second-tier Acts of Malice	Faked identities Absence of consent
Slights	Unsolicited attention Exclusion from groups Making fun of someone

It's in this area where most of the flexible reasoning is applied

Results/Key Findings

Based on thousands of text responses, the respondents did not regard online behavior as black and white – even behavior they identified as unpleasant or troubling. Rather, they used "flexible reasoning" to evaluate nuances based on context, the inferred intent of other actors, the degree of malice, and what "feels right" under each scenario. Even so, respondents' assessments became more punitive when situations crossed a line into severe malice and a clear intent to harm.

In their responses, participants described a rough hierarchy of offenses: first-tier acts of malice, second-tier acts of malice, and slights. The hierarchy serves as a guide to the "outrage threshold": the point at which respondents report to authorities (adults or the platform). Although flexible reasoning appears to guide respondents when faced with certain situations, crossing the outrage threshold triggers a more rigid response. They draw a red line when the scenario involves:

o Demand for nude photos

o Exposure of personal information

Almost universally, participants said they would report scenarios including these behaviors (e.g., "what if an impostor account sends Natalie a direct message asking Natalie for naked or nude photos?").

In contrast, when respondents experience "slights," or situations where the degree of malice appeared to be minimal, they generally felt comfortable with handling the situation themselves (e.g., blocking or ignoring). In these situations, respondents were more tolerant. When the intent to harm was unclear or ambiguous, as in second-tier acts of malice, respondents tended to

ask questions about the situation (e.g., "it depends on if [the perpetrator] was intending to hurt [the victim]"). Some resorted to criticism in these cases, but it was not as uniform as in the first tier. Respondents were pragmatic rather than doctrinaire, granting others a good deal of leeway when the degree of malice was arguable (e.g., John is using social media to attract the interest of Stephanie).

When asked how comfortable they currently felt in reporting on a platform they used frequently, less than half answered "a little comfortable" or "very comfortable." Over 60% shared that they would feel more comfortable if the reporting button was easily locatable, the platform took precautions during report investigations, and that the platform would take further action if appropriate. Overwhelmingly, respondents shared their desire for anonymity.

Respondents desired to have multiple modalities when learning how to report on a platform (text, video, and interactive), accompanied by comments noting different learning styles. Participants also frequently shared a belief that reporting should be intuitive, and not a process that must be learned.

Two characteristics of the qualitative sample should be taken into account when evaluating these results: a) participants were drawn from the HBN sample and thus are much more likely to have mental health or learning disorder diagnoses than the general population; and b) the sample is composed largely of young people who have some degree of parental guidance, and we don't know what responses we would receive among a sample of young people who are more independent due to age, demographics, or mental health status.

Recommendations and Policy Implications

Taken together, the quantitative and qualitative arms of this research suggest that youth in general, and youth with mental health symptoms in particular, are far more likely to experience negative events online than they are to report them.

1. Diagnostic status, social aptitude, and family environment contribute to four basic barriers described by respondents: a) uncertainty about the reporting process, b) embarrassment about making a report, c) uncertainty about reporting policy, and d) uncertainty about what will happen after a report is made.

2. Flexible reasoning about which online behaviors deserve to be reported, and what could or should be done about them, means that youth err on the side of not reporting.

The broad confusion with the details and design of the reporting process expressed by young people — including what outcomes to expect for themselves and those they might report — is a clear target for intervention. Below, we provide five design recommendations based on this research that may improve acceptability of reporting processes, increase reporting, and build bridges to youth users.

DESIGN FEEDBACK FROM PARTICIPANTS

"PLEASE RATE HOW COMFORTABLE YOU WOULD FEEL REPORTING AN UNPLEASANT EXPERIENCE TO THE PLATFORM IF YOU KNEW THAT THE STATED SITUATION WAS TRUE ABOUT THAT PLATFORM."

Reporting button is easy to locate on all screens: **66%**

Platform takes precautions during a reporting investigation: **60%**

If needed, the platform will contact the police: **60%**

Consequences for violations are clearly stated: **57%**

When reporting, you have to select which rule was violated: **53%**

Option to choose if the user knows you reported them: **27%**

The report requires your name: **11%**

Reporting process design recommendations for improving safety

1. **Provide accessible information about reporting policy, process, and outcomes**

   o Create different types of materials that allow for differences in learner styles/preferences e.g., text, video, interactive

   o Use tailored language for different target populations (e.g., by age or mental health or learning disorder)

   o Make reporting as easy and intuitive for users as possible

2. **Increase discoverability of resources**

   o Suggest these pages/resources to accounts that are used by youths

   o Use influencer creators to make videos encouraging kids to protect themselves e.g., Child Mind Institute "My Younger Self" Campaign (childmind.org/awareness-campaigns/)

3. **Increase ease of reporting**

   o Make report button easy to locate

   o Focus on intuitiveness of reporting for different age groups

4. **Provide more examples of reportable behaviors**

   o Links to explain each behavior that is allowed vs not allowed e.g., Interactive videos/quizzes on real world scenarios

   o 5. **Make all reporting anonymous** e.g., consider including a lag time between reporting and repercussions so that reports cannot be easily traced back to the reporter.

   o 6. **Use age-appropriate language** e.g., Instead of "report here," consider "tell the platform"

Policy implications

Adolescent social media users are typically pragmatic, flexible, and situational in their assessments, and platforms should be aware of this bias when seeking to gain their trust or inspire action (for instance, an increase in incident reporting). They have high tolerance for uncertainty; they dislike confrontation and inflexible rules; they have compassion for offenders' internal psychological needs or motivations; and their primary desire is to report in relative safety.

All of this suggests that responses to cyberbullying should be as open, flexible, and accommodating as possible. Otherwise, young people will not engage with programmatic responses to the problem.

Conclusion

The small but growing literature concerning attitudes, opinions, and actions of youth around cyberbullying and disturbing online experiences suggests that young people face many internal and external barriers to reporting these events to parents and platforms. The present research confirms this assessment and adds valuable nuance. The youth most at-risk for being bullied or targeted online – those with behavioral, emotional, and learning challenges and unstable home environments – are also more likely to struggle with reporting than their typical peers. Specifically, youth with anxiety were significantly more likely than typical peers to cite uncertainty about the outcome of the reporting process as a reason not to report; and youth with lower socioeconomic status, lower social aptitude, and less positive parenting were significantly more likely than their peers to not report because of embarrassment.

Furthermore, young people in our sample have a fluid and even compassionate view of "bad" behavior online and why their peers act out. This leads youth to err on the side of not contacting authorities in any but the most outrageous situations.

Taken together, this shifting landscape is difficult to navigate. However, the design recommendations in this report could go a long way to increasing reporting and collaboration between platforms and the young users they want to keep safe. And if adults and online platforms adopt some of the more flexible attitudes exhibited by the participants in this study, perhaps they can find other ways to work together with young users towards a shared goal of reducing online harms, abuse, and criminality.

References

Achenbach, T. (1991). Manual for the child behavior checklist/4-18 and 1991 profile. *University of Vermont, Department of Psychiatry*. https://cir.nii.ac.jp/crid/1570009749384841472

Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., Litke, S., O'Hagan, B., Andersen, J., Bronstein, B., Bui, A., Bushey, M., Butler, H., Castagna, V., Camacho, N., ... Milham, M. P. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data*, 4, 170181.

Barratt, W. (n.d.). Social class on campus. Retrieved September 15, 2023, from http://socialclassoncampus.blogspot.com/2012/06/barratt-simpli-fied-measure-of-social.html

Craig, W., Pepler, D., & Blais, J. (2007). Responding to Bullying: What Works? School Psychology International, 28(4), 465–477.

Elgar, F. J., Waschbusch, D. A., Dadds, M. R., & Sigvaldason, N. (2007). Development and Validation of a Short Form of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies*, 16(2), 243–259.

Faraci, P., Craparo, G., Messina, R., & Severino, S. (2013). Internet Addiction Test (IAT): which is the best factorial solution? *Journal of Medical Internet Research*, 15(10), e225.

Gámez-Guadix, M., De Santisteban, P., Wachs, S., & Wright, M. (2021). Unraveling cyber sexual abuse of minors: Psychometrics properties of the Multidimensional Online Grooming Questionnaire and prevalence by sex and age. Child abuse & neglect, 120, 105250. https://doi.org/10.1016/j.chiabu.2021.105250

Hollingshead, August B. (1957). Two factor index of social position. Unpublished manuscript, Department of Sociology, Yale University, New Haven, Connecticut

Hollingshead, A. B. (1975). Four factor index of social status. Unpublished manuscript, Department of Sociology, Yale University, New Haven, CT.

Jacobs, Goossens, L., Dehue, F., Vollink, T., & Lechner, L. (2015). Dutch Cyberbullying Victims' Experiences, Perceptions, Attitudes and Motivations Related to (Coping with) Cyberbullying: Focus Group Interviews. *Societies (Basel, Switzerland)*, 5(1), 43–64.

Juvonen, J., & Gross, E. F. (2008). Extending the school grounds?--Bullying experiences in cyberspace. *The Journal of School Health*, 78(9), 496–505.

Liddle, E. B., Batty, M. J., & Goodman, R. (2009). The Social Aptitudes Scale: an initial validation. *Social Psychiatry and Psychiatric Epidemiology*, 44(6), 508–513.

Li, Q. (2007). New bottle but old wine: A research of cyberbullying in schools. *Computers in Human Behavior*, 23(4), 1777–1791.

Li, Q. (2010). Cyberbullying in High Schools: A Study of Students' Behaviors and Beliefs about This New Phenomenon. *Journal of Aggression, Maltreatment & Trauma*, 19(4), 372–392.

Machmutow, K., Perren, S., Sticca, F., & Alsaker, F. D. (2012). Peer victimisation and depressive symptoms: can specific coping strategies buffer the negative impact of cybervictimisation? *Emotional and ehavioural Difficulties*, 17(3-4), 403–420.

Moore, R., Guntupalli, N. T., & Lee, T. (2010). Parental Regulation and Online Activities: Examining factors that influence a Youth's potential to become a Victim of Online Harassment. *International Journal of Cyber Criminology*, 4(1), 685–698.

Nixon, C. L. (2014). Current perspectives: the impact of cyberbullying on adolescent health. *Adolescent Health, Medicine and Therapeutics*, 5, 143–158.

Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, and Allied Disciplines,* 49(4), 376–385.

Thorn. (2021). Responding to Online Threats: Minors' Perspectives on Disclosing, Reporting, and Blocking. https://teamthorn.co/3uKqRDj

Wachs, S., Jiskrova, G. K., Vazsonyi, A. T., Wolf, K. D., & Junger, M. (2016). A cross-national study of direct and indirect effects of cyberbullying on cybergrooming victimization via self-esteem. Psicología Educativa, 22(1), 61–70. https://doi.org/10.1016/j.pse.2016.01.002

Wachs, S., Whittle, H. C., Hamilton-Giachritsis, C., Wolf, K. D., Vazsonyi, A. T., & Junger, M. (2018). Correlates of mono-and dual-victims of cybergrooming and cyberbullying: Evidence from four countries. Cyberpsychology, Behavior and Social Networking, 21(2), 91–98. https://doi.org/10.1089/cyber.2016.0733

Wechsler, D. (2005). Wechsler Individual Achievement Test 2nd Edition (WIAT II). London: The Psychological Corporation.

Whittle, Hamilton-Giachritsis, C., Beech, A., & Collings, G. (2013). A review of online grooming: Characteristics and concerns. Aggression and Violent Behavior, 18(1), 62–70. https://doi.org/10.1016/j.avb.2012.09.003

Williams, K. R., & Guerra, N. G. (2007). Prevalence and predictors of internet bullying. *The Journal of Adolescent Health*, 41(6 Suppl 1), S14–S21.

Wolak, J., Mitchell, K. J., & Finkelhor, D. (2006). Online Victimization of Youth: Five Years Later.

Zlamal, R., MacHackova, H., Smahel, D., Abramczuk, K., Ólafsson, K., Staksrud, E., & Department of Media and Communications. (2020). *EU Kids Online 2020* : technical report. EU Kids Online. https://doi.org/10.21953/LSE.04DR94MATPY7

Child Mind Institute

The Child Mind Institute is dedicated to transforming the lives of children and families struggling with mental health and learning disorders by giving them the help they need to thrive. We've become the leading independent nonprofit in children's mental health by providing gold-standard evidence-based care, delivering educational resources to millions of families each year, training educators in underserved communities, and developing tomorrow's breakthrough treatments. Together, we truly can transform children's lives. **Learn more at childmind.org**

New York City
101 East 56th Street
New York, NY 10022
(212) 308-3118

San Francisco Bay Area
2000 Alameda de las Pulgas, Suite 242
San Mateo, CA 94403
(650) 931-6565