On the Role and Form of Personal Information Disclosure in Cyberbullying Incidents

Sadiq Aliyu, Kavous Salehzadeh Niksirat, Kévin Huguenin, Mauro Cherubini

On the Role and Form of Personal Information Disclosure in Cyberbullying Incidents

Sadiq Aliyu
University of Lausanne
Switzerland
sadiq.aliyu@unil.ch

Kavous Salehzadeh Niksirat
University of Lausanne
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland
kavous.salehzadehniksirat@unil.ch

Kévin Huguenin
University of Lausanne
Switzerland
kevin.huguenin@unil.ch

Mauro Cherubini
University of Lausanne
Switzerland
mauro.cherubini@unil.ch

ABSTRACT

Disclosing personal information significantly increases the likelihood of incidents of cyberbullying. This highlights the significance of investigating the relationships between various stakeholders in cyberbullying incidents. Our objective is to gain insight into the roles of the stakeholders, types, and typical paths of personal information in cyberbullying incidents. To achieve this, we conducted a large-scale survey with a representative sample of internet users from the United States and Nigeria (N = 1555). Our findings indicate that cyberbullying is often fueled by personal information that becomes known, directly or through social media, to other stakeholders. Cyberbullying incidents involve more than just attackers and victims; they can involve other stakeholders as third-parties '*disclosers*.' Both strangers and friends typically engage in such activities. Cyberbullying incidents are twice as common in Nigeria as in the United States. Our findings have implications for design, social-media literacy programs, and policy.

KEYWORDS

cyberbullying, personal information, information disclosure, privacy

1 INTRODUCTION

Cyberbullying[1] (i.e., henceforth CB) is becoming a widespread problem [69]. Over 40% of adult internet users in the United States have been victims of CB, with at least 36% being among middle school and high school students [40]. The consequences of CB have been reported in the form of mental health issues [85], increased stress and anxiety [30], depression [26], violent behavior [51], and low self-esteem [62], to name a few. Recent research [2, 6, 47, 55] shows that the disclosure of personal information can play a crucial role in the occurrence of CB. For instance, an unsafe[2] online conversation between two individuals can enable the disclosure of sensitive information (i.e., information about the *owner*), such as by sharing private family information or by exchanging intimate photos [4]. Ultimately, such information can be used maliciously by the individual who receives the information (i.e., information *co-owner*[3]) directly from the victim or through disclosers and can cause CB victimization. For instance, the information co-owner (i.e., the *attacker*) might use the information acquired during the conversation

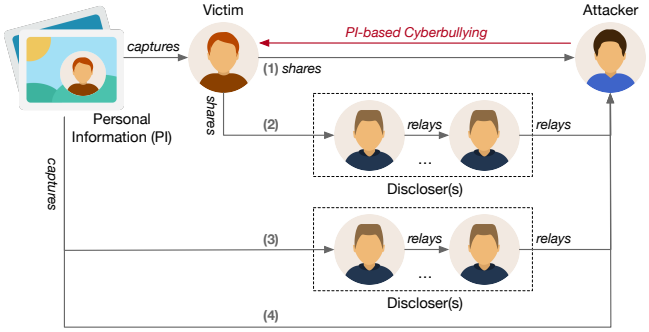

Figure 1: Path of personal information (PI) used in a CB incident, from the victim to the attacker. Here, PI is a photo of the victim. The individual who captures the PI can be the victim themselves, the attacker, or some third-party individual. There are several possible paths: (1) The victim captures their PI and shares it directly with the attacker. (2) The victim captures their PI and shares it with another individual(s)—called discloser. The PI is then relayed from the discloser(s) to the attacker. (3) A discloser captures the PI and relays it to the attacker, possibly through other disclosers. (4) The attacker captures PI *directly*. In all cases, the attacker ultimately uses the PI to cyberbully the victim.

[1]The use of cell phones, instant messaging, e-mail, chat rooms, or online social networks (OSNs) to harass, threaten, or intimidate someone [63].

[2]That involves sensitive topics or harmful behaviors.

[3]Not to be confused with the term of "co-owner" in the context of interdependent privacy (e.g., a photo depicting multiple people as co-owners). We use this term to refer to any user that jointly owns and controls data. Note that, although it is interesting, in this paper, we do not distinguish between the cases where the attacker and/or disclosers are featured in the data used for CB.

Proceedings on Privacy Enhancing Technologies 2023(4), 468–483

https://doi.org/10.56553/popets-2023-0120

to blackmail the information owner (i.e., the victim). This phenomenon can be explained by communication-privacy management (CPM) theory [67], and it is indeed an interdependent privacy situation [38] where a person's privacy can be affected by the actions of other individuals—after the disclosure of the information.

Although previous research has extensively studied CB from different angles, such as the mitigation and prevention of CB (e.g., [7]), methods for the automatic detection of CB (e.g., [58]), or solution-focused therapy (e.g., [39]), the understanding of the underlying dynamics of information disclosure and its relation to CB is limited. For instance, the difference between the way information disclosure from *self* can cause CB victimization compared to the way information disclosure from *others* needs to be clarified. In other words, it is unclear whether any individual, apart from the victim and attacker, is involved in CB incidents. We speculate that third-party individuals (i.e., henceforth '*disclosers*') also engage in CB incidents where they can facilitate the flow of information from victims to attackers and favor the occurrence of CB (see Figure 1). Understanding the role of disclosers is crucial as it can inform the design of mitigation/prevention tools.

In this work, we investigate the role of information disclosure in CB and, more specifically, we look at concrete *real* CB incidents to understand (i) *how often* CB is caused by information disclosure, (ii) the *type* of information, and (iii) the *path* of information (e.g., origin, destination, relays) that causes CB (see Figure 1). Through a comparative lens, we also seek to understand the extent of the phenomenon in the United States and Nigeria, two culturally, technologically, and economically different countries. In particular, we address the following research questions:

- **RQ1.** What features of online privacy disclosure are linked to CB? In particular, how often does CB rely on personal information disclosure? What types of personal information play a role in CB? What are the typical paths of the personal information used for in CB?
- **RQ2.** To what extent are such incidents related to information disclosure by a third-party discloser?
- **RQ3.** To what extent can existing strategies, behaviors, and practices reduce the incidence of CB caused by the online disclosure of private information?

To provide answers, we deployed a large-scale survey among a wide range of respondents from the United States and Nigeria, involving three different stakeholders: (1) victims (i.e., information owners), (2) attackers (information co-owners), and (3) disclosers (third-party information co-owners). We distributed 3939 questionnaires; among them, we collected $N = 1617$ complete answers (i.e., respondents with CB experiences). The principles of the communication-privacy management (CPM) theory [67] (i.e., information ownership, co-ownership, and turbulence) and two parameters of the contextual integrity (CI) theory [59] (i.e., types of information and transmission principles) inform our study. The former helped us to explore how users manage personal information in an interdependent communication context and how they developed an effective survey questionnaire targeted at the different stakeholders. The latter helped us to understand the flow and type of information involved in the CB incidents and to categorize the incidents according to the type of information involved and the appropriateness of the information flow (i.e., violations of CI).

We found a significant number of respondents who reported disclosing information about others that fueled CB. Indeed, almost one out of four respondents who participated in a CB incident was a discloser. Interestingly, we noticed a violation of trust in many cases, where the information co-owners further spread the personal information and caused CB. We found that CB is not only committed by strangers but also by friends. We identified key differences between the populations from Nigeria and the United States. For instance, though the incident rate was high in both countries, it was more than double in Nigeria. We also found that victimization using photos, memes, or screenshots occurs more frequently in Nigeria than in the United States; consensual information disclosure conducive to CB is slightly higher in the United States than in Nigeria; information disclosure with the intention to bully is slightly higher in Nigeria than in the United States, and attackers leverage personal information to bully victims twice as frequently in Nigeria as in the United States.

To our knowledge, this is the first work that systematically studies the involvement of '*disclosers*' in CB. Our findings have implications for design, particularly from the perspective of the disclosers, and for the development of social media literacy programs. We contribute generalized findings to the body of CB literature by revealing how closely CB incidents are related to the disclosure of personal information. This is the first study to employ the CPM theory [67] to understand the role of the different stakeholders in CB incidents, and the CI theory as a lens [59] to examine the means by which the information is disclosed.

2 RELATED WORK

We review related studies on (i) current strategies and mechanisms for CB prevention and support and (ii) the link between information disclosure and CB.

2.1 CB Prevention and Support

Several studies examine the characteristics and methods of CB prevention [48, 57, 94]. For instance, Ashktorab and Vitak [7] concentrated on designing CB-prevention solutions, through participatory design with teenagers. They reveal that attempts to prevent CB on OSNs failed mainly due to the complexity and nuance with which young people bully others online. More recently, Xiao et al. [91] conducted an interview study and designed activities where they asked adolescents about the need to address CB. They argue that designing and implementing approaches beyond content moderation can support victims in CB contexts. Another line of research explored technology design features to obtain a nuanced view of CB prevention. Lowry et al. [52] show that technology design could create a robust CB-prevention method for at-risk individuals. Other researchers recognize the current limitations in CB prevention research and propose avenues for future researchers. For example, Perren et al. [66] highlight a few limitations by reviewing successful CB prevention methods among students, parents, and school staff.

Prior studies also tested interventions to raise awareness about CB in order to address the risk associated with using social media platforms. For example, Wright et al. [90] examine how virtual

environments can be used to educate and to raise CB awareness among adolescents. Using virtual avatars that conceptualize and visualize the severity of CB in the school environment can educate parents and adolescents and can be a promising tool for preventing CB. Calvo-Morata et al. [17] experiment with validating a serious game named `Conectado` and show that the game could raise CB awareness among target users.

2.2 CB and Information Disclosure

Individuals directly supply their personal information online in exchange for service or information access [37]. For example, an online shopping platform collects names, e-mail addresses, and credit card information in return for providing users with a service. OSN platforms might even collect more sensitive information such as personal opinions and religious views [44, 82]. Such types of information consequently create a breeding ground for various undesirable online behaviors. For example, access to and disclosure of personal information can be gained in order to harm. The consequences for those who own the information can be drastic [36, 88].

Users constantly leave their footprints on the Internet with their posts, likes, blogs, followers, retweets, reposts, or comments [45]. However, private information can also flow directly from one user to another via direct personal interactions (e.g., direct messages or video chats) as social media enables users to participate in the process of identity representation and management [39]. Individuals' desires for self-reflection and self-representation gradually increase the disclosure of personal information, with users disclosing their personally identifiable details such as names, addresses, social-security numbers, etc. According to Benson et al. [12], social media activities bring about an unprecedented level of information disclosure. Users who partake in information disclosure have less control over their information [12], thus exposing themselves to CB. Nevertheless, users continue to disclose information in the quest to attract more followers and visibility [14]. Green et al. [31] examine how the LGBTQ+ community and straight allies have used video-mediated communication to discuss their experiences. They also demonstrate how non-anonymously they disclose personal information about themselves to strangers in order to find friendship, support, and empathy.

A survey conducted by Burke et al. [16] among 1200 Facebook users revealed that disclosing information such as personal messages is more satisfying to users than anything else. This behavior can increase social media gratification by attracting more likes. At worst, it can lead to misusing personal facts, potentially paving the way for CB. Aizenkot [2] surveyed more than 5000 secondary school students and found that the online disclosure of personal details is closely associated with CB victimization. They assessed the predictability of CB victimization from online self-disclosure and internet and OSN activity. Jain and Agrawal [42] showed that over-reliance on social-media security features propelled users to disclose their personal information online, thus exposing them to CB. Unsafe conversations are more likely to include sensitive personal information, such as images about others, that can be used to threaten them online [4].

Ashktorab et al. [6] surveyed youngsters on a `ASK.fm` platform and identified the forms of interaction and disclosure that occur on the platform by evaluating the motivation of users to interact and post on that platform. The special features of the platform, such as visibility and anonymity, triggered users to engage in various forms of online self-disclosure [6]. In another study, Peddinti et al. [64] studied the prevalence and behavior of identifiable vs. anonymous users on Twitter. The study found that anonymous users, compared with identifiable ones, are generally less inhibited from being active participants, as they tweet more and are more willing to expose their sensitive information. Several studies on disclosing personal information and CB focused on adolescents [2, 47, 55]. Kopecký [47] studied the incidence of CB among undergraduate students in Germany. They focused on risky online behavior among students and found that the most common type of CB is blackmail that originates after victims frequently expose personal information.

Most of our knowledge about CB is derived from studies that are focused on victims, which does not consider the multifaceted dimension of CB. We explore incidents of CB that are fueled by the disclosure of personal information, we investigate how information disclosure and interdependent privacy are connected with CB and what role disclosers and attackers play during the unfolding of CB events. Furthermore, social media users are exposed to CB differently, depending on their age and geographical location. In order to comprehensively understand the differences in information disclosure activities among social media users from two distinct populations, we collect and examine large-scale data from a wide range of samples

3 METHOD

We conducted a large-scale survey by using the critical incidents technique (CIT) [25], a technique for prompting respondents to recall incidents that might have occurred in the recent past and any additional details related to those incidents. Existing literature has used this method to investigate multi-party privacy conflicts on social media [20, 79], to understand the process of updating software by users [86], and to study CB policies [35]. Open-ended responses were collected to identify themes that further help us to understand how privacy-preserving strategies and practices and self-awareness can play an effective role in preventing the disclosure of private information and, subsequently, preventing CB.

The data was collected from the United States and Nigeria, through a survey questionnaire. We will refer to these two countries as US and NG, according to their ISO code. There are several reasons for choosing the US and NG. Prior research suggests that most existing research on CB focuses on Western, Educated, Industrialized, Rich, and Democratic (WEIRD) countries, and very little research has been conducted in (non-WEIRD) countries [33, 76], mostly African countries. To bridge this gap, we focus on the most populated and arguably most diverse non-WEIRD African country, NG, alongside a quite diverse WEIRD country, US. The socio-economic and cultural differences could also translate to differences in CB experiences and behaviors, especially when the following parameters are considered: gross domestic product (GDP) per capita (US: $70,200, NG: $2,064), Internet penetration (US: 91.8%, NG: 51.0%), mobile penetration (US: 81.6%, NG: 38.1%), unemployment (US: 3.4%, NG: 33%), and criminality rates (US: 41.1%, NG: 65.5%) [60, 61, 77, 78]. Figure 2 depicts the overall procedure of data collection (see details

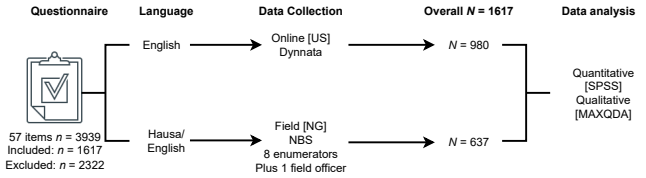

Figure 2: Data collection and analysis pipeline.

at Figure 6 in Appendix). A total number of $N = 3939$ questionnaires were distributed, out of which $n = 3022$ were distributed in the US and $n = 917$ were distributed in NG. The population is not limited to adolescents or students but is extended to include internet users in the 13-60 years age range and across all genders.

3.1 Ethics

Before deploying the survey, an ethics approval was requested and approved by the institutional review board (IRB) of the University of Lausanne. Due to the sensitive nature of our study, respondents' confidentiality and privacy were taken very seriously, by following the recommendations of Badillo-Urquiola et al. [10] on giving autonomy and respect to respondents, prioritizing data protection and privacy, and providing helpful resources. For respondents aged 13-17, the data was collected through their parents. This was ensured by the contracted vendor in the US (Dynata). In addition to our IRB, Dynata's ethical board approved the study questionnaire. Dynata used age verification and committed to enforcing the disclaimer. In NG, the school administration approved the ethics of the study for data collection, and the enumerators enforced the disclaimer. To ensure confidentiality and anonymity, we verified that no personally identifiable information was collected (e.g., name, home and e-mail address) to keep the risk of re-identification minimal and acceptable. The consent section of the survey provided helpful resources such as helplines on risk mitigation and psychological support for the respondents. The respondents were informed that participation was optional and that they could withdraw their participation at any time during the data-collection phase. During the recruitment, potential respondents were informed upfront that their responses would be used to improve the safety of social media users. This information helped them, especially those with sensitive accounts of CB (e.g., attackers), to understand the potential impact of their participation; this possibly motivated them—beyond the monetary incentives—to participate. Also, we established a safe and trusted environment for them to answer the questions by stressing that their responses will be treated confidentially.

3.2 Design of the Survey Questionnaire

The questionnaire consisted of 57 items in three sections. The full questionnaire is available as Supplementary Material 1.[4] At the beginning of the questionnaire, we presented the definition of the three stakeholders of CB incidents (i.e., victims, disclosers, and attackers) [83]. The rest of the questionnaire comprises three sections.

Section A. Demographics: We collected demographic information, such as age, gender, and region, and we measured the frequency of online platform use and screen time. We also assessed users' general privacy concerns by using Global Information Privacy Concern (GIPC) with a seven-point Likert scale [54].

Section B. Role Identification: To identify the respondent's role in the CB incident, we asked three differentiating questions on how many times the individual was involved in CB, either as a victim, discloser, or attacker (*never*, *once*, *twice*, *three times*, or *more than three times*). Note that we defined CB broadly and did not specifically target *doxxing*,[5] a closely related phenomenon to CB; yet, a few respondents mentioned it in their open-ended responses. Respondents who indicated they never had any CB experiences were excluded from the study. Respondents were directed to the different sections of the survey, based on the frequency of their experiences. For example, respondents with more victimization than disclosing or attacking experiences were directed to the victimization part of the questionnaire. Respondents who were equally victims, disclosers, or attackers were randomly assigned to sections.

Respondents under 13 and above 60 were excluded from participating in the questionnaire. For respondents between 13-17, an instruction was provided indicating that *"you are required to answer the rest of this survey together with one of your parents."* To collect more data from youth respondents, we asked the respondents between 18-60 if they had any children and if one of those children had ever been involved in CB. For the respondents with children involved in a CB incident, we asked them to fill out the survey, together with their children, and to answer the questions from their child's perspective. If the adult respondents reported having no children or if their children had no CB experiences, then we asked the adult respondents to answer the survey from their own perspectives.

Section C. Incident Report: In this section, we focus on CB incidents. The beginning of the section contained similar questions for the three different stakeholders, including the platform where the incident occurred, content type, when, where, and how the incident occurred. This section was followed by a few differentiating questions for each stakeholder. For example, we asked victims *"What is your relationship with the person who posted harmful content about you?"*; we asked disclosers *"Who was the person you disclosed his/her information to?"*; and we asked attackers *"Who was the person you offended online?"*. To determine the appropriate flow of information between stakeholders (RQ1 and RQ2), we ended the section with several questions about information transmission principles such as consent.

Respondents were also asked to describe the incident in a few sentences, paying attention to the following points: (a) What triggered the incident, and why was the victim a target for such an attack? (b) What happened? (c) Why was the victim unhappy with the attack? (d) What were the consequences? We asked this open-ended question to collect information that might not have been captured in the previous questions. The open-ended questions also help us to focus on the strategies for addressing CB (RQ3), i.e., methods of prevention, support strategies, and levels of self-awareness.

[4] All supplementary materials are available in the open science framework (OSF) at DOI: 10.17605/osf.io/9xtpc

[5] Doxxing is the act of publicly publishing personal or private information about an individual on the Internet, often with malicious intent [24].

The design of several questions was informed by the contextual integrity (CI) theory [59]. CI is particularly important for modeling and understanding privacy-related situations and information transmission among users. According to CI, information should be transmitted in accordance with specific norms such as (a) norms of appropriateness and (b) norms of distribution. Whereas the first set of norms determines whether a specific type of personal information is appropriate to disclose within a given context, the second set of norms limits the flow of information within and across contexts. A violation of privacy occurs when either of these norms is violated. To measure violations of CI transmission principles, we asked the three different stakeholders, using seven-point Likert scale questions from *strongly disagree* to *strongly agree*, about the ways information was transmitted in the context of the CB incident. For example, for victims, *"I gave my consent to the person who I disclosed my personal information to further share this information"*, and for disclosers, *"I asked for permission before sharing the victim's personal information with others"*.

3.3 Recruitment

For the US sample, respondents were recruited through `Dynata`, a vendor [6] that gives online access to a panel of adults and underage respondents. The sample is representative of the US Internet users [84]. In NG, before conducting the survey, we reached out to the representative office of the National Bureau of Statistics (NBS)[7] in both the North and South; but only the Northern office agreed to collaborate. We decided to collect data from higher education institutions (HEI) as they have students and staff from all parts of NG and they provide access to underage respondents through the HEI staff's school (i.e., elementary school under HEI's management). The respondents were recruited from nine HEIs through the NBS. The questionnaire was translated into the local language (i.e., Hausa), with English as an option. The respondents represent the different age groups of the NG population data.

Before the data collection in NG, we held a series of Zoom meetings with the NBS administrative assistants to present and explain the technicalities of the questionnaire. The NBS was not able to contact the respondents via e-mail so they opted for distributing the questionnaire through enumerators. They assigned eight enumerators to liaise with the institutions for requesting and obtaining their approval for data collection. Before the data collection, the NBS conducted a training workshop to train the enumerators on the technicalities of the survey questionnaire on Qualtrics such as the skip logic and inclusion/exclusion criteria for study participation.

The decision to go for a large sample (i.e., 3939 respondents) is consistent with similar studies on CB. For example, Yoo [93] collected data from 4000 respondents, Priebe [71] from 3432 respondents, and Bai [11] 3322 respondents.

3.4 Procedure

The questionnaire was pre-tested to determine the average completion time and to identify potential problems regarding measurement items and the wording of questions. Our colleagues participated in the cognitive pre-test, during which several areas that needed improvement were identified for adjustments. For example, one of the volunteers suggested that the researchers should consider multiple platforms instead of asking respondents for the name of a single platform on which the incident occurred. The second pre-test was conducted via a soft launch with 338 US respondents (21%) of the total response. These initial responses were used to fine-tune the questionnaire. In NG, respondents received ~2.4 USD (1000 Naira) for completing the survey. There is no average salary per hour in NG, hence the respondents were paid based on recommended quotations by the NBS as compared to those of the US. We paid ~3.1 USD to respondents who completed the survey in the US, (compared to ~7.2 USD average wage per hour) [72]. Note that money might not have been the only incentive for respondents to disclose information about incidents as sensitive as CB. For instance, respondents could have seen an opportunity to reflect on their actions in the context of a significant event in their personal life, especially for attackers and disclosers (e.g., to make amends for the harm they caused). Also, helping research in a trusted context (as stressed by the enumerators in NG) could have encouraged some respondents.

3.5 Data Reliability and Coding Process

To ensure the validity and consistency of the research findings, we followed a well-established rigorous quality assurance (QA) process [3]. In line with the Goldammer et al. [29]'s recommendations for using specific indirect indices, such as intra-individual response variability, long-string (i.e., straight-liners), response time, and response consistency in separating careless from careful respondents, we examined the data to identify any outliers or inconsistencies that could indicate the respondents' carelessness or inattention while answering the survey. Respondents that were deemed careless were removed from the data set. We identified the careless respondents by looking at open-ended comments (e.g., *"er g ergrer"*) and by looking at deviations from the expected response time (i.e., completing the survey under 3 minutes). In NG, we took further steps to ensure the quality of survey administration and management. In particular, to identify and correct any errors or ambiguities before the actual data collection, we ensured that enumerators understood the correct methodology and purpose of the study and pre-tested the survey items via Zoom. Moreover, we provided clear instructions on how to ask questions and record answers, and we developed and implemented quality control procedures, such as supervising enumerators and conducting random spot-checks. We addressed any issue identified during quality control and made necessary corrections to improve the quality of the collected data. As a result, 62 careless respondents were removed from the data set.

To analyze the open-ended responses, we carried out a collaborative coding process at the initial stage of the coding process [73]. In the first step, the codebook was developed by the lead coder and handed to the second coder. The second coder independently coded the data by using the codebook. Respondents in NG had Hausa and English options to complete the survey. Only a few respondents completed it in Hausa. The first coder and author of the paper is fluent in Hausa and translated the Hausa responses into English. The second coder had tools at his disposal to translate comments in Hausa.

[6] See https://www.dynata.com, Last access: June 2023.
[7] See https://nigerianstat.gov.ng, Last access: June 2023.

We measured the level of agreement between the codes assigned by the different coders and calculated the agreement score following B. Miles et al. [9]. We achieved a high degree of agreement between coders with values of at least 0.7 for all questions and over 0.8 for some questions. In the second step, the two coders worked collaboratively to clarify the code book inconsistencies and to resolve disagreements [53]. We inductively identified several themes, as outlined by Braun and Clarke [15]. These themes provided answers to (RQ3). When reporting the findings, to differentiate different stakeholders, we used the following identifiers: V for victims, D for disclosers, and A for attackers.

3.6 General Statistics of the Respondents

Among 3939 respondents, 2322 indicated that they had no CB experiences and were, therefore, excluded from further analysis. Among the $N = 1617$ respondents who experienced CB, $n_{US} = 980$ were from the US and $n_{NG} = 637$ from NG. **This shows that 32.4% of the respondents (or their children) from the US and 69.5% of the NG respondents were involved in CB incidents. The high incident rate in NG might have been due to the low level of awareness and few education initiatives about CB.** We removed 62 speeders and straight-liners from our data set, and the answers from the remaining 1555 (96.2%) were reported in the paper (see Figure 6). There are no speeders from NG because the enumerators recorded the data. According to their answers, 60.9% of the respondents were assigned the questions for victims ($N_{US} = 580$, $N_{NG} = 367$), 22.6% to the questions for disclosers ($N_{US} = 191$, $N_{NG} = 161$), and 16.4% to the questions for attackers ($N_{US} = 147$, $N_{NG} = 109$).

In the US sample, 409 respondents were male (44.5%), 478 were female (52%), and 16 respondents (1%) were non-binary or preferred to not respond. In terms of age, 13% of the respondents were between 13-17 years old, 33.3% were 18-34, 18.7% were 35-44, and 33% were 45-60. Parents who completed the survey on behalf of their children with age lower that 13 years old were excluded (2%). We collected responses from the four major regions of the US: Southern 40%, Northeast 22%, Midwest 21%, and the West 17%.

In the NG sample, 353 were male (55.4%), 284 were female (44.6%), and none were classified as others. In terms of age, 24.9% of the respondents were between 13-17 years old, 42.6% were 18-34, 19.4% were 35-44, and 9.2% were 45-60. Most respondents between 45-60 were personnel at the institutions where data was collected. Parents who completed the survey on behalf of their children, and indicated the age of their children as below 13 years were excluded (3.9%).

Our sample is comparable to the US and NG census quotas in terms of age and gender,[8] but it differs in terms of geographic distribution. A total of 199 respondents ($N_{US} = 113$, $N_{NG} = 86$) filled the survey on behalf of their children representing 12.7% of the sample. We instructed this group of respondents to indicate the age and gender of their children, instead of their own age and gender.[9] The GIPC scores (i.e., privacy concerns) of the respondents were as follows : US: $\mu = 4.7$, $\sigma = 1.8$, $\alpha = -0$, $\beta = 6$, NG: $\mu = 7.0$, $\sigma = 2.3$, $\alpha = -0$, $\beta = 6$), with μ as the mean score, σ the standard deviation, and α and β the bounds [54]. **Respondents from NG are more concerned about their information privacy than those from the US**. The high level of privacy concerns among respondents from NG could be due to the lack of data protection laws.

[8]See Computer and Internet access in the United States 2012, US Census https://www.census.gov/.../computer-use-2012.html and Nigerian census data https://www.populationpyramid.net/nigeria/2019/, Last access: June 2023.

[9]Gender and age information for $N = 15$ respondents is missing. In the soft launch version, respondents who reported on behalf of their children indicated their own age rather than that of their children. We later corrected this error in the full launch.

4 FINDINGS

We provide a highlight of the results. We found that CB is often fueled by personal information that becomes known to attackers, directly or through social media (45%, overall). CB is about more than just attackers and victims; there are other stakeholders that are significantly involved. Almost one in every four respondents said they participated in CB as a discloser (22.6%, overall). CB is not only likely to come from strangers, but it is also likely to be carried out by friends (20%, overall). We identify several key differences between the US and NG, In particular, the following:

- The incident rate of CB is twice as low in the US compared to the NG (32% vs. 69%).
- The rate of victimization by photos, memes, or screenshots is twice as low in the US as in NG (8% vs. 18%).
- The rate of consensual information disclosure conducive to CB incidents is slightly higher in the US than in NG (49% vs. 37%).
- The rate of attackers that leverage personal information to bully victims is twice as low in the US as in NG (33%. vs. 63%).

Now, we delve deeper into the detailed findings. Note that the CB incidents reported by respondents were slightly higher in the 1–3 months prior to the time of the survey (victims: US: 26.8%, NG: 17.2%; attackers: US: 24.0%, NG: 22.0%).

4.1 Characteristics of CB Incidents—RQ1

We first summarize the characteristics of the CB incident, *disregarding the relevance to information disclosure.*

Platforms. Table 1 shows the platform where the incident occurred. Most victims and attackers stated that the incident occurred on OSNs or on instant messaging apps. According to a victim from NG: *"Facebook, Then it turns to Tik Tok. I understand the misunderstanding, she was seriously mean even after she realized it's not my fault."* [V128NG]. **Interestingly, the occurrence of CB in instant messaging apps was almost four times lower in the US compared to NG (US: 8.3%, NG: 28.1%). There were relatively more incidents in online gaming platforms, online forums, and via e-mail in the US than in NG.**

With regards to the actual service where the incident occurred (see Table 2), the majority of the *victims* indicated that it was on *Facebook* (US: 52.2%, NG: 48.2%). The high incident rate on Facebook could be associated with the limited control of public posts or comments, as well as with the anonymity that makes it easier for users to hide their identity and commit CB. Instagram was the second most often reported platform (US: 8.5%, NG: 7.9%). Interestingly, we found a huge difference on the reports about WhatsApp. Although only 1% indicated being victims in the US, this number rose to 33% for NG. The efficient features of WhatsApp, such as voice notes and the ability to use it with little Internet connection, make it popular among users in NG. Our results on some 'platforms' (e.g., Facebook and WhatsApp) are confirmed by the data on the penetration of

Table 1: Where it happened.

Items	Victims		Disclosers		Attackers	
Where did the incident happen?	**US**	**NG**	**US**	**NG**	**US**	**NG**
Online social networks	395 (68.1%)	210 (57.2%)	81 (42.4%)	86 (53.4%)	74 (50.3%)	65 (59.6%)
Instant messaging apps	47 (8.1%)	123 (33.5%)	23 (12.0%)	64 (39.8%)	13 (8.8%)	40 (36.7%)
Online gaming platforms	26 (4.5%)	4 (1.1%)	18 (9.4%)	0 (0.0%)	11 (7.5%)	0 (0.0%)
Blogging websites	10 (1.7%)	1 (0.3%)	5 (2.6%)	0 (0.0%)	9 (6.1%)	1 (0.9%)
E-mail services	23 (4.0%)	8 (2.2%)	15 (7.9%)	2 (1.2%)	12 (8.2%)	0 (0.0%)
Video sharing platforms	24 (4.1%)	6 (1.6%)	10 (5.2%)	5 (3.1%)	3 (2.0%)	1 (0.9%)
Online forums or chat rooms	18 (3.1%)	0 (0.0%)	15 (7.9%)	0 (0.0%)	12 (8.2%)	1 (0.9%)
Others	15 (2.6%)	10 (2.7%)	6 (3.1%)	0 (0.0%)	7 (4.8%)	0 (0.0%)
I don't remember	20 (3.4%)	5 (1.4%)	16 (8.4%)	4 (2.5%)	5 (3.4%)	1 (0.9%)

Table 2: Type of specific platform.

Items	Victims		Disclosers		Attackers	
Platforms	**US**	**NG**	**US**	**NG**	**US**	NG
Facebook	302 (52.2%)	177 (48.2%)	64 (33.5%)	71 (44.1%)	59 (40.4%)	58 (53.3%)
WhatsApp	6 (1.0%)	121 (33.0%)	2 (1.1%)	63 (39.1%)	1 (0.7%)	37 (33.9%)
Instagram	49 (8.5%)	29 (7.9%)	25 (13.2%)	13 (8.1%)	9 (6.2%)	6 (5.5%)
Snapchat	29 (5.0%)	1 (0.3%)	10 (5.3%)	0 (0.0%)	7 (4.8%)	0 (0.0%)
Twitter	29 (5.0%)	2 (0.8%)	10 (5.3%)	1 (0.6%)	10 (6.8%)	0 (0.0%)
TikTok	24 (4.2%)	9 (2.5%)	7 (3.7%)	6 (3.7%)	11 (7.5%)	6 (5.5%)
YouTube	24 (4.2%)	4 (1.1%)	13 (6.9%)	4 (2.5%)	5 (3.4%)	1 (0.9%)
Messagger	23(4.0%)	5 (1.4%)	15 (7.9%)	1 (0.6%)	8 (5.5%)	1 (0.9%)
Gmail	13 (2.2%)	4 (1.1%)	9 (4.8%)	1 (0.6%)	7 (4.8%)	0 (0.0%)
Xbox cloud gaming	6 (1.0%)	1 (0.3%)	1 (0.5%)	0 (0.0%)	2 (1.4%)	0 (0.0%)
Others	51 (12.5%)	13 (3.6%)	32 (17.5%)	1 (0.6%)	27 (18.6%)	0 (0.0%)

instant-messaging platforms in the two countries. According to Jacob [41], after WhatsApp, Facebook is the most used social media platform in NG. In US, WhatsApp is one of the least popular platforms [8]. **CB incidents on Snapchat, Twitter, YouTube, and Messenger were low overall, but more common in the US than NG (e.g., Twitter; US: 5%, NG: 0.8%).** Similarly, the majority of *attackers* reported Facebook as the most common place to begin an attack (US: 40.4%, NG: 53.3%). TikTok was platform the second most commonly reported by attackers (US: 7.5%, NG: 5.5%).

Types of Media. Table 3 shows the different media chosen to begin the attack. The majority of the *victims* reported being victimized by an online post. Comments under a post, direct messages, and instant messages either private or shared in groups were the other frequent means of attack. The most common reported attack (by *attackers*) was through online posts, direct (private) messages, and finally, instant messages shared in a group. **These findings show that most CB incidents stem from posting on OSNs.** Table 4 shows the form of harmful content used for CB. Most of the *victims* reported that a *text*, either through a message or comments under a post, was the most common means of CB. The second form most mentioned was via *image*, either photos, memes, or screenshots. The number of reported incidents involving photos was twice as low in the US compared to the NG. Videos and audios were low overall, but slightly higher in NG. **These findings show that information that causes CB is usually conveyed either through text, directly to the victims or publicly in a comment section.**

Types of Bullying. Table 5 shows the different strategies of bullying used by attackers. Some of the *victims* stated that they were offended through messages/images that contain hurtful comments, as indicated in a response made by a victim from the US: *"A former coworker posted mocking comments on a Facebook post that I had made."* [V011US]. This was followed by unkind comments/rumors about them or images of them publicly posted online, and messages/images that contain threats of spreading gossip or damaging the victim's reputation. For example, *"I broke up with my girlfriend and her family started posting stuff about me. It made me uncomfortable because I am a private person."* [V004US], or *"When I and my ex broke up, she ended up having a bunch of our old mutual friends over to drink. I guess after they all had some drinks they decided to start sending me messages on Snapchat talking crazy to me."* [V385US]. One of the victims stated that: *"my neighbor shares in a WhatsApp group rumors that I am not contributing to sanitizing the community."* [V151NG]. Threatening messages were more often reported by victims in NG such as the threats of damaging a person's reputation and of causing physical harm. For example, *"Cyberbullying can cause anything ranging from misunderstanding to physical fight between people."* [V164US].

On the *attacker* side, they stated that it was messages/images that contain threats of spreading gossip, damaging a person's reputation, or of instigating an argument. Messages/images that, by containing hurtful comments, unkind comments/rumors, or images about victims, are publicly posted online and that are circulated personal/private information online were also mentioned by some attackers. For instance, *"I uploaded a picture of my girlfriend on a social networking site and the picture was porn. She was not happy about it and there were problems."* [A026US]. **These findings show that information containing hurtful content was often times**

Table 3: How it happened.

Items	Victims		Disclosers		Attackers	
How did they incident happened?	**US**	**NG**	**US**	**NG**	**US**	**NG**
Via a direct message (private)	69 (11.9%)	53 (14.4%)	46 (24.3%)	26 (16.1%)	22 (15.1%)	22 (22.9%)
Via an instant message (private)	42 (7.3%)	38 (10.4%)	22 (11.6%)	11 (6.8%)	14 (9.6%)	15 (13.8%)
Via an instant message (shared in a group)	49 (8.3%)	103 (28.1%)	29 (15.3%)	60 (37.3%)	13 (8.9%)	21 (19.3%)
Via e-mail (private)	14 (2.4%)	4 (1.1%)	9 (4.8%)	1 (0.6%)	11 (7.5%)	0 (0.0%)
Via e-mail (CC'ed to other people)	14 (2.4%)	0 (0.0%)	10 (5.3%)	1 (0.6%)	4 (2.7%)	0 (0.0%)
Via online posts (seen by other people)	252 (43.6%)	86 (23.4%)	32 (16.9%)	39 (24.2%)	47 (32.2%)	28 (25.7%)
Via online stories (seen by other people)	26 (4.5%)	22 (6%)	3 (1.6%)	7 (4.3%)	4 (2.7%)	11 (10.1%)
Via commenting under a post (seen by other people)	72 (12.5%)	32 (9%)	13 (6.9%)	10 (6.2%)	14 (9.6%)	8 (7.3%)
Via commenting in a chatroom (e.g., in gaming platforms) (seen by others)	4 (0.7%)	8 (2.2%)	6 (3.2%)	2 (1.2%)	3 (2.1%)	0 (0.0%)
Others	7 (1.2%)	9 (2.5%)	6 (3.2%)	0 (0.0%)	1 (0.7%)	0 (0.0%)
I don't remember	29 (5%)	11 (3%)	13 (6.9%)	2 (2.5%)	13 (8.9%)	1 (0.9%)

Table 4: Media types.

Items	Victims		Disclosers	
Media	**US**	**NG**	**US**	**NG**
Text - a message	125 (21.6%)	78 (21.3%)	54 (28.6%)	25 (15.5%)
Text - a comment	276 (47.2%)	136 (37.1%)	50 (26.5%)	75 (46.6%)
Audio - a recorded voice or speech	9 (1.6%)	6 (1.6%)	9 (4.8%)	2 (1.2%)
Audio - an insulting song or music	6 (1.0%)	5 (1.4%)	6 (3.2%)	1 (0.6%)
Audio - an argument in social audio apps	4 (0.7%)	5 (1.4%)	6 (3.2%)	1 (0.6%)
Image - a taken photo	48 (8.3%)	68 (18.5%)	15 (7.9%)	33 (20.5%)
Image - a meme or an edited (photoshopped) photo	20 (3.5%)	12 (3.3%)	16 (8.5%)	4 (2.5%)
Image - a screenshot from message or post	13 (2.2%)	9 (2.5%)	7 (3.7%)	8 (5.0%)
Video - a recorded footage	24 (4.2%)	18 (4.9%)	7 (3.7%)	5 (3.1%)
Video - an edited video such as filtered or DeepFake video	5 (0.9%)	5 (1.4%)	3 (1.6%)	5 (3.1%)
Others	29 (5.0%)	20 (5.4%)	4 (2.1%)	1 (0.6%)
I don't remember	19 (3.3%)	5 (1.4%)	12 (6.3%)	1 (0.6%)

used for CB incidents. This could be due to the social-media disinhibition effect where users feel comfortable expressing themselves in a more offensive way than they do in face-to-face interactions [80]. Particularly in NG, where some users might not be familiar with online netiquettes (i.e., the rules and conventions of polite and respectful online behavior) compared to the US. A previous study found that users from NG make privacy-disclosure decisions based on benefits rather than on-sensitivity [92]. Whereas, users from the US make privacy decisions based on sensitivity rather than on benefits [28].

Underlying Reasons. Analyzing the open-ended responses, we identified the reasons that ignited the CB. *Victims* reported **political arguments on OSNs** as the most common reason. A victim indicated being targeted by political opponents in order to tarnish his reputation: *"Political interests from my rivalry, to tarnish my image because I am into politics. Posted my picture upside down to show to the world that I am defeated."* [V079NG]. Such arguments were the main reasons reported by *attackers*. This was particularly true in NG, where cultural and political factors in expressing strong and hurtful comments can be seen as a sign of strength and assertiveness. The second most frequent reason was **relationships**. For example, *"She was my friend, and I am jealous of her love for her boyfriend because I liked him. She is cheating on him, while he is parading himself as her only lover. When I disclosed the incident, they broke up. Because I revealed her secrets. It ended their relationship."* [A058NG].

Victims also reported that **offline meetings or events** triggered the CB. One respondent stated that, *"It started because they were talking nonsense about me in real life that was jeopardizing my job and my friends and then they posted a picture of me that was inappropriate. I got extremely offended by it because I did not approve them to post it and it was very embarrassing and humiliating."* [V124US]. *Victims* also reported the **disclosure of health-related information** as a type of CB, especially among respondents with disabilities. For example, *"People say rude comments because of what I am going through. I have anxiety, [...] The consequences were me being upset and feeling unseen."* [V158US] **These findings show that disclosing health-related information about others can be harmful.**

Relationships. CB is not committed only by **strangers**, it can also be carried out by **friends** and **peers**. In the US, victims reported occurrences of CB more from ex-partners (9.2%) than from current partners (1.9%) (see Figure 3). However, in NG, the inverse trend was found (partners: 5.4%, ex-partners: 4.4%). **These findings show that CB incidents can be carried out not only by strangers but also by people in close relationships.** Also, though in US attacks were made more frequently by ex-partners, perhaps for revenge [56], in NG, attacks were more frequently made by current partners, perhaps to express possession of the other person. Almost half of the attackers in NG reported that their victims were their **friends** (see Figure 4). However, in the US the attackers mostly

Table 5: Type of bullying.

Items	Victims		Disclosers		Attackers	
Types of bullying	**US**	**NG**	**US**	**NG**	**US**	**NG**
Messages/images that contain unkind and hurtful comments about me	191 (22.2%)	71 (15.7%)	29 (11.2%)	31 (16.2%)	27 (12.1%)	19 (14.7%)
Unkind comments/rumours or images about me publicly posted online for others to see	157 (18.2%)	41 (9.1%)	22 (8.5%)	23 (12.0%)	24 (10.7%)	14 (10.9%)
Messages/images that contain threats to spread gossip or damage reputation	94 (10.9%)	77 (17.1%)	34 (13.1%)	135 (8.3%)	28 (12.5%)	24 (18.6%)
Personal/private information or image about me been circulated and publicly posted online for others to see	89 (10.3%)	45 (10.0%)	32 (12.4%)	26 (13.6%)	19 (8.5%)	16 (12.4%)
Messages/images that contain threats to physical harm	70 (8.1%)	61 (13.5%)	15 (5.8%)	14 (7.3%)	19 (8.5%)	8 (6.2%)
Messages that contain unwanted rude or sexual explicit comments or images	64 (7.4%)	29 (6.4%)	22 (8.5%)	7 (3.7%)	19 (8.5%)	9 (7.0%)
Involved in an unfriendly argument with someone from whom I received unkind messages/images	51 (5.9%)	37 (8.2%)	25 (9.7%)	14 (6.3%)	27 (12.1%)	11 (8.5%)
Rude comments from online gamers	47 (5.5%)	21 (4.7%)	20 (7.7%)	19 (9.9%)	10 (4.5%)	11 (8.5%)
Excluded from a social network group or Facebook friends list	35 (4.1%)	20 (4.2%)	17 (6.6%)	9 (4.7%)	16 (7.1%)	7 (5.4%)
I don't remember	32 (3.7%)	18 (4.0%)	26 (10.0%)	6 (3.1%)	19 (8.5%)	3 (2.3%)
Secretly used my identity to send unkind messages or images to others	32 (3.7%)	31 (7.1%)	17 (6.6%)	9 (4.7%)	16 (7.1%)	7 (5.4%)

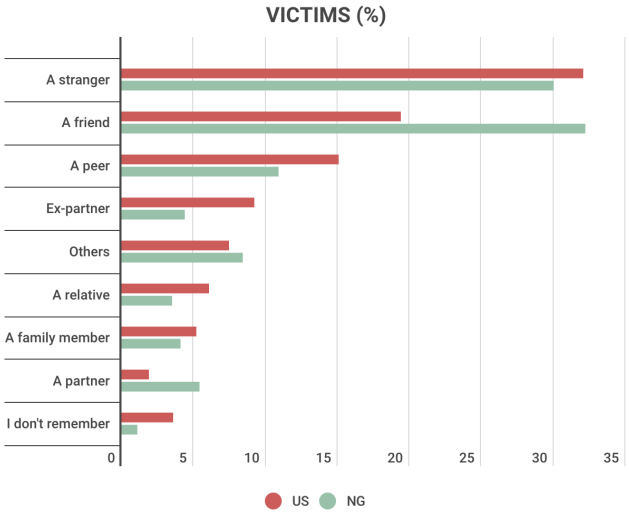

Figure 3: Relationship with attackers—reported by victims

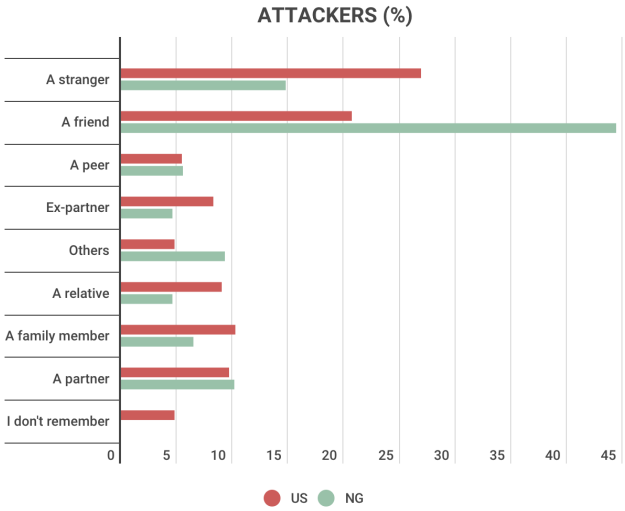

Figure 4: Relationship with victims—reported by attackers

mentioned **strangers**. **In summary, disclosing personal information to both friends and strangers puts users at risk of being victimized.** This finding is consistent with a study that found a significant relationship between sharing personal information with friends on social media and CB victimization [21].

Violation of Contextual Integrity. Contextual integrity is a framework that describes how information flow should conform to certain norms or expectations, depending on the specific contexts in which they occur [59]. In this study, a victim is an individual who has ownership and control over a particular piece of information, and an information discloser is someone who shares partial ownership or control over the same information. Whereas, an attacker is someone who is neither the victim nor the discloser of the information but who has access to it in some way. Contextual integrity violations occur when there is a mismatch between the norms or expectations surrounding the flow of information in a particular context and the actual behavior of the victim, discloser, or attacker.

We found that fewer victims in the US agreed to share their personal information (i.e., to give consent for sharing data) with others compared to in NG (US: 13.8%, NG: 33.7%). This shows that victims in the US are generally more aware that once information is released, it is difficult to control and can be used to create harm. Similarly, we found that disclosers in the US are more aware of seeking consent before sharing information compared to in NG (US: 51%, NG: 40%). However, attackers from both countries, but more so in NG, perceived information disclosure as a means to cyberbully victims, where the attackers violated the contextual integrity after acquiring the personal information of victims, either directly or through disclosers (US: 34%, NG: 63%). These findings show that victims (i.e., information owners) lose control of their personal information after sharing it with others, and there is a

need for interdependent privacy management on OSNs [59]. A similar view was suggested by prior research that examined the implementation of user information control on Facebook and ways to limit information sharing with other users and third-party apps (e.g., [87]); another relevant study investigated the user's view on the information collected by Facebook third-party applications (e.g., [81]).

4.2 Characteristics of CB Incidents Induced by Third-Party Information Disclosure - RQ2

Here, we first summarize the responses provided by *disclosers* about the platform, content type, type of bullying, and the relationship between the disclosers and the information owners (victims). Next, to better understand how information disclosure and CB are related, we focus on the specific questions posed to the three stakeholders.

Platforms Where Information Was Disclosed. As shown in Table 1, most *disclosers* reported OSNs, for example, Facebook, and instant messaging apps, such as WhatsApp. CB incidents on Messenger, Twitter, and e-mail were reported more in the US than in NG. A discloser stated that: *"I screenshot the WhatsApp status of one of my friends, I share it with my family members to see. Later she found out what I did to her from one of my friends, and in the end, it resulted in a fight between us."* [D013NG]. **These findings show that Facebook, Instagram, and WhatsApp are more convenient for information disclosure activities.**

Types of Media. Most *disclosers* stated that the disclosure of information occurred through private direct messages (DMs), instant messages (either private or shared in a group), and online posts seen by other people (see Table 3 for details). DMs were meant to ease private interactions between users, whereas disclosers misuse this for harmful purposes. Often, these DMs contain *text* messages, *comments*, *images*, *videos*, and *audio recordings*. A discloser stated that a well-meaning comment was misunderstood as CB by the recipient: *"A truthful comment that the person did not like."* [D057US]. (see Table 4 for details). Victims similarly reported these same types of media. **These findings show that direct private messages and posting on OSN feeds are more convenient for information disclosure activities.**

Types of Bullying. Concerning the types of CB that the disclosure of information causes, the majority of *disclosers* reported the following categories: messages/images that contain threats of spreading gossip or damaging a person's reputation, of circulating personal/private information or images about the victim by posting this information online, of posting messages/images that contain unkind and hurtful comments and unkind comments/rumors or images about the victims (see Table 5). A disclosers explained: *"The thing that triggered the incident was about smoking; my motivation was to help him stop smoking, but he was unhappy with it because I published it in a WhatsApp group."* [D021NG]. **This highlights the different CB behavior between disclosers and attackers. Attackers bully people mostly by spreading rumors, arguments, or gossip, whereas disclosers circulate personal information that can harm victims' reputations.**

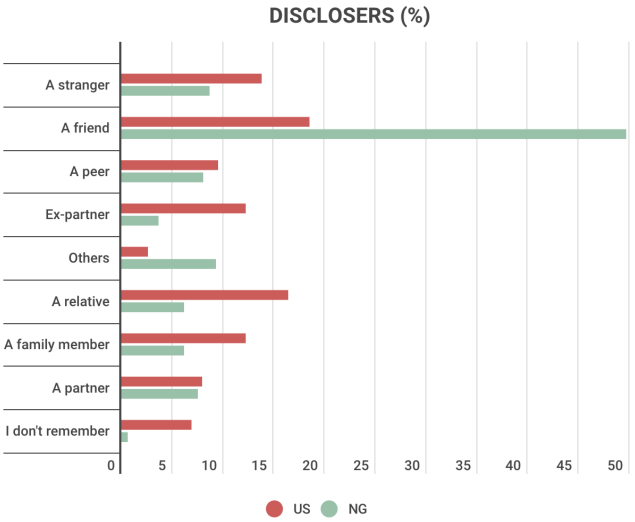

Figure 5: Relationship with victims—reported by disclosers.

Underlying Reasons. Using the open-ended answers, we identified the reasons for information disclosure. **Offline misunderstanding**, such as misunderstanding among classmates within a school environment, was reported by *disclosers* as being the most common reason, followed by **social and political arguments**. **Disclosure of health-related information** was reported by some *disclosers*. A respondent stated that he disclosed personal information about someone in the interest of public health: *"What triggered the incident was about society's health. What motivated me to do so was to protect lives."* [D116NG]. **These findings show that offline disagreement between users spills over into online interactions, where users amplify negative opinions about others.** This could be due to the ease with which people can reach peers online and the general difficulty of feeling empathy through online media.

Relationship between Disclosers and Victims. Most *disclosers* revealed that they disclosed information about a *friend*. *Relatives* and *strangers* were among other common answers (see Figure 5). Disclosing information about friends was much more common in NG compared with the US. Interestingly, where *victims* often reported 'strangers,' *disclosers* reported 'friends' as the third most common type of relationship. According to one discloser: *"I found the information because she gave it to me."* [D024NG]. **Our findings show that *victims*, most of the time, (over)trust their friends, and sometimes, even after CB incidents, they mistakenly think their friends would not share their private information with others.**

Types of Disclosed Information. We asked the *disclosers* what type of information they disclosed (see Table 6). The most frequent types of information disclosed by the *disclosers* from the US were information about *health status* and *childhood*. In NG, the most common types were information about *socioeconomic status* and *family affairs*. **The disclosure of personal information was much more common in NG compared to the US** (US: 2.1%, NG: 11.2%). This could be due to the lack, in NG, of programs that educate people on the proper use of the Internet.

Table 6: Type of personal information (a.k.a. "information type" in the CI theory) used in CB incident.

Items	Disclosers	
Types of personal information	**US**	**NG**
Information about my sexual orientation (e.g., gender identity)	16 (8.4%)	7 (4.3%)
Information about my family affairs (e.g., parenting issues)	25 (13.1%)	27 (16.8%)
Information about my socioeconomic status (e.g., income)	22 (11.5%)	30 (18.6%)
Information about my childhood memory (e.g., childhood trauma)	29 (15.2%)	10 (6.2%)
Information about my health status (e.g., medical background)	36 (18.8%)	9 (5.6%)
Information about my school records (e.g., grades)	18 (9.4%)	21 (13.0%)
Information about my personal care (e.g., beauty care)	27 (14.1%)	12 (7.5%)
Information about my personal identifiers (e.g., social security number)	4 (2.1%)	18 (11.2%)
Others	14 (7.3%)	27 (16.8%)

The Relation Between Information Disclosure and CB. We asked the *victims* whether the CB incident occurred due to the disclosure of personal information. Almost one-fourth of the *victims* stated that CB occurred after information disclosure. Although some were not sure, around half reported that the incident was not the result of disclosing personal information.

Victims who said the incidents occurred because of information disclosure stated that it was information related to personal identifiers such as name and e-mail address. Only 13.8% of *victims* in the US said they gave their consent to the discloser to share their information. However, in NG, almost one-third of the victims (33.7%) stated that they disagreed. Despite the high privacy concerns reported by victims in NG, respondents in NG reported that they disclose more information. There could be several underlying reasons for this, such as trust, social validation, self-expression, relational development, and/or entertainment [46]. Moreover, 32.1% of the *victims* from the US agreed that the person who offended them was the same person they had disclosed the information to (i.e., the attacker discloses personal information about the victim). In NG, more *victims* (42.5%) agreed that *disclosers* and *attackers* were the same person. **The results show that in most cases, the CB was induced with information disclosure, where another person, in most cases a friend, facilitated the CB by further disclosing the private information about the victims**.

Attackers were asked questions about the effects of information disclosure on their attack. Around one-third of the *attackers* from the US and half of those from NG agreed that before being involved in CB, they had gained access to the personal information of the victim, and that having this information helped them to post harmful content about the victims. For example, *"What triggered the incidents is about my boyfriend who broke my heart. What motivated me to do so is love. He was unhappy because this personal information was shared online."* [A067NG]. Finally, 40.8% of the attackers in the US and 67.7% of those in NG agreed that they had received the information directly from the victim and not a *discloser*. This confirms the previous findings. Furthermore, around half of the attackers agreed that they had posted harmful content intentionally to harm the victim or because they did not care about the victim. **This shows that the rate that attackers leverage personal information to bully victims is as twice as low in the US as in NG**.

4.3 Strategies to Address CB—RQ3

We asked the respondents about the strategies they would use to address CB. *Victims* mentioned they would **avoid disclosing sensible information with others in order to retain control over delicate aspects of their life and to avoid being exposed again.** A respondent said: *"I will not disclose my personal information to anyone again."* [V026NG]. They would also refrain from commenting on posts made by other people online to avoid being misunderstood and triggering their reactions: *"I would not have commented at all."* [V483US]. Finally, some reported that they would report the incident on social media platforms so that the platforms would take appropriate action, such as reviewing and removing the content: *"I try to report to the Facebook management to allow me to retrieve my hacked social media account and stop the hackers from sending money requests to my close friends."* [V043NG]. Using social media moderation features was acknowledged to be an effective practice to avoid being bullied again. Several victims reported blocking unfriendly people or toxic friends, deleting content that exposed them to CB, and even deleting their own accounts: *"If I see someone who is weird and they start messaging me, I delete their stuff and block them."* [V044US]. A few respondents indicated that they limit interaction with strangers by switching accounts to private: *"I try to close the account from my side to deactivate it and stop them from offending me."* [V043NG].

They also highlighted a range of mitigation strategies such as *talking to someone* after an incident occurred, *ignoring* the attackers, *reporting CB to authorities*, seeking support from peers or parents, or *counseling*: *"Yes, I did something to control the negative consequence of the incident: I reported it to the police so that they can trace him."* [V081NG]. They thought such strategies could help them to reduce the harm or at least avoid the risk of being bullied again. Other victims, who mentioned their social media accounts were hacked and used to post unwanted content, said they recovered their accounts and apologized to their followers for the unwanted content posted by the hackers. Similarly, most *disclosers* stated that to prevent CB, they would **stop sharing personal information about others** on social media platforms. This would reduce the risk of CB incidents and make online interactions safer. They also mentioned refraining from negatively commenting on other people's posts to prevent arguments that lead to unpleasant experiences. One of the respondents regretted the response made to a comment that resulted in a given incident: *"I shouldn't have said her name so that nobody could identify who the story belonged to."* [D020NG].

Disclosers also reported other strategies that could prevent them from disclosing personal information about others, such as *avoiding social or political arguments* on social media or *asking permission* from owners of the information. A respondent said: *"To avoid asking personal information about others or rather not to engage yourself in any political arguments."* [D049NG].

As effective strategies for preventing CB, *attackers* stated that they could have avoided making negative comments on other people's posts, using personal information about others, and expressing personal opinions that will hurt others: *"I could not have taken a screenshot of the group chat and posted it."* [A053NG]. **These comments show that disclosers and attackers are concerned about the incidents and their potential negative outcome on the victims**.

Self-Awareness of Disclosers and Attackers—RQ3

In many cases, CB puts a social stain on the victims and spreads inaccurate information about them. Examples from disclosers and attackers point to the disclosure of employment losses and **the misinterpretation that people in the community derived from these:** *"Damage their reputations and affect their personalities."* [A015NG]. Other disclosers and attackers mentioned **mental and physical illness**, as some of the consequences related to the disclosure of personal information. For example, *"They may experience negative emotions like frustration, anger, or depression and can also increase suicidal ideation."* [A077NG].

Disclosers were mostly concerned about **the implications of the incidents on the victim's family reputation** as indicated in this response: *"They may get gossip in society, and it may lead to damage of family reputation."* [D063NG]. A few *disclosers* mentioned emotional stress or mental illness as some of the consequences of the incidents. *Attackers* mentioned that sharing personal information about others can ruin their victim's reputation, as stated by a respondent: *"Loss of integrity and respect in the society."* [A013NG]. They also mentioned the effect of incidents on the mental well-being of victims. *"... mental effects such as stress, depression, and you may act violently."* [A076NG]. **These findings show that some disclosers and attackers are aware of some of the consequences that victims face as a result of CB. Nonetheless, they still engaged in the incidents; perhaps they believed they could get away with it.**

5 DISCUSSION

Using the communication privacy management theory [67] and the contextual integrity theory [59], this study investigates the extent to which CB is fueled by the disclosure of personal information.

The pattern of this phenomenon was further assessed based on the types of information disclosure involved in the CB incidents, the types of personal information used, the typical path used, and based on when, where, and how the incidents occurred. Contrary to previous research that indicates that OSN users typically exercise self-disclosure restraint due to the possibility of negative personal consequences [46], we find that users disclose information about others hence resulting in CB. Almost one out of every four respondents in our study said they had taken part in CB as a discloser. We found that CB is often fueled by personal information that becomes known to an attacker, directly or through third-party disclosers. Our findings are in line with prior research that, was carried out mainly on adolescents (e.g., [2]) found that young OSN users prioritize disclosing their personal information at the expense of their privacy in order to gain popularity or intimacy, and other benefits associated with social rewards [18, 27, 43, 65]. Unfortunately, such disclosures can be further used for CB.

Our study not only reveals the extent to which CB is fueled by information disclosure but also the *modalities* in which the information is disclosed. Users often disclose information either about themselves or others through direct messages (DMs), commenting under posts and through online stories, etc. These findings are in line with prior literature that revealed hurtful conversations that involve sensitive topics and behaviors, occur through direct messaging on Instagram [4]. We extend this literature by uncovering the types of information disclosed through these media, for example, photos, screenshots, and memes. The highly interactive nature of OSNs was found to stimulate the sharing of sensitive information [13]. Previous research has found texting, aggressive comments, and forum posts as the most commonly used method of CB [89]. Our study adds to this literature by identifying common types of information contained in posts, comments, and stories such as disclosing health status, childhood memories, marital affairs, etc. These findings bear implications for design solutions that reflect the perspective of disclosers and attackers. In particular, these solutions can use a range of participatory design techniques. For instance, Ashktorab and Vitak [7] conducted a study on designing CB prevention and CB mitigation strategies through a participatory design with teens. The study proposed a solution mainly for victims. However, our study reveals the need for soliciting design ideas on prevention and mitigation from other stakeholders (i.e., disclosers and attackers). Future work might look at how technical solutions for tracking path and preventing further disclosure can mitigate this incident.

Our findings also indicate that non-consensual information disclosure conducive to CB is high overall, but it is slightly higher in NG. In response to non-consensual information sharing, the US has used a variety of approaches, including victim responsibility, educational programs, and/or formal judicial intervention [5, 34, 74]. Apparently, these programs bring a positive effect on limiting non-consensual information sharing. These findings bear implications for a social-media literacy program, especially in NG where currently there is no program that educates citizens about proper online behavior such as consensual information sharing. Awareness programs have been established in US through digital citizenship with a specific focus on adolescents [68]. However, as CB exists within other age groups, there is a need to extend this program to adults. In NG, there is a need for teachers, school administrators, and community leaders to be informed about CB and the related mechanisms and for them to be equipped with mediatory and preventive skills, as they are responsible for offering guidance in this regard. Moreover, although there are currently no official federal laws regarding CB in the US, many states have enacted anti-cyberbullying legislation. According to the Cyberbullying Research Center,[10] all

[10]See Cyberbullying Research Center, Last access: June 2023.

50 states have a law specifically addressing bullying [22]. Educating offenders about this legislation could make them less likely to offend.

The existing laws and practices in the US, such as child pornography law [23] and other related programs, have been shown to have practical implications in taming CB, given the low incident rate compared to NG (32% vs. 69%, approx). None of these laws or practices exist in NG. As punitive measures are non-existent or unknown, the lack of specific laws and awareness programs can make offenders more likely to commit cyber-related crimes. These findings bear policy implications. The greater chance of addressing CB is to see it not only as a problem that affects individuals or schools but also as a community problem because it affects individuals and schools hence the community at large. To combat the problem of CB, there is a need for policy deliberation. For instance, in the US, legislation and legal action are becoming part of the landscape in addressing CB [19]. However, in NG, with a very high incident rate (69%), there is no legislation specific to CB. CB is punishable under the Cybercrime Prohibition and Prevention Act [1]. This law is not flexible for the different modalities identified in the study. Specific legislation should mandate schools to have a CB policy that will help identify risky online behavior and to discuss the possible formal and informal disciplinary or, preferably, restorative responses that can follow [75]. As CB also occurs outside schools in NG, the law should also empower law enforcement agencies, rehabilitation, and correctional centers with the mandate to address the problem of CB and to raise awareness about the prevention and consequences of CB.

5.1 Strategies for Mitigation

Our findings further reveal that individuals still rely on the conventional prevention and mitigation mechanisms provided by OSNs. Most respondents from the US have shown good familiarity with OSN moderation features such as blocking and reporting malicious users and deleting unwanted content. Unfortunately, respondents in NG were not familiar with these mechanisms. This reveals that the lack of knowledge about OSNs' mitigation features contributes to exacerbating the problem of CB.

Research has shown that the researchers from the United States have conducted the most CB research, with input from different stakeholders, including families, schools, and teachers; they are followed by researchers from Europe. In contrast, Africa and South America are still lagging behind [76]. Future research and design should propose better mitigation mechanisms for CB. Ideally, such mechanisms should involve the disclosers in the cycle of information circulation and should limit their ability to non-consensual data sharing. One potential direction is watermarking [49] personal information on OSN platforms in order to track and possibly block the relaying of personal information captured. Such a mechanism should mark personal information directly (e.g., photos) or indirectly (screenshots of conversations) by including the watermark in the UI of the app [32, 50, 70], thus enabling the tracking of the information and determining whether it has been disclosed with or without the consent of a data subject (i.e., the potential victim). If implemented successfully, this approach could make disclosers accountable and could reduce the risk of using personal information for CB. However, some challenges and considerations need to be taken into account before implementing such a system: Future research should indeed investigate the feasibility, usability, and adoption of such systems. Also, OSN platforms can ensure that users are fully informed about the watermarking system and are given the option to opt-out if they want to.

5.2 Limitations

Our study has limitations. First, respondents' answers depend on the category they placed themselves in at the beginning of the survey (i.e., victims, disclosers, or attackers). Participation depends on the willingness of respondents to admit they were involved in a particular incident. We addressed this by defining the role of the three different stakeholders at the beginning of the survey, as suggested by Tokunaga [83]. We also made sure respondents chose, based on their experience, the correct category of the survey. Second, our study restricts each stakeholder to only one role in the CB incidents, without considering the possibility of belonging to multiple roles in other CB incidents. For example, an individual can be a victim in one incident and an attacker in another. Future research could possibly examine what motivates users to commit CB after they themselves have been on the receiving end. Third, we collected data from respondents between 13-17 years, through their parents. Data collected through parents might not reflect every detail about the incident, but it helps gain perspective. Fourth, we examined only the relationship between victims and disclosers; future research should examine the relationship between disclosers and attackers. Fifth, our recruitment in Nigeria mainly focused on members of higher education institutions; these individuals are privileged to own a smartphone, have access to the Internet, and have active social media accounts. Our rationale for selecting this group of users is the fact that CB mostly occurs when users interact with these technologies. This has, however, exposed our study to a potential non-representative gap in terms of education, socio-economic status, and the availability of communication channels.

6 CONCLUSION

Social media have enabled us to interact with families and friends worldwide. It is also true that interaction on social media exposes individuals to various privacy threats. This study examines the extent to which the disclosure of personal information fuels cyberbullying and which strategies for prevention and awareness could be used. The findings from our study contribute to the body of cyberbullying literature by revealing insights regarding information-disclosure activities on social media and their connection to CB. We hope these results will contribute to raising awareness about the CB problem and building safer online environments.

ACKNOWLEDGMENTS

We thank Holly Cogliati, Vincent Vandersluis, and James Tyler for their great editing jobs and for providing insightful feedback. We thank our colleagues who took part in the cognitive pre-tests. This work was partially funded by the Swiss National Science Foundation with grant NCCR LIVES—'Overcoming Vulnerability: Life Course Perspectives'.

REFERENCES

[1] Adejoke O Adediran. 2021. Cyberbullying in Nigeria: Examining the adequacy of legal responses. *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique* 34, 4 (2021), 965–984.

[2] Dana Aizenkot. 2020. Social networking and online self-disclosure as predictors of cyberbullying victimization among children and youth. *Children and Youth Services Review* 119 (2020), 105695.

[3] Pertti Alasuutari, Leonard Bickman, and Julia Brannen. 2008. The SAGE handbook of social research methods. *Torrossa online digital bookstore* 15, 4 (2008), 336–355.

[4] Shiza Ali, Afsaneh Razi, Seunghyun Kim, Ashwaq Alsoubai, Joshua Gracie, Munmun De Choudhury, Pamela J Wisniewski, and Gianluca Stringhini. 2022. Understanding the Digital Lives of Youth: Analyzing Media Shared within Safe Versus Unsafe Private Conversations on Instagram. , 3–10 pages.

[5] Steven Angelides. 2013. 'Technology, hormones, and stupidity': The affective politics of teenage sexting. *Sexualities* 16, 5-6 (2013), 665–689.

[6] Zahra Ashktorab, Eben Haber, Jennifer Golbeck, and Jessica Vitak. 2017. Beyond cyberbullying: self-disclosure, harm and social support on ASKfm. , 3–10 pages.

[7] Zahra Ashktorab and Jessica Vitak. 2017. Designing cyberbullying mitigation and prevention solutions through participatory design with teenagers. , 3–10 pages.

[8] Brooke Auxier and Monica Anderson. 2021. Social media use in 2021. *Pew Research Center* 1 (2021), 1–4.

[9] Matthew B. Miles, A. Michael Huberman, and Johny Saldana. 2019. Qualitative Data Analysis: A Methods Sourcebook. https://us.sagepub.com/en-us/nam/qualitative-data-analysis/book246128.

[10] Karla Badillo-Urquiola, Zachary Shea, Zainab Agha, Irina Lediaeva, and Pamela Wisniewski. 2021. Conducting risky research with teens: co-designing for the ethical treatment and protection of adolescents. *CSCW* 4, CSCW3 (2021), 1–46.

[11] Qiyu Bai, Shan Huang, Fang-Hsuan Hsueh, and Taofu Zhang. 2021. Cyberbullying victimization and suicide ideation: A crumbled belief in a just world. *Computers in human behavior* 120 (2021), 106679.

[12] Vladlena Benson, George Saridakis, and Hemamaali Tennakoon. 2015. Information disclosure of social media users: does control over personal information, user awareness and security notices matter? *Information Technology & People* 15, 21 (2015), 3–7.

[13] Ina Blau. 2011. Application use, online relationship types, self-disclosure, and Internet abuse among children and youth: Implications for education and Internet safety programs. *Journal of Educational Computing Research* 45, 1 (2011), 95–116.

[14] Steven W Bradley, James A Roberts, and Preston W Bradley. 2019. Experimental evidence of observed social media status cues on perceived likability. *Psychology of Popular Media Culture* 8, 1 (2019), 41.

[15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[16] Moira Burke, Robert Kraut, and Cameron Marlow. 2011. Social capital on Facebook: Differentiating uses and users. *SIGCHI conference* 2, 1 (2011), 2–6.

[17] Antonio Calvo-Morata, Cristina Alonso-Fernández, Manuel Freire, Iván Martínez-Ortiz, and Baltasar Fernández-Manjón. 2021. Creating awareness on bullying and cyberbullying among young people: Validating the effectiveness and design of the serious game Conectado. *Telematics and Informatics* 60 (2021), 101568.

[18] José A Casas, Rosario Del Rey, and Rosario Ortega-Ruiz. 2013. Bullying and cyberbullying: Convergent and divergent predictor variables. *Computers in Human Behavior* 29, 3 (2013), 580–587.

[19] Wanda Cassidy, Chantal Faucher, and Margaret Jackson. 2013. Cyberbullying among youth: A comprehensive review of current international research and its implications and application to policy and practice. *School psychology international* 34, 6 (2013), 575–612.

[20] Mauro Cherubini, Kavous Salehzadeh Niksirat, Marc-Olivier Boldi, Henri Keopraseuth, Jose M Such, and Kévin Huguenin. 2021. When forcing collaboration is the most sensible choice: Desirability of precautionary and dissuasive mechanisms to manage multiparty privacy conflicts. *CSCW* 5, 1 (2021), 1–36.

[21] Kyung-Shick Choi, Sujung Cho, and Jin Ree Lee. 2019. Impacts of online risky behaviors and cybersecurity management on cyberbullying and traditional bullying victimization among Korean youth: Application of cyber-routine activities theory with latent class analysis. *Computers in Human Behavior* 100 (2019), 1–10.

[22] laws cyberbullying. 2016. Bullying laws and cyberbullying laws across america. https://cyberbullying.org/bullying-laws

[23] Alexa Dodge and Dale C Spencer. 2018. Online sexual violence, child pornography or something else entirely? Police responses to non-consensual intimate image sharing among youth. *Social & Legal Studies* 27, 5 (2018), 636–657.

[24] Stine Eckert and Jade Metzger-Riftkin. 2020. Doxxing. *The international encyclopedia of gender, media, and communication* 162, 10 (2020), 1–5.

[25] John C Flanagan. 1954. The critical incident technique. *Psychological bulletin* 51, 4 (1954), 327.

[26] Manuel Gámez-Guadix, Izaskun Orue, Peter K Smith, and Esther Calvete. 2013. Longitudinal and reciprocal relations of cyberbullying with depression, substance use, and problematic internet use among adolescents. *Journal of Adolescent Health* 53, 4 (2013), 446–452.

[27] Saswati Gangopadhyay and Debarati Dhar. 2014. Social Networking Sites and Privacy Issues Concerning Youths. *Global Media Journal: Indian Edition* 5, 1 (2014), 3–9.

[28] Reza Ghaiumy Anaraky, Kaileigh Angela Byrne, Pamela J Wisniewski, Xinru Page, and Bart Knijnenburg. 2020. To disclose or not to disclose: examining the privacy decision-making processes of older vs. younger adults. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* 1, 18 (2020), 3–9.

[29] Philippe Goldammer, Hubert Annen, Peter Lucas Stöckli, and Klaus Jonas. 2020. Careless responding in questionnaire measures: Detection, impact, and remedies. *The Leadership Quarterly* 31, 4 (2020), 101384.

[30] Joaquin González-Cabrera, Esther Calvete, Ana León-Mejía, Carlota Pérez-Sancho, and José M Peinado. 2017. Relationship between cyberbullying roles, cortisol secretion and psychological stress. *Computers in Human Behavior* 70 (2017), 153–160.

[31] Michael Green, Ania Bobrowicz, and Chee Siang Ang. 2015. The lesbian, gay, bisexual and transgender community online: discussions of bullying and self-disclosure in YouTube videos. *Behaviour & Information Technology* 34, 7 (2015), 704–712.

[32] Wei Gu, Ching-Chun Chang, Yu Bai, Yunyuan Fan, Liang Tao, and Li Li. 2023. Anti-Screenshot Watermarking Algorithm for Archival Image Based on Deep Learning Model. *Entropy* 25, 2 (2023), 3–8. https://doi.org/10.3390/e25020288

[33] Alaa HDaffalla, Lucy Simko, Tadayoshi Kohno, and Alexandru G Bardas. 2021. Defensive technology use by political activists during the Sudanese revolution. , 3–10 pages.

[34] Nicola Henry and Anastasia Powell. 2015. Embodied harms: Gender, shame, and technology-facilitated sexual violence. *Violence against women* 21, 6 (2015), 758–779.

[35] Sheila Finley Hilton. 2018. Selected South Carolina School Leaders' Experiences in Addressing Cyberbullying. , 538–546 pages.

[36] Sameer Hinduja and Justin W Patchin. 2010. Bullying, cyberbullying, and suicide. *Archives of suicide research* 14, 3 (2010), 206–221.

[37] Donna L Hoffman, Thomas P Novak, and Marcos Peralta. 1999. Building consumer trust online. *Commun. ACM* 42, 4 (1999), 80–85.

[38] Mathias Humbert, Benjamin Trubert, and Kévin Huguenin. 2019. A survey on interdependent privacy. *CHI Conference* 52, 6 (2019), 1–40.

[39] Netta Iivari, Leena Ventä-Olkkonen, Sumita Sharma, Tonja Molin-Juustila, and Essi Kinnunen. 2021. CHI Against Bullying: Taking Stock of the Past and Envisioning the Future. *CHI Conference* 2, 3 (2021), 1–17.

[40] Clements J. 2022. Cyberbullying statistics and facts. https://www.statista.com/topics/1809/cyber-bullying/

[41] Jacob Jacob. 2021. Uses and Abuses How increased Social Media usage threatens Nigeria's democracy. https://www.premiumtimesng.com/news/headlines/499276-uses-and-abuses-how-increased-social-media-usage-threatens-nigerias-democracy.html?tztc=1

[42] Shilpi Jain and Soni Agrawal. 2020. Perceived vulnerability of cyberbullying on social networking sites: effects of security measures, addiction and self-disclosure. *Indian Growth and Development Review* 65, 18 (2020), 3–9.

[43] Zhenhui Jiang, Cheng Suang Heng, and Ben CF Choi. 2013. Research note—privacy concerns and privacy-protective behavior in synchronous online social interactions. *Information Systems Research* 24, 3 (2013), 579–595.

[44] Adam N Joinson, Ulf-Dietrich Reips, Tom Buchanan, and Carina B Paine Schofield. 2010. Privacy, trust, and self-disclosure online. *Human–Computer Interaction* 25, 1 (2010), 1–24.

[45] Jan H Kietzmann, Kristopher Hermkens, Ian P McCarthy, and Bruno S Silvestre. 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons* 54, 3 (2011), 241–251.

[46] Mehrdad Koohikamali, Daniel A Peak, and Victor R Prybutok. 2017. Beyond self-disclosure: Disclosure of information about others in social network sites. *Computers in Human Behavior* 69 (2017), 29–42.

[47] Kamil Kopecký. 2014. Cyberbullying and other risks of internet communication focused on university students. *Procedia-Social and Behavioral Sciences* 112 (2014), 260–269.

[48] Robin M Kowalski, Gary W Giumetti, Amber N Schroeder, and Micah R Lattanner. 2014. Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. *Psychological bulletin* 140, 4 (2014), 1073.

[49] Sanjay Kumar, Binod Kumar Singh, and Mohit Yadav. 2020. A Recent Survey on Multimedia and Database Watermarking. *Multimedia Tools Appl.* 79, 27–28 (jul 2020), 20149–20197. https://doi.org/10.1007/s11042-020-08881-y

[50] Li Li, Rui Bai, Shanqing Zhang, Chin-Chen Chang, and Mengtao Shi. 2021. Screen-Shooting Resilient Watermarking Scheme via Learned Invariant Keypoints and QT. *Sensors* 21, 19 (Sept. 2021), 6554. https://doi.org/10.3390/s21196554

[51] Brett J Litwiller and Amy M Brausch. 2013. Cyber bullying and physical bullying in adolescent suicide: the role of violent behavior and substance use. *Journal of youth and adolescence* 42, 5 (2013), 675–684.

[52] Paul Benjamin Lowry, Gregory D Moody, and Sutirtha Chatterjee. 2017. Using IT design to prevent cyberbullying. *Journal of management information systems* 34, 3 (2017), 863–901.

[53] Kathleen M MacQueen, Eleanor McLellan, Kelly Kay, and Bobby Milstein. 1998. Codebook development for team-based qualitative analysis. *Cam Journal* 10, 2 (1998), 31–36.

[54] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[55] Alice E Marwick and Danah Boyd. 2014. Networked privacy: How teenagers negotiate context in social media. *New media & society* 16, 7 (2014), 1051–1067.

[56] Clare McGlynn, Erika Rackley, and Ruth Houghton. 2017. Beyond 'revenge porn': The continuum of image-based sexual abuse. *Feminist legal studies* 25 (2017), 25–46.

[57] Ersilia Menesini and Christina Salmivalli. 2017. Bullying in schools: the state of knowledge and effective interventions. *Psychology, health & medicine* 22, sup1 (2017), 240–253.

[58] Miljana Mladenović, Vera Ošmjanski, and Staša Vujičić Stanković. 2021. Cyber-aggression, cyberbullying, and cyber-grooming: a survey and research challenges. *CHI Conference* 54, 1 (2021), 1–42.

[59] Helen Nissenbaum. 2019. Contextual integrity up and down the data food chain. *Theoretical Inquiries in Law* 20, 1 (2019), 221–256.

[60] author no. 2023. Crime comparison Crime Between Nigeria and United States. https://www.numbeo.com/crime/compare_countries_result.jsp?country1=Nigeria&country2=United+States

[61] author no. 2023. Top countries/markets by smartphones users. https://newzoo.com/insights/rankings/top-countries-by-smartphone-penetration-and-users

[62] Justin W Patchin and Sameer Hinduja. 2010. Cyberbullying and self-esteem. *Journal of school health* 80, 12 (2010), 614–621.

[63] Justin W Patchin and Sameer Hinduja. 2015. Measuring cyberbullying: Implications for research. *Aggression and Violent Behavior* 23 (2015), 69–74.

[64] Sai Teja Peddinti, Keith W Ross, and Justin Cappos. 2014. " On the internet, nobody knows you're a dog" a Twitter case study of anonymity in social networks, In Proceedings of the second ACM conference on Online social networks. *conference on online social networks* 60, 53, 83–94.

[65] Joy V Peluchette, Katherine Karl, Christa Wood, and Jennifer Williams. 2015. Cyberbullying victimization: Do victims' personality and risky social network behaviors contribute to the problem? *Computers in Human Behavior* 52 (2015), 424–435.

[66] Sonja Perren, Lucie Corcoran, Helen Cowie, Francine Dehue, Conor Mc Guckin, Anna Sevcikova, Panayiota Tsatsou, Trijntje Völlink, et al. 2012. Tackling cyberbullying: Review of empirical evidence regarding successful responses by students, parents, and schools. *International Journal of Conflict and Violence (IJCV)* 6, 2 (2012), 283–292.

[67] Sandra Petronio. 2010. Communication privacy management theory: What do we know about family privacy regulation? *Journal of family theory & review* 2, 3 (2010), 175–196.

[68] Buchhole Poushter. 2022. Which countries spend the most time on social media. https://www.weforum.org/agenda/2022/04/social-media-internet-connectivity/

[69] Jacob Poushter. 2016. Internet Access Growing Worldwide but Remains Higher in Advanced Economies. https://www.pewresearch.org/global/2016/02/22/internet-access-growing-worldwide-but-remains-higher-in-advanced-economies/

[70] K. Prabha and I. Shatheesh Sam. 2022. A Survey of Digital Image Watermarking Techniques in Spatial, Transform, and Hybrid Domains. *Int. J. Softw. Innov.* 10, 1 (sep 2022), 1–21. https://doi.org/10.4018/IJSI.309113

[71] Gisela Priebe and Carl Göran Svedin. 2012. Online or off-line victimisation and psychological well-being: a comparison of sexual-minority and heterosexual youth. *European child & adolescent psychiatry* 21, 10 (2012), 569–582.

[72] Michael Reich. 2015. The ups and downs of minimum wage policy: The fair labor standards act in historical perspective. , 538–546 pages.

[73] Johnny Salda. 2021. The coding manual for qualitative researchers. *Qualitative research in organizations and management: an international journal* 10, 2 (2021), 3–6.

[74] Michael Salter, Thomas Crofts, and Murray Lee. 2013. Beyond criminalisation and responsibilisation: Sexting, gender and young people. *Current Issues in Criminal Justice* 24, 3 (2013), 301–316.

[75] Sarita Schoenebeck, Carol F Scott, Emma Grace Hurley, Tammy Chang, and Ellen Selkie. 2021. Youth Trust in Social Media Companies and Expectations of Justice: Accountability and Repair after Online Harassment. *CHI Conference* 5, CSCW1 (2021), 1–18.

[76] Peter K Smith. 2019. Research on cyberbullying: strengths and limitations. , 9–27 pages.

[77] research department statista. 2022. Internet usage in Nigeria. https://www.statista.com/topics/7199/internet-usage-in-nigeria/

[78] research department statista. 2022. Internet usage in the United States. https://www.statista.com/topics/2237/internet-usage-in-the-united-states/

[79] Jose M Such, Joel Porter, Sören Preibusch, and Adam Joinson. 2017. Photo privacy conflicts in social media: A large-scale empirical study. *CHI conference* 22, 3 (2017), 2–7.

[80] John Suler. 2004. The online disinhibition effect. *Cyberpsychology & behavior* 7, 3 (2004), 321–326.

[81] Iraklis Symeonidis, Gergely Biczók, Fatemeh Shirazi, Cristina Pérez-Solà, Jessica Schroers, and Bart Preneel. 2018. Collateral damage of Facebook third-party applications: a comprehensive study. *Computers & Security* 77 (2018), 179–208.

[82] Monika Taddicken. 2014. The 'privacy paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of computer-mediated communication* 19, 2 (2014), 248–273.

[83] Robert S Tokunaga. 2010. Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in human behavior* 26, 3 (2010), 277–287.

[84] CB US. 2014. Computer and internet access in the united states 2012. https://www.census.gov/data/tables/2012/demo/computer-internet/computer-use-2012.html

[85] Tracy Vaillancourt, Robert Faris, and Faye Mishna. 2017. Cyberbullying in children and youth: Implications for health and clinical practice. *The Canadian journal of psychiatry* 62, 6 (2017), 368–373.

[86] Kami Vaniea and Yasmeen Rashidi. 2016. Tales of software updates: The process of updating software. *CHI conference* 21, 1 (2016), 3–10.

[87] Na Wang, Heng Xu, and Jens Grossklags. 2011. Third-party apps on Facebook: privacy and the illusion of control. *CHI Conference* 5, 1 (2011), 3–9.

[88] Mathias Weber, Marc Ziegele, and Anna Schnauber. 2013. Blaming the victim: the effects of extraversion and information disclosure on guilt attributions in cyberbullying. *Cyberpsychology, Behavior, and Social Networking* 16, 4 (2013), 254–259.

[89] Elizabeth Whittaker and Robin M Kowalski. 2015. Cyberbullying via social media. *Journal of school violence* 14, 1 (2015), 11–29.

[90] Vivian H Wright, Joy J Burnham, T Inman Christopher, and N Ogorchock Heather. 2009. Cyberbullying: Using virtual scenarios to educate and raise awareness. *Journal of Computing in Teacher Education* 26, 1 (2009), 35–42.

[91] Sijia Xiao, Coye Cheshire, and Niloufar Salehi. 2015. Sensemaking, Support, Safety, Retribution, Transformation: A Restorative Justice Approach to Understanding Adolescents' Needs for Addressing Online Harm. *CHI Conference* 23 (2015), 2–12.

[92] Victor Legbo Yisa, Reza Ghaiumy Anaraky, Bart P Knijnenburg, and Rita Orji. 2023. Investigating Privacy Decision-Making Processes Among Nigerian Men and Women. *Proceedings on Privacy Enhancing Technologies* 1, 18 (2023), 3–9.

[93] Changmin Yoo. 2021. What are the characteristics of cyberbullying victims and perpetrators among South Korean students and how do their experiences change? *Child Abuse & Neglect* 113 (2021), 104923.

[94] Izabela Zych, Rosario Ortega-Ruiz, and Rosario Del Rey. 2015. Systematic review of theoretical studies on bullying and cyberbullying: Facts, knowledge, prevention, and intervention. *Aggression and violent behavior* 23 (2015), 1–21.

A APPENDIX

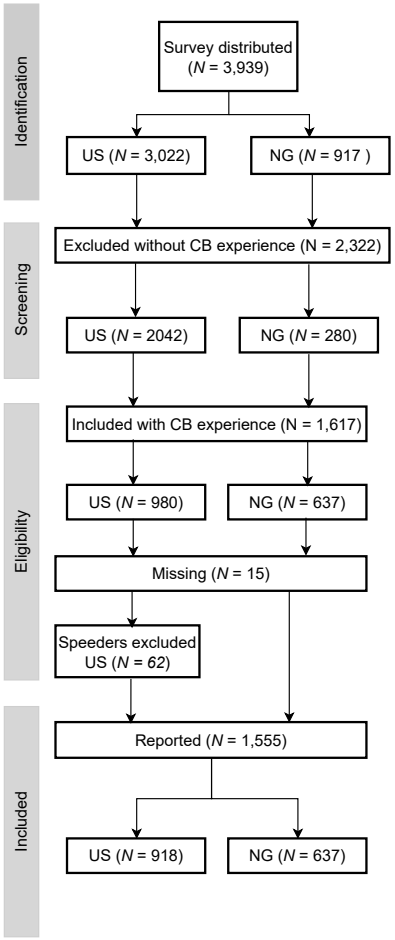

Figure 6: Data collection and analysis pipeline.