Sensemaking, Support, Safety, Retribution, Transformation: A Restorative Justice Approach to Understanding Adolescents' Needs for Addressing Online Harm

Sijia Xiao
University of California, Berkeley
Berkeley, USA
xiaosijia@berkeley.edu

Coye Cheshire
University of California, Berkeley
Berkeley, USA
coye@berkeley.edu

Niloufar Salehi
University of California, Berkeley
Berkeley, USA
nsalehi@berkeley.edu

ABSTRACT

Online harm is a prevalent issue in adolescents' online lives. Restorative justice teaches us to focus on those who have been harmed, ask what their needs are, and engage in the offending party and community members to collectively address the harm. In this research, we conducted interviews and design activities with harmed adolescents to understand their needs to address online harm. They also identified the key stakeholders relevant to their needs, the desired outcomes, and the preferred timing to achieve them. We identified five central needs of harmed adolescents: sensemaking, emotional support and validation, safety, retribution, and transformation. We find that addressing the needs of those who are harmed online usually requires concerted efforts from multiple stakeholders online and offline. We conclude by discussing how platforms can implement design interventions to meet some of these needs.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing**.

KEYWORDS

Online harassment, content moderation, online governance, social media, survivors

ACM Reference Format:
Sijia Xiao, Coye Cheshire, and Niloufar Salehi. 2022. Sensemaking, Support, Safety, Retribution, Transformation: A Restorative Justice Approach to Understanding Adolescents' Needs for Addressing Online Harm. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3491102.3517614

1 INTRODUCTION

Online harm such as harassment is prevalent on social media platforms. According to Pew Research Center, social media is by far the most common online venue for harassment in the United States — 75% targets of online abuse, which equals 31% Americans say their most recent harm experience was on social media [60]. Social media platforms tend to address these harms through the framework of content moderation: the review and removal of content that violates the platform's rules, and banning of repeat offenders [16, 44]. Though research has found some impact of content moderation in reducing offenders and offending behaviors [22, 23], the framework leaves out victims' experiences and needs for addressing harm [49]. Research has found that the current form of content moderation leaves victims out of the decision making process [49] and fails to adapt to their individual experiences [5].

In recent years, the HCI and CSCW communities have explored a victim-centered perspective to address online harm. Researchers have examined victims' strategies for dealing with harm [7, 59], engaged victims in designing interventions to address online harm [2], and studied their notions of justice [4, 49]. Our research builds on this line of work and is inspired by restorative justice – a victim-centered justice approach – to understand people's needs for addressing online harm. Content moderation follows a punitive justice approach, where it responds to harm by centering the offending party and regulating their offending behavior through punishment. Restorative justice, on the other hand, centers victims' experiences and desired outcomes. Through communicating with victims, a restorative justice process aims to support victims to reflect on their needs for addressing harm, and also engaging offenders and community members to help victims meet those needs [65].

We focus our investigation on adolescents (10-20 years old) [47], which are a particularly vulnerable group for a variety of harms in the online space. The vast majority of teens (90%) in the United States believe online harassment is a problem that affects people their age, and they mostly think teachers, social media companies and politicians are failing at addressing this issue [1]. Restorative justice has been successfully applied to address harm among adolescents in schools [19]. In recent years, researchers have seen the potential for applying restorative justice principles and practices in the online space [4, 20, 49]. We follow this line of work and explore how restorative justice helps us understand adolescents' needs in addressing online harm. Human needs and motivations are the driving force of behavior [37]. Without understanding needs and motivations, we may presuppose why certain actions (e.g., moderation) are important for addressing harm, but not understand why they are important for victims [42].

We examine adolescents' needs for addressing online harm from three interrelated perspectives: *what* needs they identify, *how* they believe their needs can be met (and by whom), and *when* they believe different needs should be met. Before we can develop specific recommendations to address the needs, we must first understand

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04.
https://doi.org/10.1145/3491102.3517614

the types of needs that adolescents identify from their own experiences. For example, adolescents may identify needs for themselves, as well as needs for their online communities. Our goal in the first research question is to better identify specific, major types of needs that come from adolescents experiences with online harm:

> RQ1. What types of needs do adolescents identify for addressing online harm?

Next, we examine *how* adolescents hope to achieve their needs, including the relevant stakeholders and the actions they perform in order to meet those needs. When harm happens, victims suffer from a lack of agency and require actions from relevant stakeholders (e.g., moderators, bystanders) to collectively address the harm [23, 56].

> RQ2. How do adolescents want to achieve different needs in addressing online harm? What specific actions can help adolescents address their needs, and by whom?

Adolescents may have multiple needs for addressing harm, which requires more than one action from a single stakeholder. Needs are not necessarily independent from one another, and some needs may have more immediacy than others. For example, social and self-esteem needs can become more important once fundamental needs of safety and security are achieved [33]. In our research, we aim to understand the immediacy of needs in addressing online harm:

> RQ3. When do adolescents hope to achieve different needs when online harms occur?

To address our research questions, we conducted interviews and design activities with 28 participants who experienced online harm during adolescence. In the interviews, participants complete a series of task on the online whiteboard to identify and reflect on stakeholders, actions, needs, and the timeline to achieve those needs. We found five major needs for addressing online harm from participants: sensemaking, support and validation, safety, retribution, and transformation. Participants identified both online and offline stakeholders that may address needs, including moderators, offenders, family and school, and proposed actions for address needs both in the short and long term.

In this paper, we examine what it would take for social media spaces to realize important social values such as supporting the safety and growth of adolescents, instead of the bare minimum of banning some types of offending content. Our findings shed light on how we may expand our understanding of victims' needs both spatially and temporally. We argue that online platforms can implement approaches beyond content moderation and can collaborate with other stakeholders to support victims both in the short and long term. In particular, we see potential for applying restorative justice approaches in addressing online harm for adolescents, such as by helping offenders realize their wrong-doing or utilizing the support of communities that victims are a part of. The design task that we created, contributes an innovative method for victims to reflect on their needs in addressing online harm. Finally, our research builds on and extends recent work that center victims' perspectives in addressing online harm [5, 48] and examines alternative justice models in online governance [4, 49].

2 RELATED WORK

2.1 Online harm and content moderation

Social computing researchers have studied a variety of harmful behaviors online, including hate speech [34], online harassment [5], and digital self-harm [41]. In this research, the harm we study falls under interpersonal violence (compared to self-directed or collective violence), where the victim believe they have been harmed by one or more individuals [29]. Online platforms usually address online harm with content moderation: when online users post content that violates platforms rules, the platform will sanction the users with punitive measures such as removal of content or bans [49]. The person who issues the sanctions may be commercial content moderators hired by the platform or the community content moderators who are end users [52]. There is also a growing use of bots, algorithms, and artificial intelligence that automatically detect content that is against the rules and make decisions about removing or limiting the reach of the content [3, 28]. Researchers have examined ways to reduce offending behaviors through moderation mechanisms, for example, through setting positive examples and social norms in moderation [9, 51], providing moderation explanations [21, 22], or designing novel algorithms and AI moderation systems [8].

Research on content moderation and related areas have presented design solutions that have been shown to reduce offenders and offensive behaviors. However, researchers have argued that content moderation does not give victims agency in addressing the harm they have experienced and fails to provide appropriate solutions to meet victims' diverse needs [5, 49]. Some researchers have explored victims' perspectives in addressing harm and center victims in the design process. Researchers have studied victims' experience and perspectives of online harm, including their perspectives on the classification of harm [5], the impact of harm [31, 48], and effective coping strategies [7, 57]. Through studying the diverse needs of victims, researchers have explored interventions beyond platforms' content moderation, for example, outsourcing the filtering of problematic content to bystanders or friends [12, 32, 56], offender's apologizing [49, 50], or providing tools to help victims gather authentic evidence of harm to share with the public [54]. More recently, researchers have applied alternative justice models, including restorative justice, to address issues of online harm. Schoenebeck et al. conducted survey studies to understand people's notion of justice and preferred design solutions of online harm under restorative justice and other justice frameworks [49]. Hughes and Roy contributed an online tool to facilitate restorative justice processes [20].

Research has found that some groups are particular vulnerable in online spaces and are more frequently and severely harmed than others, including women and gender minorities [14, 57, 59], racial minorities [18], and young people [2, 40, 50]. Some researchers have worked to center the views of vulnerable populations by including them in the design process. For example, Ashktorab et al. explored cyberbullying mitigation and prevention through participatory design with teenagers [2]. Niksirat et al. explored the interventions of non-consensual image or video sharing through participatory design with young adults [40]. In our research, we focus on adolescents as the study population, and involve them in a series of design

tasks to support them in reflecting on their needs and potential interventions to help address those needs.

2.2 Restorative justice

In this research, we see restorative justice as both a philosophy that centers victims' needs, and as a set of established practices and knowledge about what victims of harm generally need and how to support them.

2.2.1 Restorative justice values. To introduce restorative justice, we first discuss its differences with punitive justice. Punitive justice, as a widely applied justice model in the Western world, holds that harm is a violation of rules and offenders should suffer in proportion to their offense [15]. Under this model, victims are often left out of the process. Victims suffer from a lack of agency in the process of addressing harm, and do not get sufficient resources to recover from the offense [55].

Restorative justice provides an alternative way to address harm. Restorative justice believes that harm is a violation of people and relationships instead of just a violation of rules. While punitive justice is concerned primarily with making sure offenders receive punishment proportionate to their offense, restorative justice begins with a concern for victims and their needs [65]. Additionally, restorative justice ask for engagement from diverse stakeholders in addressing harm. Under a restorative justice model of addressing harm, people are all interconnected in a community, or society at large. Thus, harm creates obligations for relevant parties (e.g., offenders, members of community where harm has happened, victims and offenders' family members) to collectively address it [65]. Our research focuses on the harmed party, the victims, and understand their needs for addressing harm from other relevant stakeholders.

2.2.2 Restorative justice practices. Restorative justice processes usually begin with communication. In a process called *pre-conference*, a restorative justice facilitator will communicate with the victim to help them reflect on their needs for addressing harm [43]. The process that follows involves communication with the offenders and other relevant parties, sometimes collectively, to reach consensus on how to address those needs [43]. Here, a widely used communication tool is *circles* [24]. In a circle, victims, offenders and other relevant stakeholders (e.g., friends and family, community members) sit together with a restorative justice facilitator to discuss three core questions [65]: (1) what has happened? (2) who has been affected and how? (3) what is needed to repair the harm? The facilitator mediates this process to ensure that victims and offenders have equal footing and helps move the parties towards reaching consensus. After the meeting, the stakeholders carry out actions to repair the harm. While we believe that establishing full procedures of restorative justice online requires great adaptation, we argue that a first step is understanding victims' needs in addressing harm following restorative justice principles and procedures.

Restorative justice practitioners have found a range of needs that are commonly addressed in a restorative justice process. According to Zehr, there are four types of needs victims may achieve in restorative justice process, that are often ignored in punitive justice [65]: (1) Information. Victims gain information about why harm happened and what has happened since. (2) Truth-telling. Victims get a chance to tell their stories and receive acknowledgement from parties such as the community and offenders. (3) Empowerment. Restorative justice provides agency in addressing the harm they experienced, which can return them power and control that is taken away by the harm. (4) Restitution or vindication. Relevant parties, such as the offender, will make amendment to repair the harm (e.g., apology, financial compensation). In our work, we use the four fundamental needs to guide our understanding of victims needs in addressing online harm.

2.2.3 Applying restorative justice to adolescents. Restorative justice has been successfully applied in a myriad of offline settings, such as criminal justice, family, and workplace settings [58, 64]. In particular, restorative justice has been applied to address harm in schools and the juvenile criminal justice systems all around the world [19]. Practitioners argue that a restorative justice approach can be beneficial to both young offenders and victims compared to a punitive one. For young offenders, restorative justice acknowledge their needs and enhances their development instead of merely punishing them for their crime [39]. For young victims and the communities where harm happens, restorative justice provide support for them to heal from harm and restore or strengthen social relationships [26]. Restorative justice considers stakeholders beyond just the offenders, including the state and schools, providing a chance for family members and schools to collectively support the growth of young people [26]. Katic et al. conducted a systematic evaluation of restorative justice practices in schools in the United States and found that the majority of studies reported positive outcomes, including improved social relationships and reductions in office discipline referrals [27]. However, Latimer et al. noted that positive findings in restorative justice can be tempered by the self-selection bias – since it is a voluntary process, those who choose it may benefit more than others [30]. We build on the research and practices from offline restorative justice to study how it may benefit adolescents' online lives.

3 METHOD

Our research aims to understand adolescents' needs for addressing online harm, including what those needs are, how to meet those needs, and when. While asking people what they need may seem like a straightforward task at first, prior work in the restorative justice literature and our own preliminary research showed that it is challenging for victims to know and express what needs they have [6]. This is particularly the case when when those needs were not met when the harm happened, or when meeting those needs seems impossible given available resources from the online platforms or other relevant stakeholders.

Victims of harm need to go through a process of sensemaking to understand the harm, its effects on them, and to decide what they need to heal from the harm [65]. In restorative justice practices, this is often done through a pre-conferencing session with a facilitator who support the victim and helps them figure out what they need [43]. For this research, we hope to design a task to support the process of sense-making and enable participants to tell us the whole range of their needs – even those that could not be immediately met given current constraints and resource limitations. In this section, we first described the process of designing the task. Next,

we presented the task procedure. We then explained our recruiting and interview process, and finally ended with a description of our data analysis method.

3.1 Designing the task

3.1.1 Designing the need-finding questions to understand types of needs and actions. The goal of our research is understanding adolescents' needs when they are harmed online from three levels: (1) what their needs are, (2) the actions to meet those needs, and (3) the timing to meet those needs. It is challenging for people to know what their needs are and how to express them. Thus, our goal is to design *need-finding questions* to help support participants' sensemaking process.

Weick argues that sensemaking is retrospective. People first come up with or perform actions, then provide explanations for their actions [61]. Thus, we focus on actions in the need-finding questions, and then ask participants to explain their actions. Through the explanation, participants can identify their needs behind those actions. The process enables us to answer RQ1 and RQ2 together; by understanding what peoples' needs are, as well as the actions that can meet those needs.

We aim to design need-finding questions that cover all the categories of actions participants may identify. We started our research design process by looking at how victims talk about needs and actions for addressing harm in the restorative justice literature. In Zehr's foundational work on restorative justice, he proposes four categories of needs that victims commonly have: the need for information, the need for truth-telling, the need for empowerment, and the need for restitution or vindication [65]. Zehr also describes example actions that can address those needs, for example, offenders' acknowledgement can meet victims' need for restitution, and understanding why the harm happened can meet victims' need for information. We rely heavily on this work in designing our research method.

First, the research team brainstormed potential actions based on the examples provided by Zehr. We then categorized the actions through pilot testing with 15 participants who we selected through convenience sampling [45]. We asked pilot participants with experiences of online harm to select from those actions, and come up with additional needs they might have. For pilot participants who hadn't experienced online harm, we asked them to group the actions in a card sorting activity [53]. We asked pilot participants to think out loud to understand their thought process. This process led to five questions which cover most actions our pilot participants mentioned: *(1) what information do you need from [the stakeholder]? (2) what do you want to share with [the stakeholder]? (3) what acknowledgement / understanding do you want from [the stakeholder]? (4) what actions is needed from [the stakeholder] to repair the harm? (5) what change do you want [the stakeholder] to do in the future?*

3.1.2 Designing a timeline to envision the story. While the need-finding questions help us answer RQ1 and RQ2, we were also interested in the temporal aspects of addressing online harm and envisioned a story line of addressing harm for RQ3. The process of participants walking through their own storylines provides more chances for them to reflect on their different needs and actions, as well as their order when a harm occurs. Inspired by previous research in speculative design [63], we decided to design and facilitate a reflection process to achieve this goal.

Our design task borrows from the Timelines speculative design activity proposed by Wong and Nguyen [63]. Timelines is designed to help participants reflect on their values and ethics around a technology. Participants complete the Timeline activity with sticky notes and a whiteboard. There are four steps in the timeline activity: (1) participants decide on an artifact (e.g., a technology) as the topic of discussion, (2) identify stakeholders around the artifact, (3) create potential news headlines and stories related to the artifact, and (4) organize the news headlines and stories on multiple timelines to create stories of events related to the artifact. Overall, through a visual board, the Timelines activity helps "the creation of an imagined world that can lead participants to critical refection" [63]. In borrowing from the Timelines activity, our goal is to help participants picture a storyline for addressing harm, while reflecting on their values and desired outcomes in the process. While our work is not entirely speculative, we encourage participants to think beyond perceived constraints while building on their own experiences.

3.2 Task procedure

Our study consists of four main stages:

(1) Participants decide on a harm case from their adolescence they'd like to talk about.
(2) Participants identify stakeholders relevant to the harm case.
(3) Participants generate actions the stakeholders might perform with five need-finding questions. Participants reflect on their needs through identified stakeholders and actions in stage 2 and 3.
(4) Participants map those actions spatially to illustrate their preferred timeline for addressing the harm.

Figure 1 provides an example of the interface where participants complete the task. In the following sections, we describe each stage in more detail.

3.2.1 Stage 1: Participants choose a harm case from their adolescence they'd like to talk about. In the first of four stages, the researcher asks participants to share a harm case they want to talk about. In pilot studies, we shared a hypothetical harm scenario with participants and asked them to imagine themselves in that situation and share their needs. However, participants found it hard to empathize with the hypothetical scenario. Therefore, we chose to use participants' lived experiences. While relying on each person's own experiences made it more difficult to control for the types of harm in our study, their personal experiences contain concrete details (e.g., what the offender said to them, their relationship with different stakeholders) that are important to determining their needs.

After participants select a case, we ask a series of questions about the case (e.g., when and where it happened, if and how they addressed the harm, and their feelings at that time). The purpose is twofold: first, in later stages of the task, we provide participants flexibility in expressing the needs based on their unique experiences. Talking about the harm they experienced helps participants recall what has happened in detail, which enables them to reflect on their needs thoroughly. Second, information about the harm case

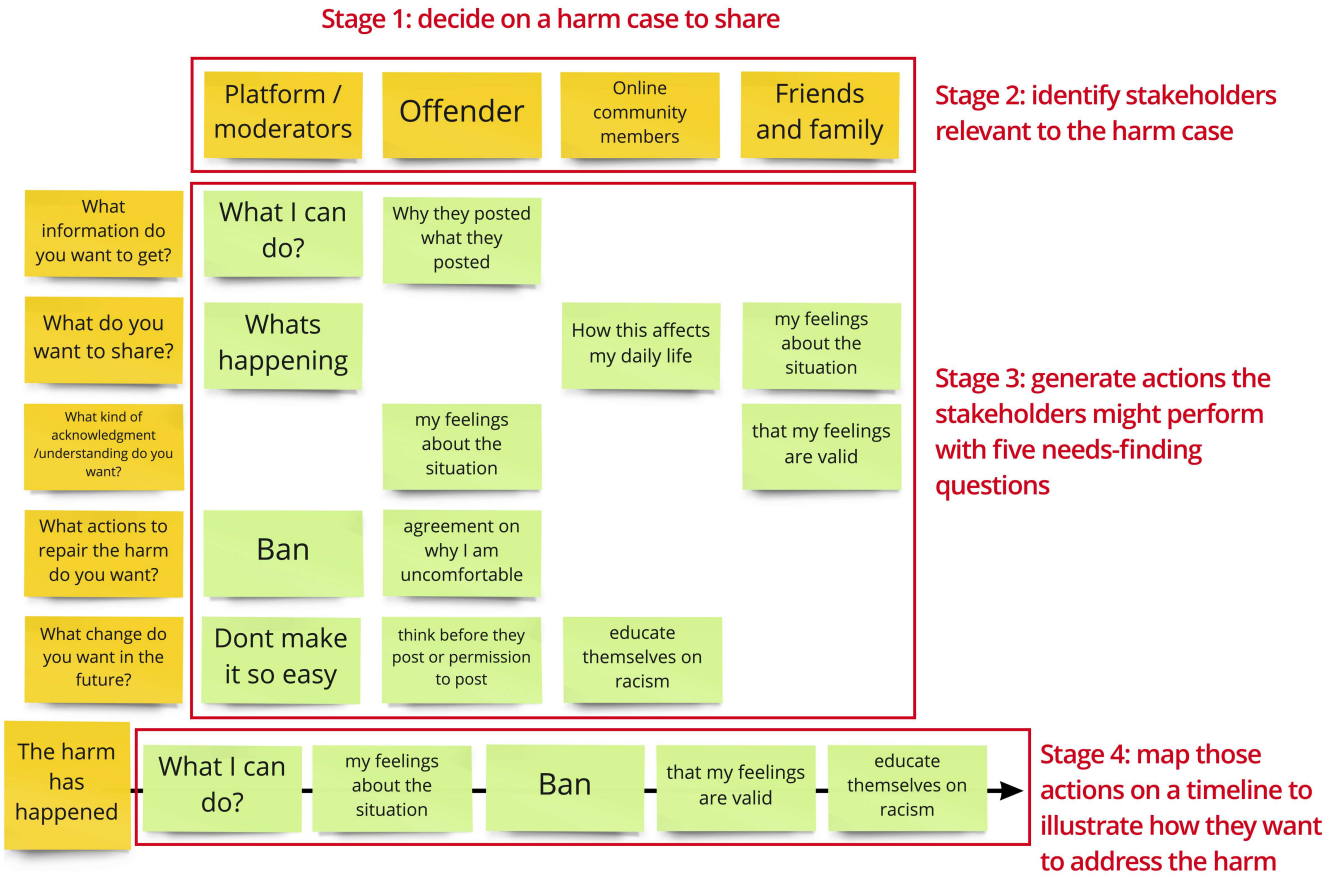

Figure 1: An example of the interface where participants complete the four-stage task. The sections marked by red frames are what that participants need to complete in each stage of the study. To protect participants' privacy, the data in the red frames comes from multiple participants.

provides context for our interpretation and understanding of their needs in data analysis.

3.2.2 Stage 2: Participants identify stakeholders relevant to the harm case. In the second stage, participants identify the stakeholders relevant to the harm. We asked two questions to facilitate their selection of stakeholders: *(1) Who is responsible to help you address the harm? (2) Regardless of responsibility, who can support or help you to address the harm?* Because some participants weren't sure how to answer this question in pilot studies, we provided them with some example stakeholders as starting points, including offenders, family and friends, online community members, and platform/moderators. Participants also have the option of adding additional stakeholders either in this stage or in later stages.

3.2.3 Stage 3: Participants generate actions the stakeholders might perform with the five need-finding questions. In the third stage, we used the stakeholders they identified and five need-finding questions discussed in section 3.1.1 to form a table, and asked participants to answer each question with respect to each stakeholder group in the table (see Figure 1). Thus, the process allows participants to identify the *actions* required from stakeholders for addressing the harm.

In stages 2 and 3, we asked the participants to think out loud and explain their rationale for selecting/writing a note. This allows us to understand the type of *needs or motivations* of choosing certain actions or stakeholders.

3.2.4 Stage 4: Participants map those actions on a timeline to illustrate how they want to address the harm. In the final stage, participants rearranged the actions they had just created spatially to reflect on an ideal timeline to address the harm. Since participants have reflected on the needs behind the actions, the series of actions on the timeline also represents the sequence of needs. We also encouraged them to create new notes to complete the timeline. In the process, the researchers asked participants to think out loud and explain their reasoning for the order of notes.

Table 1: Participant demographics and experiences of online harm

	Age	Gender[1]	Race/ethnicity[1]	When harm happened	Online Platform	Offline Site	Number of offender(s)	Relationship	Description of harm[2]
P1	19	Female	Asian	Middle school	Instagram, Snapchat	School in U.S	1	Friends	Racist comments, public shaming, physical harm
P2	20	Male	White	17-18	Twitter	N/A	1	Stranger	Physical threat
P3	20	Female	Asian	15	Discord	N/A	1	Friends	Sexual harassment, non-consensual image sharing
P4	19	Female	Asian	16	Instagram	N/A	1	Schoolmate	Racial discrimination, public shaming
P5	20	Male	Asian	20	League of Legends	N/A	1	Stranger	Offensive name-calling
P6	20	Female	Asian	Middle school	Instagram post	School in U.S	3 to 4	Friend	Public shaming, non-consensual image sharing
P7	20	Female	Asian	High school	Instagram	N/A	1	Classmate	Body shaming
P8	18	Female	Indigenous	18	Twitter	N/A	3 to 4	Strangers	Racist comments, offensive name-calling
P9	19	Female	Asian	First year of high school	Instagram	N/A	1	Classmate	Body shaming
P10	19	Male	Asian	End of middle school	WhatsApp	School in India	10 to 15	Classmate	Public shaming, physical threats, physical harm
P11	19	Male	Hispanic white	18	Instagram	N/A	1	Friend	Making fake profile of me
P12	19	Female	Asian	19	Tiktok	N/A	Multiple	Strangers	Racist comments
P13	19	Male	Asian	19	Grindr, tinder	N/A	Multiple	Strangers	Financial fraud with fake account
P14	20	Male	Asian	20	Reddit	N/A	Multiple	Strangers	Trolling, harassment
P15	20	Male	Asian	High school	Instagram, Facebook Messenger	School in U.S	Multiple	Friends	Racist comments, making jokes of my disability
P16	19	Female	Asian	High school	Instagram, Snapchat	N/A	Multiple	Friends	Non-consensual image sharing
P17	20	Male	Asian	First year of high school	Instagram	School in India	1	Friend	Making fake account for public shaming
P18	20	Female	Asian	High school	Instagram, Snapchat	School in China	Multiple	Classmate, strangers	Racist comments
P19	20	Female	Asian	13	Tumblr, Insagram	N/A	Multiple	Strangers	Offensive comments of my arts
P20	19	Female	Asian	High school	Instagram	N/A	Multiple	Strangers	Racist comments
P21	19	Female	Asian	18	Slack, email	N/A	1	Stranger	Non-consensual image sharing
P22	20	Female	Asian	18	Instagram, Twitter	N/A	Multiple	Strangers	Racist comments
P23	20	Female	Black	Elemetary to high school	ASKfm	N/A	Multiple	Schoolmate	Sexual harassment
P24	18	Male	Asian	High school	Facebook	School in India	Multiple	Classmate	Body shaming
P25	19	Female	Asian	High school	Weibo	School in China	Multiple	Classmate, strangers	Trolling, harassment
P26	20	Male	Asian	High school	Twitter	School in U.S	Multiple	Schoolmate	Racist comments, offensive name-calling
P27	19	Female	Black	10th grade	Instagram	School in U.S	Multiple	Schoolmate	Public shaming, harassment
P28	20	Female	Asian	20	Instagram	N/A	1	Stranger	Racist comments

[1, 2]Participants' gender and race/ethnicity are self-identified. The majority of participants are Asian, yet they come from diverse cultural background, including Afghan, Chinese, Filipino, and Indian. [3]Here, we did not follow a strict definition of types of harm but rather stay close to participant's description and categorization of their experiences.

3.3 Recruitment and Interviews

We recruited 28 students from a University on the West Coast of the United States. We used the university's internal recruiting platform to reach potential participants from a pool of students who agreed to be contacted about paid social research opportunities. We focused our recruiting message to late adolescent students between 18-20 years old, and indicated that we were looking for participants who have experienced online harm on social media during their adolescence. We also provided some examples of online harm (e.g., offensive name-calling, public shaming, stalking, harassment, physical threats) to help them reflect on potentially relevant experiences.

We show participants' demographic information and the information about their online harm experiences in table 1. Participants reported a wide range of harm experiences including non-consensual image sharing, body-shaming, sexual harassment and physical threat. For many participants, the harm cases had an offline part in school, or happened online with their classmates or schoolmates. Due to the restriction of IRB, the participants we interviewed are in the late adolescence group, but the harm cases they shared happened from early to late adolescence.

We conducted all the interviews between July and August 2021. The interviews were within one hour and in the form of video or voice call on Zoom (www.zoom.us). We used an online whiteboard tool, Miro (http://miro.com), to facilitate the task in remote sessions.

Participants received a $25 gift card as compensation. The study is approved by the institutional review board.

3.4 Data Analysis

We transcribed the interview recordings with an online transcription service (www.rev.com). We exported the data on the online whiteboards into an excel sheet, and also referred back to the original board for spatial information during analysis. We analyzed the interview transcript and data from online whiteboards together.

We conducted the data analysis in an iterative process. We applied interpretative qualitative coding to the data [38]. We began with initial coding, where we applied short phrases as codes [46]. The first round of coding was done on a line by line basis so that the codes stayed close to the data. Some example codes include "need empathy" and "design automatic moderation tools." We then conducted focused coding by identifying themes that appeared repeatedly to form higher-level descriptions [46]. Examples of second-level codes include "prevention of harm" and "acknowledgement of responsibilities." Throughout the analysis, we not only paid attention to the needs victims have, but also the actions and stakeholders that were proposed to meet the needs. In analyzing the timelines data, we paid attention to the order and time span in which participants wanted to address those needs.

4 FINDINGS: STAKEHOLDERS AND ACTIONS

Our first research question concerns the types of needs that adolescents identify for addressing online harm. Our second research question explores their preferred actions and stakeholders to address those needs. As we noted in methodology, it is challenging for participants to directly identify their needs. Thus, we first asked participants to identify their preferred actions and stakeholders for addressing harm, and then asked them to explain the needs behind the actions and stakeholders retrospectively. Since needs, actions and stakeholders are three interrelated concepts, we answer the two research questions together in this section. We presented our major findings in table 2. We found five major needs that participants frequently mentioned in our study: sensemaking, support and validation, retribution, safety and transformation. Next, we detail the actions and the relevant stakeholders related to those needs.

4.1 Need for sensemaking

Through the reflective task, all participants indicated that the offenders had done something wrong. However, some participants told us that they were not as certain about the wrongdoing immediately after the the incident had occurred. P3 reflected that her offender *"[repeatedly] sent me unwanted pictures then brushed it off as a mistake."* She was unsure about the situation: *"That were almost in a gray zone [. . .] I was not sure if that was just how he normally is and I'm just overreacting or is he actually making unwanted advances towards me."* Additionally, even when some participants were sure that they were harmed, they were less sure how to address the problem. To make sense of the situation, participants hope to understand why offenders did what they did, and get instructions or advice on how to deal with the harm.

4.1.1 Seeking information on offenders' motives for conducting harm. Participants hope to get information on why the harm happened, and whether they are actually responsible for the harm instead. Several participants mentioned that they want to get the offenders view, and to understand their actions. P28 explained that it would help her to *"understand that the problem is within themselves [offenders] and not with me."* P5 expressed that understanding offenders' motives helped him rationalize offenders' behavior: *"maybe they're having a really bad day then it makes more sense I feel for them to behave that way."* Understanding where the offender is coming from is also a step towards addressing the problem. P6 said, *"I wanted to know why it happened, so we were able to talk about it."*

4.1.2 Seeking advice on how to address the harm. Many participants mentioned that they need advice from others in dealing with the harm. Some of them chose to resort to their family and friends. P18 believes that family and friends can give *"personalized advice"*. Other participants want to look for people who have expertise in addressing harm. P10 explained that mental health professionals can help with his emotional sufferings: *"they just know techniques that people can use to cope with any sort of suffering and how to alleviate it."* P6 described the needs for online platforms to help address some technical challenges: *"I didn't know how to protect myself. Even now I don't fully understand how to do things on Facebook, prevent people from tagging me in photos that I don't want to be tagged in or stuff like that."* P10 also identified the guidance counselor from school as someone who could help with harmful situations, particularly since many of the harms happened among classmates: *"They're used to dealing with school kids. They know how bullying happens. They know the triggers."*

4.2 Need for emotional support and validation

Almost all participants mentioned the need for getting emotional support and validation. Participants look for emotional support from family, friends and online supporting groups, while hoping platforms and online community members will acknowledge their responsibility in addressing the harm. In some cases, they hope the offenders can acknowledge their mistake and apologize.

4.2.1 Emotional support from family, friends and online users with similar experiences. When a harm occurs, participants are charged with negative emotions. Participants explain that they need to vent their feelings and hope to get emotional support. Some resort to friends and family for those conversations: *"Because with friends and family, you can really open up about how you're feeling...whether you just want to rant, or you want advice, then they can either offer that support. Because they probably care more than most people" (P2).* Some participants also turn to other online users who have similar experiences. Here, participants gain support not only through sharing their own stories, but also by listening to others. P3 said, *"Support communities are a good place to share experiences without fear of being judged, and survivors tend to feel more solidarity when hearing about others' stories."*

4.2.2 Platforms and online community members can show their stance against harm. Some participants believe that online community members and online platforms can provide support and validation by standing in solidarity with the victims against online

Table 2: The table presents participants' needs for addressing harm, actions to meet the needs, and the stakeholders to perform the actions from left to right. For example, to meet the need of sensemaking, participants hope to seek information on offenders' motives for conducting harm from their offenders.

Needs	Actions	Stakeholders				
		offenders	family and friends	online platform	online community members	school
Sensemaking	Seek information on offenders' motives for conducting harm	x				
	Seek advice on how to address the harm		x	x		x
Emotional support and validation	Emotional support		x		x	
	Show stance against harm			x	x	
	Acknowledge wrong-doing and issue apology	x				
Retribution	Content moderation			x		
	Report or call out offenders				x	
	In-school repercussion					x
Safety	Acknowledge wrong-doing and issue apology	x			x	
	Show stance against harm				x	
	Content moderation			x		
	In-school repercussion					x
Transformation	Improve design and moderation			x		
	Raise public awareness				x	x

harm. P19 explained that she hopes online users realize that they are connected: *"It's important to know that we're all one big community that's sharing something...it's important to build each other up."* P20 expected online community members to express *"kind of a collective understanding that what was happening was wrong and there should be preventative action against it."* Several participants hoped that platforms could issue direct statements to show their stance towards the incident (or similar incidents). P7 provided an example: *"[I hope the platform can state that] 'these types of behaviors and comments are not appropriate in any setting.'"* Participants explained that acknowledgement of online harm is one step towards addressing the issue: *"I guess to repair the harm, acknowledging that there is a problem is the first step" (P18).*

4.2.3 Offenders can acknowledge their wrong-doing and issue an apology. Some participants explained that they want the offenders to acknowledge the harm that they created. For example, P2 hoped the offender might understand that *"it was really hurtful. It was hateful. It was unnecessary."* Participants not only hope to share their feelings and frustrations, they often want an acknowledgement of wrongdoing from the offender. In particular, several participants explained that the offenders can (or should) apologize to them. P1 wants to tell her offender *"this is how you made me feel, and this is how interpreted the situation. please understand my side"*, and in return, the offender should *"[apologize] and clear the air."*

4.3 Need for retribution

Participants often explained that they hope offenders receive consequences for their negative behaviors: *"you'd want to send a message [to offenders] that hate will not be tolerated...that actions have consequences" (P14).* Some participants specifically mentioned that they want offenders to receive punishment as consequence. After being harmed repeatedly, P27 admitted that she hoped the offender would suffer in return: *"I used to be really angry and I just wanted bad things to happen to them...they shouldn't get away with it."* While some participants expressed a need for punishment, restorative justice explicitly seeks to create alternatives to punitive justice [65]. We will discuss this conflict, as well as the difference between accountability and punishment, in the discussion section.

Participants described different authority figures who might administer retributive actions. Some believe that online platforms can use moderation (e.g., bans) to hold offenders accountable. Participants explained that online community members could help report offenders to moderators, or call out offenders at the time of the offense. Since many participants receive harm from their classmates or schoolmates, some believe the school administrators should hold their students accountable.

4.3.1 Platforms: issue punitive moderation decisions to the offender. Some participants believe that platform moderation (e.g., banning, muting) is a form of punishment to offenders or a way to hold them accountable. P23 explained her rationale: *"I think that your presence on social media is a privilege that can be taken away...if you don't follow the rules or the guidelines of the platform."* P18 thinks

that banning is a form of denial to offenders: *"suspension of a rude account isn't really a big thing, but...personally, I think that would make me feel better that the people who were rude to me, someone is telling them that what you did [is] wrong."*

4.3.2 Bystanders: Reporting offenders or calling out offenders for their actions. Participants also mentioned online community members, in particular, bystanders' role in holding offenders accountable. Some participants believe that bystanders should report the offenders to the platform. P6 hopes bystanders know that *"It is important and very simple online to report things that you see that are harmful to others."* In addition to officially reporting harm, some participants expressed that bystanders could call out the offending behavior when it occurs. P28 described it as *"a form of positive online peer pressure"*, while P23 phrased it as *"public backlash."*

4.3.3 School Administration: Punishing students for their online behavior. Some participants believe that the school should hold their students accountable for their online behavior. P24 explained, *"It's your [the school's] duty to ensure that the students of your school behave in a good way...and then train them to be good citizens. So that's why I feel like, even if it's online, they [the responsibilities] still go to your school."* Participants believe that the school should give students academic repercussions for their online behavior. P27 described it as *"getting detention or something showing up on their records to show that they have poor behavior."* P18 stated that the school should *"Talk to their [offenders'] parents, maybe even suspend them."*

4.4 Need for safety

We find that sometimes participants need to deal with an ongoing harm. Even when participants think that a harm has stopped for the moment, they are often unsure if the harm will resume in the future. In such situations, one's safety from continuing harm is a priority. Previously, we explained how a variety of actions can meet participants' need for support and validation, as well as a desire to get retribution from offenders. Importantly, we find that these *same* actions can serve another purpose – to stop the continuation of harm and help individuals feel safe.

4.4.1 Participants hope that acknowledgement can stop the harm. Earlier we described how participants need emotional support and validation after the experience of harm. They get comfort when online community members show their stance against online harm, or when offenders acknowledge their wrong-doing and apologize. Some participants believe those actions also stop the continuation of harm. P19 believes that online community members' stance against harm can reduce offending behavior: *"If it's publicly announced that, 'Oh, this is not the behavior that we're going to tolerate,' I feel like people would be more ashamed to act out like that."* Some participants also expect their conversation with the offender can prevent continuous harm from them: *" [If offender] apologize and clear the air between us, and then any acts of harm should stop because we are done with it" (P1).*

4.4.2 Retributive actions to stop the harm. Some participants hope that retributive actions will teach the offender a lesson, while also stopping the harm. We find that participants often put the onus on platforms to enact some type of retribution which will also stop the harm: *"If they [offenders] are not willing to change, it's kind of the responsibility of the platform to kick them off"(P19).* P24 explained why having the harmful post reported and removed can stop the escalation of harm: *"you want to ideally reduce the number of views that it gets and prevent it from growing even bigger."* In addition, some participants expect the school can intervene: *"the school should step in and say, 'stop taking people into this [the harm]"' (P6).*

4.4.3 Uncertainty about stopping the harm. Participants in our study do not always expect that having a conversation with offenders will force a change. P7 explained, *"The offender would always try to defend themselves I think, and not really address anything."* She believes that *"people don't need don't really change overnight."* P15 expressed reluctance to face the offender and worried if he will be disappointed by the response: *"There is the sort of fear that maybe they won't understand...if you don't get the response that you're ultimately looking for, then it can just be uncomfortable."*

Even when participants rely on others for help, they sometimes cannot specify the actions they want others to perform, or know whether those actions will effectively stop the harm. Individuals may know that they want something to stop, but they do not know who should actually take action. P6 expressed this type of frustration, *"I feel like someone should put a stop to that."* P10 put his need as a inquiry: *"I would want to know what they [moderators] can do to stop these kind of hateful messages from being spread, so that they could put a halt to the situation and at least [lower] its severity."* P6 was also unsure whether the school can hold their students accountable for online behaviors: *"ultimately they don't have the physical capability to make it stop...if those students were not willing to be compliant, I'm not really sure what actions you could expect the school to take."*

4.5 Need for transformation

Some participants believe that addressing online harm not only mean working on their individual cases, but also fundamentally change or transform the online environment. Participants suggest that it is important to bring more attention and resources to the issue of online harm. These individuals believe that online platforms should use more resources to improve their moderation procedures, and emphasize online environments that identify and stop harms from occurring. At a broader level, these participants indicate that it is important to raise public awareness of online harm.

4.5.1 Online platforms should improve design and moderation to address harm. Several participants indicated that the platforms should moderate content before a harm occurs. For example, several participants believe that platforms can improve their automatic detection mechanisms to filter out hateful comments before they reach online users. P7 thinks the current detection only works for explicit hateful words: *"Some comments they filter are usually only addressed for inappropriate things as in rated R things inappropriate."* He hopes platforms can *"filtering comments that are close to hate or bullying."* P18 also expressed that human moderators can *"keep an eye out for certain rude words or phrases."*

Participants explained that platforms can provide more information, tools and resources in response to people's individualized

experiences. P11 wanted to correct a fake account someone made of him, and he hopes the platform can demonstrate *"the process of investigating and compiling a report on a duplicate social media account."* P10 argues that every harm is different: *"like every person's experience is nuanced and sometimes it really just doesn't fit in one category. In my case I actually knew the person, so it'd be nice for you to have more of a place to talk. "* He told us that only human moderators can provide customized responses: *"unless and until a human being looks at it and figures out what's going on, I don't think that's very accurate."*

4.5.2 Raising public awareness about harm can improve online environments. Many participant believe that educating the public about the importance of online harms and how to deal with them can fundamentally improve online environments. P18 talked about how she resorted to her family for help when online harm happened but wasn't given enough attention. She thinks that online harm was a new concept to her elderly family members, which she wanted them to understand: *"For them, bullying was something in person and rude comments on Instagram isn't even considered bullying for a lot of people."* Other participants explained that they hope people will learn that online harm can elicit as much pain as in-person harm: *"If you think that saying something bad to someone in person is bad, then you should also just assume the same for social media, it's not any different"(P26).*

Some participants expressed the belief that people need to educate themselves about the dangers of online harms, while others expressed a need for influencers and other celebrities to get involved in public education about these issues. P8 experienced racist comments and thinks that people should educate themselves on the topic: *"I would like to see this community educating themselves more and uplifting Indigenous peoples rather than invalidating our experiences."* P7 hopes that celebrities or other public online figures can utilize their influence to *"empowering all types of individuals... and speaking up about the issues [online harm]."*

Education about online harm does not necessarily have to come from online sources. Some participants believe that it is particularly important for their school to educate students about online harms. For example, P18 was bullied online by her classmates and she wondered if the school could have prevented it with more education: " *If in eighth grade the school constantly talks to their students about social media bullying, online appropriation, using words correctly, not saying rude things online... if they do that from a younger age, then that would solve the issue from the beginning itself."*

5 FINDINGS: TIMELINES OF NEEDS

Our third research question asks about the order and timing people want to meet the needs. In the first several stages of the design activity, participants first identify the actions and relevant stakeholders for addressing harm, then reflect on their needs behind the choice. In the final stage, we asked participants to place the sticky notes representing different actions onto a timeline. Since the actions are attached to specific needs, we were then able to analyze the preferred temporal order of meeting the needs. We summarize the patterns that we found in Figure 2. Next, we present participants' perceived immediacy of the five identified needs: sensemaking, support and validation, safety, retribution, and transformation.

5.1 Need for sensemaking comes first

We find that when participants need to make sense of the situation, they usually do it before meeting other goals. As we discussed in 4.1, some participants were not sure if they overreacted, or they don't know how to proceed with addressing the harm. Thus, understanding offenders' motives and getting instructions for addressing harm are prerequisites for participants' next moves. The two outliers (P4, P16) chose to stop the continuation of harm before making sense of it.

5.2 Need for emotional support is dominant and happens at an early stage

The need for receiving emotional support and validation is dominant on the timelines. Many participants also wish to start to address it earlier in the process. About half participants place it as the first need to accomplish. P10 explained why she would like receive support first: *"I feel like the time after which the harm happens is when you're the most emotionally charged by this situation. So then, a comfortable space where you can ease in and grieve is pretty important."* P27 believes that she has an urgent need to get "kind words and maybe a hug" from friends and family: *"When I saw it [the offending post], I was upset for the longest time. And I really considered ways of trying to avoid going to school and miss class so I wouldn't have to see them [offenders] again."*

Some participants hope to have a conversation with offenders to get their acknowledgement or stop the continuation of harm. We find that those participants usually need to receive emotional support before facing offenders. P3 explained that emotional support should come before confronting the offender: *"Because going straight to the offender and be like, 'Hey, what you did was wrong,' isn't going to happen if the survivor doesn't feel supported enough."*

5.3 Timing to meet safety needs depends on the types of actions

We noticed that some participants deal with harm that is ongoing, and hope to gain safety as soon as possible: *"once online harm happen, I guess the first thing to do is to stop it (P21)."* For participants who hope to stop the harm immediately, they usually rely on platform moderation (P4, P12, P16, P22, P24, P25). P4 thinks that removal of content as soon as possible can stop the spread of harm: *"Firstly, I want the Instagram moderators to delete the comments and give the Instagram users who posted the disrespectful comments some warning so these negative comments won't let more people to see it and let this discriminative mood to spread around individuals."* When participants hope to stop the harm through talking with offenders and gain their understanding (P1, P3, P8, P15, P27) it usually happens (or is finally achieved) at a later stage of the timeline. Some participants believed that it takes time for offenders to understand the harm, thus they hope to talk with offenders later: *"The offenders won't realize that what they've done was wrong at the time after they post the bad information. I will give them some space to let them understand what they've done was wrong."* Additionally, we talked about how P3 needs emotional support before talking with offenders (5.2).

As we mentioned in 4.4.3, sometimes participants are not sure how they can stop the harm. We find that some participants place

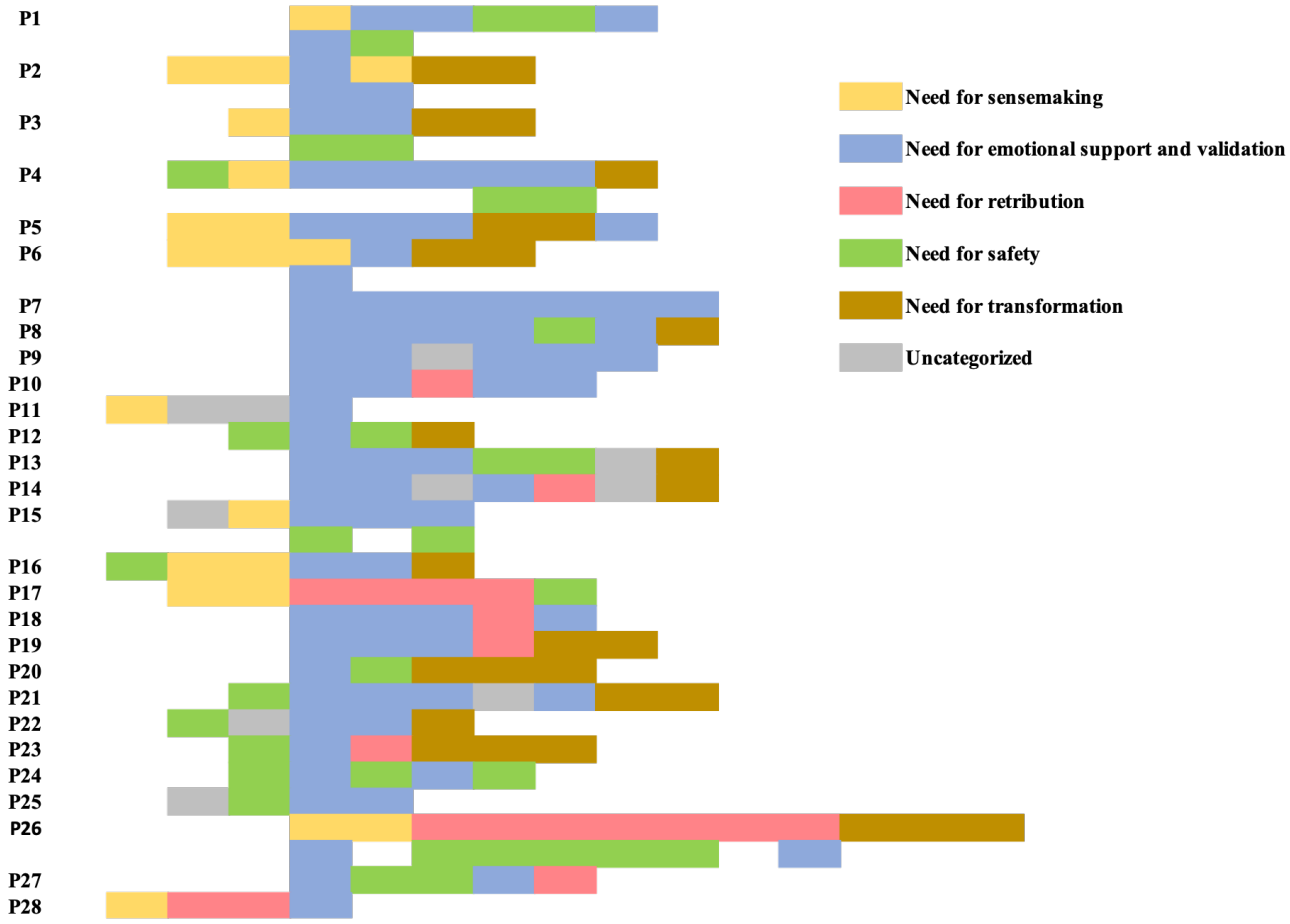

Figure 2: Participants' timelines to address the needs. Each rectangle represents an action on the timeline. We color-coded the actions according to which category the participant's need falls into – for an action that represents more than one need, we used stacked notes. We aligned the timelines according to the need for emotional support and validation.[1]

actions to stop the harm at multiple stages of the timeline. For example, P24 hopes to first inform the platform and to get the offensive post removed, but he also hopes the school can tell offenders to stop their actions later. He admitted that ideally the offenders should stop the continuation of harm immediately, but *"It will take some time to talk to the students and all of that."*

5.4 Needs for retribution and transformation come last

We noticed that retribution is not the central goal for many participants. Less than half of our participants include the need for retribution in their timeline of addressing harm, and they usually place the need after meeting the need for sensemaking, emotional support and validation, and safety.

Finally, participants placed their need for transformation toward the end of their timelines. Participants told us that they hope the transformation of online environments is not only to address their harm case, but instead will prevent future harm: *"This one [the transformation need] definitely comes more at the end because this is more of trying to prevent future things from happening"(P3).* P2 believes that besides addressing past harm, people should "learn from this experience." Some participants indicated that the need for transformation are less for themselves and the harm they have experienced. Instead, they hope that fundamentally changing the online environment can benefit people they care about: "I hope that there are less and less victims that will be harmed."
The need for transformation that our participants mentioned echoes calls for transformative justice [11, 25]. Rather than focusing on cases of harm individually, transformative justice seeks to reveal and address root causes and cultures of violence and harm in society. Similarly, our participants discussed a longer term need to change the conditions that enabled harm to happen in the first place.

[1]For a grayscale version, please visit https://applexiao.com/images/timelines.jpeg

6 DISCUSSION

6.1 Expansion of the scope of needs, stakeholders and actions in addressing online harm

Starting from victims' needs in the restorative justice literature, we identified five needs adolescents have for addressing online harm: sensemaking, support and validation, safety, retribution, and transformation. We also find that moderation actions, such as content removal or bans, meet some participants' needs for safety or retribution. This last finding aligns with existing research that shows that moderation grants participants safety [48] and can hold offenders accountable through retribution [49].

While content moderation is currently the major tool online platforms use to address harm, our findings suggest other ways platforms can help. In fact, participants mentioned online platforms' role in addressing all five needs we identified above. Participants believe that platforms can help them make sense of what has happened, validate their experiences of harm, and transform online environment to prevent future harm from happening. Further, participants proposed specific actions from online platforms that could help, such as providing instructions and advice on how to address harm, and showing their stance against harm. These proposed actions are useful steps towards concrete design solutions to online harm.

While platforms have a responsibility to their users to address harm, it is also clear from our interviews that victims need more than what platforms alone can offer. Social media companies, as well as current research, typically consider individual platforms or communities as the primary (or only) site for addressing harm. However, the restorative justice approach conceptualizes harm as an interconnected web of relationships, creating obligations for stakeholders (including offenders and members of related social circles) to address the harm collectively [65]. When we asked participants to choose which stakeholder(s) to engage with in order to address a harm, they often identify multiple stakeholders online and offline. This broader network of stakeholders requires us to think about online harm not as isolated incidents, but events that connect individuals experience with their relevant communities beyond a particular online platform. This recognition further emphasizes the importance of customization and flexibility when providing support [5, 49].

The multi-stakeholder perspective also reveals the importance of utilizing available social capital and resources that people have in their social circles [10, 36]. Kretzmann and McKnight explain that, "It is the capacities of local people and their associations that build powerful communities" [35]. For many victims we interviewed, their process of addressing harm involves stakeholders and resources from multiple social circles which may or may not directly relate to the community where a specific harm occurred. It is important to note that the multi-stakeholder approach doesn't alleviate the responsibility of online platforms to protect their users. For example, one concrete suggestion is for platforms to directly point victims to the internal and external social resources they might need for addressing harm.

The involvement of multiple stakeholders can be particularly important for adolescents, our focal population. We find that their online harm experiences may happen between schoolmates, or even include an offline component at school or other extracurricular events. As a vulnerable group, adolescents often need the help of their parents or schools in dealing with harm. When online and offline harm intersect in schools, it creates a grey area of obligations between the school, the platform, and parents. Existing restorative justice practices for adolescents are often a collaborative effort between school and parents [19]. In the HCI and CSCW literature, researchers show how involvement of parents may benefit adolescents in dealing with online safety issues [13, 62]. Our findings support this line of research, and emphasize the importance of interpersonal and familial relationships for adolescents in order to prevent or respond to online harm.

Our research provides insight into how online platforms and communities might implement procedures that enact restorative justice values and processes. Participants identified many actions that can potentially embed restorative justice values. For example, support from bystanders, online community members, or society at large is important for acknowledging the harm and empowering victims. Instead of punishing the offenders to stop the harm, some participants explained that it might be possible to stop further harm through offenders' growth: offenders can learn from the harm and grow through conversations with victims and other online community members, or learn from schools and society at large. We believe these restorative measures are particularly important for the health and growth of adolescents who have experienced harm, or who caused harm to others. While restorative justice has demonstrated success in achieving those goals within schools [19], we believe online platforms are an important social context where even small, visible changes in our response to harm can have a positive effect on adolescents.

6.2 Understanding harm and need through a temporal perspective

When do online harm victims start to address the harm, and when does the process end? Our research shows that participants often deal with ongoing harm, or expect the harm to happen again even if it has temporarily stopped. While current research often conceptualizes harm as discrete incidents and designs solutions, it is important to take continuous or ongoing harm into consideration as well.

Participants' process of addressing harm often involves a series of actions over time. They also show preferences for addressing some needs in a particular order. Participants may hope to talk with offenders, but it's only after they gained emotional support; or they need to ensure safety before addressing other needs. Thus, it is important to consider both the relation of needs and the timing of meeting each need. For example, participants mentioned that sometimes they often need to make sense of what happened before deciding how to address the harm. Thus, besides designing solutions that focuses on the outcome, it is also important to design ways to help users make informed decisions on how to address the harm. We also find that several participants see moderation as an efficient way to ensure safety right after a harm occurs. While

much research examines *how* to moderate, our research shows the importance of *when* to moderate (i.e., moderation efficiency). Our participants expressed that they hoped for efficiency among both human moderators or automatic moderation tools. For this reason, we believe this is also relevant to the studies of labor and human moderators [44], and automation in moderation [8, 17].

In addition to short term needs and actions to address a specific harm, many participants explained that they wanted transformation of online environments in the long term. The broader need for transformation may be less apparent (and seem less urgent) compared to more immediate needs in an individual harm case. However, our interviews reveal that larger transformative changes are no less important to some individuals. In fact, many participants put it at the end of the timeline but describe the need as fundamental. In some ways, this is akin to wanting justice for a specific crime, but also wanting to change the laws and social norms that allow such crimes to regularly occur.

Since the focus of restorative justice is on interpersonal relationships [65], its effect on transformation may be limited. Participants' need for transformation include issues such as raising public awareness about online harms, and changing the platform's design to prevent future harms. These insights raise the potential of focusing on transformative justice in future research on online harm. Transformative justice aims to address structural conditions and root causes that enable harm to happen [11]. The challenge with taking a transformative justice approach is that unlike restorative justice, transformative justice is not well codified and does not have a set of established practices to build on. Instead, transformative justice is an open-ended, community based approach to understanding and addressing root causes while changing dominant, harmful cultures [25].

6.3 Reflection on our method: what we learn from victims' process of identifying needs

Our research process centers on the experiences of victims, and gives them agency to explain how they would want to deal with specific experiences of online harm. Of course, victims' proposed actions should not be conflated with specific, actionable implications that should be implemented directly. Instead, we argue that we should consider their suggestions as they relate to the values and norms within the relevant online communities and platforms.

There are many existing approaches to addressing online harms, and the most common actions such as banning, muting, and removing content are primarily punitive. Such punitive actions tend to focus on after-the-fact removal of offending content or the person who has broken the rules. In our research, some participants proposed such punitive actions, and some explicitly expressed the need to punish the people who hurt them. This should be expected when punitive approaches are the dominant way to address harm both on online platforms as well as in society more broadly [15]. The dominance of punitive actions creates a dilemma for people who want to grant agency to survivors, but also aspire to the values of restorative justice. Thus, we must emphasize the difference between taking a survivor-centered approach, with one that is survivor-led. Taking a survivor-led approach means acting on exactly what a survivor of harm asks us to do; however, in a survivor-centered approach, we listen to survivors and work to meet their need to heal from the harm within the framework of established values. Therefore, it is important to establish agreement on shared values before implementing processes like restorative justice.

Our interviews used a multi-stage design task to capture participants' needs for addressing harm. We found that participants usually do not know how to address the harm immediately. They constantly reflected and came up with needs throughout the task, went back and forth in the procedures to add stakeholders or actions, and re-arranged the items on a timeline. In fact, many identified sensemaking as one of their needs in addressing harm — indicating that they may not know what has happened or know how to address the harm at the onset. We also found that the same action participants identified may relate to several different needs: for example, content moderation can satisfy retribution or safety; offenders' acknowledgement may either provide validation or act as a way to stop the continuation of harm (see Table 2). Therefore, identifying needs and actions is labor intensive and instead of expecting victims to provide us with direct, actionable solutions, it is important for us to provide time, support, and resources to them. Besides asking what actions they prefer, it is important to work to understand their underlying motives and needs. We believe that building on restorative justice values and processes can enable researchers, technologists, policy makers, and platforms to engage with victims of online harm in respectful ways that offer support and compensation as they work together toward designing new ways to effectively address online harm.

6.4 Limitation and Future Work

Our recruitment and interview processes have some limitations. First, the need-finding questions and examples, as well as the order they are presented in, have the potential to influence the responses of participants. Second, online harm is a broad topic; the cases participants shared do not cover all types of harm cases. Third, our participants may not be representative of all adolescents. Most participants are Asian. All participants are in their late adolescence (18-20) and study in a university in the United States. In future work, we plan to conduct large-scale surveys among adolescents to cover a wide range of online harm experiences and demographics. We also plan to use the survey to examine how people's experiences of harm relates to the needs, actions and stakeholders for addressing harm.

Our research has identified resources victims can mobilize in addressing harm, including stakeholders and what victims need from them. However, as we mentioned in 6.3, researchers and platform designers need to continue to work with victims to transform those ideas into design solutions. In particular, our research has shown potential of using restorative justice in addressing online harm. Applying restorative justice principles to online platforms is not easy or straightforward. Restorative justice processes can be time consuming, and may require participation and effort from diverse parties including restorative justice facilitators, offenders and other community members[6]. In addition, a restorative justice process is often voluntary and consensual, thus it does not concern offenders who do not plan to engage [65]. Those issues present new challenges to the current online landscape which experiences

insufficient moderation labor and expanding communities [16, 44]. While our research has showed the potential of online restorative justice, it is important to bear those constraints in mind in future design and implementation.

7 CONCLUSION

In this research, we identified adolescents' need for addressing online harm, including sensemaking, support and validation, safety, retribution, and transformation. Our findings shed light on how online platforms may support victims beyond moderation, and show how we can design for victims' needs beyond the scope of online platforms, or short term solutions. Additionally, we see the potential of restorative justice in understanding and addressing adolescents' needs in online harm. How we may design for those needs and implement restorative justice principles in online platforms is challenging, yet important future work.

ACKNOWLEDGMENTS

We would like to thank our research participants and the early participants who helped us pilot the study. This work was supported by the National Science Foundation award IIS-1948067.

REFERENCES

[1] Monica Anderson. 2018. A majority of teens have experienced some form of cyberbullying. (2018).
[2] Zahra Ashktorab and Jessica Vitak. 2016. Designing cyberbullying mitigation and prevention solutions through participatory design with teenagers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3895–3905.
[3] Reuben Binns, Michael Veale, Max Van Kleek, and Nigel Shadbolt. 2017. Like trainer, like bot? Inheritance of bias in algorithmic content moderation. In *International conference on social informatics*. Springer, 405–415.
[4] Lindsay Blackwell, Tianying Chen, Sarita Schoenebeck, and Cliff Lampe. 2018. When online harassment is perceived as justified. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 12.
[5] Lindsay Blackwell, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. 2017. Classification and its consequences for online harassment: Design insights from heartmob. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–19.
[6] Jane Bolitho and Jasmine Bruce. 2017. Science, art and alchemy: Best practice in facilitating restorative justice. *Contemporary Justice Review* 20, 3 (2017), 336–362.
[7] Jie Cai and Donghee Yvette Wohn. 2019. What are Effective Strategies of Handling Harassment on Twitch? Users' Perspectives. In *Conference companion publication of the 2019 on computer supported cooperative work and social computing*. 166–170.
[8] Eshwar Chandrasekharan, Chaitrali Gandhi, Matthew Wortley Mustelier, and Eric Gilbert. 2019. Crossmod: A cross-community learning-based system to assist reddit moderators. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–30.
[9] Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. 2018. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 32.
[10] Alexander Cho, Roxana G Herrera, Luis Chaidez, and Adilene Uriostegui. 2019. The" Comadre" Project: An Asset-Based Design Approach to Connecting Low-Income Latinx Families to Out-of-School Learning Opportunities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
[11] Erin Daly. 2001. Transformative justice: Charting a path to reconciliation. *Int'l Legal Persp.* 12 (2001), 73.
[12] Dominic DiFranzo, Samuel Hardman Taylor, Franccesca Kazerooni, Olivia D Wherry, and Natalya N Bazarova. 2018. Upstanding by design: Bystander intervention in cyberbullying. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.
[13] Lee B Erickson, Pamela Wisniewski, Heng Xu, John M Carroll, Mary Beth Rosson, and Daniel F Perkins. 2016. The boundaries between: Parental involvement in a teen's online world. *Journal of the Association for Information Science and Technology* 67, 6 (2016), 1384–1403.
[14] Jesse Fox and Wai Yen Tang. 2017. Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. *New media & society* 19, 8 (2017), 1290–1307.
[15] David Garland. 2012. *Punishment and modern society: A study in social theory*. University of Chicago Press.
[16] Tarleton Gillespie. 2018. *Custodians of the Internet*. Yale University Press.
[17] Tarleton Gillespie. 2020. Content moderation, AI, and the question of scale. *Big Data & Society* 7, 2 (2020), 2053951720943234.
[18] Kishonna L Gray. 2012. Deviant bodies, stigmatized identities, and racist acts: Examining the experiences of African-American gamers in Xbox Live. *New Review of Hypermedia and Multimedia* 18, 4 (2012), 261–276.
[19] Belinda Hopkins. 2002. Restorative justice in schools. *Support for Learning* 17, 3 (2002), 144–149.
[20] Maggie Hughes and Deb Roy. 2020. Keeper: An Online Synchronous Conversation Environment Informed by In-Person Facilitation Practices. In *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing*. 275–279.
[21] Shagun Jhaver, Darren Scott Appling, Eric Gilbert, and Amy Bruckman. 2019. "Did You Suspect the Post Would Be Removed?": Understanding User Reactions to Content Removals on Reddit. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 192 (Nov. 2019), 33 pages. https://doi.org/10.1145/3359294
[22] Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 2019. Does transparency in moderation really matter? User behavior after content removal explanations on reddit. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–27.
[23] Shagun Jhaver, Sucheta Ghoshal, Amy Bruckman, and Eric Gilbert. 2018. Online harassment and content moderation: The case of blocklists. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 2 (2018), 1–33.
[24] Gerry Johnstone and Daniel Van Ness. 2013. *Handbook of restorative justice*. Routledge.
[25] Mariame Kaba and Naomi Murakawa. 2021. *We Do this' til We Free Us: Abolitionist Organizing and Transforming Justice*. Haymarket Books.
[26] David R Karp and Beau Breslin. 2001. Restorative justice in school communities. *Youth & Society* 33, 2 (2001), 249–272.
[27] Barbara Katic, Laura A Alba, and Austin H Johnson. 2020. A systematic evaluation of restorative justice practices: school violence prevention and response. *Journal of school violence* 19, 4 (2020), 579–593.
[28] Charles Kiene and Benjamin Mako Hill. 2020. Who Uses Bots? A Statistical Analysis of Bot Usage in Moderation Teams. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
[29] Etienne G Krug, James A Mercy, Linda L Dahlberg, and Anthony B Zwi. 2002. The world report on violence and health. *The lancet* 360, 9339 (2002), 1083–1088.
[30] Jeff Latimer, Craig Dowden, and Danielle Muise. 2005. The effectiveness of restorative justice practices: A meta-analysis. *The prison journal* 85, 2 (2005), 127–144.
[31] Amanda Lenhart, Michele Ybarra, Kathryn Zickuhr, and Myeshia Price-Feeney. 2016. *Online harassment, digital abuse, and cyberstalking in America*. Data and Society Research Institute.
[32] Kaitlin Mahar, Amy X Zhang, and David Karger. 2018. Squadbox: A tool to combat email harassment using friendsourced moderation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.
[33] Abraham Harold Maslow. 1943. A theory of human motivation. *Psychological review* 50, 4 (1943), 370.
[34] Binny Mathew, Anurag Illendula, Punyajoy Saha, Soumya Sarkar, Pawan Goyal, and Animesh Mukherjee. 2020. Hate begets hate: A temporal study of hate speech. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–24.
[35] Alison Mathie and Gord Cunningham. 2003. From clients to citizens: Asset-based community development as a strategy for community-driven development. *Development in practice* 13, 5 (2003), 474–486.
[36] Alison Mathie and Gord Cunningham. 2005. Who is driving development? Reflections on the transformative potential of asset-based community development. *Canadian Journal of Development Studies/Revue canadienne d'études du développement* 26, 1 (2005), 175–186.
[37] David C McClelland. 1987. *Human motivation*. CUP Archive.
[38] Sharan B Merriam and Robin S Grenier. 2019. *Qualitative research in practice: Examples for discussion and analysis*. Jossey-Bass.
[39] Allison Morris and Gabrielle M Maxwell. 1993. Juvenile justice in New Zealand: A new paradigm. *Australian & New Zealand Journal of Criminology* 26, 1 (1993), 72–90.
[40] Kavous Salehzadeh Niksirat, Evanne Anthoine-Milhomme, Samuel Randin, Kévin Huguenin, and Mauro Cherubini. 2021. "I thought you were okay": Participatory Design with Young Adults to Fight Multiparty Privacy Conflicts in Online Social Networks. In *Designing Interactive Systems Conference (DIS)*.
[41] Jessica Pater and Elizabeth Mynatt. 2017. Defining digital self-harm. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1501–1513.
[42] Richard Stanley Peters. 2015. *The concept of motivation*. Routledge.
[43] Kay Pranis. 2015. *Little book of circle processes: A new/old approach to peacemaking*. Simon and Schuster.
[44] Sarah T Roberts. 2019. *Behind the screen*. Yale University Press.

[45] Oliver C Robinson. 2014. Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative research in psychology* 11, 1 (2014), 25–41. Publisher: Taylor & Francis.
[46] Johnny Saldaña. 2021. *The coding manual for qualitative researchers.* sage.
[47] Susan M Sawyer, Peter S Azzopardi, Dakshitha Wickremarathne, and George C Patton. 2018. The age of adolescence. *The Lancet Child & Adolescent Health* 2, 3 (2018), 223–228.
[48] Morgan Klaus Scheuerman, Jialun Aaron Jiang, Casey Fiesler, and Jed R Brubaker. 2021. A Framework of Severity for Harmful Content Online. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–33.
[49] Sarita Schoenebeck, Oliver L Haimson, and Lisa Nakamura. 2021. Drawing from justice theories to support targets of online harassment. *new media & society* 23, 5 (2021), 1278–1300.
[50] 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. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–18.
[51] Joseph Seering, Robert Kraut, and Laura Dabbish. 2017. Shaping pro and anti-social behavior on twitch through moderation and example-setting. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 111–125.
[52] Joseph Seering, Tony Wang, Jina Yoon, and Geoff Kaufman. 2019. Moderator engagement and community development in the age of algorithms. *New Media & Society* 21, 7 (2019), 1417–1443.
[53] Donna Spencer. 2009. *Card sorting: Designing usable categories.* Rosenfeld Media.
[54] Sharifa Sultana, Mitrasree Deb, Ananya Bhattacharjee, Shaid Hasan, SM Raihanul Alam, Trishna Chakraborty, Prianka Roy, Samira Fairuz Ahmed, Aparna Moitra, M Ashraful Amin, et al. 2021. 'Unmochon': A Tool to Combat Online Sexual Harassment over Facebook Messenger. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.
[55] Zenon Szablowinski. 2008. Punitive justice and restorative justice as social reconciliation. *The Heythrop Journal* 49, 3 (2008), 405–422.
[56] Samuel Hardman Taylor, Dominic DiFranzo, Yoon Hyung Choi, Shruti Sannon, and Natalya N Bazarova. 2019. Accountability and empathy by design: Encouraging bystander intervention to cyberbullying on social media. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.
[57] Jirassaya Uttarapong, Jie Cai, and Donghee Yvette Wohn. 2021. *Harassment Experiences of Women and LGBTQ Live Streamers and How They Handled Negativity.* Association for Computing Machinery, New York, NY, USA, 7–19. https://doi.org/10.1145/3452918.3458794
[58] Daniel W Van Ness. 2016. An overview of restorative justice around the world. (2016).
[59] Jessica Vitak, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab. 2017. Identifying women's experiences with and strategies for mitigating negative effects of online harassment. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1231–1245.
[60] Emily Vogels. 2021. The state of online harassment. *Pew Research Center* (2021).
[61] Karl E Weick. 1995. *Sensemaking in organizations.* Vol. 3. Sage.
[62] Pamela J Wisniewski, Heng Xu, Mary Beth Rosson, and John M Carroll. 2014. Adolescent online safety: the" moral" of the story. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1258–1271.
[63] Richmond Y Wong and Tonya Nguyen. 2021. Timelines: A World-Building Activity for Values Advocacy. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.
[64] William R Wood and Masahiro Suzuki. 2016. Four challenges in the future of restorative justice. *Victims & Offenders* 11, 1 (2016), 149–172.
[65] Howard Zehr. 2015. *The little book of restorative justice: Revised and updated.* Simon and Schuster.