

Teens, Privacy, and Algorithms: Navigating and Co-Designing Solutions for Interpersonal Boundary Management on Social Media

Jinkyung Katie Park
School of Computing
Clemson University
Clemson, South Carolina, USA
jinkyup@clemson.edu

Renkai Ma
Vanderbilt University
Nashville, Tennessee, USA
renkai.ma@vanderbilt.edu

Naima Samreen Ali
Department of Computer Science
Vanderbilt University
Nashville, Tennessee, USA
naima.samreen.ali@vanderbilt.edu

Naulsberry Jean Baptiste
College of Engineering and Computer
Science
University of Central Florida
Orlando, Florida, USA
naulsberry@knights.ucf.edu

Zainab Agha
Department of Computer Science
San Francisco State University
San Francisco, California, USA
zainab.gha@vanderbilt.edu

Pamela J. Wisniewski
Socio-Technical Interaction Research
Lab
Orlando, Florida, USA
pamwis@stirlab.org

Abstract

On social media, teens must manage their interpersonal boundaries not only with other people, but also with the algorithms embedded in these platforms. In this context, we engaged seven teens in an Asynchronous Remote Community (ARC) as part of a multi-year Youth Advisory Board (YAB) to discuss how they navigate, cope, and co-design for improved boundary management. Teens had pre-conceived notions of different platforms and navigated boundaries based on specific goals; yet, they struggled when platforms lacked the granular controls needed to meet their needs. Teens enjoyed the personalization afforded by algorithms, but they felt violated when algorithms pushed unwanted content. Teens designed features for enhanced control over their discoverability and for real-time risk detection to avoid boundary turbulence. We provide design guidelines for improved social media boundary management for youth and pinpoint educational opportunities to enhance teens' understanding and use of social media privacy settings and algorithms.

CCS Concepts

• Human-centered computing → Empirical studies in HCI.

Keywords

Adolescent Online Safety, Social Media, Interpersonal Boundary, Privacy, Algorithms, Asynchronous Remote Community, Participatory Design

ACM Reference Format:

Jinkyung Katie Park, Renkai Ma, Naima Samreen Ali, Naulsberry Jean Baptiste, Zainab Agha, and Pamela J. Wisniewski. 2025. Teens, Privacy, and Algorithms: Navigating and Co-Designing Solutions for Interpersonal Boundary Management on Social Media. In *Interaction Design and Children*



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

IDC '25, Reykjavik, Iceland

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1473-3/25/06

<https://doi.org/10.1145/3713043.3728840>

(IDC '25), June 23–26, 2025, Reykjavik, Iceland. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3713043.3728840>

1 Introduction

Increased social media usage among teenagers has made them vulnerable to various online risks ranging from unwanted sexual solicitations and cyberbullying to privacy violations and exposure to risky content [87]. As a result, managing interpersonal boundaries has become crucial for teens to help define who they interact with online, how these interactions take place, and when they need interpersonal boundaries [92]. *Interpersonal boundary management*, a concept from privacy research [11], describes how individuals manage their interactions by adjusting their level of openness with others. Adolescence represents a distinct developmental phase bridging childhood and adulthood, with a primary objective of growing toward independence and autonomy,¹ where one of their most important developmental tasks is to acquire proficiency in managing online interactions to be resilient against online risks [34, 90]. From a developmental perspective, some level of risk-taking, autonomy-seeking, and self-regulation is a natural and necessary part of adolescence [15, 36, 102]. Preventing such experiences in the name of protecting teens from potential mild risks may impede their healthy growth toward adulthood, which teens also perceived as privacy-invasive [90]. Thus, we need a better understanding of how teens' interpersonal boundaries are violated, how they can independently set and manage boundaries, and how social media platforms can support their boundary management.

In the recent work in the SIGCHI community, including IDC, researchers have involved teens in co-designing for privacy management [4, 50, 100] or preventing unwanted contact [13, 14] on social media. Yet, there remains a need for more studies to explore teens' real-world experiences—how they navigate interpersonal

¹Developmentally, adolescents are characterized as between the ages of 10 and 19. In the digital privacy literature in the U.S. context, however, the age range for adolescents is often described as between the ages of 13 and 17, focusing on the legal implications of being a minor in the U.S. [24]

boundaries in their everyday interactions across social media platforms. Moreover, privacy research highlights that managing interpersonal boundaries on social media involves not only people but also algorithms that shape what content is visible to them [62, 64]. While boundary management has been explored in the context of peer disclosure or parental monitoring [47, 91, 96], few studies have examined how teens understand and regulate these boundaries with respect to the automated decisions made by social media algorithms. Our work addresses this gap by investigating how teens define, experience, and respond to boundary challenges involving both human and algorithmic actors through the lens of interpersonal boundary regulation [11, 70], while foregrounding teens' voices in co-designing safety features that reflect their needs for agency. We ask three research questions:

- **RQ1:** *How do teens navigate their interpersonal boundaries with 1) people and 2) algorithms on social media?*
- **RQ2:** *What interpersonal boundary violations occur and how do teens cope with violations?*
- **RQ3:** *How do teens envision re-designing social media to overcome interpersonal boundary violation challenges?*

To explore these questions, we engaged seven teens in an eight-week Asynchronous Research Community (ARC) on a private Discord server (February–March 2023). Teens completed online safety activities at their own pace, beginning with reflections on how they navigate interpersonal boundaries with people and algorithms on social media, followed by design tasks where they proposed their own solutions. We found that teens actively manage these boundaries while balancing different goals across social contexts and platforms (RQ1). They encountered boundary violations from people through unwanted contact or exposure, and from algorithms due to opaque rule-setting and lack of control (RQ2). In response, they proposed features such as audience-specific privacy controls and real-time feedback mechanisms to address algorithmic harms, offering actionable insights for designing more supportive and autonomy-preserving online experiences (RQ3). Based on our findings, we suggest design implications that can support teens with healthy boundary-setting with people and algorithms on social media. We make the following contributions to the IDC and broader SIGCHI community:

- We offer an integrated understanding of how teens navigate boundaries with both people and algorithms to support their online safety on social media, drawing on the lens of *interpersonal boundary management*.
- We demonstrate the value of multi-year research with teens to capture insights into their lived online experience of interpersonal boundary management and generate design ideas and artifacts.
- We provide evidence-based educational strategies and design guidelines to help social media platforms better support teens' autonomy and safe boundary management with both people and algorithms.

The remainder of the paper is structured as follows: we begin by reviewing related work on adolescent privacy, boundary management, and human-algorithm interaction. We then describe our methodological approach using an Asynchronous Research Community (ARC) to engage teens in reflective and participatory design

activities. The findings section presents how teens navigate boundary tensions and the strategies they propose to manage them. We conclude with a discussion of theoretical implications, design recommendations, and future directions for supporting youth digital resilience.

2 Background: Promoting Adolescent Online Safety through HCI Research

Teens's online safety encompasses addressing a wide range of risks, including exposure to harmful content [53, 80], cyberbullying [13], and interpersonal privacy risks [8, 33, 58]; as online interactions become integral to youth's digital life, HCI research has increasingly focused on addressing these diverse risks to enhance adolescent online safety through developing evidence-based interventions (e.g., [3, 7, 29, 58, 78, 80]). Early work on adolescent online safety has emphasized *restrictive mediation*, such as parental control apps that parents use to limit their children's access to social media or establish rules limiting media consumption [48, 79]. While such approaches can be effective for young children and early adolescents (e.g., aged 8-12) [26], as adolescents mature, they require opportunities to develop skills in risk assessment, problem-solving, and independently managing potential online risks [69, 93, 95]. In response, adolescent online safety research has shifted toward a teen-centered approach where teens have more autonomy in making their own online safety decisions. For instance, 'just-in-time' nudge-based interventions (e.g., subtle cues intended to influence behavior) have been proposed to support teens at the moment when they face risks online [2, 3, 58]. Researchers have also explored ways to work with teens to collect ecologically valid online risk data to build teen-centric machine-learning (ML) systems (e.g., [10, 68, 74]) to design teen-centered, context-aware interventions to enhance adolescent online safety.

To empower teens to self-regulate and cope in the face of online risks, researchers have often adopted synchronous research approaches such as participatory design (PD) to reflect on teens' online experiences and design teen-centered online safety solutions (e.g., [6, 13, 22]). To overcome the challenges in such a synchronous approach, such as conflicting schedules to partake in research, inconsistent attendance [6], and limited time for teens to articulate thoughtful responses [31], researchers have increasingly recognized the value of the Asynchronous Remote Community (ARC). ARC, an alternative participatory design approach, offers an asynchronous online space for researchers and participants to engage in research activities at their own pace, providing greater flexibility and thoughtful contributions [45, 55]. Recently, researchers started to apply the ARC methods to co-design solutions with teens for stress management and depression [17], and voice technology [37]. Yet, the potential of applying it to explore teens' interpersonal boundary management on social media is under-explored. To fill this gap, we designed an ARC as an engaging environment for teens to actively discuss their boundary management on social media.

3 Conceptualizing Interpersonal Boundary Management through Privacy Research

Interpersonal boundary management, refers to a process by which a person or group regulates interaction with others by altering

the degree of openness, as conceptualized by Altman [11], where individuals open and close their boundaries to achieve a desired level of privacy and social interaction. Petronio [71] further expands on Altman's framework by formulating the communication privacy management (CPM) theory and proposes that managing privacy involves boundary coordination around personal information. Using a metaphor of a "boundary," CPM presents that boundary management is a dynamic and rules-based process [84]. CPM explicates that individuals' decisions on how interpersonal boundaries are managed follow the rules of "boundary opening" and "boundary closure" [84], depending on the degree of risk associated with privacy. CPM also uses the concept of "boundary turbulence" to describe the breakdown of interpersonal boundaries when expectations or rules around privacy (e.g., information disclosure) are violated [71]. Therefore, individuals re-negotiate their boundaries with other parties following their experiences with the boundary turbulence. Our study builds on this by examining teens' interpersonal boundary management on social media, specifically on how teens (re)negotiate boundaries with others and algorithms.

3.1 Interpersonal Boundary Management in Teens' Privacy on Social Media

Foundational work on teens' privacy on social media described teens' boundary management as a dynamic and co-constructed process [57]. In a broader sense, interpersonal boundaries help teens define who they are, who they interact with, how, and when [92], and these boundaries are influenced by social and technical contexts such as platform affordances, family, or school environments [39]. Teens manage such boundaries through various strategies, such as deleting friends, using fake names, deleting content, withholding information, etc. [30, 42, 56]. These practices reflect teens' understanding of leveraging different platform features (e.g., privacy settings [52]) to selectively share their information online [49, 60, 98], often vary across different platforms [86]. For instance, Instagram's asymmetrical relationships can lead to riskier behaviors, while Snapchat's symmetrical model can encourage more careful curation of friends for teens [99]. Therefore, examining how teenagers manage their interpersonal boundaries across individual social media platform is critical.

Research has shown teens' privacy management practices on specific social media platforms, particularly revealing unexpected directions [1, 27, 65, 97]. A notable example is the creation of "Finsta" accounts on Instagram—secondary accounts for teens to share content they prefer not to associate with their public identity and family members, allowing for intimate self-expression and identity exploration [27, 97]. Managing privacy is less critical when teens coordinate boundaries with peers than when dealing with parental involvement [96]. Older teens often adopt a "nothing to hide" mindset, diminishing the perceived importance of privacy [1]. Despite existing efforts, how teens make sense of boundary management across different platforms and how platforms can better support boundary management remain underexplored. Our study aims to address this gap by exploring teens' real-world experiences in inform platform designs that support them in managing interpersonal boundaries with people on social media.

3.2 Interpersonal Boundary Management in Teens' Algorithmic Experience

As social media content consumption becomes the most popular media activity among children aged four to eighteen [63], researchers are increasingly examining how algorithms shape teens' experiences on those platforms. Algorithms can enhance teens' online experiences, such as supporting youth's informal learning [44], self-expression through content creation and sharing [61, 77], and engaging contemporary culture and creating online communities with peers [103]. At the same time, algorithms can pose significant risks to teens. For instance, YouTube's and TikTok's recommendation algorithms often suggest inappropriate content to youth, such as animal cruelty and violence [41, 53, 67]. These conflicting algorithmic experiences underscore the importance of teens setting interpersonal boundaries with algorithms related to what content they engage with; as social media algorithms influence users' content feeds, users also shape these algorithms through their interactions with content [12, 81]. Prior research shows that adult social media users are aware that algorithms shape the content they see, so they make sense of how algorithms work [35, 73]. However, it remains unclear how teens make sense of this and manage their interpersonal boundaries with algorithms.

A few recent studies have explored how young children understand and manage their interpersonal boundaries with recommendation algorithms [46, 88]. While studies on younger children highlight their limited understanding of algorithmic decision-making [49, 88, 101], research on teenagers suggests they have more agency in their interactions with algorithmic systems, navigating both their benefits and limitations [19], positioning them as active navigators of algorithms on social media. These create opportunities to explore how curated content feeds and algorithmic practices shape teens' interpersonal boundary management on social media, highlighting the need for further research. For instance, teens view social media algorithms as accurate reflections of themselves and are willing to engage with them to shape their online experience, suggesting that adolescents are aware of and interact with algorithmic recommendations [59]. In addition, prior work showed that teens appreciate customized social media experiences and believe that algorithms, such as TikTok's, have become sophisticated enough to accurately reflect their interests, indicating that adolescents are aware of how their interactions affect content recommendations [25]. Yet, gaps remain in understanding how teens conceptualize and manage their interpersonal boundaries with algorithms, particularly in relation to content curation, data privacy, and risk exposure on social media. This study addresses these gaps by examining how teens negotiate their algorithmic boundaries and co-design solutions to regain agency over their social media experiences.

4 Methods

4.1 Study Overview

We engaged with seven teens in the Asynchronous Research Community (ARC) as part of a Youth Advisory Board (YAB) program to understand how they navigate, cope, and co-design for improved boundary management. YAB is a multi-year participatory research

effort (May 2022–August 2023) to empower teens by engaging them in online safety research to learn essential skills and knowledge to create their own online safety solutions. We conducted the ARC study towards the end of the YAB program (February 2023–March 2023), where we carried out asynchronous activities with teens (see Table 3) on Discord for eight weeks to discuss teens’ boundary management on social media and co-design solutions catering to their needs and preferences. Before participating in the ARC activities, as part of the larger YAB program, teens participated in two UX workshops (Aug. 2022–Nov. 2022) to learn key UX concepts including storyboarding, user personas, mind mapping, and using Figma to create interactive prototypes. We also held a couple of synchronous meetings with the teens (Oct. 2022–Nov. 2022) via Zoom where we introduced the concept of ARC methodology and discussed teens’ interest in online safety topics to discuss in the ARC before implementing the method. After participating in the above workshops and synchronous meetings, teens carried out ARC activities that researchers posted on Discord weekly and biweekly (February 2023–March 2023). Below we provide the details of our ARC environment and the activities/prompts for the the ARC study.

4.2 Asynchronous Research Community (ARC)

4.2.1 Design of Asynchronous Research Community. We chose “Discord”² as the medium for our ARC. Discord is popular among teens and supports various media types including pictures, GIFs, videos, images, etc., through private chats [32]. It also offers security features to create private communities, called ‘servers’, to organize activities, making it a prime choice for asynchronous research. For our study, we first created a private Discord server for our youth advisory board. Then, we created various channels to 1) facilitate coordination among the team (e.g. scheduling meetings, getting updates on tasks, etc.), and 2) carry out research activities (e.g. ARC activities on social media privacy, discussing research-related topics, etc.) (see Fig 1). We assigned the role “YAB Teen Members” to the teens, and the role “Teenovate Staff” to researchers, to make notifying and organizing posts made only for the teens easier on the server. Finally, following the recommendation from Bhattacharya et al. [16], we included a #random channel in the server meant for socializing and posting comments and topics outside of the board’s regular activities.

4.2.2 Activities Conducted on Asynchronous Research Community. We first conducted a brainstorming session with teen members where we introduced the ARC methodology, discussed their considerations for an ARC environment, and developed a list of online safety topics and activities they were interested in exploring. Based on teens’ input, we identified that teens were mostly interested in discussing topics pertaining to interpersonal privacy concerns on social media platforms. In accordance with our research questions, we finalized two sets of activities to 1) explore how they manage their interpersonal boundaries with people, and algorithms, along with the challenges they face (RQ1 and RQ2), and 2) design for interpersonal boundary management based on identified problems (RQ3). Each activity was further divided into prompts with detailed instructions for the teens to complete the tasks (see Table 3). We

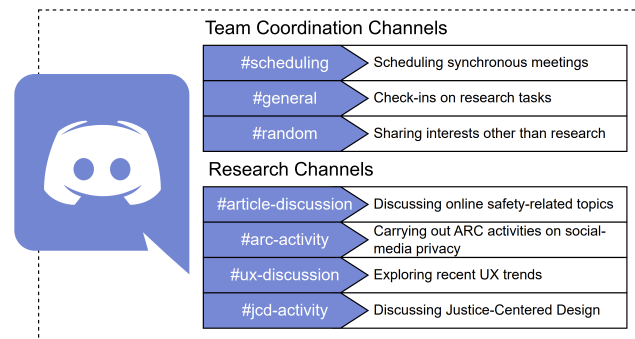


Figure 1: ARC environment on Discord

conducted the activities in our ARC channel via asynchronous weekly discussions on Discord and synchronous design sessions on Zoom. For RQ1 and RQ2, we conducted asynchronous activities using text responses and screenshots to explore teens’ opinions and perspectives on the safety and privacy features offered by their preferred social media platforms to manage interpersonal boundaries with both people and algorithms. For RQ3, we carried out a design activity on Figma where teens created their own design solutions to navigate interpersonal boundaries on social media. Asynchronous follow-up discussions were done in Discord for group feedback and suggestions on their continuing design work.

4.3 Participant Recruitment and Demographics

To ensure ethical standards, we obtained Institutional Review Board (IRB) approval from the last author’s institution. All team members completed CITI human subject research and Protection of Minors (POM) training. YAB participants were recruited through outreach campaigns targeting youth-serving organizations, high schools, and social media. Applicants completed a pre-screening survey to confirm they were 13–17 years old, fluent in English, had internet access, and could commit time to YAB activities. Researchers reviewed applications and conducted 15–30 minute interviews to assess applicants’ motivation, experience with online safety, creativity, and awareness of online trends. Finalists were selected, notified via email, and provided instructions to join a private Discord server. Each participant received a certificate and up to \$450 in gift cards after completing all activities. More details regarding considerations for participant recruitment and data ethics are explained in Section 8.

We had seven participants between the ages of 15–17, out of which four identified as male and three as female. Our participants self-reported as Caucasian ($n = 3$), East Asian ($n = 1$), Hispanic ($n = 1$), African American ($n = 1$), and South Asian ($n = 2$). The majority of the participants reported using technology “almost constantly” ($n = 4$), followed by “several times a day” ($n = 3$), with most of them using a laptop computer and their phones to go online. Table 2 in the Appendix shows our teen participants’ demographic information and technology use.

²<https://discord.com/>

4.4 Qualitative Data Analysis Approach

Data was collected through audio/video recordings from the initial ARC design session and its transcription, researcher observational notes on participants' engagement and interaction in the Discord server, and saved copies of each activity response (texts, screenshots, and designs) from the participants. We employed a reflexive approach to the thematic analysis [20]. We first familiarized ourselves with the data by reading through the posts and teens' design artifacts on Figma. This process included going back to the meeting recordings, inspecting their corresponding transcripts, and making notes of concepts and ideas. Through this initial process, the first author identified several topics for each research question and grouped them into themes. Once the initial themes were identified, the first two authors worked with the last author to conceptually organize themes that aligned with our research questions and theoretical framing. This process was iterative, involving constant discussion among the researchers. Finally, we reviewed our themes alongside our dataset to make sure that they actually captured important meanings within the data and named the themes. **Table 4** in the Appendix presents the themes along with illustrative quotations for each research question. Illustrative quotes from the participants are presented when describing the findings. Each quote is presented with the Participant ID (T= Participant ID) followed by their age and gender to ensure privacy and confidentiality.

5 Results

In this section, we present how teens navigate interpersonal boundaries with 1) people and 2) algorithms on social media (RQ1), what interpersonal boundary violations happen and how teens cope with those violations (RQ2), and how teens re-design social media to overcome interpersonal boundary violations (RQ3).

5.1 How Teens Navigate Their Interpersonal Boundaries with 1) People and 2) Algorithmic on Social Media (RQ1)

5.1.1 How Teens Manage Interpersonal Boundaries with People.

Teens categorized social media platforms based on their preconceived notions of the platforms to manage interpersonal boundaries with people. As active social media users, teens in our study accumulated preconceived notions of the platform and categorized the platforms based on the level of privacy afforded by the platforms. For instance, teens conceptualized popular platforms such as TikTok and Twitter as public platforms that allow users to share content and socialize publicly while affording limited choices to stay private. They perceived Instagram or Reddit as a semi-public platform that supports public content sharing while managing their privacy on granular levels (e.g., real name vs. anonymous username). Teens perceived Discord servers or Snapchat chatrooms as private platforms as they support users to socialize, share content, and communicate with a controlled and pre-set audience.

Teens made decisions on which platform to engage with people based on their goals of using social media platforms. We observed three primary goals: 1) to create and share content, 2) to consume online content, and 3) to socialize (1-to-1 or 1-to-many). For example, some teens like to actively create content as "influencers" and

seek ways to promote it on public platforms that they considered to support wide content reach (i.e., public content search on Twitter) over private platforms, as one teen shared:

"I am an influencer. I mainly use public social media sites, because I can reach more people. For example, a message or picture posted on a discord channel will only receive 30 likes max, but an Instagram post can receive tens of thousands of likes." – T1 (16-year-old, male)

Teens who liked to consume content preferred to use public or semi-public platforms that support the consumption of a variety of content. Teens who enjoy 1-to-many interaction also tended to choose public platforms with a large user base and various social features, whereas teens who like to have 1-to-1 interaction preferred using private platforms that support direct messaging. While teens often had multiple goals in choosing social media platforms, they prioritized certain goals considering platform affordances. For instance, a teen shared:

"Although I mainly use private social media sites, the time I do spend on public social media sites is to get updates on what peers are doing. I wouldn't consider Discord the best place for that." – T2 (16-year-old, male)

In this case, T2 was aware that public platforms support greater visibility and access to peer updates, while a private platform like Discord was perceived as affording more intimate interactions. Thus, platform affordances influenced teens' choices of platforms for their specific goals.

Teens gradually adapted to negotiating their boundaries with people on different platforms. Initially feeling socially pressured by peers to join public platforms, teens adapted by negotiating their interpersonal boundaries with people and gradually becoming comfortable with using these platforms. Some teens shared that they chose to be on popular social media platforms (mostly public platforms) even if those platforms did not support their desired privacy levels. For instance, a teen said:

"I think its become the norm to have platforms like Instagram or Snapchat, especially in high school. I feel like there is also a stereotype behind the moderated apps. Some of my friends think that the platform Discord is weird and that it's for "weird people." – T6 (15-year-old, female)

In this case, peer pressure to be on the popular platforms created a mismatch between what platform teens felt comfortable being on and what platform they were actually on. This fear of missing out motivated teens to "just go with whatever their peers are using" (T6). However, once teens joined public platforms, they learned how to use the features on those platforms to stay private over time. They outweighed the benefits of being on public platforms, such as being connected with a wider audience and consuming new content, over worrying about privacy concerns, as one teen shared:

"Up until a few months ago, I was still private on everything but as I've grown and opened up more, I've been comfortable with sharing my platform publicly. Through public platforms, I've found people with the same interests as me, same beliefs as me, and same

goals as me, so that's been extremely encouraging to experience." -T7 (16-year-old, female)

As such, sometimes, teens decided to be on certain social media platforms even if they initially did not afford the desired level of privacy. Yet, teens gradually learned to negotiate their interpersonal boundaries with people on those platforms by leveraging various platform features.

5.1.2 How Teens Navigate Boundaries With Social Media Algorithms.

Teens consumers enjoyed personalized content discovery through algorithms. The first thing that came to teens' minds when discussing algorithms on social media was recommendation algorithms. They showed a good understanding of how algorithms work by describing them as sets of rules that were learned from past data to achieve certain goals. All teens, as content consumers, described positive experiences with recommendation algorithms that curated content based on their past activities to align with their interests, perceiving them as effective in organizing topics and suggesting content based on their past activities on the platforms:

"An algorithm can provide additional benefits with use. It can provide the user with information based on search history to save time and overall be effective and efficient." -T6 (15-year-old, female)

Teens liked discovering new content that algorithms suggested, especially those pieces they had not encountered elsewhere. While teens recognized that discovering new content could make them come across inappropriate content, many found the excitement of content discovery outweighed or neutralized this risk, as T3 shared:

"Sometimes I see new content that I never really knew about and this creates some excitement if I like it. If I don't like it, then it's either a neutral or negative reaction but that's the risk of exploring new things." -T3 (15-year-old, male)

Overall, teens had a good understanding of how content recommendation algorithms work. They enjoyed interactions with content recommendation algorithms that suggest content aligning with their personal interests.

Teen creators outweigh the benefits over harm in maintaining loose interpersonal boundaries with algorithms for their content reach. We observe that some teens felt less sensitive or protective about how their personal data is collected and referenced by recommendation algorithms, particularly those creating and sharing content on social media as teen creators. These teen creators prioritized the benefits over concerns about privacy or exposure to negative interactions. For example, T1 shared:

"Exposure is good for me as a content creator, but it can come with some drawbacks, such as me getting more spam and hate comments, since anyone who is popular have haters. However, exposure is beneficial, since it allows me to get more influence and followers on the platform. Overall, I think algorithms on Instagram to promote a post are more helpful than harmful." - T1 (16-year-old, male)

In this case, T1 valued gaining followers and content reach over the risks of receiving "spam and hate comments" by loosening interpersonal boundaries with algorithms that actively collect and reference their personal data to suggest content to them or suggest their content to others. As such, teen creators shared experiences of a deliberate calculus of the benefits and drawbacks of loosening boundaries with recommendation algorithms on popular social media. Additionally, teen creators believed algorithms tend to prioritize high-quality content, motivating platform users to create and share better content for visibility. Therefore, they saw that recommendation algorithms help maintain or enhance the quality of content creation and consumption on the platforms.

5.2 Interpersonal Boundary Violations and How Teens Cope with Violations (RQ2)

5.2.1 Boundary Violations by People on Social Media.

Teens perceived losing control of their interpersonal boundaries when known or unknown contacts violated their identity and relationship boundaries. Relationship and identity boundaries [66], as core parts of interpersonal boundaries, establish when users decide with whom to connect and how those connections are managed over time (i.e., relationship boundary) [11] and the management of presentation of identities during information exchanges (i.e., identity boundary) [66]. Our teens expressed losing control of both types of boundaries sometimes by unknown and other times by known contacts. For instance, among teens, one of the most criticized features on public platforms such as Twitter and Instagram was the search feature that allowed any user to find teens by their account names and view their personal data, even if those accounts were set to private. To cope with such boundary violations, teens utilized privacy features on the platforms. For instance, teens employed settings that block strangers or non-followers from sending direct messages to them, as T2 shared:

"Instagram allows for the usage of 'Requests' and limiting if non-followers can message you [...] I find the feature particularly useful as an effort to deter potential privacy breaches before an event even happens." - T2 (16-year-old, male)

When known contacts like friends or acquaintances mentioned them without permission, teens leveraged platform features (e.g., disabling Tags and Mentions on Instagram) to cope with undesired intrusions. While teens actively used different platform features to manage interpersonal boundaries with people, they also found such features often provided only partial support. For instance, T3 highlighted:

"The problem with normal blocking is if you block someone and they find out they might continue to do whatever they were doing to you by making another account or finding another way to contact you" -T3 (15-year-old, male)

In this case, interpersonal boundary violation occurred when platform features could not fully protect him from encountering undesired contacts without consent. Although teens used existing privacy features on platforms to cope with boundary violations

with contacts, they still found such features often provided only partial support over their interpersonal boundary management.

Teens perceived their interpersonal boundaries were violated when boundary turbulence occurred in their daily content sharing. Teens frequently encountered privacy turbulence in their daily content sharing when their private content reached unexpected audiences. For example, a teen shared: *"If someone posts on their private account and a friend were to repost it, then a different audience than the desired one would see the post"* (T6, 15-year-old, female). This was a major challenge for teens because in many cases, there is personal information associated with content (e.g., location, personal connections) that teens would not want to share with the unintended audience.

To cope with such boundary turbulence, teens actively used privacy features that help control content visibility across audience groups. For instance, some teens used Instagram's "Close Friends" feature to ensure only a select group can view their personal stories. Other teens employed settings that only allowed their followers to mention them in posts, preventing strangers from interacting with their content without permission. These features helped teens balance their desire to be public and the need to maintain interpersonal boundaries with different audiences. Meanwhile, teens who were open to sharing content more broadly expressed frustration with platforms' social features such as "tagged" on public platforms (i.e., Instagram) as these features conflicted with the privacy features to manage boundaries with unexpected audiences by showing their social connections and location information by default. T1 shared:

"The biggest fail when it comes to social media privacy features is probably the 'tagged' feature on Instagram because when you look at someone's profile, you can see who tagged them on which post, which seriously invades privacy because can reveal private relationships and detail about someone's location." – T1 (16-year-old, male)

As such, teens used platform features to maintain desired avoid boundary turbulence while actively sharing content with different audience groups. Yet, similar to boundary violation by contacts, teens felt that platforms failed to fully support their interpersonal boundary management when boundary turbulence occurred in their content sharing.

5.2.2 Boundary Violation by Social Media Algorithms.

Teens lost control over rule-setting in interpersonal boundaries with algorithms. Teens felt unable to establish clear rules about how algorithms collect and reference their input/data to suggest content. Some teens shared that a single accidental click could lead to a surge of unwanted content recommendations. Teens felt that this is because algorithms prioritized "profit" from user engagement over users' well-being. For example, T1 perceived that even though some platforms are aware of potentially risky content recommended by algorithms to teens, they do nothing to address the issue, as he shared:

"Minors on Instagram should not be able to view explicit content, but Instagram does not care and the algorithm usually encourages that type of content if it gets good engagement." – T1 (16-year-old, male)

Even when teens tried to manage content suggested by algorithms through feedback features such as "Not-Interested," they often found these mechanisms ineffective, feeling their efforts to set rules with algorithms were ignored. For those unfamiliar with such features, the lack of control was even more pronounced; they coped this by either waiting for unwanted content to disappear or uninstall the apps altogether to reclaim their personal agency. One teen shared:

"There's been risky content in my feeds before. I noticed it more on Snapchat which is why I got rid of it. I never interacted with content like it before but I've seen it and I couldn't seem to get it to go away so I got rid of the app." -T3 (15-year-old, male)

T3 here felt his attempts to set rules for how algorithms shape his content preferences were disregarded by algorithms on Snapchat, resulting in boundary turbulence. After such failed attempts, the teen chose to terminate interpersonal relationships with algorithms by uninstalling the app.

Teens experienced negative outcomes of boundary turbulence with algorithms. Teens experienced violations of their interpersonal boundaries as boundary turbulence occurred: the lack of clear rule-setting between teens and algorithms resulted in harmful outcomes. For example, a teen explained that clicking on a post about a model could trigger a stream of body image-related content, exacerbating mental health issues such as body dysmorphia:

"For example, if someone were to look at a model and have past negative feelings for themselves, they may develop body dysmorphia. But since they did click on a post, a social media may misunderstand and think the user is interested in that said post. It may recommend more posts of that model and others which will further feed into the mental health of the user." -T6 (15-year-old, female)

The above example showed how boundary turbulence happened between co-owners of information (teens and platforms), leading to not only privacy breaches but also negative mental health impacts. While teens were aware that their data and online activities were constantly collected and analyzed by recommendation algorithms, they felt powerless against the consequences of these. For example, some teens shared they experienced a "filter bubble," as one teen said: *"Of course, other limitations exist, such as the filter bubble that it creates where you are only recommended what you want, but ultimately that is the goal of all of these recommendation algorithms"* (T2, 16-year-old, male). Other teens were particularly concerned that the negative impact of this boundary turbulence could worsen if the content pushed by algorithms was risky, as T5 shared:

"If you click on a post that promotes negativity, the algorithm might think you are interested in it, and continue to give you similar posts to it." -T5 (17-year-old, female)

Some teens referred to this situation as falling down a "rabbit hole," recalling how they were exposed to and internalized unhealthy beliefs during their early teenage due to persistent boundary turbulence by recommendation algorithms. They described ending up on the "bad side of the internet" (T4, 17-year-old, male)

at younger ages, where they adopted the beliefs that they now recognize as harmful.

5.3 Teens Re-designing Social Media to Overcome Interpersonal Boundary Violations (RQ3)

5.3.1 Design for Control Over the Cycle of Interpersonal Boundary Management with People.

Teens designed “invisibility” features that give them control over discoverability and relationship boundary. One of the biggest challenges was relationship boundary violation, where teens were followed or requested to be connected by strangers on public platforms through a public search. Therefore, teens suggested design ideas to manage social connections more privately on those platforms. For instance, some teens liked the fact that social connections are formed by invitation only on some private platforms, so they envisioned repurposing the same features for public platforms by removing public search features and implementing invitation-based mechanisms (e.g., inviting connections through a link or QR code), as a teen shared:

“They could share their account through methods such as a link or a QR code. This allows people to stay private and interact with content, making this a mix of a private and public social media platform.” -T3 (15-year-old, male)

Moreover, to manage teens’ identity boundaries, teens suggested features that support control over discoverability. These include completely hiding account information to make their user accounts anonymous or removing search features (e.g., search by account name) on public platforms so that strangers cannot reach out to them. Similarly, on private platforms such as Discord, teens feared that blocking or reporting features would not work for severe situations where strangers keep trying to connect with them. Therefore, they envisioned enhanced privacy features that make them completely invisible:

“Present in both Discord and Instagram is the ability to hide that you are currently using the platform. This is especially useful in situations in which any teen is at an at-risk situation that requires a more complicated approach than simply blocking.” - T2 (16-year-old, male)

However, teens also recognized that such solutions could entail additional challenges for users on public platforms. For instance, some teens valued features that supported content curation on their feed based on previous interactions and connections (e.g., influencers), which may be lost with “invisibility.” Therefore, the reasonable middle ground teens suggested was the option to opt in/out of the public search. By having such options, teens envisioned having an agency to decide whether they need a personalized safety net to manage interpersonal boundaries with people on social media.

Teens designed features to granularly control content visibility among different audiences to manage their interpersonal boundaries. To control content visibility shared with known or unknown contacts, either expectedly or unexpectedly, teens wanted

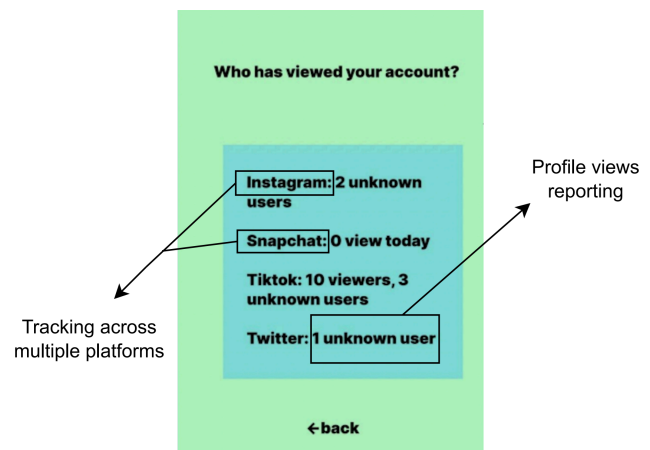


Figure 2: New app designed by teens to help them keep track of who viewed their accounts on public platforms

to first have full transparency by getting informed about who the audiences were so that they could reactively manage their interpersonal boundaries with them. For instance, one teen designed a new app that can be used along with social media apps (Figure 2). With this separate app, the teen wanted to keep track of who has viewed their account or content across the social media platforms so that they can further use privacy features on those platforms to block or filter unwanted connections for interpersonal boundary management. With transparency provided by this new app, teens envisioned that they could be aware of and cope with extreme boundary violations, such as being stalked by strangers or hacked by malicious users, as T7 shared:

“Any time someone sees your profile, it will keep track of who, when, and if they tried to hack into your account. That way you are able to keep track of your accounts and know who to block, or even create a different account safe from that hacker.” - T7 (16-year-old, female)

Second, to avoid boundary turbulence in their content sharing, teens designed features to manage sharing at an audience level. For instance, one teen designed a feature to allow users to categorize their connections into multiple groups, and then selectively share content within different groups or allow different levels of interactivity (Figure 3). With this design idea, teens can avoid the hassle of creating two different accounts (e.g., one public and one private account) for different groups of audiences on one platform:

“My design was to alleviate the issues of having a private Instagram by being able to create more “private” follower groups with full customizability, even being able to edit the ring color around your story indicator.” -T2 (16-year-old, male).

Last, teens suggested privacy features that allow users to decouple personal information from content (Figure 4). The idea was to address the loopholes reported in Section 5.2.2, where teens found that existing social features like “tagged” reveal personal information in their content when teens’ known contacts interact



Figure 3: Design idea for customizable follower groups with advanced controls for content sharing

with the content. Thus, teens designed a feature that allows them to set their posts public while they can set their personal information (e.g., user name, age, or location) and interaction information (e.g., comments, tags, mentions) as private. Teens envisioned this design can not only help maintain their interpersonal boundaries but also boost content reach by getting likes from wider audiences since they do not need to worry about unexpected audiences.

5.3.2 Design for Interpersonal Boundary Management with Algorithms.

Teens designed interactive feedback loops with algorithmic training for desired content recommendation. Feedback loops refer to interactive mechanisms, where algorithms allow teens' input to shape the training of recommendation algorithms, and this eventually affects how content is recommended or curated [12, 81]. While teens had clearly indicated whether they liked or disliked certain content pushed by algorithms through existing features like "Not interested" on Instagram or YouTube, they observed those feedback features never worked for them. Therefore, they designed feedback features with settings beyond interest or disinterest. For example, T3 shared:

"So potentially it would look like a list of said topics ranking what the algorithm would think you'd like most to what they think you'd like less. Within this list you would ideally be able to take out and or add certain topics and arrange their placement. This allows for a more personalized algorithm that can both satisfy users as well as provide safety." - T3 (15-year-old, male)

Here, T3 proposed an interactive feedback mechanism, similar to a "card sorting" approach, that supports teens with the flexibility to add, remove, or sort topics of content according to their preferences. The rationale behind this design was rooted in teens' positive experiences with recommendation algorithms, such as discovering new and personalized content they might not have otherwise encountered. Moreover, teens understood that algorithms are a black box. Hence, they did not expect to fully understand the algorithms' decision-making process. Rather, teens expected that with the interactive feedback feature (i.e., topic ranking feature), they would be able to understand algorithmic decisions and further adjust them to meet their needs and preferences.

Meanwhile, some teens shared concerns over the unintended consequences of such interactive feedback features, so they asked for other stakeholders to provide feedback. For example, one teen recognized the abusive use of content flagging features that malicious users intentionally flag safe content as "risky" to harass teen creators, and *"it can be used as a negative tool to target certain creators"* (T2, 16-year-old, male).

To mitigate this, T2 proposed that human moderators can help review flagged content and/or users before reducing its content from recommendation or removing or banning content.

Teens designed real-time risk detection algorithms that proactively filter risky content. The major motivation for such a design idea was that teens perceived reactive approaches such as reporting or blocking risky content as ineffective since those were only possible after teens have already encountered risky content, as shared by one teen:

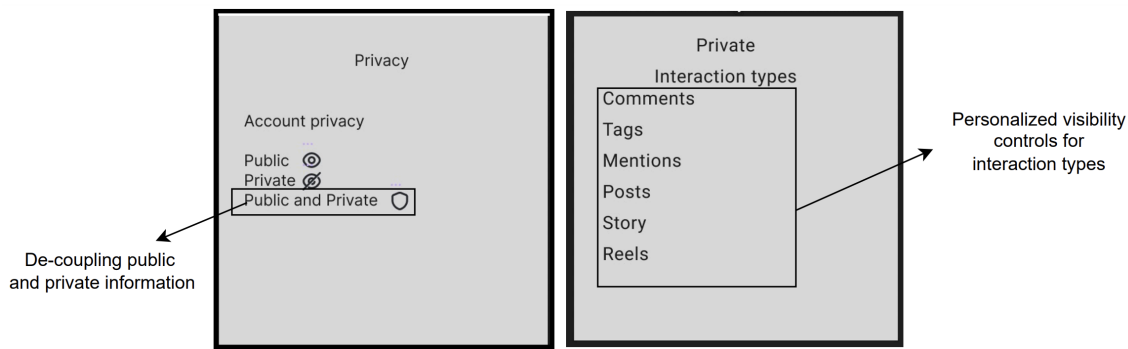


Figure 4: Design solution for visibility through different privacy levels for content and personal information

“I would want to prevent them [fake news and misinformation] from being like spread like go on trending, maybe through like filter or something. I may spread that like unknowingly, just spread of like false information without like knowing or like fact-checking it.” - T6 (15-year-old, female)

When designing real-time risk detection algorithms, some teens designed algorithms that can trigger contextualized nudges explaining risks in granularity. For instance, instead of having binary “fake” or “true” labels for misinformation detection, they wanted to have gray areas highlighted through labels (e.g., possibly misleading, controversial) to inform teens about the potential risk of the flagged content so that they can re-think or fact-check before they unintentionally share misinformation on social media:

“Put guidelines and nudges for content that is identified as potentially harmful, perhaps putting a “controversial” or “possible misleading” flag to content that is majorly flagged by the community.” - T2 (16-year-old, male)

When designing risk detection algorithms, we observed that contrary to recommendation algorithms, teens were willing to let the risk detection algorithms step into their interpersonal boundaries if they could help keep teens safe from online risk.

For instance, teens shared that they were okay with the algorithm using their account activities, such as the number of posts, comments, likes, and followers, to identify suspicious accounts and ban those accounts. One teen explicitly mentioned that tracking users’ log data would not violate their online privacy as they perceived that digital trace data is being collected and used by algorithms, not by “people” monitoring their online activities.

“When I say past social media usage I mean accounts you’ve followed, content you’ve liked, etc. Since it’s done by an algorithm instead of a human I don’t really think it negatively impacts people’s privacy but I can see how some would argue it does.” - T3 (15-year-old, male)

The underlying assumption for opening boundaries with risk detection algorithms was the trust in algorithmic decisions, where teens perceived that algorithms are less prone to making errors than human moderators. In addition, teens perceived that sharing their personal data with risk detection algorithms could help algorithms work better in risk mitigation. Hence, most teens were willing to let

algorithms step into their interpersonal boundaries by collecting and analyzing their online activities to keep them safer on social media platforms.

6 Discussion

Drawing on the framework of *interpersonal boundary management* [11, 71], we explored how teens navigate interpersonal boundary management with people and algorithms and how they propose redesigns to overcome boundary violations on social media platforms, as summarized in Table 1. Below, we discuss our key findings and suggest educational and design opportunities to support safe and autonomous online experiences for teens.

6.1 Teens’ Online Privacy as Interpersonal Boundaries on Social Media

Contextualizing Altman’s definition of privacy as an “interpersonal event” [11] on social media through teens’ experiences, we found that teens actively categorize platforms and select certain platforms to use based on their primary goals (e.g., content creation vs. socializing). Teens are well aware of the risks and benefits associated with different platforms, hence, teens employ varied features on the platforms to strike a balance between meeting their goals and mitigating risks. Our findings showed that teens often experience boundary turbulence in their content-sharing practices on public platforms (e.g., Instagram) as they try to balance between social connectivity through content and full control over personal information associated with the content. Our findings suggest the need to support teens with opportunities to develop skills in risk assessment, problem-solving, and independently managing potential online risks [69, 93, 95].

Moreover, we found that some teens choose to be discoverable on popular social media platforms due to peer pressure, even though the level of privacy supported by those platforms does not match their desire. Prior work well documented how average adult users cope with boundary violations (e.g., [23, 89, 94]). We extend this line of work by discovering teens’ coping mechanisms, including using privacy features on different platforms, segmenting audience groups, and proposing designs to reduce their discoverability. All these demonstrate teens’ digital skills and increasing independence as they mature into adulthood [95]. Once again, our findings suggest the need for strength-based design practices that can empower teens

Table 1: Mapping of teens' interpersonal boundary violations and their design suggestions

Interpersonal Boundary Violations (RQ2)	Teen's Design Suggestions (RQ3)
Relationship and identity boundaries as core parts of interpersonal boundaries were violated by unknown and known contacts (Section 5.2.1)	Teens designed features that give them control over the discoverability for their identity and relationship boundaries (Section 5.3.1)
Teens perceived their interpersonal boundaries were violated when boundary turbulence occurred in their daily content sharing (Section 5.2.1)	Teens designed features to granularly control content visibility among different audience (Section 5.3.1)
Teens' loss of control over rule-setting in interpersonal boundaries with algorithms (Section 5.2.2)	Teens designed interactive feedback loops with algorithmic training for desired content recommendation (Section 5.3.2)
Teens experience negative outcomes of boundary turbulence with algorithms (Section 5.2.2)	Teens designed real-time risk detection algorithms that proactively filter risky content (Section 5.3.2)

to manage online risks in meaningful ways, rather than trying to protect them by restricting technology use. As such, our work contributes to the recent shift in adolescent online safety research that supports teens' autonomy and independence.

Our findings suggest that while teens are aware of the existing features on social media platforms for interpersonal boundary management, these features often fall short. Especially, teens proposed designs to manage the audience in granularity (almost "micromanage" in their language) for full customization of audience-specific interactions, ranging from who can reply to their posts to the color theme of each group. By extending the recent effort in the IDC [14, 50, 100] and broader SIGHCI communities [4, 5, 9, 13, 45] that involve teens in co-designing safety solutions, our study shows two unintended consequences of the solutions. First, while teens in our study challenged the default discoverability features on social media platforms and suggested changes that allow them to disappear from public searches, they considered that such an "invisibility" feature could be misused by perpetrators or stalkers wishing to cause harm by being invisible to victims. Second, our teens designed an app to keep track of their account information across social media platforms to block or filter unwanted connections. Yet, such apps can pose serious privacy risks if not managed securely. Therefore, education on critical thinking about the unintended consequences should be provided to empower teens with a comprehensive understanding of online safety solutions.

6.2 Teens' Boundary Management with Social Media Algorithms

Our findings show that boundary turbulence occurs when social media algorithms, acting as co-owners of information [71], disregard teens' preferences and input, suggesting content in an undesired way. As content consumers, teens preferred to close interpersonal boundaries with recommendation algorithms to avoid filter bubbles and excessive content suggestions. They expressed interest in interactive tools—like a "card sorting" feature—to reorganize their feeds based on preferred content, contrasting with the current automated rankings. In contrast, as content creators, teens were more willing to accept algorithmic boundary turbulence in exchange for increased visibility and reach. While their design ideas reflected a strong desire for personalization and control, this level

of autonomy may inadvertently reinforce filter bubbles, especially if teens prioritize familiar topics. To balance personalization with exposure to diverse content, platforms could prompt teens to select interests across a range of categories by default. Future co-design research can explore how to navigate this tension and identify support features that best serve teens' evolving needs.

Interestingly, teens demonstrated more openness in sharing granular data with online risk detection algorithms to prevent boundary turbulence. Building on recent efforts in teen-focused risk detection (e.g., [68, 75, 83]), our study is among the first to show that teens are willing to loosen boundaries with these systems to avoid harms from recommendation algorithms—such as exposure to explicit or inappropriate content. Teens expressed a strong need for proactive risk detection tools, especially on popular platforms with wider audiences, where undetected risks could spread quickly and affect their broader networks. While existing features like Instagram's generic sensitivity filters provide basic safeguards, teens called for systems that address more subjective and nuanced risks [5]. Their design ideas included flexible and open-ended risk labels that go beyond current filters to reduce false positives and better reflect the ambiguous nature of online harm.

Additionally, our findings demonstrated that while teens have a certain level of trust in AI-based systems, they still believe that these systems can work best in combination with collective effort (e.g., other stakeholders like human moderators on social media). While prior work has called for collective online safety efforts [3, 7], we build upon this idea for collective algorithmic decision-making to ensure fairness and safety through moderators, teens, and platform involvement. Finally, teens may benefit from educational efforts related to regaining control of their boundaries in terms of time spent online (e.g., unintended "rabbit holes"). While filter bubbles are mostly studied in the context of adult users [51, 76], our results showed that teens were not only aware of the phenomenon but also reflected on their experiences at younger ages. One way to support teens to avoid "rabbit holes" is through the intentional and planned use of social media platforms (e.g., to watch a certain video on a certain topic during planned time) [28, 54] as self-reflective approaches have been shown to reinforce teens' agency in their technology use [28] and help them in making goal-directed choices as planned [43]. Therefore, social media platforms could promote

educational content on how to set screen time to help teens renegotiate their boundaries with social media algorithms.

6.3 Theoretical Implications for Interpersonal Boundary Management on Social Media

Our study extends theories of interpersonal boundary management [11] and communication privacy management (CPM) [70] by applying them to teens' boundary negotiation within social media environments. A key novelty of our work lies in demonstrating how teens manage boundaries not only in dyadic peer interactions but across complex, networked platforms shaped by shifting norms and evolving affordances. Unlike prior work focused on parental or peer influence [47, 91, 96], we show how teens strategically regulate boundaries across public, semi-public, and private spaces using tools such as "Close Friends" to manage audience visibility. We also document their lived experiences of boundary turbulence, such as unintended tagging or content exposure, and highlight how they proactively anticipate such violations by designing layered boundary strategies, moving beyond ad hoc solutions such as "Finstas" [27, 97]. These practices position teens as active agents who manage co-ownership of personal information across platforms, extending CPM's conceptualization of boundary turbulence [71].

In addition to interactions with human contacts, our findings expand the theories of interpersonal boundary management [11] and communication privacy management [70] by exploring algorithms as opaque, non-human co-owners of personal information. Our work sheds light on teens' understanding of social media algorithms as commercial entities to maximize engagement and monetization, amplifying boundary turbulence when they lose control over how their data is collected, interpreted, and used for content curation. We showed that this asymmetry challenges a foundational assumption in interpersonal boundary management with algorithms: that boundary rules are co-negotiated [57]. Instead, our findings call for the need to address how social media platforms' business models (re)shape boundary management, often undermining teens' efforts to set or enforce their privacy preferences.

Finally, we argue for an interdisciplinary approach to understanding interpersonal boundary management in the age of algorithms. Teens' experiences, particularly those involving algorithmic amplification of harm, underscore the importance of situating boundary work within broader sociotechnical, economic, and political systems [21, 72, 85]. Future research could examine how commercial incentives shape algorithmic decisions that affect teens' boundaries and how legal or regulatory frameworks contribute to these dynamics. Such interdisciplinary efforts can help design more transparent and equitable systems that empower teens as emerging digital citizens navigating increasingly complex online environments.

6.4 Design Implications

Based on our findings, we provide guidelines for designing social media platforms to promote safer interpersonal boundary management for teens.

- **Balancing visibility with privacy protection:** Social media platforms can offer flexible options for (in)visibility features to manage the broad searchability of their identity and information. This will allow teens to benefit from public

presence while protecting their personal information by decoupling it from content, particularly on public platforms that afford wide content reach (e.g., Twitter, TikTok).

- **Customizing content and privacy for different audiences:** Teens require greater granularity in controlling content visibility among different audience groups on the same platform rather than creating Finsta accounts [27, 97]. Therefore, platforms can offer customizable privacy features that allow teens to adjust visibility based on relational contexts. This could particularly benefit teen content creators who strive to reach a wider audience with their public content while maintaining interpersonal boundaries.
- **Providing interactive feedback for algorithms for tailored content recommendations:** Teens expressed a strong preference for interactive feedback mechanisms that allow them to shape their recommendation feeds dynamically, aligning with recent studies advocating for user-driven algorithmic transparency [38, 82]. Social media platforms can provide teens with interactive ways to share feedback to algorithms based on their content preferences, e.g., features to manage ranking topics of content by interest. Setting a minimum number of different topics to include in their content preference list can help teens consume diverse content.
- **Collective efforts between platforms, moderators, and teens for verifying risk detection:** Unlike previous research that highlights concerns over AI moderation's potential biases [18, 40], our teens are open to AI involvement in detecting risky content, provided that such systems incorporate a verification process with human moderators and peer networks. This suggests an opportunity to design hybrid moderation systems where AI and human feedback coalesce to enhance safety without undermining user autonomy.
- **Intentional boundary setting for content consumption:** Building on prior work on intentional social media use [28, 43], platforms can implement features that help teens set intended screentime before they consume social media content. This could help reduce unintentional engagement with "rabbit holes" and excessive use. Content curation algorithms could also be leveraged to deliver educational content on online safety and addictive consumption patterns to raise awareness, particularly on media-heavy platforms such as TikTok and Instagram.

6.5 Limitations and Future Research

We acknowledge potential biases that may have shaped the data and findings. Participants in our study were recruited through youth-serving organizations and an application process that included interviews to assess motivation, experience, and creativity. As such, our sample may reflect a group of highly engaged, articulate, and digitally literate teens who are already motivated to reflect on online safety and contribute to design processes. In addition, their exposure to UX concepts through workshops may have further influenced their ability to conceptualize and communicate nuanced design ideas. While this enriched the quality of their co-design ideas, it may not fully represent the broader population of teens with varying levels of interest, digital access, or socio-technical

awareness. Additionally, participants self-selected into our study and were compensated for their time, which may introduce participation biases. Future research should explore broader and more diverse teen populations, including those less digitally inclined or from underrepresented communities, to better understand varied experiences with boundary management on social media.

Reflecting on our use of the Asynchronous Research Community (ARC) method, we conducted a follow-up interview study with teen participants to better understand their experiences and critically examine our methodological approach (forthcoming publication). Teens appreciated the flexibility of the ARC format, which allowed them to complete activities at their own pace and on their own time. They noted that the asynchronous environment gave them space to read prompts carefully, understand the tasks, reach out to researchers with questions, and respond thoughtfully—making them feel more confident in their contributions. However, we also encountered challenges in sustaining consistent engagement over the eight-week period. Despite these limitations, the ARC approach offered several methodological strengths. The asynchronous format supported flexibility and reflection that would have been difficult to achieve in synchronous settings. It also allowed us to build rapport with a small group of teens, track their progress over time, and encourage active participation, enabling us to capture rich insights into their boundary management practices and co-design ideas. Our findings suggest that ARC is a promising method for engaging youth in online safety and participatory design research.

7 Conclusion

In this study, we explored how teens manage interpersonal boundaries with both people and algorithms on social media. Using an Asynchronous Research Community (ARC), we engaged teens in reflecting on their experiences and co-designing solutions to support safer, more autonomous online interactions. Our findings extend theories of interpersonal boundary management and communication privacy management by illustrating how teens navigate boundary coordination not only with peers but also with algorithmic systems driven by social media platforms. We also demonstrate the value of ARC in capturing teens' nuanced perspectives over time and supporting meaningful participation. By centering teens' voices, we provide actionable insights for designing social media features that align with youth values and agency. Future work should build on these insights to examine how boundary practices evolve within the broader sociotechnical and economic structures of digital platforms.

8 Selection and Participation of Children

Our participants were recruited through outreach to youth-serving organizations, high schools across the U.S., and social media posts. Interested teens completed a pre-screening survey confirming they were 13–17 years old, fluent in English, had internet access via a smartphone or computer, and could commit time to our study. Selected applicants were invited to a 15–30 minute interview where researchers assessed their motivation, experience with online risks, creative problem-solving skills, and awareness of current social media safety trends. Finalists were notified via email, provided with instructions for creating a Discord account, and asked to confirm

their contact and payment details. Participants received a certificate and up to \$450 in gift cards for completing all YAB and ARC activities. All study procedures were approved by the Institutional Review Board (IRB) at the last author's institution. All team members completed CITI human subjects and Protection of Minors (POM) training. YAB participant information was not made public and was only accessible to trained researchers listed on the IRB protocol. To protect privacy, we removed all identifiable data and paraphrased participant quotes.

Acknowledgments

This research was supported by the National Science Foundation under grants IIS-2333207 and IIP-2329976, William T. Grant Foundation #187941, and the University of Wisconsin-Madison's Technology and Adolescent Mental (TAM) Wellness program. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of our sponsors.

References

- [1] Michael Adorjan and Rosemary Ricciardelli. 2019. Student perspectives towards school responses to cyber-risk and safety: the presumption of the prudent digital citizen. *Learn Media Technol* 44, 4 (Oct. 2019), 430–442. <https://doi.org/10.1080/17439884.2019.1583671>
- [2] Zainab Agha, Naima Samreen Ali, Jinkyung Park, and Pamela J Wisniewski. 2024. A systematic review on design-based nudges for adolescent online safety. *International Journal of Child-Computer Interaction* (2024), 100702.
- [3] Zainab Agha, Karla Badillo-Urquiola, and Pamela J Wisniewski. 2023. "Strike at the Root": Co-designing Real-Time Social Media Interventions for Adolescent Online Risk Prevention. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–32.
- [4] Zainab Agha, Kelsey Miu, Sophia Piper, Jinkyung Park, and Pamela J. Wisniewski. 2023. Co-Designing User Personas and Risk Scenarios for Evaluating Adolescent Online Safety Interventions. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*. 249–253. <https://doi.org/10.1145/3584931.3606964>
- [5] Zainab Agha, Jinkyung Park, Ruyuan Wan, Naima Samreen Ali, Yiwei Wang, Dominic Difranzo, Karla Badillo-Urquiola, and Pamela J Wisniewski. 2024. Tricky vs. Transparent: Towards an Ecologically Valid and Safe Approach for Evaluating Online Safety Nudges for Teens. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.
- [6] Zainab Agha, Zinan Zhang, Oluwatomisin Obajemu, Luke Shirley, and Pamela J. Wisniewski. 2022. A Case Study on User Experience Bootcamps with Teens to Co-Design Real-Time Online Safety Interventions. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–8.
- [7] Mamtaj Akter, Amy J Godfrey, Jess Kropczynski, Heather R Lipford, and Pamela J Wisniewski. 2022. From Parental Control to Joint Family Oversight: Can Parents and Teens Manage Mobile Online Safety and Privacy as Equals? *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–28.
- [8] J. Alemany, E. del Val, J. Alberola, and A. García-Fornes. 2019. Enhancing the privacy risk awareness of teenagers in online social networks through soft-paternalism mechanisms. *Int J Hum Comput Stud* 129 (Sept. 2019), 27–40. <https://doi.org/10.1016/j.ijhcs.2019.03.008>
- [9] Naima Samreen Ali, Teulebit Ahn, Zainab Agha, Jinkyung Katie Park, and Pamela J Wisniewski. 2024. From Ideas to Impact: Cracking the Code for Effectively Engaging Teens in Long-Term Online Safety Research. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.
- [10] Shiza Ali, Afsaneh Razi, Seunghyun Kim, Ashwaq Alsoubai, Chen Ling, Munmun De Choudhury, Pamela J Wisniewski, and Gianluca Stringhini. 2023. Getting Meta: A Multimodal Approach for Detecting Unsafe Conversations within Instagram Direct Messages of Youth. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–30.
- [11] Irwin Altman. 1975. *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding*. Brooks/Cole Publishing Company, Monterey, CA.
- [12] Oscar Alvarado and Annika Waern. 2018. Towards algorithmic experience: Initial efforts for social media contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12. <https://doi.org/10.1145/3173574.3173860>

- [13] 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.
- [14] Karla Badillo-Urquiola, Diva Smriti, Brenna McNally, Evan Golub, Elizabeth Bonsignore, and Pamela J Wisniewski. 2019. Stranger danger! social media app features co-designed with children to keep them safe online. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 394–406.
- [15] Diana Baumrind. 1987. A developmental perspective on adolescent risk taking in contemporary America. *New directions for child and adolescent development* 1987, 37 (1987), 93–125.
- [16] Arpita Bhattacharya, Calvin Liang, Emily Y Zeng, Kanishk Shukla, Miguel ER Wong, Sean A Munson, and Julie A Kientz. 2019. Engaging teenagers in asynchronous online groups to design for stress management. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 26–37.
- [17] Arpita Bhattacharya, Ria Nagar, Jessica Jenness, Sean A Munson, Julie A Kientz, et al. 2021. Designing asynchronous remote support for behavioral activation in teenagers with depression: formative study. *JMIR Formative Research* 5, 7 (2021), e20969.
- [18] Reuben Binns, Michael Veale, Max Van Kleek, and Nigel Shadbolt. 2017. Like trainer, like bot? Inheritance of bias in algorithmic content moderation. In *Social Informatics: 9th International Conference, SocInfo 2017, Oxford, UK, September 13–15, 2017, Proceedings, Part II* 9. Springer, 405–415.
- [19] Lindsay Blackwell, Emma Gardiner, and Sarita Schoenebeck. 2016. Managing expectations: Technology tensions among parents and teens. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work, CSCW*. 1390–1401. <https://doi.org/10.1145/2818048.2819928>
- [20] V. Braun, N Clarke, V and Hayfield, and G Terry. 2019. *Thematic Analysis*. Springer, Singapore.
- [21] Xavier V. Caddle, Sarvech Qadir, Charles Hughes, Elizabeth A. Sweigart, Jinkyung Katie Park, and Pamela J. Wisniewski. 2025. Building a Village: A Multi-stakeholder Approach to Open Innovation and Shared Governance to Promote Youth Online Safety. *arXiv preprint arXiv:2504.03971* (2025).
- [22] Neeraj Chatlani, Arianna Davis, Karla Badillo-Urquiola, Elizabeth Bonsignore, and Pamela Wisniewski. 2023. Teen as research-apprentice: A restorative justice approach for centering adolescents as the authority of their own online safety. *International Journal of Child-Computer Interaction* 35 (2023), 100549.
- [23] Hichang Cho and Anna Filippova. 2016. Networked privacy management in facebook: A mixed-methods and multinational study. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 503–514.
- [24] Federal Trade Commission. 2013. Children's Online Privacy Protection Rule. <https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa/>
- [25] Nancy Costello, Rebecca Sutton, Madeline Jones, Mackenzie Almassian, Amanda Raffoul, Oluwadunni Ojumu, Meg Salvia, Monique Santoso, Jill R Kavanaugh, and S Bryn Austin. 2023. ALGORITHMS, ADDICTION, AND ADOLESCENT MENTAL HEALTH: an interdisciplinary study to inform state-level policy action to protect youth from the dangers of social media. *American Journal of Law & Medicine* 49, 2-3 (2023), 135–172.
- [26] Melissa J Cox, Joy Gabrielli, Tim Janssen, and Kristina M Jackson. 2018. Parental restriction of movie viewing prospectively predicts adolescent alcohol and marijuana initiation: Implications for media literacy programs. *Prevention Science* 19 (2018), 914–926.
- [27] Christopher R. Darr and Erin F. Doss. 2022. The Fake One is the Real One: Finstas, Authenticity, and Context Collapse in Teen Friend Groups. *Journal of Computer-Mediated Communication* 27, 4 (July 2022). <https://doi.org/10.1093/JCMC/ZMAC009>
- [28] Arianna J Davis. 2020. Co-Designing "Teenovate": An Intergenerational Online Safety Design Team. (2020).
- [29] Katie Davis, Petr Slovak, Rotem Landesman, Caroline Pitt, Abdullatif Ghajar, Jessica Lee Schleider, Saba Kawa, Andrea Guadalupe Perez Portillo, and Nicole S Kuhn. 2023. Supporting Teens' Intentional Social Media Use Through Interaction Design: An exploratory proof-of-concept study. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. 322–334.
- [30] Vanessa P. Dennen, Stacey A. Rutledge, Lauren M. Bagdy, Jerrica T. Rowlett, Shannon Burnick, and Sarah Joyce. 2017. Context collapse and student social media networks: Where life and high school collide. In *Proceedings of the 8th International Conference on Social Media & Society*. F129683. <https://doi.org/10.1145/3097286.3097318>
- [31] Jill P Dimond, Casey Fiesler, Betsy DiSalvo, Jon Pelc, and Amy S Bruckman. 2012. Qualitative data collection technologies: A comparison of instant messaging, email, and phone. In *Proceedings of the 2012 ACM International Conference on Supporting Group Work*. 277–280.
- [32] Discord. 2024. About Discord | Our Mission and Values. <https://discord.com/company>
- [33] Nico Ebert, Tim Geppert, Joanna Strycharz, Melanie Knieps, Michael Hönig, and Elke Brucker-Kley. 2023. Creative beyond TikToks: Investigating Adolescents' Social Privacy Management on TikTok. *Proceedings on Privacy Enhancing Technologies* 2023, 2 (Jan. 2023), 221–235. <https://doi.org/10.56553/popets-2023-0049>
- [34] 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. *J Assoc Inf Sci Technol* 67, 6 (June 2016), 1384–1403. <https://doi.org/10.1002/ASL.23450>
- [35] Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I "like" it, then I hide it: Folk theories of social feeds. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2371–2382. <https://doi.org/10.1145/2858036.2858494>
- [36] Theodore W Gardner, Thomas J Dishion, and Arin M Connell. 2008. Adolescent self-regulation as resilience: Resistance to antisocial behavior within the deviant peer context. *Journal of abnormal child psychology* 36 (2008), 273–284.
- [37] Radhika Garg. 2021. Engaging Parents and Teens in an Asynchronous, Remote, Community-Based Method for Understanding the Future of Voice Technology. In *Interaction Design and Children*. 224–235.
- [38] Anna Gausen, Ce Guo, and Wayne Luk. 2024. An approach to sociotechnical transparency of social media algorithms using agent-based modelling. *AI and Ethics* (2024), 1–19.
- [39] Cami Goray and Sarita Schoenebeck. 2022. Youths' Perceptions of Data Collection in Online Advertising and Social Media. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–27.
- [40] Robert Gorwa, Reuben Binns, and Christian Katzenbach. 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* 7, 1 (2020), 2053951719897945.
- [41] Asia Grace. 2023. I pretended to be a girl, 14, on TikTok — here's what happened. *New York Post*. <https://nypost.com/2023/03/08/i-pretended-to-be-a-girl-14-on-tiktok-heres-what-happened/>
- [42] Wannes Heirman, Michel Walrave, Anne Vermeulen, Koen Ponnet, Heidi Vandebosch, Joris Van Ouytsel, and Ellen Van Gool. 2016. An open book on Facebook? Examining the interdependence of adolescents' privacy regulation strategies. *Behaviour & Information Technology* 35, 9 (Sept. 2016), 706–719. <https://doi.org/10.1080/0144929X.2016.1181210>
- [43] Alexis Hiniker, Bongshin Lee, Kiley Sobel, and Eun Kyoung Choe. 2017. Plan & play: supporting intentional media use in early childhood. In *Proceedings of the 2017 conference on interaction design and children*. 85–95.
- [44] Ikhti Imaniah, Nurul Fitria, Kumala Dewi, and Akhmad Zakky. 2020. Youtube kids channels in developing young children's communication skills in English: parents' beliefs, attitudes, and behaviors. *IJLECR (International Journal of Language Education and Cultural Review)* 6, 1 (June 2020), 20–30. <https://doi.org/10.21009/IJLECR.061.03>
- [45] Naulsberry Jean Baptiste, Jinkyung Park, Neeraj Chatlani, Naima Samreen Ali, and Pamela J Wisniewski. 2023. Teens on Tech: Using an Asynchronous Remote Community to Explore Adolescents' Online Safety Perspectives. In *Companion publication of the 2023 conference on computer supported cooperative work and social computing*. 45–49.
- [46] Juho Kahila, Henriikka Vartiainen, Matti Tedre, and Sonsoles López Pernas. 2024. Children's Data Privacy Fears on Social Media. *SSRN Electronic Journal* (April 2024). <https://doi.org/10.2139/SSRN.4794092>
- [47] Hyunjin Kang, Wonsun Shin, and Junru Huang. 2022. Teens' privacy management on video-sharing social media: the roles of perceived privacy risk and parental mediation. *Internet Research* 32, 1 (2022), 312–334.
- [48] Atika Khurana, Amy Bleakley, Amy B Jordan, and Daniel Romer. 2015. The protective effects of parental monitoring and internet restriction on adolescents' risk of online harassment. *Journal of youth and Adolescence* 44 (2015), 1039–1047.
- [49] Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L. Clegg, and Jessica Vitak. 2017. "No Telling Passcodes Out Because They're Private". In *Proc ACM Hum Comput Interact*, Vol. 1. <https://doi.org/10.1145/3134699>
- [50] Priya Kumar, Jessica Vitak, Marshini Chetty, Tamara L Clegg, Jonathan Yang, Brenna McNally, and Elizabeth Bonsignore. 2018. Co-designing online privacy-related games and stories with children. In *Proceedings of the 17th ACM conference on interaction design and children*. 67–79.
- [51] Mark Ledwich and Anna Zaitsev. 2019. Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. *arXiv preprint arXiv:1912.11211* (2019).
- [52] Eden Litt. 2013. Understanding social network site users' privacy tool use. *Comput Human Behav* 29, 4 (July 2013), 1649–1656. <https://doi.org/10.1016/J.CHB.2013.01.049>
- [53] Sonia Livingstone, Lucyna Kirwil, Cristina Ponte, and Elisabeth Staksrud. 2014. In their own words: What bothers children online? *Eur J Commun* 29, 3 (March 2014), 271–288. <https://doi.org/10.1177/0267323114521045>
- [54] Kai Lukoff, Ulrik Lyngs, Himanshu Zade, J Vera Liao, James Choi, Kaiyue Fan, Sean A Munson, and Alexis Hiniker. 2021. How the design of youtube influences user sense of agency. In *Proceedings of the 2021 CHI Conference on Human Factors*

- in *Computing Systems*. 1–17.
- [55] Haley MacLeod, Ben Jelen, Annu Prabhakar, Lora Oehlberg, Katie Siek, and Kay Connelly. 2017. A guide to using asynchronous remote communities (ARC) for researching distributed populations. *EAI Endorsed Transactions on Pervasive Health and Technology* 3, 11 (2017).
 - [56] Mary Madden, Amanda Lenhart, Sandra Cortesi, and Urs Gasser. 2013. Teens and mobile apps privacy. Pew Research Center. <http://www.pewinternet.org/Reports/2013/Teens-and-Tech/Main-Findings/Teens-and-Technology.aspx>
 - [57] Alice E Marwick and Danah Boyd. 2014. Networked privacy: How teenagers negotiate context in social media. *New media & society* 16, 7 (2014), 1051–1067.
 - [58] Hiroaki Masaki, Kengo Shibata, Shui Hoshino, Takahiro Ishihama, Nagayuki Saito, and Koji Yatani. 2020. Exploring nudge designs to help adolescent sns users avoid privacy and safety threats. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–11.
 - [59] Nora McDonald. 2024. Teens see social media algorithms as accurate reflections of themselves, study finds. <https://theconversation.com/teens-see-social-media-algorithms-as-accurate-reflections-of-themselves-study-finds-226302>
 - [60] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. 2017. Toys that listen: A study of parents, children, and internet-connected toys. In *Proceedings of the Conference on Human Factors in Computing Systems*. 5197–5207. <https://doi.org/10.1145/3025453.3025735>
 - [61] Sarah McRoberts, Elizabeth Bonsignore, Tamara Peyton, and Svetlana Yarosh. 2016. "Do it for the viewers!" Audience engagement behaviors of young YouTubers. In *Proceedings of the 15th International Conference on Interaction Design and Children*. 334–343. <https://doi.org/10.1145/2930674.2930676>
 - [62] Xiaoxiao Meng. 2024. From Interpersonal Privacy to Human Technological Privacy: Communication Privacy Management Theory Revisited. *Technical Communication* 71, 2 (2024), 72–88.
 - [63] Melinda Wenner Moyer. 2022. Kids Are Using Social Media More Than Ever, Study Finds. New York Times. <https://www.nytimes.com/2022/03/24/well/family/child-social-media-use.html>
 - [64] Patricia A Norberg and Daniel R Horne. 2024. When the bridge is not human: Algorithmic interference in forming social relationships through the manipulation of weak ties. *Journal of Consumer Affairs* (2024).
 - [65] Patricia A. Norberg, Daniel R. Horne, and David A. Horne. 2007. The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. *Journal of Consumer Affairs* 41, 1 (June 2007), 100–126. <https://doi.org/10.1111/J.1745-6606.2006.00070.X>
 - [66] Leysia Palen and Paul Dourish. 2003. Unpacking "privacy" for a networked world. In *Proceedings of the Conference on Human Factors in Computing Systems*. 129–136. <https://doi.org/10.1145/642611.642635>
 - [67] Kostantinos Papadamou, Antonis Papasavva, Savvas Zannettou, Jeremy Blackburn, Nicolas Kourtellis, Ilias Leontiadis, Gianluca Stringhini, and Michael Sirivianos. 2020. Disturbed YouTube for Kids: Characterizing and Detecting Inappropriate Videos Targeting Young Children. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14. 522–533. <https://doi.org/10.1609/ICWSM.V14I1.7320>
 - [68] Jinkyung Park, Joshua Gracie, Ashwaq Alsoubai, Gianluca Stringhini, Vivek Singh, and Pamela Wisniewski. 2023. Towards Automated Detection of Risky Images Shared by Youth on Social Media. In *Companion Proceedings of the ACM Web Conference 2023*. 1348–1357.
 - [69] Jinkyung Katie Park, Mamta Akter, Pamela Wisniewski, and Karla Badillo-Urquiola. 2024. It's Still Complicated: From Privacy-Invasive Parental Control to Teen-Centric Solutions for Digital Resilience. *IEEE Security & Privacy* (2024).
 - [70] Sandra Petronio. 1991. Communication boundary management: A theoretical model of managing disclosure of private information between marital couples. *Communication theory* 1, 4 (1991), 311–335.
 - [71] Sandra Petronio. 2002. *Boundaries of Privacy: Dialectics of Disclosure*. SUNY Press. https://books.google.com/books/about/Boundaries_of_Privacy.html?hl=&id=gTCsf8zVXgC
 - [72] Sarvech Qadir, Andy Niser, Xavier V Caddle, Ashwaq Alsoubai, Jinkyung Katie Park, and Pamela J Wisniewski. 2024. Towards a Safer Digital Future: Exploring Stakeholder Perspectives on Creating a Sustainable Youth Online Safety Community. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–10.
 - [73] Emilee Rader and Rebecca Gray. 2015. Understanding user beliefs about algorithmic curation in the facebook news feed. In *Proceedings of the Conference on Human Factors in Computing Systems*. 173–182. <https://doi.org/10.1145/2702123.2702174>
 - [74] Afsaneh Razi, Ashwaq Alsoubai, Seunghyun Kim, Shiza Ali, Gianluca Stringhini, Munmun De Choudhury, and Pamela J Wisniewski. 2023. Sliding into my DMs: Detecting uncomfortable or unsafe sexual risk experiences within Instagram direct messages grounded in the perspective of youth. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–29.
 - [75] Afsaneh Razi, Seunghyun Kim, Ashwaq Alsoubai, Gianluca Stringhini, Tamar Solorio, Munmun De Choudhury, and Pamela J Wisniewski. 2021. A human-centered systematic literature review of the computational approaches for online sexual risk detection. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–38.
 - [76] Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgílio AF Almeida, and Wagner Meira Jr. 2020. Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 131–141.
 - [77] Alexandra Ruiz-Gomez, Tama Leaver, and Crystal Abidin. 2022. Playing YouTube: How the Nancy YouTuber doll and app position children as aspiring YouTube influencers. *International Journal of Cultural Studies* 25, 2 (March 2022), 121–140. <https://doi.org/10.1177/13678779211063222>
 - [78] Tara L Rutkowski, Heidi Hartikainen, Kirsten E Richards, and Pamela J Wisniewski. 2021. Family communication: examining the differing perceptions of parents and teens regarding online safety communication. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–23.
 - [79] Diane J Schiano and Christine Burg. 2017. Parental controls: Oxymoron and design opportunity. In *HCI International 2017—Posters' Extended Abstracts: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part II* 19. Springer, 645–652.
 - [80] 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. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 2 (apr 2021), 18 pages. <https://doi.org/10.1145/3449076>
 - [81] Donghee Shin. 2020. How do users interact with algorithm recommender systems? The interaction of users, algorithms, and performance. *Comput Human Behav* 109 (2020). <https://doi.org/10.1016/J.CHB.2020.106344>
 - [82] Donghee Shin, Kerk F Kee, and Emily Y Shin. 2022. Algorithm awareness: Why user awareness is critical for personal privacy in the adoption of algorithmic platforms? *International Journal of Information Management* 65 (2022), 102494.
 - [83] Vivek K Singh, Souvik Ghosh, and Christin Jose. 2017. Toward multimodal cyberbullying detection. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2090–2099.
 - [84] Jeffrey M Stanton. 2003. Information technology and privacy: A boundary management perspective. In *Socio-technical and human cognition elements of information systems*. Igi Global, 79–103.
 - [85] Elizabeth A Sweigart, Jinkyung Katie Park, and Pamela J Wisniewski. 2025. It Takes a Village: Youth Online Safety Research Highlights Need for Interdisciplinary, Multistakeholder Solutions. *Journal of Advances in Information Technology* 16, 1 (2025).
 - [86] Edson C. Tandoc, Chen Lou, and Velyn Lee Hui Min. 2019. Platform-swinging in a poly-social-media context: How and why users navigate multiple social media platforms. *Journal of Computer-Mediated Communication* 24, 1 (Jan. 2019), 21–35. <https://doi.org/10.1093/JCMC/ZMY022>
 - [87] Emily A Vogels. 2022. Teens and cyberbullying 2022. <https://www.pewresearch.org/internet/2022/12/15/teens-and-cyberbullying-2022/>
 - [88] Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2022. "Don't make assumptions about me!": Understanding Children's Perception of Datafication Online. *Proc ACM Hum Comput Interact* 6, CSCW2 (Nov. 2022). <https://doi.org/10.1145/3555144>
 - [89] Jason Watson, Andrew Besmer, and Heather Richter Lipford. 2012. + Your circles: sharing behavior on Google+. In *Proceedings of the eighth symposium on usable privacy and security*. 1–9.
 - [90] Pamela Wisniewski, Arup Kumar Ghosh, Heng Xu, Mary Beth Rosson, and John M Carroll. 2017. Parental control vs. teen self-regulation: Is there a middle ground for mobile online safety? In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 51–69.
 - [91] Pamela Wisniewski, Haiyan Jia, Heng Xu, Mary Beth Rosson, and John M Carroll. 2015. "Preventative" vs. "Reactive" How Parental Mediation Influences Teens' Social Media Privacy Behaviors. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 302–316.
 - [92] Pamela Wisniewski, Heather Lipford, and David Wilson. 2012. Fighting for My Space: Coping Mechanisms for SNS Boundary Regulation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2207676.2207761>
 - [93] Pamela Wisniewski, Jinkyung Park, Karla Badillo-Urquiola, Joy Gabrielli, Jennifer L Doty, and Heidi Hartikainen. 2025. Moving Beyond Fear and Restriction to Promoting Adolescent Resilience and Intentional Technology Use. In *Handbook of Children and Screens*. Springer, Cham, 403–410.
 - [94] Pamela J Wisniewski, Bart P Knijnenburg, and Heather Richter Lipford. 2017. Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of human-computer studies* 98 (2017), 95–108.
 - [95] Pamela J Wisniewski, Jessica Vitak, and Heidi Hartikainen. 2022. Privacy in adolescence. In *Modern socio-technical perspectives on privacy*. Springer International Publishing Cham, 315–336.
 - [96] Ralf De Wolf. 2019. Contextualizing how teens manage personal and interpersonal privacy on social media. *New Media Soc* 22, 6 (Sept. 2019), 1058–1075. <https://doi.org/10.1177/1461444819876570>

- [97] Sijia Xiao, Danaë Metaxa, Joon Sung Park, Karrie Karahalios, and Niloufar Salehi. 2020. Random, Messy, Funny, Raw: Finstas as Intimate Reconfigurations of Social Media. In *Proceedings of the Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376424>
- [98] Leah Zhang-Kennedy, Christine Mekhail, Sonia Chiasson, and Yomna Abdelaziz. 2016. From nosy little brothers to stranger-danger: Children and parents' perception of mobile threats. In *Proceedings of the 15th International Conference on Interaction Design and Children*. 388–399. <https://doi.org/10.1145/2930674.2930716>
- [99] Dorothy Zhao, Mikako Inaba, and Andrés Monroy-Hernández. 2022. Understanding Teenage Perceptions and Configurations of Privacy on Instagram. *Proc ACM Hum Comput Interact* 6, CSCW2 (Nov. 2022). <https://doi.org/10.1145/3555608>
- [100] Jun Zhao, Blanche Duron, and Ge Wang. 2022. KOALA hero: Inform children of privacy risks of mobile apps. In *Proceedings of the 21st Annual ACM Interaction Design and Children Conference*. 523–528.
- [101] Jun Zhao, Ge Wang, Carys Dally, Petr Slovak, Julian Edbrooke-Childs, Max Van Kleek, and Nigel Shadbolt. 2019. I make up a silly name' Understanding Children's Perception of Privacy Risks Online. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [102] Melanie J Zimmer-Gembeck and W Andrew Collins. 2006. Autonomy development during adolescence. *Blackwell handbook of adolescence* (2006), 174–204.
- [103] Dorothy Zulli and David James Zulli. 2022. Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media Soc* 24, 8 (Aug. 2022), 1872–1890. <https://doi.org/10.1177/1461444820983603>

A Participant Information

Table 2: Participants' demographic information and technology use

Participant ID	Gender	Age	Race/Ethnicity	Technology Usage	Device Usage
T1	Male	16	East Asian	Several times a day	Smartphone, Tablet, Laptop
T2	Male	16	White/Caucasian, Latino/Hispanic	Several times a day	Smartphone, Desktop
T3	Male	15	White/Caucasian	Almost constantly	Smartphone, Tablet, Laptop
T4	Male	17	White/Caucasian	Almost constantly	Smartphone, Laptop, Desktop
T5	Female	17	South Asian	Almost constantly	Smartphone, Laptop
T6	Female	15	South Asian	Almost constantly	Smartphone, Tablet, Laptop, Desktop
T7	Female	16	Black/African-American	Several times a day	Smartphone, Tablet, Laptop

B ARC Activities and Prompts

Table 3: The ARC activities and prompts to explore research questions

Research questions 1 & 2	Activities	Prompts
Interpersonal boundary management with people: What social media platforms do you use and why? How do you manage your privacy on these platforms? What features of social media platforms do you use and why?	Scoping the topic and defining the problem-space through text responses and screenshots (weeks 1-2)	ARC Prompt 1.a: Based on our previous conversations, it seems like you all consider there is a differentiation between "public" and "private" platforms, and you all tend to move towards more private social media platforms to socialize (e.g., Discord, Reddit, YikYak, etc.), instead of public social media (e.g., Instagram, Snapchat, etc.). Is this observation correct, why or why not?
		ARC Prompt 1.b: Screenshot and share some of the different privacy features that you like and use OR dislike and don't use on different social media platforms. Explain why they are useful/not useful in protecting your online safety.
		ARC Prompt 1.c: Describe one of the biggest "failures," if any, when it comes to social media privacy features (on either more public or more private platforms). Explain why this problem could negatively impact your online safety.
Interpersonal boundary Management with Algorithms: How do you understand algorithms on social media? How was your experience with algorithms with specific platforms, if any? What do you think about the positive or negative influence of algorithms on teens' safety?	Scoping the topic and defining the problem-space through text responses (weeks 5-6)	ARC Prompt 2.a: Based on the ARC meeting we had on Zoom, you all seemed to have a pretty good idea of what algorithms are and how they work across social media platforms, but I would still like to have a discussion with you guys about your thoughts on algorithms. What do you think your relationship with algorithms? Do you encounter any issues when interacting with algorithms?
		ARC Prompt 2.b: Describe one of the best/worst ways you think an algorithm could impact your online safety, if any, and why.
Research question 3	Activities	Prompts
Re-designing to for Interpersonal Boundary Management	Designing for problems or concerns identified earlier (design for people: weeks 3-4, design for algorithms: weeks 7-8)	ARC Prompt 3.a: Use Figma to design one social media feature that would solve one of the problems we defined earlier. It can be a problem that you came up with or one from another YAB member in our Discord server.
		ARC Prompt 3.b: Share your design with a description of how it would work in our Discord server, and engage in discussions.

C Summary of key findings

Table 4: Themes and illustrative quotes identified under each dimension

Dimensions	Themes	Quotes
How teens manage interpersonal boundaries with people (RQ1)	Teens had preconceived notions of platform affordances to manage personal boundaries with people	<i>"I would like a mix of that and that's sort of what I try to do with Instagram. I attempt to stay private and close to just my friends. But I also like to go onto my explore page and look at posts of content I am interested in"</i> - T3
	Teens gradually adapted to negotiating their boundaries with people on different platforms	<i>"Up until a few months ago, I was still private [...] I've been comfortable with sharing my platform publicly."</i> - T7
How teens navigate boundaries with social media algorithms (RQ1)	Teen consumers enjoyed personalized content discovery through algorithms	<i>"It [content algorithm] curated feeds to better suit a user's interests and make the individual experience more unique and fulfilling."</i> - T4
	Teen creators outweigh the benefits over harms in maintaining loose boundaries with algorithms for their content reach	<i>"Exposure is good for me as a content creator, but it can come with some drawbacks, such as me getting more spam [...] Overall, I think algorithms on Instagram to promote a post are more helpful than harmful"</i> - T1
Boundary violations by people on social media (RQ2)	Teens perceived their interpersonal boundaries were violated when known or unknown contacts led to boundary turbulence	<i>"The biggest fail would be the ability to search people by name, as it is able to find specific people and any data that is easily accessible on their account"</i> - T2
	Teens perceived their interpersonal boundaries were violated when boundary turbulence occurred in their daily content sharing	<i>"When you look at someone's profile [on Instagram], you can see who tagged them on which post, which seriously invade privacy because can reveal private relationships and details about someone's location"</i> - T1
Boundary violation by social media algorithms (RQ2)	Teens lost control over rule-setting in interpersonal boundaries with algorithms	<i>"There's been risky content in my feed before [...] I couldn't seem to get it go away so I got rid of the app"</i> - T3
	Teens experienced negative outcomes of boundary turbulence with algorithms	<i>"If you click on a post that promotes negativity, the algorithms might think you are interested in, and continue to give you similar posts to it."</i> - T5
Design for interpersonal boundary management with people (RQ3)	Teens designed "invisibility" features that give them control over discoverability and relationship boundary management	<i>"Allowing them to hide accounts from people unless they have shared their accounts with someone. They could share their account through methods such as a link or a QR code"</i> - T3
	Teens designed features to granularly control content visibility among different audiences to manage their interpersonal boundaries	<i>"My design was to alleviate the issues of having a private Instagram by being able to create more "private" follower groups with full customizability, even being able to edit the ring color around your story indicator"</i> - T2
Design for interpersonal boundary management with social media algorithms (RQ3)	Teens designed interactive feedback looks with algorithmic training for the desired content recommendation	<i>"It would look like a list of said topic ranking what the algorithms would think you'd like [...] Within this list you would ideally be able to take out and or add certain topics and arrange"</i> - T3
	Teens designed real-time/early detection algorithms that proactively filter risky content	<i>"I would want to prevent them [fake news and misinformation] from being spread like go on trending, maybe through like filter or something]"</i> - T6