

University of Central Florida

**STARS**

---

Electronic Theses and Dissertations, 2020-

---

2022

## Identifying Challenges and Opportunities for Designing Social Media Nudges for Adolescents

Oluwatomisin Obajemu  
*University of Central Florida*



Part of the [Computer Sciences Commons](#), and the [Social Media Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd2020>

University of Central Florida Libraries <http://library.ucf.edu>

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Obajemu, Oluwatomisin, "Identifying Challenges and Opportunities for Designing Social Media Nudges for Adolescents" (2022). *Electronic Theses and Dissertations, 2020-*. 1421.

<https://stars.library.ucf.edu/etd2020/1421>

IDENTIFYING CHALLENGES AND OPPORTUNITIES FOR DESIGNING SOCIAL MEDIA  
NUDGES FOR ADOLESCENTS

by

OLUWATOMISIN OBAJEMU  
B.S. Babcock University, 2020

A thesis submitted in partial fulfillment of the requirements  
for the degree of Master of Science  
in the Department of Computer Science  
in the College of Engineering and Computer Science  
at the University of Central Florida  
Orlando, Florida

Fall Term  
2022

© 2022 Oluwatomisin Obajemu

## **ABSTRACT**

With the prevalence of online risks encountered by youth online, strength-based approaches such as nudges have been recommended as a potential solution to subtly guide teens toward safer decisions. However, most nudging interventions to date have not been designed to cater to teens' unique needs and online safety concerns. To address this gap, this study aimed to better understand adolescents' perceptions and feedback on online safety nudges to inform the design of more effective online safety interventions. We conducted 12 semi-structured interviews and 3 focus group sessions with 21 teens (13 – 17 years old) to get their feedback on three types of nudge designs from two opposing perspectives (i.e., risk sender and victim) and for two different online risks (i.e., information breaches and cyberbullying). We found that teens preferred actionable nudge approaches, with the action based on the specific risk scenario. Additionally, for both the risk sender and victim, teens wanted nudges to emphasize warnings, making them harder to ignore. They desired actionable nudges that intervene early and extend beyond a bare warning notice to suggested safe responses. Teens also wanted nudges that prevent the risk sender from perpetuating harm by restricting or penalizing them. Finally, teens wanted personalized and controlled nudges that confirmed their final actions, did not interrupt their regular online activity and had no possibility of escalating the risk. Overall, we found that nudges need to be contextualized to teens' risk experiences, risk medium, personal preferences, and user perspectives (e.g., victim vs. sender).

## **ACKNOWLEDGMENTS**

I would like to thank Dr. Pamela Wisniewski, Zainab Agha, and Farzana Chowdhury, and the other researchers at the STIR Lab for helping me with the project from start to finish. I would also like to thank my committee (Dr. Charles Hughes, Dr. Pamela Wisniewski, Dr. Mary Jean Amon, Dr. Yao Li), for putting in their years of expertise to provide meaningful feedback to improve this research.

This research was supported by the William T. Grant Foundation (#187941, #190017). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of our sponsor. I would also like to appreciate every teen that participated in this study.

# TABLE OF CONTENTS

LIST OF FIGURES .....	vii
LIST OF TABLES .....	viii
CHAPTER ONE: INTRODUCTION.....	1
CHAPTER TWO: LITERATURE REVIEW.....	4
Nudge Based Interventions against Unsafe Online Experiences.....	4
Towards Teen-Centric Online Safety Design.....	5
From Abstinence-Only to Intelligent Assistance.....	7
CHAPTER THREE: METHODOLOGY .....	10
Study Overview .....	10
Risky Scenarios and Nudge Approaches .....	11
Information Breach Risky Scenario ( <i>IBs</i> ).....	11
Cyberbullying Risky Scenario ( <i>CYs</i> ).....	11
General Warning Nudge ( <i>GWN</i> ) .....	13
Sensitivity Filter Nudge ( <i>SFN</i> ).....	13
Guided Actions Nudge ( <i>GAN</i> ).....	13
Session Procedure .....	15
Participant Recruitment and Demographics .....	17
Data Analysis Approach .....	18
CHAPTER FOUR: RESULTS .....	20

Most effective nudges in dealing with unsafe interactions online (RQ1).....	20
Sensitivity filters were generally considered to be the most effective nudge type (RQ1a):.	20
Guided Actions were considered more appropriate for Information Breaches only .....	21
How the teen’s rankings varied by user’s perspective.....	23
Factors that may alter the perception of nudges (RQ2) .....	24
The medium of the risk.....	24
Age of the teen .....	25
User awareness of the risk .....	25
Perceived severity of the risk.....	26
Nudge challenges and recommendations offered by the teens (RQ3).....	27
Designing Actionable and Convincing Nudges.....	27
Nudging for Timely Risk Prevention.....	28
Solving the risk with nudges.....	29
CHAPTER FIVE: DISCUSSION.....	32
Enforced Risk Prevention for the Perpetrator vs Control for the Victim .....	32
Overcoming the challenges of effective nudging for positive behavioral change.....	33
Tailoring nudge to provide autonomy and personalization .....	34
Nudge Design Recommendations based on Empirical Evidence .....	34
Limitations and Future Work.....	35
CONCLUSION.....	37

APPENDIX A IRB APPROVAL ..... 38

APPENDIX B IRB APPROVAL CONNECTION STATEMENT ..... 40

APPENDIX C FULL VIEW RISKY SCENARIOS ..... 42

APPENDIX D FULL VIEW VICTIM NUDGES (INFORMATION BREACH) ..... 44

APPENDIX E FULL VIEW VICTIM NUDGES (CYBERBULLYING) ..... 47

APPENDIX F FULL VIEW PERPETRATOR NUDGES (INFORMATION BREACH) ..... 49

APPENDIX G FULL VIEW PERPETRATOR NUDGES (CYBERBULLYING) ..... 52

REFERENCES ..... 54



## LIST OF FIGURES

Figure 1: Risky Scenarios .....	12
Figure 2: Information Breach Victim Nudges .....	14
Figure 3: Cyberbullying Victim Nudges.....	14
Figure 4: Information Breach Sender Nudges .....	14
Figure 5: Cyberbullying Risk Sender Nudges .....	15
Figure 6: Rankings by User perspective (21 teens x2 scenarios) .....	21
Figure 7: Rankings by Risk Type (21 teens x2 scenarios).....	23
Figure 8: Rankings by User perspective (21 teens x2 scenarios) .....	24

**LIST OF TABLES**

Table 1: Participants’ Demographic Information ..... 17

Table 2: Nudge Challenges and Suggestions Table..... 19

Table 3: Provided Responses to Sensitivity Filters..... 29

## CHAPTER ONE: INTRODUCTION

Many teens experience risks on social media, and they often prefer having control over the decision to handle online risks autonomously as opposed to with the help of a third party in the form of a parent, guardian, or respected adult (Badillo-Urquiola, et al., 2019; Schaeffer, 2019; Mayo Clinic Staff, 2022). These teens have shown an affinity for “just-in-time” interventions for their online safety, which helps them (in that exact moment) when a risk occurs instead of safety features that can restrict or limit their online activity (Badillo-Urquiola, Shea, Agha, Lediaeva, & Wisniewski, 2021; Badillo-Urquiola, et al., 2019). These interventions can be implemented via nudging, which involves the use of any feature that can alter a user's decision without removing their autonomy (Thaler & Sunstein, 2009). Recent research suggests that nudges seem to be a way forward in ensuring adolescents' online privacy and safety (Masaki, et al., 2020).

Nudging has been successfully implemented in other fields (of privacy, security, and other disciplines) to promote positive choices to people (Hartwig & Reuter, 2021; Masaki, et al., 2020; Mele, Russo Spena, Kaartemo, & Marzullo, 2021; Rozin, et al., 2011). In the context of this study and online safety, nudges are used to make indirect suggestions to positively influence a teens’ response to a risky scenario (Badillo-Urquiola, Shea, Agha, Lediaeva, & Wisniewski, 2021). Very few studies involve teens in the process of designing and evaluating online safety nudges, the majority of feature development is usually done by older experts belonging to a different population group. As such, it is important to interact directly with teens to better understand how nudges could be made to be effective for them and cater to their unique experiences. There is also a noticeable trend in adolescent online safety where online safety features for teens are transitioning from a parent-based or restrictive approach to a teen-centric approach that involves giving them the ability to handle their online safety issues autonomously

(Cameron, 2018; Wisniewski, Ghosh, Xu, Rosson, & Carroll, 2017), This philosophy supports the goals the nudges implemented in this study aims to achieve. (Alderman, 2018; Badillo-Urquiola, Shea, Agha, Lediaeva, & Wisniewski, 2021; Mele, Russo Spena, Kaartemo, & Marzullo, 2021) .

Based on the results of a previous study where teens co-designed online safety interventions (nudges) (Agha, Zhang, Obajemu, Shirley, & J. Wisniewski, 2022), as well as common nudge interventions implemented on social media, we identified major themes across nudges and developed these designs to represent those themes. The designs include General Warning, where there is a provided warning statement for the teen via a popup, the Sensitivity Filter, where the detected risk content is automatically censored and the teen receives a warning before consenting to see the risky content, and the Guided Actions, where the platform provides automated actions and responses to the risk as a way of assisting the teens. Because online risks occur in two perspectives (risk sender and risk victim), we added a new dimension to this study by implementing all nudges from those two perspectives. All curated nudges were applied in two risk settings: an information breach risky scenario, where a stranger is making an attempt to convince the teen to disclose their personal information, and a cyberbullying risky scenario, where a stranger is harassing the teen with unwarranted and rude messages.

To understand teens' evaluation of the three online safety nudges, we ask the following research questions.

**RQ1:** What nudges are considered most effective by teens for dealing with unsafe interactions online based on a) nudge approach, b) risk type, c) user perspective?

**RQ2:** What other key differences emerged that influenced teens' impressions of online safety nudges?

**RQ3:** What were the challenges identified by teens when evaluating the nudges and what were their recommendations to address them?

To answer these research questions, we conducted three focus groups with three teens each, and twelve 1-on-1 interviews. During the sessions, the teens were presented with two different online risks- information breach and cyberbullying, each scenario had an implementation of the guided actions, sensitivity filter, and general warning nudge approaches. The teens were then prompted with a series of semi-structured questions as a means of collecting their feedback on the nudges from two opposite perspectives (sender and receiver). The questions were asked to determine: (a) their impressions and perceptions of the nudges, (b) their suggestions to improve the nudges, (c) their concerns with the nudges, and (d) their overall preferences. We found out that teens prefer different nudges based on the context of the risk, the teens also found nudge attempting based on multiple reasons such as being previously nudged improperly, and a subjective intention that cannot be easily changed. We also discovered that adolescents want online safety nudges to be more convincing by providing more specific and contextual information, and finally, the teens expressed a desire for online safety warnings to not only warn the user, but to provide an accessible and appropriate response or solution to solve the risk. Based on our methods and results, we make the following contributions to adolescent online safety research.

1. A teen-centric approach to designing and evaluating online safety nudges.
2. Actionable recommendations to improve online safety nudges for the adolescent population.
3. Consideration of both victim and risk sender perspectives in designing online safety features.

## CHAPTER TWO: LITERATURE REVIEW

Extensive literature has explored the concept of nudging in relation to adolescent online safety. We started by reviewing prior works about Nudge-based interventions against unsafe online experiences. Then we analyzed how the standard online safety mechanisms for teens are transitioning from a dependent parent-centric approach to a more independent teen-centric approach. Then, we explored further materials that suggest intelligent assistance as a solution to online risks rather than approaches that simply limit the teen's usage without providing any other benefits.

### Nudge Based Interventions against Unsafe Online Experiences

Social media is popular among teenagers of which 95% in the US carry at least one smartphone and ~45% are always active online (Anderson, 2018). The unsafe online experiences faced by these adolescents include sexual risks, information breaches, and cyberbullying (Duggan, 2017). According to a study by The Pew Research Center, it was discovered that 60% of teenagers encountered online threats and violence (Anderson, 2018), and nearly 50% of teenagers are prone to sharing personal information online (Gainsbury, Browne, & Rockloff, 2019; Harriman, et al., 2020). With the number of constantly online risk-prone adolescents, it is important to ensure their safety and welfare as a vulnerable population.

There has been a general adoption of online safety nudges by common social media platforms like Twitter, Instagram, and Snapchat as a way of improving the online safety of their users. On Twitter, tweets exposing violent images or hateful comments can be censored and marked as sensitive, with more rigorous and unyielding censoring policies put in place for younger (<18 years) users to restrict their access to risky content (Twitter Help, 2022). Instagram

flags unusual adult activity of frequently sending friend requests and direct messages to teenagers (João Carrasqueira, 2021), and also tracks down and keeps vilifying messages in a separate folder (Usman Khan Lodhi, 2021). With a significant number of features engaged to nudge users away from danger, social media users have a plethora of safety features available to them.

Several investigations have been carried out to evaluate the effectiveness of nudges in ensuring the online safety of adolescents as they are of a unique population group. Similar to this study, Masaki et al. considered eleven (11) online safety interventions for nine (9) different risky scenarios, and they gathered teenagers' feedback on them. They discovered that adolescents act more carefully when they receive nudges promoting consciousness (Mele, Russo Spina, Kaartemo, & Marzullo, 2021). Another study by Alemany-Bordera et al. supplements that finding by demonstrating that teens are specifically targeted by risk perpetrators on social networks and the warnings sent by paternalistic interventions or nudges (through bold texts and images) may empower them to understand the associated risks and reconsider their actions (Alemany-Bordera, Del Val Noguera, Alberola Oltra, & García-Fornes, 2019). In addition to influencing individuals' actions, nudge intervention does also function autonomously by blocking inappropriate content, taking legal action, providing mental support, etc. (Weinmann, Schneider, & Brocke, 2016).

### Towards Teen-Centric Online Safety Design

According to Wisniewski et al., most of the research on online risks with teens assumes a risk-centric perspective to the risks, and the common approach taken is done with the view that teen users are highly risk-prone and those risks (as well as the experiences leading up to them) are negative and should be avoided (Wisniewski, Ghosh, Xu, Rosson, & Carroll, 2017; Wisniewski,

Xu, Rosson, & Carroll, 2017). The studies cover trends in current teen technology features, as well as their impact and ultimately concludes that these risk-centric models may hinder the personal development of the teens and have negative implications for the parent-adolescent relationship. Instead, the authors suggest that risk-taking is a necessary element of developing into a young adult and online safety features need to encapsulate a risk-taking element.

There are multiple examples of online applications that utilize this risk-centric approach in dealing with adolescent online risks i.e., a study determined that 89% of 75 mainstream mobile applications utilized a form of risk-centric approach in the form of parental controls (Cameron, 2018). Most of these apps were discovered to be privacy-invasive to the young users as they provided direct access and control of their intimate social experiences to their parents (Chouhan, et al., 2019) which further discourages adolescents from openly sharing their issues and challenges with their parents (Statista Research Department, 2022).

More recent research confirms the need of reducing parental involvement in online apps as a means of enhancing teen online safety (Cameron, 2018). However, there still exists a belief that skews much of the research recommendations and best practices to parental involvement (Wisniewski P. , *The Privacy Paradox of Adolescent Online Safety*, 2018), and it poses a challenge in the long-term effectiveness of online safety features for adolescents. This is further backed up by multiple studies such as by Agatston (Agatston, Kowalski, & Limber, 2012) where it is stated that the youth have an anti-authoritarian preference as a cyberbullying coping strategy and would prefer seeking help from their peers or by extension, autonomously. As a result, the research landscape has transitioned into a more teen-centric approach to regulating risky online experiences for adolescents (João Carrasqueira, 2021).



In conclusion, it has been established that self-regulation and teen-centric approaches to online safety is the way forward, and there is a need to transition away from the commonplace risk-centric approaches. This study aims to address the lack of resources and systems that hold the teen-centric approach by giving the teens complete control of their online safety. This concept invokes a favourable opinion from the teens (Mele, Russo Spena, Kaartemo, & Marzullo, 2021).

### From Abstinence-Only to Intelligent Assistance

Information breaches, sexual solicitations, and cyberbullying are among the harmful elements teens face on the internet (Kim, Colwell, Kata, Boyle, & Georgiades, 2018). The risk-centric approach to online risks described in the previous session suggests that unsafe online experiences can be mitigated by abstinence and by reducing access to the internet. This idea is implemented in our sensitivity filter nudge where the teens' access to the risky message is limited or reduced. People make decisions on digital platforms, and according to Weinmann et al., user interface designers have evolved into choice architects who affect people's decisions in digital settings (Weinmann, Schneider, & Brocke, 2016). As a result, a successful nudge considers the context of the digital environment in which the user is present in (Alderman, 2018). This is followed-up in a study by Rodriguez et al. where the authors evaluated ways to integrate a nudge intervention mechanism with an artificial intelligence (AI) system to successfully recognize learners at risk of failing or dropping out early (Rodriguez, Guerrero-Roldán, Baneres, & Karadeniz, 2022), this bespoke approach to nudging helps classify the Guided Actions nudge.

In cases when security-related decisions are ambiguous, Yevseyeva et al. propose nudging by considering the context of the decision making environment and the fact that an individual may be in a better position to make a more appropriate decision rather than following

strict regulations - which informs the goals of nudging in general, which is why all nudges in this study have an option for the user to opt out and make their own decision (Yevseyeva, et al., 2014).

Conventional victimization and criminology models have focused on the standardization of two kinds of mutually restrictive involvement to comprehend online delinquent behavior (victims and perpetrators) (Badillo-Urquiola, Chouhan, Chancellor, De Choudhary, & Wisniewski, 2020). As previously stated by Vale et al., these roles are rarely static, and many teens assume a cyber-double role, that is, victim-perpetrators (Vale, Pereira, Gonçalves, & Matos, 2018). Control, social-learning, and social-information reception theories have received a lot of support as justifications for this double cyber participation compared to the actual world (Hummel & Maedche, 2019).

As a result, it is possible that adolescents would be unable to resist the opportunity to offend. Thus, it makes them vulnerable to their cyber-risky behaviors, may reduce adolescents' self-control, and increase their propensity to be short-sighted, risk-takers, and be exposed to deviant models, justifications, and reinforcements. To prevent cyber aggression victims from becoming perpetrators, Vale et al. argue that there is a pressing need to ramp up efforts to raise awareness and educate people about how to use the internet responsibly. In this commitment, adolescents, parents, educators, formal sources, and society all play a role (Badillo-Urquiola, Chouhan, Chancellor, De Choudhary, & Wisniewski, 2020). For this purpose, data gathering should be triangulated across teens, parents, and adolescent-parent dyads to uncover new risk variables.

In conclusion, rather than focusing efforts on avoiding minors from being exposed to all digital risks, it may be more advantageous to teach them more effective risk-coping methods.

Instead of using abstinence-based strategies to prevent teens from using social media, it may be more useful to teach them how to report abusers to the appropriate authorities and effectively manage the risk in other ways (Badillo-Urquiola, Chouhan, Chancellor, De Choudhary, & Wisniewski, 2020) .In particular, guided actions that help youth participate in more proactive measures may minimize future risk exposure and help them cope with post-traumatic stress (Lehner, Mont, & Heiskanen, 2016). The ultimate objective for teens should be to enjoy the benefits of social media activity while avoiding the long-term detrimental impacts of risk exposure by following a set of goal-directed actions.

## CHAPTER THREE: METHODOLOGY

In this section, we describe the methods used to carry out this study, including the study design, nudge descriptions, recruitment strategy, and data analysis approach.

### Study Overview

We conducted 15 sessions (3 groups and 12 interviews) with 21 teens to get their feedback on three online safety nudges which were implemented in two risky scenarios. The three nudges were designed based on teens' ideas from previous co-design sessions (Agha, Zhang, Obajemu, Shirley, & J. Wisniewski, 2022), as well as current trends in adolescent online safety research and relevant platforms (Masaki, et al., 2020). The nudges include a) General Warning – a pop-up nudge warning the teen about the risk with an option to dismiss, b) Sensitivity Filter - a nudge that censors the received risky content while also giving a warning, and c) Guided Actions - a nudge that suggests risk responses to the user. The purpose of implementing them in multiple risk scenarios is to provide a result that is inclusive of more than one risk type. These nudges were also implemented from the perspectives of both the risk victim and risk sender because the teens from the co-design sessions suggested them over the risk victim nudges.

To carry out the evaluation process, we conducted feedback sessions with 21 participants (aged 13-17) via a Zoom video call and a shared virtual whiteboard (FigJam) to understand their assessment and impressions of these nudges. The sessions were designed in a way that made the teens able to express their ideas beyond words by combining the traditional discussion elements of an interactive user study (interviews, focus groups) with the whiteboarding style of participatory design where they were able to mockup and sketch their feedback (Agha, Zhang, Obajemu, Shirley, & J. Wisniewski, 2022; Badillo-Urquiola, Shea, Agha, Lediaeva, &

Wisniewski, 2021; Badillo-Urquiola, et al., 2019; Badillo-Urquiola, Chouhan, Chancellor, De Choudhary, & Wisniewski, 2020). The sessions were run virtually on Zoom in conjunction with FigJam (a collaborative virtual whiteboarding tool) with 21 teens ranging from 1-3 teens per session, which lasted for approximately 2 hours. Around three researchers moderated each focus group session.

### Risky Scenarios and Nudge Approaches

The risky scenarios used in this study are based on teens' unsafe experiences which they shared during the previous research to ensure our examples were based on a realistic scenario a teen might experience and relate to. We made use of 2 low-level risky scenarios to implement all three nudges. The full implementations can be found in appendices C-G.

#### *Information Breach Risky Scenario (IB<sub>s</sub>)*

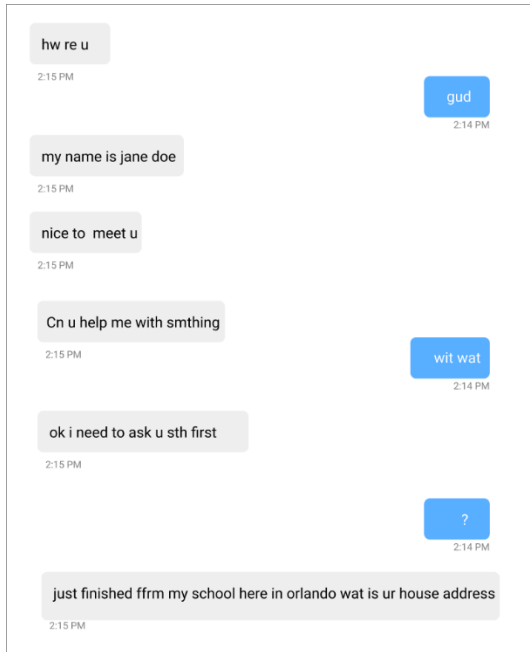
This is the first risky scenario presented to the teens. As shown in Figure 1, it is an information breach risk type where a stranger sends a private message to the teen and requests them to disclose their personal information (i.e., house address).

#### *Cyberbullying Risky Scenario (CY<sub>s</sub>)*

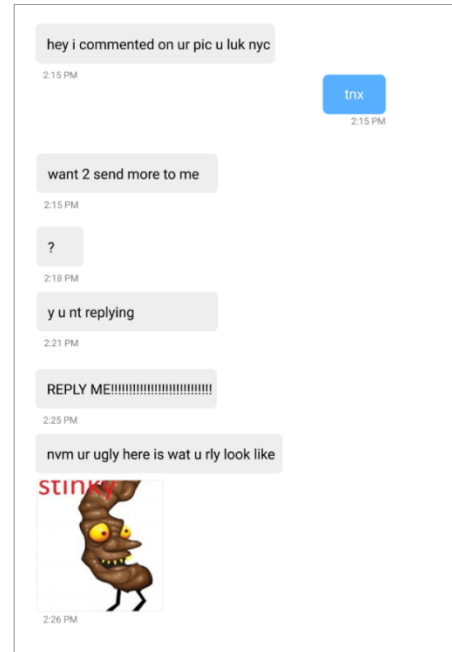
This is the second risky scenario presented to the teens. Figure 1 shows a cyberbullying risk scenario where a stranger sends multiple messages to harass and insult the teen.

A major aspect of the feedback process involves the user as a victim in a hypothetical risky scenario to assess the effectiveness of the nudges in that context. However, for the same scenario, we also took a preventive approach of using nudges to deal with online risks by injecting the nudge to the risk sender as a means of preventing the detected risk from being sent, this was the more common perspective in the co-designed nudges by the teens in prior work. We implemented both perspectives by introducing the teens to the risky scenario as a victim, giving

them the prepared prompts, getting their feedback, before finally asking them to provide feedback from the perspective of the risk sender, i.e., if they think the nudge would prevent the risk sender from sending the risky content.



Information Breach Risky Scenario



Cyberbullying Risky Scenario

*Figure 1: Risky Scenarios*

All nudges from this study were derived from prior co-design research by Agha et al. (Agha, Zhang, Obajemu, Shirley, & J. Wisniewski, 2022), where 21 teens were taught how to create a UX high-fidelity prototype, they were able to design a feature to help solve a risky scenario of their choice. The ideas provided by the teens in their solutions informed the nudge designs in this study where we present three ways to classify nudges based on their approach to handling online risks. They include Guided Actions, General Warning, and Sensitivity Filter.

### *General Warning Nudge ( $GW_N$ )*

This was a common nudge method suggested by the teens from the Bootcamps. The General Warning nudge performs the action of letting the teen know that they are experiencing a risky scenario and urges them to be careful. This is implemented in Figures 2 and 3 as a pop-up warning to a potential victim as soon as the risky message is detected. In Figures 4 and 5, the risk perpetrator receives a similar pop-up message with a snippet of their message, letting them know the message has been detected as risky while recommending them to review it.

### *Sensitivity Filter Nudge ( $SF_N$ )*

This was also a common nudge method suggested by the teens from the bootcamp where the nudge censors the risky content and notifies the teen about the presence of a detected risk while giving them the option to show the content or hide it, as shown in Figures 2 and 3. In Figures 4 and 5, Based on the assumption that there isn't much reason to censor a message from the originator, the risk sender receives the sensitivity filter as a pop-up reprimand that the message has been tagged as risky, and the receiver would see a censored version.

### *Guided Actions Nudge ( $GA_N$ )*

From the UX Bootcamp study, 7 out of 21 teens suggested Guided Actions in the form of an intelligent or virtual assistant that suggest safe responses as a method of helping teen victims deal with risky scenarios online. Generated responses have also been seen as a valid element of nudges in prior literature (Alderman, 2018). This nudge involves the platform generating safe responses and actions for the teens to deal with the risky scenario. When the teen receives a detected risky scenario, they are presented with a list of safe auto responses as a suggestion, with an option to respond directly. This is shown in Figures 2 and 3. For the risk perpetrator, an autocorrect-based implementation is taken where each detected risky content by the teen is

underlined and marked red while they are typing with suggested replacement messages and a clear message option above the text box as it is in 4 and 5.

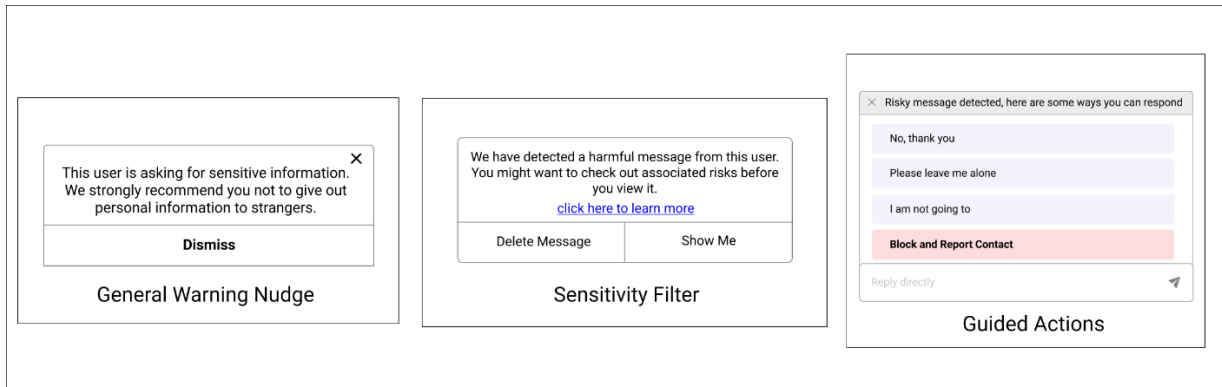


Figure 2: Information Breach Victim Nudges

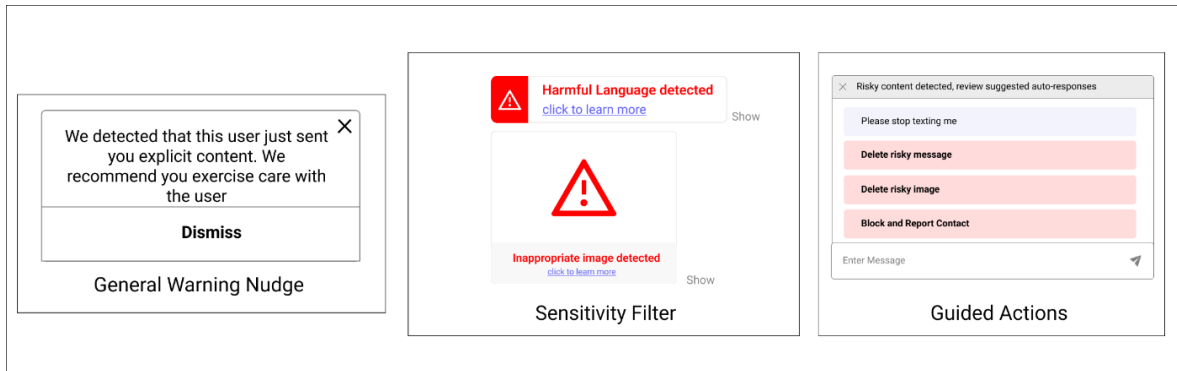


Figure 3: Cyberbullying Victim Nudges

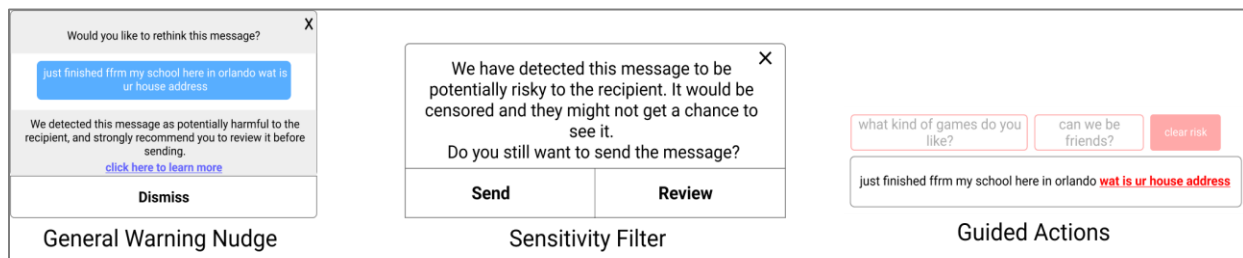


Figure 4: Information Breach Sender Nudges





Figure 5: Cyberbullying Risk Sender Nudges

### Session Procedure

The sessions were conducted online via a Zoom video conference and lasted approximately 120 minutes. At the start of each session, each participant was introduced to the researchers, the concept of nudging, how nudging is used in other domains (food ordering, video streaming, etc.), and an ice-breaking activity was performed where each participant had to introduce themselves and to identify common interests with the researchers as a means of building rapport.

The teens were then given a crash course on adolescent online safety and nudges to understand the goals of the session and what is expected of them. Considerable effort was also put in making the teens aware of the subjectivity of their responses. To promote interaction, the participants were asked to discuss what online safety means to them as a teen and to share an instance when a nudge has influenced their decisions.

The feedback process included 12 nudges spread across (a) 2 perspectives, (b) of which each had 2 scenarios, (c) with 3 nudges each. The scenarios and their respective nudges were presented to the teens through a click-through high-fidelity prototype on Figma, this was done to allow them clearly understand the flow of the scenarios and the functions of each nudge. To request feedback for the nudges, a researcher walked the teens through a prototype of the risky scenarios (in the order of: Information Breach, Cyberbullying) and their associated nudges (in the order of: General Warning, Sensitivity Filter, and Guided Actions), after which was a switch

to the whiteboarding and discussion activity on FigJam to allow the teens provide their feedback. The researchers used sticky notes, drawing tools, and shapes to mockup the teen's feedback and ideas to improve the nudges. The responses were a mixture of verbal and design elaboration over the nudge prototype. The presentation and co-design activity processes were carried out for each nudge, and the teens were asked to rank the nudges in order of their personal preferences.

We followed a semi-structured approach in giving the participants question prompts to generate feedback. This was done to allow the participants express themselves while giving us room to get the needed feedback. The general prompts given are highlighted below.

1. What do you think about this nudge?
2. How, if at all, do you think this nudge addresses the risky situation?
3. Why yes/no?
4. How can they be improved? If you were the designer, what would you change about them?
5. What could go wrong with this nudge? Can the nudge make the risk worse?
6. Which, if any, of these nudges do you think works best for ensuring a safer interaction?

After receiving feedback from the teens, they were given the opportunity to ask further questions about the logistics of the sessions, nudging, and any other thing that required clarification, and they were also given a link to a feedback survey. In the feedback survey, the teens were asked a series of Likert scale and open-ended questions about their demography and multiple aspects of the sessions which include tools used, pacing, ability to express feedback and future suggestions. The feedback survey results were taken into consideration to revise the following sessions.

### Participant Recruitment and Demographics

After receiving IRB approval, the participants were recruited from youth-serving organizations, social media, and middle/high schools around the United States. The mode of recruitment includes flyer distribution, phone calls, and emails. We supplemented our open recruitment strategy with existing contacts in the STIR lab participant database from previous studies. All participants were aged 13-17, based in the United States, and able to communicate in English. Majority of the teens were 16 years (N=10, 47.6%) with the mean age and standard deviation being 15.6 and 1.29 respectively. The identified racial identities of the participants are as follows: Asian (47.6%), Black/African American (23.8%), Hispanic/Latino (9.52%), and White/Caucasian (28.5%). The participants comprised of 8 males (38%) and 13 females (64%) (Table 1). They were compensated with a \$20 amazon gift card on completion of the study.

*Table 1: Participants' Demographic Information*

<b>Session</b>	<b>ID</b>	<b>Sex</b>	<b>Age</b>	<b>Ethnicity</b>
Session 1	P1	F	14	Black/African
Session 2	P2	F	15	Black/African
	P3	M	16	White/Caucasian
	P4	M	13	White/Caucasian
Session 3	P5	F	13	Black/African
Session 4	P6	M	17	White/Caucasian
	P7	F	16	Asian
	P8	F	16	Asian
Session 5	P9	M	17	Asian
Session 6	P10	M	13	Asian
Session 7	P11	F	15	Asian
Session 8	P12	F	17	Asian
Session 9	P13	M	16	White/Caucasian
	P14	M	16	White/Caucasian, Hispanic/Latino
	P15	F	16	Black/African
Session 10	P16	F	16	Black/African
Session 11	P17	F	17	Asian
Session 12	P18	F	16	Asian
Session 13	P19	F	16	Asian
Session 14	P20	M	17	Asian
Session 15	P21	F	16	White/Caucasian, Hispanic/Latino

### Data Analysis Approach

The data obtained from the teens' feedback via the whiteboard annotations and session transcripts were analyzed using Braun and Clarke's thematic analysis guide (Braun & Clarke, 2006). The recorded sessions were transcribed to text using Otter.ai. The primary source of data is from the responses and discussions stemming from the question prompts given during the presentation of each nudge. The co-design whiteboard artefacts were also considered as some of the annotations and sketches over the nudge design contained valuable information. Two researchers participated in the data analysis process where there was an initial coding of the data to note emerging ideas which were grouped into the major themes. The codebook was coded along the dimensions of nudge response, nudge feedback, and areas for improvement. The two researchers had consistent meetings to merge their individual codes and resolve conflicts. The nudge ranking data was treated quantitatively and analyzed using cross tabulation. It helped answer the question of which nudges were preferred, while the qualitative thematic data answered why those nudges were preferred. Our final codebook is described in the results section.

Table 2: Nudge Challenges and Suggestions Table

Themes	Subthemes	Count	Exemplar
Nudges are easy to Ignore	The user’s intent can’t be changed by nudging	(15, 71%)	“I don't think this would be helpful at all, because they already have harmful intentions.” P7
	Nudges might be prone to warning fatigue	(7, 33%)	“After a while, teens will get so used to that pop up that they'll sort of ignore it.” P4
Nudges might have usability issues	Nudges might interfere with regular use	(15, 71%)	“It would get tedious to have that [nudge], just like in every conversation.” P21
	Nudge may happen too late	(9, 43%)	“Prevent repeated unwanted messages before you got to the point of like, this image.” P6
Nudges can have an adverse effect	Censoring might make the risk more appealing	(15, 71%)	“I don't feel like teens are going to ignore this like, 'Oh, harmful, harmful'...It might even get some teens to see it more.” - P3
	Nudges might escalate the risk they address	(12, 57%)	“The idea of like continuing the interaction [with the risk sender] could also be harmful.” P21
Nudges should warn users better	Emphasize risk harm using design cues	(20, 95%)	“Capitalize harmful...If they don't look at the message maybe [they'll] look at the bold words.” P11
	Explain the need for a warning	(17, 81%)	“It could elaborate on like why they don't recommend it...why should I care?” P9
Nudging to solve the risk optimally	Provide safe actions for the user	(19, 90%)	“I feel like it would [should] give you an option to block this user [risk perpetrator].” P4
	Limit who safety features apply to	(15, 71%)	“I would like [the nudge] to like not be there when I'm talking to my friend. ” P10
	Require the user to confirm their action	(10, 48%)	“There could be another pop up like, 'are you sure?' like, 'this is sensitive information'” P2
Nudging to prevent the risk	Prevent the perpetrator from sending risk	(17, 81%)	“You want to prevent them [risk perpetrator] from like ever sending it out instead of, they send it out and like someone [the risk victim] just blocks it.” P20
	Include reprimands for the risk perpetrator	(17, 81%)	“Make you aware that the receiver end might delete it and not get a chance to see it.” P12
	Penalize perpetrator for their harmful actions	(5, 24%)	“Maybe like a strike, and like [in] three strikes, you're banned.” P9

## CHAPTER FOUR: RESULTS

In this section, we discuss the results gotten from teens' responses across all focus groups and interviews according to the research questions.

### Most effective nudges in dealing with unsafe interactions online (RQ1)

In the presented risk instances, the teens preferred the  $SF_N$  or  $GA_N$  for dealing with their online safety issues for both risk perspectives and they mostly considered the  $GW_N$  to be least effective in all instances. Their nudge preferences are based on the nudge's ability to provide a safe function (censor, risk response) which goes beyond warning only. Their feedback is covered in more details in the following sections. In (RQ1a), we looked at the nudges being ranked against one another, then we looked at how the nudge rankings changed based on the risk type (RQ1b), and lastly, we compared nudges based on the two user perspectives (RQ1c).

*Sensitivity filters were generally considered to be the most effective nudge type (RQ1a):*

All the nudge approaches ranked were against one another without any external considerations, and it was discovered that the teens liked all nudges for providing a degree of risk awareness, but they explained that the differences that affect how these nudges are perceived lie within how the nudges function in addition to the risk awareness. The  $GW_N$  was seen as depending on the teen knowing what to do without implying any real solution, which they did not like because they felt the nudge could do more, and as a result, most of the participants ranked it worst. On the positive side, the  $GA_N$  was found to provide actionable prompts which could be of great help to the user, especially in a situation where they do not know how to respond to the risky scenario. However, it was often criticized for possibly making the risk worse for the user by encouraging victim-perpetrator communication, therefore prolonging the risky encounter. The nudge considered to

be most effective by the teens is the  $SF_N$  because they found censoring the risk to provide control to the user by preventing them from seeing what they are not supposed to.

*“[ $SF_N$ ] would actually be the top just because you don't really see the image or the message, [...] and then I would do maybe the third one [ $GA_N$ ] next because the options of having like being able to choose, [...] and probably just the first one, [ $GW_N$ ] just because you can still see everything in the chat.” - P20. 17, M ( $CY_s$ )*

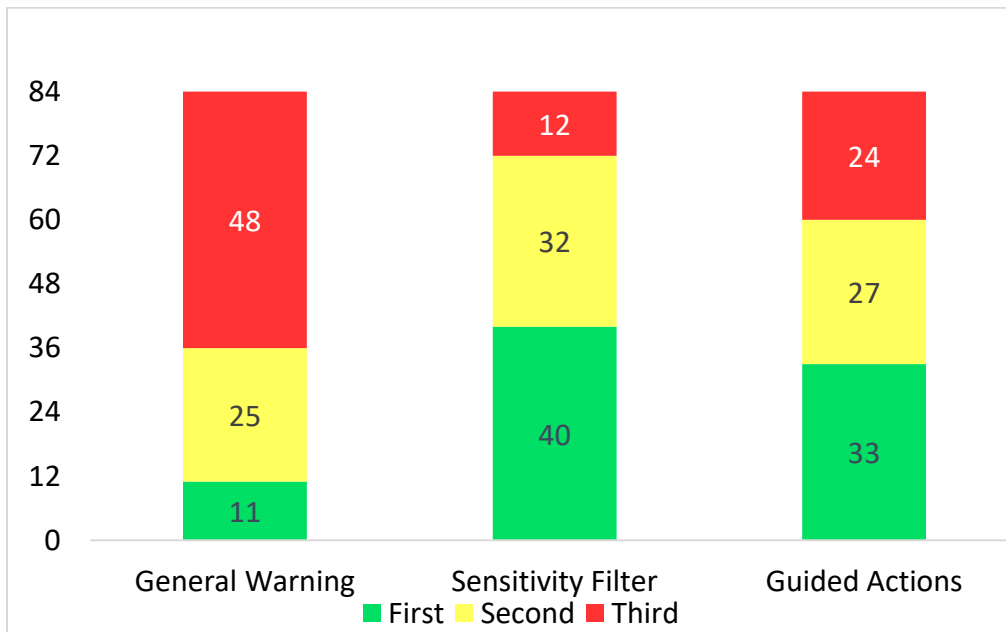


Figure 6: Rankings by User perspective (21 teens x2 scenarios)

*Guided Actions were considered more appropriate for Information Breaches only*

Both risk scenarios which the nudges aim to solve are different. The danger of an Information Breach Scenario ( $IB_s$ ) is by responding and divulging personal information to the risk perpetrator, while the danger of the Cyberbullying Scenario ( $CY_s$ ) lies in viewing or the exposure of the victim to the risky messages. While the teens generally preferred the  $SF_N$  overall, there was a difference in their preferred nudge for each scenario. However, their impressions of the  $GW_N$  remained negative for both the  $IB_s$  and the  $CY_s$  for not implying a risk solution remained.

For the  $IB_s$ , the teens did not mind being exposed to the risky message and their main goal was to not disclose their personal information to the risk perpetrator. They found an appropriate response to the risk sender being to ignore or block them, and this led to their most preferred nudge for this risk type being the  $GA_N$  which provides those options. However, the nudge was considered polarizing because more teens ranked it worst when compared to the second ranked nudge  $SF_N$ . This was because of the message responses provided by the nudge possibly prolonging the risky scenario. Teens had a different preference for the  $CY_s$ . In most cases, they did not want to be exposed to the risky scenario due to its explicit nature and tendency to have negative emotional consequences, they liked the  $GA_N$  for easily giving them an option to block the user, but they preferred the  $SF_N$  for censoring the harmful messages sent to the victim.

*“If someone says ‘Hey, what's your address?’ That’s not like affecting me directly, like I could just ignore it, [or] block, like it's not that big of a deal, but if they are genuinely sending like inappropriate pictures, it would be better to just not see that and not even have to make that choice [to see].” – P15. 16, F.*



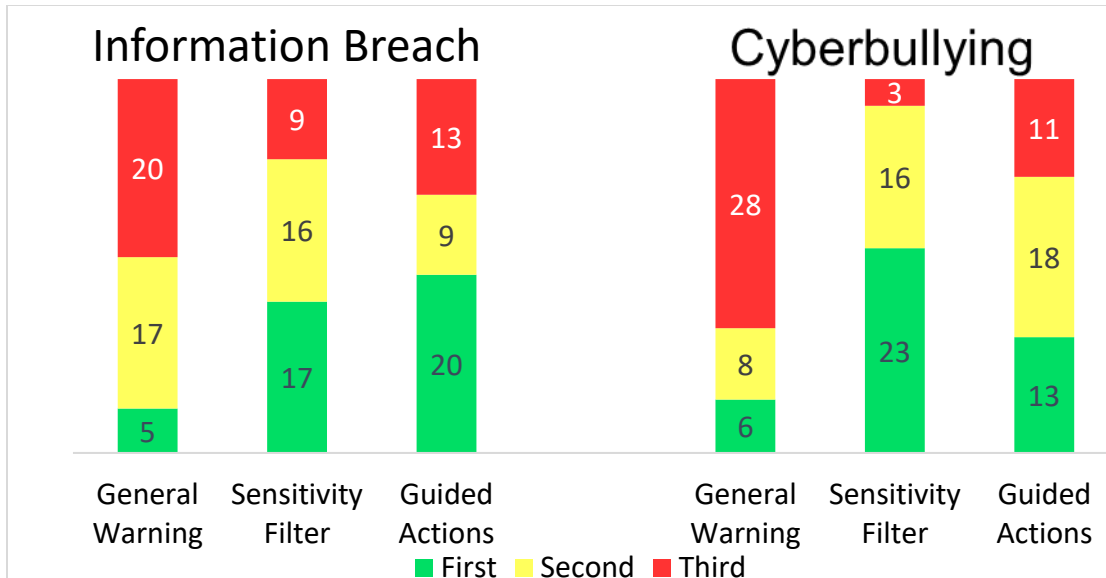


Figure 7: Rankings by Risk Type (21 teens x2 scenarios)

How the teen's rankings varied by user's perspective

With the nudges being implemented from the perspective of both the risk sender and risk victim, most of the feedback remained the same, although they were for different reasons that are unique to the target perspective. For example, the  $GW_N$  was ranked worst for both parties. It was ranked worst for the victim because it was considered easy to ignore without providing an actionable response to the risk, but it was ranked worst for the risk perpetrator because they have a strong intent to cause harm which cannot be changed with a warning nudge. The theme of their harmful intent carries over to the other nudge rankings, such as the  $GA_N$  which was ranked poorly for the risk perpetrator. Meanwhile, the  $SF_N$  was ranked best as the reprimanding language in the text was seen as the most effective tool in making the risk sender reconsider their actions motivated by their harmful intent. Majority of the motives for the victim nudges were covered in the previous section, their strongest nudge preference was the  $GA_N$  due to the responses provided, followed by the  $SF_N$  because they found censoring to be effective at protecting the teen from being exposed to harmful content.

“I like the victims respective because there were all the options to like delete it and block the person, whereas for this one [sender’s perspective], the sender can still send the message and there’s no way of like protecting the person who received the message.” – P17. 17, F (GA<sub>N</sub>)

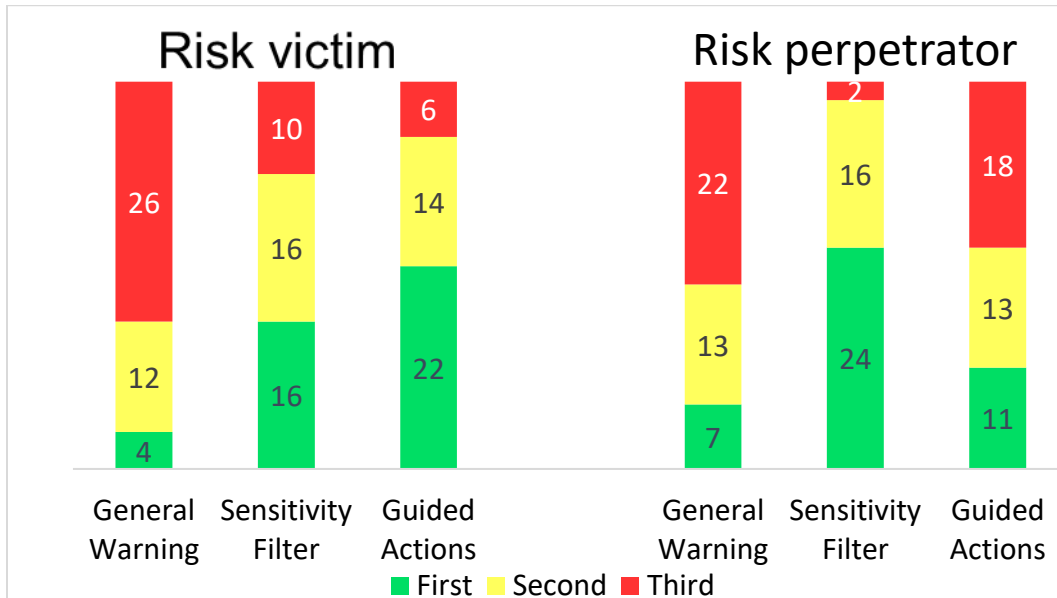


Figure 8: Rankings by User perspective (21 teens x2 scenarios)

Factors that may alter the perception of nudges (RQ2)

Based on the teens’ responses, we were able to identify some external factors or edge cases that might alter how a single online safety nudge might behave or be perceived by adolescents. These factors can be categorized in the following sections.

*The medium of the risk*

A very interesting factor was brought up by the teens in the *CY<sub>s</sub>* where the risk occurred via multiple mediums: (1) offensive language, (2) offensive image, which prompted them to express they would be more wary of an image-based risk over a text-based risk. Noticing this disparity, we followed up with them by asking how they would respond to both risk mediums as they were censored in the displayed *CY<sub>s</sub>*. 55% of the teens said they would view the censored text, while 37% said they would view the censored image.

*“Usually if I see like a harmful photo, I probably wouldn't want to see that. Then I would just ignore that text [censor], I think I would be curious to see what it says.” - P19. 16, F*

For the  $IB_s$ , which does not include an image-based risk, 62% of participants said they would view the censored text. This suggests that different risk mediums might affect how well nudges address them when applied,

#### *Age of the teen*

While our participants included teens aged 13-17, the participants mentioned a disparity on how teens on both ends of that age range might behave in an online risk instance. The trends suggested that older teens might have a greater degree of maturity and competence with dealing with their online safety, while the younger teens might be more vulnerable to these online risks and would benefit from stricter safety measures because they might not have enough experience to learn from risky encounters.

*“Older teens have somewhat of an idea what to do [...], younger teens can feel pressure to be put in a corner to answer to adult [perpetrator] authority.” - P2. 15, F ( $IB_s$ ,  $GA_N$ )*

The participants also suggested generic high-level behavioral differences between younger and older teens which manifests in differing views of their online safety issues and preferences.

#### *User awareness of the risk*

This finding suggests a learning element in risk management, the teens implied that online safety nudges might be viewed differently based on the teen's ability to handle the risk in the form of identifying or knowing how to respond to it. It implies that if the nudge does not offer a novel idea, it might come off as obnoxious or unnecessary to the teen just like P10's quote below.

*“It tells you to be careful. But I guess it's just common sense to be careful that if you just see someone texting you something that looks like that, that they're just weird people or someone that's stupid.” P10. 13, M*

Likewise, a risk perpetrators perception of a nudge would depend on if their awareness of the risky content about to be sent.

#### *Perceived severity of the risk*

The nudge rankings by the teens in Figure 8 shows some general trends in their interests and disinterests. More importantly, there is an inconsistency in the top ranked nudge which shows that teens prefer different nudges for different risk types, and it boils down to different perceived levels of intensity.

*“If someone says, ‘hey what's your address?’ that's not like affecting me directly, like I could just ignore it because like it's not that big of a deal, but if they are genuinely sending inappropriate pictures, maybe like explicit pictures, I feel like the youth can be a lot more susceptible to that.”*

- P15. 16, F

The teens felt that low-level risks might require more passive safety features while the higher-level risks might require a stronger intervention, which is why the teens did not prefer the sensitivity filters until the risk was considered severe enough.

### Nudge challenges and recommendations offered by the teens (RQ3)

Over the course of our interactions with the teens during the sessions, they were able to identify some challenges with online safety nudges and offer recommendations to solve them based on the nudges and scenarios presented. The challenges and recommendations are displayed in the codebook shown in Table 2. The themes in the Nudge challenges and recommendations sections are put in a 1-1 mapping and explained in the following subsections.

#### *Designing Actionable and Convincing Nudges*

The teens (N=15) believed that the nudges might not be doing enough to alter a user's action. For example, a risk sender with a desire and incentive to cause harm to a teen might not be easily nudged away from their action. Likewise, a victim who is convinced that they are not in a risky situation, either due to naivety or believing the risk detection is a false-positive. In addition to nudges being easy to ignore due to the user's intent, the teens said nudges are also easy to ignore due to a preconception ranging from a subjective desire to "not want to read" to more.

Specifically, the teens were concerned about warning fatigue (N=7), where the nudge loses its impact after being triggered multiple times for the user. This is made worse by majority of users having an experience of being nudged improperly in the past, with some teens (N=2), even likening to nudge popup format to a virus or a fake JavaScript popup scam (Kaspersky, 2022).

As a solution to nudges being easy to ignore, the most suggested idea was to emphasize the risk harm using design cues (N=20), like in the following quote from P11, they suggest the use of visual cues and design elements such as bold text as an alternative to text in drawing the teen's attention to the risk.

*"Right now, this just looks like a passage, it's just boring, I don't want to read. You should capitalize harmful, or somehow convince them how dangerous this image might be...If they don't*

*look at the message, maybe [they'll] look at the bolded words and think to themselves that something is really wrong.” – P11. 15, F*

Other suggested design cues include: colour coding, account flagging, bold text, and danger signs. With the emphasis on design cues over text, many of teens (N=17) still wanted the warning message on those nudges to be revised to be non-generic and tailor-made to address the intricacies of the specific detected risk; with the goal being to clearly communicate the underlying threats identified by the nudge to the user. The quote from P10 explains the scenario comprehensively, they express a desire of teens to not want to read warning messages and suggest grabbing their attention with bold text.

#### *Nudging for Timely Risk Prevention*

Many of the teens (N=15) considered online safety features to be intrusive to regular use and would either have the option to disable them for contacts they trust enough to be risk-proof or have them designed in a way that does not detract from the user flow. Additionally, the teens also wanted those solutions to be triggered in a timely manner, a notable amount of teens (N=9) criticized the nudges implemented in the  $CY_{s(v)}$  for being triggered too late because they believed the nudges should have been triggered at the initial point of contact as a means of preventing the risky scenario from reaching a more serious level.

*“Have filters in place to prevent repeated unwanted messages before you got to the point of like, this image.” - P6. 17, M*

As explained by P8 in the following quote, the teens (N=19) expressed a dislike for some of the nudges that provided no actions for the user to take in response to the risk- which suggests an emphasis for online safety nudges to provide effective actions or safe responses in addition to notifying the user about the risk. This ideology is also reflected in their rankings from the

previous section. Finally, the teens (N=10) suggested a follow-up “are you sure?” nudge to confirm the user’s action in case they ignore the nudge and continue engaging with the risk or to have a second thought.

*“There's like no proactive option that you can take, like okay, ‘This is risky,’ you tell us that, but there is no option to act on it.”*– P8. 16, F. ( $IB_{s(v)}$ ,  $SF_N$ )

Some teens (45.5%, N=5) suggested adding a follow up-nudge to any nudge that requires a crucial action to make it more convincing. The motivation behind this suggestion is that some teens might select an option by accident, or without thinking too much about it, so a supplementary nudge asking them to confirm their actions or possible doubts and reminding them of the risks involved was recommended.

*“I guess that's kind of a downside to it. because if they tap it [dismiss] and it sends [without any other safeguards]. Then like, it doesn't really do much.”* - P1. 14, F

### *Solving the risk with nudges*

The participants mentioned that nudges might inadvertently worsen the risk being addressed. An example of this is censor-induced curiosity (N=15) with the sensitivity filter nudges where censoring the risk can make the teen even more interested in what was censored. The table below explains how the participants stated they would respond to the  $SF_N$  as a victim, and the results show that in multiple cases, a teen victim is more likely to uncensor a detected risk than take the safer approach by leaving it censored.

*Table 3: Provided Responses to Sensitivity Filters*

	$IB_s$	$CY_{s(text)}$	$CY_{s(image)}$
Leave risky content censored	5	5	7
Uncensor risky content	13	6	4
Undecided	3	10	

$IB_s$  – Information Breach,  $CY_s$  - Cyberbullying

*“I don't feel like teens are going to ignore this and be like, 'Oh, harmful, harmful, oh I'm not going to look at this.' ... I'm being totally honest with you, like I'm a teen, if I see that, I'm clicking show like, I want to see.” - P3. 16, M*

Another challenge was raised with the  $GA_N$  for both the victim and sender. For the victim, they felt generally responding to an aggressor (which is suggested by the  $GA_N$  nudge) could lead to the risk escalating or being unnecessarily prolonged, either by the risk sender forming a harmful bond with the victim, or the victim suggesting they are interested in continuing the risky conversation. For the risk sender, the teens felt (especially for an adversarial risk sender), the  $GA_N$  could worsen the scenario by letting the risk sender know what is wrong with their messages and teach them how to be better or more covert at perpetrating risks which can make it more difficult for the victim and platform to identify their risks.

*“The one where it's like highlighted in red, like what part of like, the text message was bad like, I feel like the creepy person could just like change the wording to bypass that system.” - P9. 17 ,M*

For the risk sender, the teens made several recommendations as methods of preventing the risk from reaching the victim. Many (N=17) teens recommended features based on restricting the risk sender, which is taking away the ability of the risk sender to send risky content to a possible victim. Another group of teens (N=17) recommended using reprimanding language to emphasize the consequences of their risky decision which is done to a degree in the  $SF_N$  victim nudges and is reflected in the rankings. Some teens (N=5) also recommended the implementation of this to a greater degree by penalizing the risk sender in ways such as flagging their accounts and punishing them either based on single or multiple infractions.



*" The nudge could get them to reconsider by mentioning 'if you send the risky messages, you could risk getting banned from our platform.' Maybe like a strike, and like [in] three strikes, you're banned." – P9. 17,M*

## CHAPTER FIVE: DISCUSSION

In this section, we discuss the implications of our results and provide our inference and synthesis of what was found in the results to contribute to the field of Adolescent Online Safety Nudging and provide direction for future work.

### Enforced Risk Prevention for the Perpetrator vs Control for the Victim

Based on prior work that address adolescent online safety nudges, The key takeaway from Agha et al.'s study is that teens want nudges to prevent the risk by designing online safety features that are targeted towards the risk perpetrator (Agha, Zhang, Obajemu, Shirley, & J. Wisniewski, 2022), As a result, both risk coping (victim) and risk prevention (perpetrator) nudges were designed. However, due to the novel approach of evaluating those co-designed nudges critically , the teens were able to identify that preventing the risk from the risk perpetrator's perspective is a lot more difficult than expected due to the risk perpetrator being intentional about the risk. As a result, the teens recommended stricter measures (i.e., reprimanding language and punishments) for the risk perpetrator to enforce compliance with the nudge. The reprimanding language preference ties into the finding by Masaki et al. that adolescents would be more considerate of negatively-framed nudges as opposed to positively framed ones (Masaki, et al., 2020).

We also expand on the results of the study given for the risk victim with the criticisms provided by the teens. We were able to identify what they liked and did not like about these nudges as well as provide a deeper understanding of the context and factors needed to be considered to ensure the nudges remain effective as outlined under the results of RQ1 and RQ2. More specifically, we identified that the risk victims need controlled forms of nudges with restriction only needed when considered necessary.

### Overcoming the challenges of effective nudging for positive behavioral change

With the overarching theme of overcome the challenges with online safety nudging in academia and industry, Masaki et al. was able to identify preferences in how teens want nudges to be framed i.e., the approach the nudges take to communicate with the teen, they discovered that teens preferred negative/consequence-driven nudges over the affirmative/positive ones. Due to the quantitative method used, the challenges were limited to what was asked in the survey (Masaki, et al., 2020). We contribute to the quantitative results with a qualitative semi-structured process that removed the boundaries on what the teens could respond to and as a result, we identified a key challenge with nudges, which is that teens do not have a positive impression of nudges, and this could bias how they would respond to them, irrespective of how effective or properly designed the nudges are. For example, warning fatigue made the nudges likely to be ignored, the teens also had personal preconceptions that made them see these nudges as a chore, such as the harmful intent of a risk perpetrator, or the risk victim not wanting to read as explained in the results section. There is also the issue of unethical nudging with some platforms including elements of deception when nudging, which can cause them to be more wary of the nudge than the risk. As a result, the long-term impressions of nudging in general need to be considered during the design stage. The goal should be to appeal to the user's preferences and make the nudges more pleasant to experience. We were able to identify the preferences the teens had for nudges in the result, and they include the use of warning symbols over text, making sure nudges don't disrupt the user flow, and so on. The teens also brought up the issue of the risk detection capabilities of the platform whereby a lot of false positives can lead to a worse form of warning fatigue, where the teen is aware of the nudge, but they don't acknowledge it due to the nudge being saturated or losing credibility.

### Tailoring nudge to provide autonomy and personalization

This section explains how tailored nudges that provide autonomy and personalization can make them more effective. It also matches the findings from Agha et al. where the teens expressed a desire to customize or have control over their nudges as a way of making them more effective (Agha, Zhang, Obajemu, Shirley, & J. Wisniewski, 2022). Tailored and customized nudges can also help address the warning fatigue issue as well as the multiple factors that can alter how nudges may be perceived in RQ2. We infer that giving teens control over their nudges (the type of nudge they receive, how often they are nudged, etc.) can make them more effective. An important consideration for nudges is to make them flexible enough to address multiple risk types, user awareness, and the personal discretion of the teen. While we have discussed that teens should have the ability to customize their nudges to make them more effective, we believe the platform should also have a part to play as well in customizing nudges for example: profiling the user by determining their age and making attempts to determine their intent and behavior to determine the best way to approach them with a nudge.

### Nudge Design Recommendations based on Empirical Evidence

Under this section, we provide a summary of guidelines for designing effective nudges for adolescent online safety based on our findings.

- Capture Attention: Nudges should be made difficult to ignore and capture the teen's attention with bold design cues.
- Direct User: Nudges should provide a clear action or direction to improve online safety beyond warnings.
- Tailor Preferences: Nudges should be tailored to the user to accommodate behavioral and other personal differences.

- Assert Risk Prevention: Nudges should provide assertive language to enforce compliance from a risk perpetrator.
- Impress User: Nudges should provide a positive impression on the user to prevent any possible negative future biases towards nudges
- Prevent risk: Nudges should prevent the risk from the perspective of the perpetrator when possible.
- Risk Tailoring: Nudges need to be tailored to the distinctness of each risk being addressed.
- Articulate Warning: Nudges should clearly articulate the warning provided to the user to avoid any possible gaps in communication.

### Limitations and Future Work

One major limitation of our study is that both online risk and proposed nudges were considered to be mutually exclusive when in reality, teens may experience a combination of multiple risk scenarios in a single risk instance and different elements from multiple nudges can also be combined into one nudge. Additionally, this project covers how nudges can be used to handle common online risks faced by teens. However, even though they are based on what was provided to us by the teens, the risky scenarios used are not a recreation of a real risk.

Another limitation of our work is that the participants were based in the United States, and had access to remote-conferencing technology, and as a result, our results may not be generalizable to all teens across the world or in all socio-economic classes. The group sessions of this study were also prone to social desirability bias or groupthink, where the participants were reluctant to disagree with one another which possible created unnatural similarities in their feedback or rankings.

Finally, the in-depth, qualitative method used to gather the teens' feedback was time-consuming and logistically restrictive that it limited the number that could be evaluated, especially when compared to a broader method such as an asynchronous survey.

For the future work, we plan to implement these nudge designs and risk scenarios in a realistic setting to have the teens evaluate them in practice. We also recommend future researchers to explore the option of using a virtual assistant or live demonstration as a method of presenting the designed-based study materials to the participants to evoke a greater degree of realism. Finally, we recommend future social moderation work to also include the perspective of the risk perpetrator in online safety instances to provide wider context on the overarching issue.

## **CONCLUSION**

Through a series of focus groups and interviews with 21 teens, this study contributes to adolescent online safety research by gathering their feedback on different approaches to online safety nudges for the purpose of gaining a better understanding of how teens perceive them to guide future implementations. We discovered that nudges are very subjective and specific to a risk instance, but the general takeaway is that teens want nudges to make them aware about the specific risk, and also provide a relevant action to mitigate it.

**APPENDIX A**  
**IRB APPROVAL**





UNIVERSITY OF CENTRAL FLORIDA

**Institutional Review Board**  
FWA00000351  
IRB00001138, IRB00012110  
Office of Research  
12201 Research Parkway  
Orlando, FL 32826-3246

APPROVAL

March 2, 2022

Dear Pamela Wisniewski:

On 3/2/2022, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title:	A Teen-based Focus group to Evaluate Targeted Online Safety Interventions
Investigator:	Pamela Wisniewski
IRB ID:	STUDY00003987
Funding:	Name: William T. Grant Fdn, Grant Office ID: 1063027, Funding Source ID: 187941
Grant ID:	1063027;
IND, IDE, or HDE:	None
Documents Reviewed:	<ul style="list-style-type: none"> <li>• CITI Certificate_Mohammed Alqadhi, Category: Other;</li> <li>• Eligibility_Survey.docx, Category: Other;</li> <li>• Feedback_survey.docx, Category: Other;</li> <li>• Flyer (Alternate web version), Category: Recruitment Materials;</li> <li>• FLYER.png, Category: Recruitment Materials;</li> <li>• Focus_Group_Prep_Worksheet.docx, Category: Other;</li> <li>• Focus_Group_IRB_Protocol.docx, Category: IRB Protocol;</li> <li>• Focus_Group_Parental_Consent.pdf, Category: Consent Form;</li> <li>• Focus_Group_Session_Script.docx, Category: Other;</li> <li>• Focus_Group_Teen_Assent.pdf, Category: Consent Form;</li> <li>• Help_Resources.docx, Category: Other;</li> <li>• Nudges+Prompt.docx, Category: Other;</li> <li>• Preparatory_Email.docx, Category: Other;</li> <li>• Recruitment_Email.docx, Category: Recruitment Materials;</li> <li>• Session_Slides.pptx, Category: Other</li> </ul>

The IRB approved the protocol from 3/2/2022.

In conducting this protocol, you are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within the IRB system. Guidance on submitting Modifications and a Continuing Review or Administrative Check-in are detailed in the manual. When you have completed your research, please submit a Study Closure request so that IRB records will be accurate.

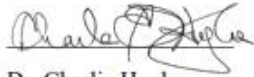
If you have any questions, please contact the UCF IRB at 407-823-2901 or [irb@ucf.edu](mailto:irb@ucf.edu). Please include your project title and IRB number in all correspondence with this office.

Sincerely,

Katie Kilgore  
Designated Reviewer

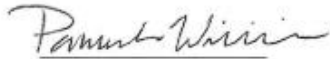
**APPENDIX B**  
**IRB APPROVAL CONNECTION STATEMENT**

We acknowledge that this MS Thesis Document “Identifying Challenges and Opportunities for Designing Social Media Nudges for Adolescents” is connected to UCF IRB STUDY00003987 titled “A Teen-based Focus group to Evaluate Targeted Online Safety Interventions.”



Dr. Charlie Hughes

Advisor

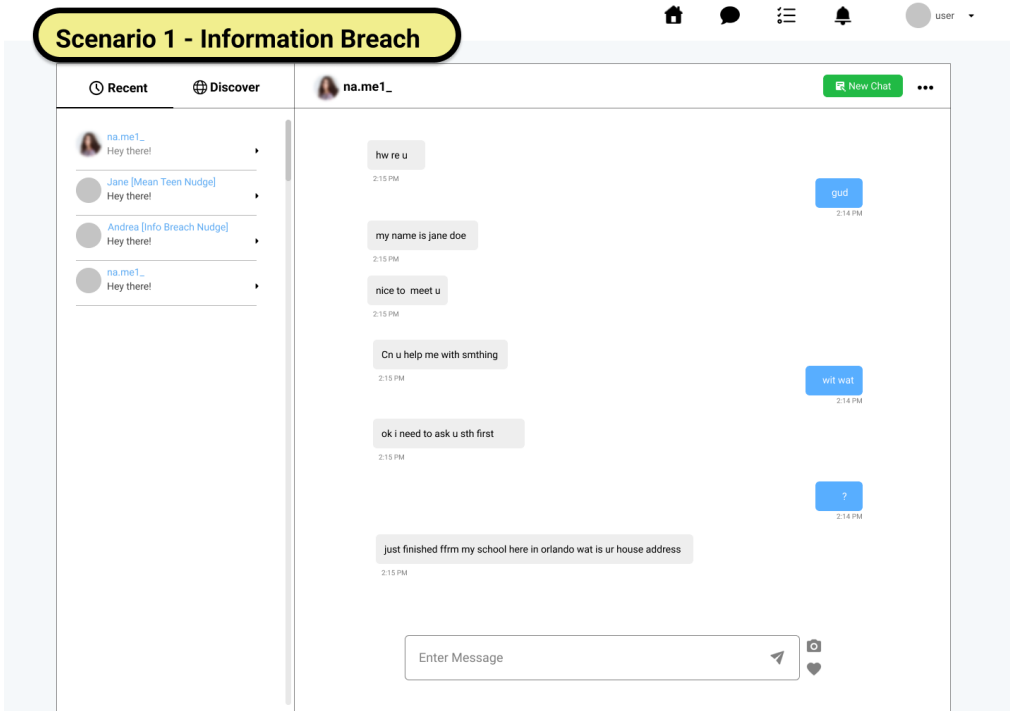


Dr. Pamela Wisniewski

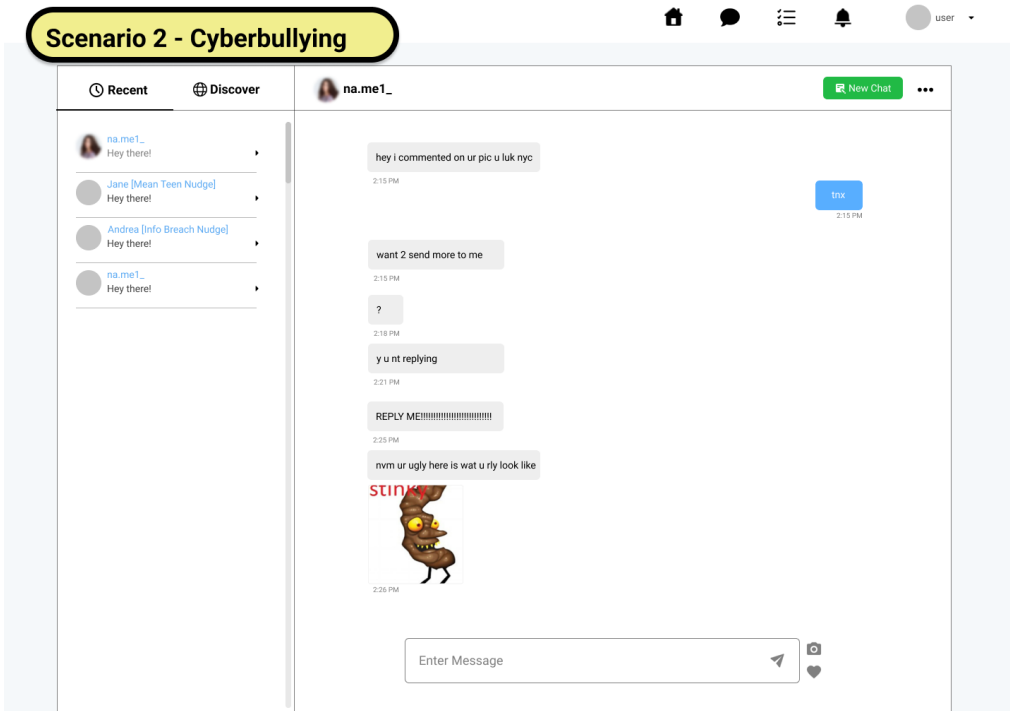
IRB Investigator

**APPENDIX C**  
**FULL VIEW RISKY SCENARIOS**

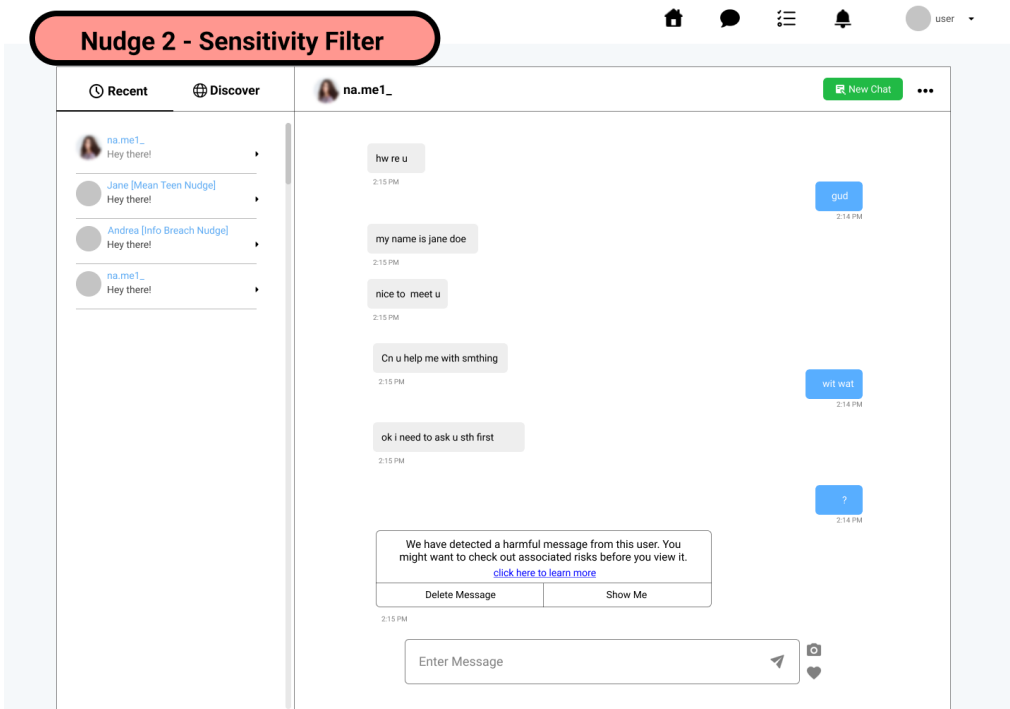
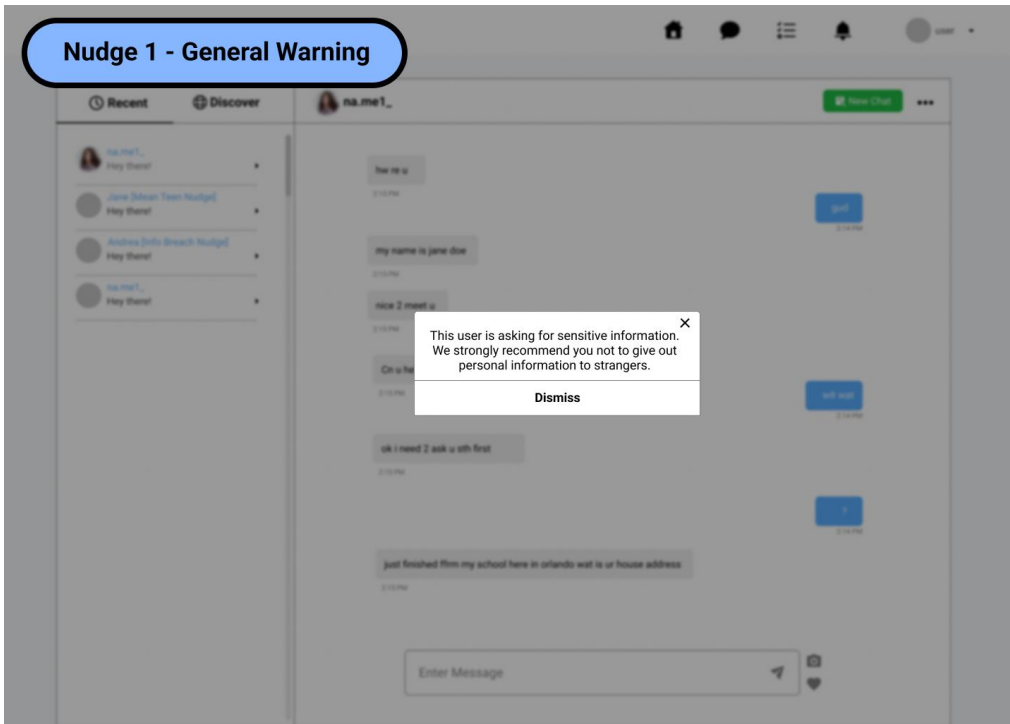
## Scenario 1 - Information Breach



## Scenario 2 - Cyberbullying



**APPENDIX D**  
**FULL VIEW VICTIM NUDGES (INFORMATION BREACH)**



### Nudge 3 - Guided Actions

Recent Discover na.me1\_ New Chat

na.me1\_ Hey there!

Jane [Mean Teen Nudge] Hey there!

Andrea [Info Breach Nudge] Hey there!

na.me1\_ Hey there!

my name is jane doe 2:15 PM

nice to meet u 2:15 PM

Can u help me with smthing 2:15 PM

ok i need to ask u sth first 2:15 PM

just finished ffrn my school here in orlando wat is ur house address 2:15 PM

wit wat 2:14 PM

? 2:14 PM

Risky message detected, here are some ways you can respond

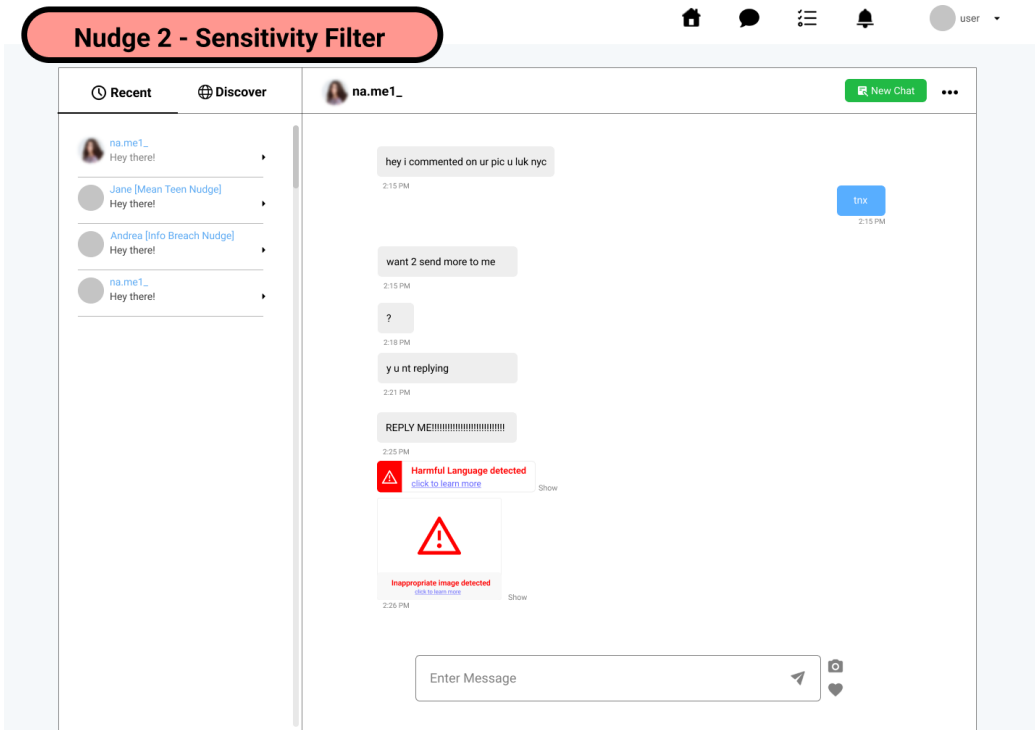
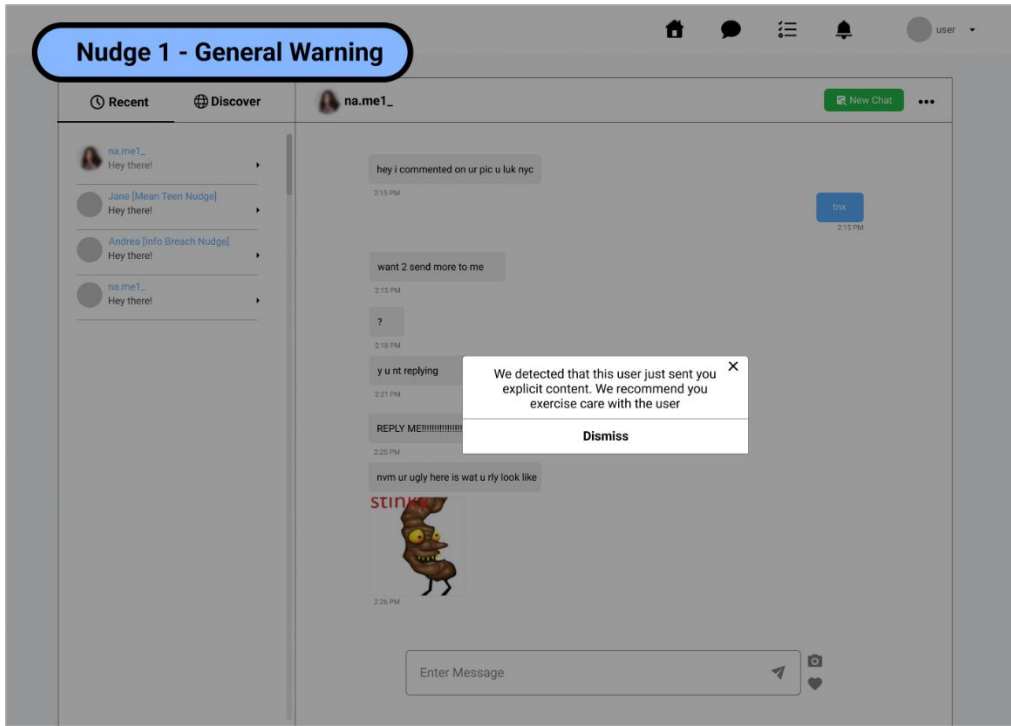
- No, thank you
- Please leave me alone
- I am not going to
- Block and Report Contact**

Reply directly



# APPENDIX E

## FULL VIEW VICTIM NUDGES (CYBERBULLYING)



### Nudge 3 - Guided Actions

Recent Discover

na.me1\_ Hey there!

Jane [Mean Teen Nudge] Hey there!

Andrea [Info Breach Nudge] Hey there!

na.me1\_ Hey there!

na.me1\_

?

2:18 PM


y u nt replying

2:21 PM

REPLY ME!!!!!!!!!!!!!!!!!!!!!!

2:25 PM

nvm ur ugly here is wat u rly look like



2:26 PM

× Risky content detected, review suggested auto-responses

Please stop texting me

Delete risky message

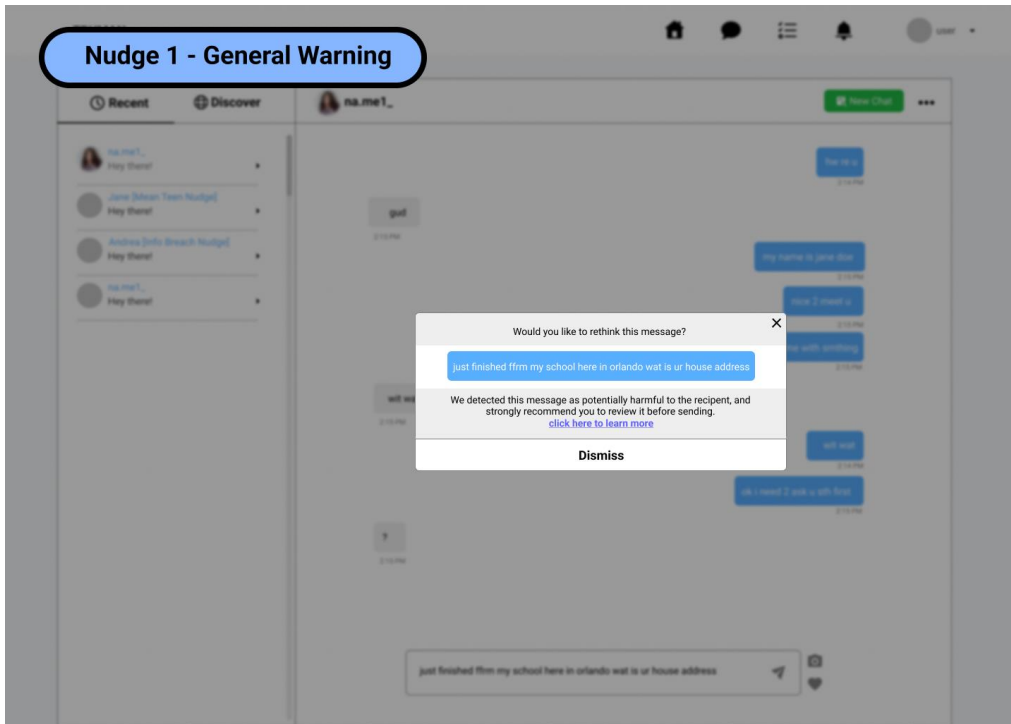
Delete risky image

Block and Report Contact

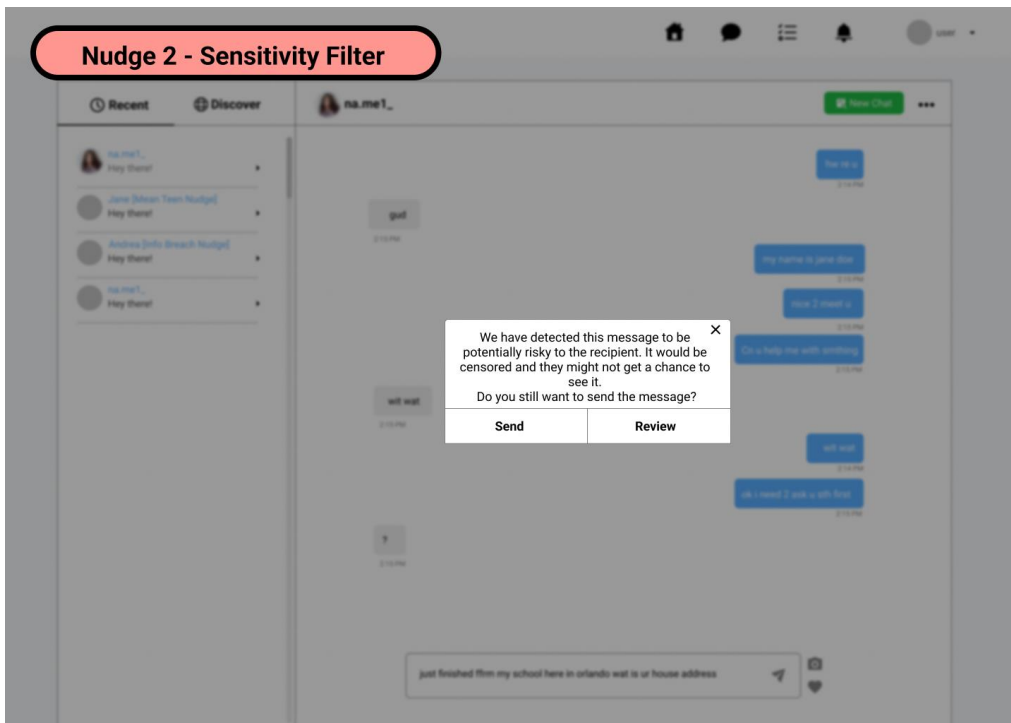
Enter Message

**APPENDIX F**  
**FULL VIEW PERPETRATOR NUDGES (INFORMATION BREACH)**

## Nudge 1 - General Warning



## Nudge 2 - Sensitivity Filter



### Nudge 3 - Guided Actions

Recent Discover

na.me1\_

New Chat

- na.me1\_ Hey there!
- Jane [Mean Teen Nudge] Hey there!
- Andrea [Info Breach Nudge] Hey there!
- na.me1\_ Hey there!

gud

2:15 PM

wit wat

2:15 PM

?

2:15 PM

hw re u

2:14 PM

my name is jane doe

2:15 PM

nice to meet u

2:15 PM

Can u help me with smthing

2:15 PM

wit wat

2:14 PM

ok i need to ask u sth first

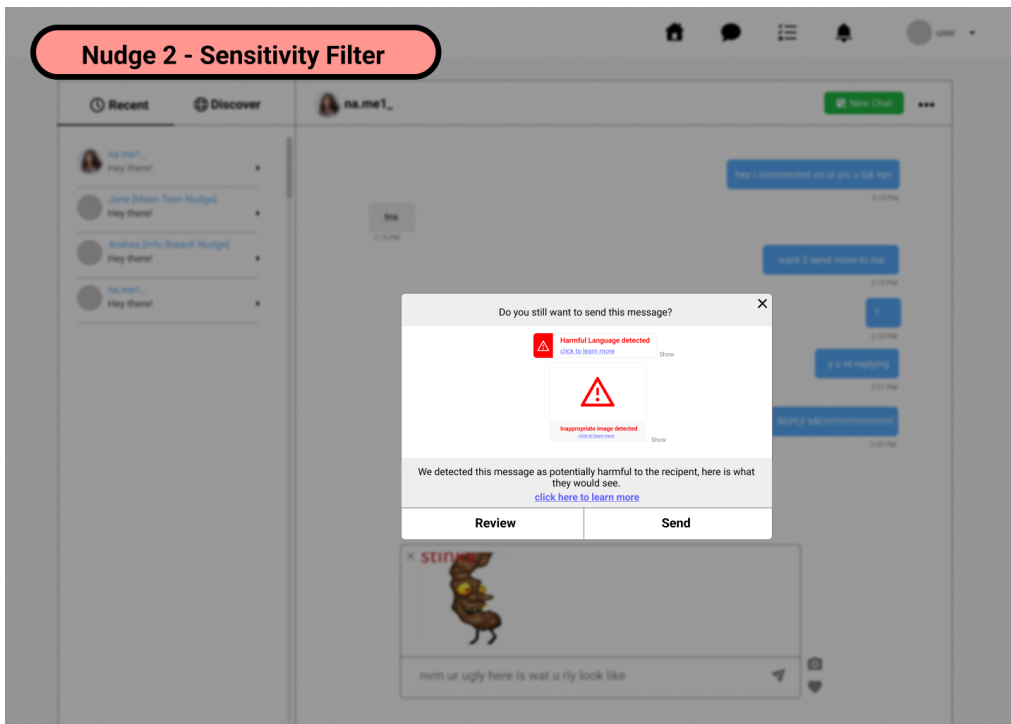
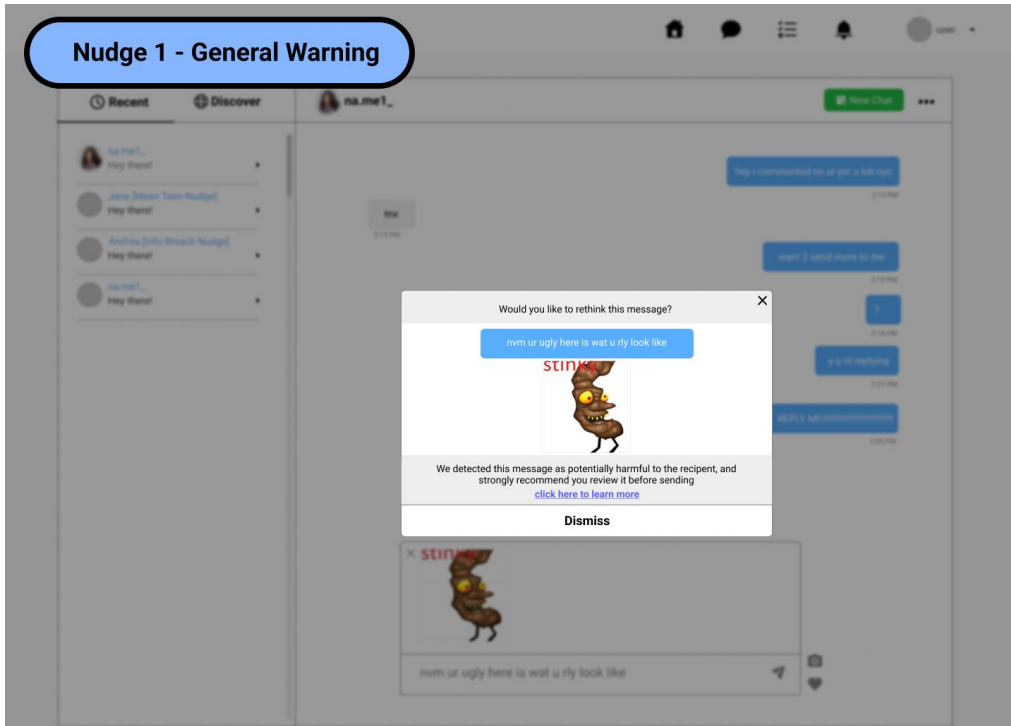
2:15 PM

what kind of games do you like? can we be friends? clear

just finished ffrm my school here in orlando wat is ur house address

# APPENDIX G

## FULL VIEW PERPETRATOR NUDGES (CYBERBULLYING)



### Nudge 3 - Guided Actions

Recent Discover

- na.me1\_ Hey there!
- Jane [Mean Teen Nudge] Hey there!
- Andrea [Info Breach Nudge] Hey there!
- na.me1\_ Hey there!

na.me1\_ New Chat

hey i commented on ur pic u luk nyc 2:15 PM

trx 2:15 PM

want 2 send more to me 2:15 PM

? 2:18 PM

y u nt replying 2:21 PM

REPLY ME!!!!!!!!!!!!!!!!!!!!!!!!!!!! 2:25 PM

STINK

clear text clear image i'm upset reply

nvm ur ugly here is wat u rly look like

## REFERENCES

- Agatston, P., Kowalski, R., & Limber, S. (2012). Youth views on cyberbullying. *Cyberbullying prevention and response: Expert perspectives*, 57–71. Retrieved from <https://vdoc.pub/download/cyberbullying-prevention-and-response-expert-perspectives-44p196lqq9m0>
- Agha, Z., Ghaiumy Anaraky, R., Badillo-Urquiola, K., McHugh, B., & Wisniewski, P. (2021). ‘Just-in-Time’ Parenting: A Two-Month Examination of the Bi-directional Influences Between Parental Mediation and Adolescent Online Risk Exposure. In A. Moallem (Ed.), *HCI for Cybersecurity, Privacy and Trust* (Vol. 12788, pp. 261–280). Cham: Springer International Publishing. doi:10.1007/978-3-030-77392-2\_17
- Agha, Z., Zhang, Z., Obajemu, O., Shirley, L., & J. Wisniewski, P. (2022, April). A Case Study on User Experience Bootcamps with Teens to Co-Design Real-Time Online Safety Interventions. *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (pp. 1–8). New Orleans LA USA: ACM. doi:10.1145/3491101.3503563
- Alderman, S. (2018, February). 5 things you can do to improve online safety for your young learners. *5 things you can do to improve online safety for your young learners*. Retrieved March 3, 2022, from <https://www.english.com/blog/5-things-to-improve-online-safety-for-young-learners/>
- Alemany-Bordera, J., Del Val Noguera, E., Alberola Oltra, J. M., & García-Fornes, A. (2019, September). Enhancing the privacy risk awareness of teenagers in online social networks through soft-paternalism mechanisms. *International Journal of Human-Computer Studies*, 129, 27–40. doi:10.1016/j.ijhcs.2019.03.008



- Anderson, M. (2018, September). A Majority of Teens Have Experienced Some Form of Cyberbullying. *A Majority of Teens Have Experienced Some Form of Cyberbullying*. Retrieved May 16, 2022, from <https://www.pewresearch.org/internet/2018/09/27/a-majority-of-teens-have-experienced-some-form-of-cyberbullying/>
- Badillo-Urquiola, K., Chouhan, C., Chancellor, S., De Choudhary, M., & Wisniewski, P. (2020, January). Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design. *Journal of Adolescent Research, 35*, 147–175. doi:10.1177/0743558419884692
- Badillo-Urquiola, K., Shea, Z., Agha, Z., Lediaeva, I., & Wisniewski, P. (2021, January). Conducting Risky Research with Teens: Co-designing for the Ethical Treatment and Protection of Adolescents. *Proceedings of the ACM on Human-Computer Interaction, 4*, 1–46. doi:10.1145/3432930
- Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E., & Wisniewski, P. J. (2019, June). Stranger Danger!: Social Media App Features Co-designed with Children to Keep Them Safe Online. *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (pp. 394–406). Boise ID USA: ACM. doi:10.1145/3311927.3323133
- Braun, V., & Clarke, V. (2006, January). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101. doi:10.1191/1478088706qp063oa
- Cameron, L. (2018). A New Approach to Internet Safety Apps Helps Teens Regulate Their Own Online Activity. *A New Approach to Internet Safety Apps Helps Teens Regulate Their Own Online Activity*. Retrieved May 16, 2022, from

<https://www.computer.org/publications/tech-news/research/teen-internet-safety-apps-parental-control>

- Chouhan, C., LaPerriere, C. M., Aljallad, Z., Kropczynski, J., Lipford, H., & Wisniewski, P. J. (2019, November). Co-designing for Community Oversight: Helping People Make Privacy and Security Decisions Together. *Proceedings of the ACM on Human-Computer Interaction*, 3, 1–31. doi:10.1145/3359248
- Davis, S., & Charisse, N. (2014). *Youth Voice Project: student insights into bullying and peer mistreatment*. Champaign, Illinois: Research Press Publishers.
- Duggan, M. (2017, July). 5. Other types of negative experiences online. 5. *Other types of negative experiences online*. Retrieved October 6, 2022, from <https://www.pewresearch.org/internet/2017/07/11/other-types-of-negative-experiences-online/>
- Farrukh, A., Sadwick, R., & Villasenor, J. (2014, October). Youth Internet Safety: Risks, Responses, and Research Recommendations. 18. Retrieved from [https://www.brookings.edu/wp-content/uploads/2016/06/Youth-Internet-Safety\\_v07.pdf](https://www.brookings.edu/wp-content/uploads/2016/06/Youth-Internet-Safety_v07.pdf)
- Gainsbury, S. M., Browne, M., & Rockloff, M. (2019, June). Identifying risky Internet use: Associating negative online experience with specific online behaviours. *New Media & Society*, 21, 1232–1252. doi:10.1177/1461444818815442
- Harriman, N., Shortland, N., Su, M., Cote, T., Testa, M. A., & Savoia, E. (2020, January). Youth Exposure to Hate in the Online Space: An Exploratory Analysis. *International Journal of Environmental Research and Public Health*, 17, 8531. doi:10.3390/ijerph17228531

- Hartwig, K., & Reuter, C. (2021, October). Nudge or Restraint: How do People Assess Nudging in Cybersecurity - A Representative Study in Germany. *European Symposium on Usable Security 2021* (pp. 141–150). Karlsruhe Germany: ACM. doi:10.1145/3481357.3481514
- Hummel, D., & Maedche, A. (2019, June). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, 80, 47–58. doi:10.1016/j.socec.2019.03.005
- João Carrasqueira. (2021). Instagram announces new efforts to protect kids from adults. *Instagram announces new efforts to protect kids from adults*. Retrieved October 6, 2022, from <https://www.neowin.net/news/instagram-announces-new-efforts-to-protect-kids-from-adults/>
- Kaiser, S., Martinussen, M., Adolfsen, F., Breivik, K., & Kyrrestad, H. (2021, November). An App-Based Intervention for Adolescents Exposed to Cyberbullying in Norway: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, 10, e31789. doi:10.2196/31789
- Kaspersky. (2022, 5 12). *What is the “ransomware detected” pop-up?* Retrieved from <https://usa.kaspersky.com/resource-center/threats/identify-and-remove-fake-pop-ups>
- Kim, S., Colwell, S. R., Kata, A., Boyle, M. H., & Georgiades, K. (2018, March). Cyberbullying Victimization and Adolescent Mental Health: Evidence of Differential Effects by Sex and Mental Health Problem Type. *Journal of Youth and Adolescence*, 47, 661–672. doi:10.1007/s10964-017-0678-4
- Lehner, M., Mont, O., & Heiskanen, E. (2016, October). Nudging – A promising tool for sustainable consumption behaviour? *Journal of Cleaner Production*, 134, 166–177. doi:10.1016/j.jclepro.2015.11.086

- Masaki, H., Shibata, K., Hoshino, S., Ishihama, T., Saito, N., & Yatani, K. (2020, April). Exploring Nudge Designs to Help Adolescent SNS Users Avoid Privacy and Safety Threats. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–11). Honolulu HI USA: ACM. doi:10.1145/3313831.3376666
- Mayo Clinic Staff. (2022, February). Teens and social media use: What's the impact? *Teens and social media use: What's the impact?* Retrieved May 16, 2022, from <https://www.mayoclinic.org/healthy-lifestyle/tween-and-teen-health/in-depth/teens-and-social-media-use/art-20474437>
- McHugh, B. C., Wisniewski, P., Rosson, M. B., & Carroll, J. M. (2018, October). When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress. *Internet Research*, 28, 1169–1188. doi:10.1108/IntR-02-2017-0077
- McLoughlin, L. T. (2021, September). Understanding and measuring coping with cyberbullying in adolescents: exploratory factor analysis of the brief coping orientation to problems experienced inventory. *Current Psychology*, 40, 4300–4311. Retrieved May 16, 2022, from <https://link.springer.com/content/pdf/10.1007/s12144-019-00378-8.pdf>
- Mele, C., Russo Spena, T., Kaartemo, V., & Marzullo, M. L. (2021, May). Smart nudging: How cognitive technologies enable choice architectures for value co-creation. *Journal of Business Research*, 129, 949–960. doi:10.1016/j.jbusres.2020.09.004
- Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. (2022, January). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, 119, e2107346118. doi:10.1073/pnas.2107346118

- Moreno, M. A., Egan, K. G., Bare, K., Young, H. N., & Cox, E. D. (2013, December). Internet safety education for youth: stakeholder perspectives. *BMC Public Health, 13*, 543. doi:10.1186/1471-2458-13-543
- Olson, J. A., Sandra, D. A., Chmoulevitch, D., Raz, A., & Veissière, S. P. (2021, January). *A nudge-based intervention to reduce problematic smartphone use: Randomised controlled trial*. preprint, PsyArXiv. doi:10.31234/osf.io/tjynk
- Patchin, J. W., & Hinduja, S. (Eds.). (2012). *Cyberbullying prevention and response: expert perspectives*. New York: Routledge.
- Rodriguez, M. E., Guerrero-Roldán, A. E., Baneres, D., & Karadeniz, A. (2022, February). An Intelligent Nudging System to Guide Online Learners. *The International Review of Research in Open and Distributed Learning, 23*, 41–62. doi:10.19173/irrodl.v22i4.5407
- Rozin, P., Scott, S. E., Dingley, M., Urbanek, J., Jiang, H., & Kaltenbach, M. (2011, June). Nudge to Nobesity I: Minor Changes in Accessibility Decrease Food Intake. *Judgment and Decision Making, 6*, 323–332. Retrieved from [https://repository.upenn.edu/marketing\\_papers/410](https://repository.upenn.edu/marketing_papers/410)
- Santora, J. (2020). 6 Reasons Why Popups, Welcome Gates and Slide-Ins Suck. *6 Reasons Why Popups, Welcome Gates and Slide-Ins Suck*. Retrieved May 16, 2022, from <https://optinmonster.com/6-reasons-pop-ups-welcome-gates-and-slide-ins-suck-and-the-solution/>
- Schaeffer, K. (2019, August). Most U.S. teens who use cellphones do it to pass time, connect with others, learn new things. *Most U.S. teens who use cellphones do it to pass time, connect with others, learn new things*. Retrieved May 17, 2022, from

<https://www.pewresearch.org/fact-tank/2019/08/23/most-u-s-teens-who-use-cellphones-do-it-to-pass-time-connect-with-others-learn-new-things/>

Spears, B. A., Taddeo, C., & Barnes, A. (2018). Online social marketing approaches to inform cyber/bullying prevention and intervention: What have we learnt? In *Reducing Cyberbullying in Schools* (pp. 75–94). Elsevier. doi:10.1016/B978-0-12-811423-0.00006-7

Statista Research Department. (2022, April 18). *Social media use during coronavirus (COVID-19) worldwide*. Retrieved from Statista:

<https://www.statista.com/statistics/1112995/social-media-platforms-usa-coronavirus/>

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: improving decisions about health, wealth, and happiness* (Rev. and expanded ed ed.). New York: Penguin Books.

Twitter Help. (2022). Twitter's sensitive media policy. *Twitter's sensitive media policy*. Retrieved October 6, 2022, from <https://help.twitter.com/en/rules-and-policies/media-policy>

Twitter Safety. (2022, June). How Twitter is nudging users to have healthier conversations. *How Twitter is nudging users to have healthier conversations*. Retrieved June 15, 2022, from <https://blog.twitter.com/content/blog-twitter/common-thread/en/topics/stories/2022/how-twitter-is-nudging-users-healthier-conversations.html>

Usman Khan Lodhi. (2021). Instagram's new keyword-based tool will filter out offensive DMs. *Instagram's new keyword-based tool will filter out offensive DMs*. Retrieved October 6, 2022, from <https://www.neowin.net/news/instagrams-new-keyword-based-tool-will-filter-out-offensive-dms/>

Vale, A., Pereira, F., Gonçalves, M., & Matos, M. (2018, October). Cyber-aggression in adolescence and internet parenting styles: A study with victims, perpetrators and victim-

- perpetrators. *Children and Youth Services Review*, 93, 88–99.  
doi:10.1016/j.chilyouth.2018.06.021
- Vandebosch, H. (2019). Cyberbullying Prevention, Detection and Intervention. In H. Vandebosch, & L. Green (Eds.), *Narratives in Research and Interventions on Cyberbullying among Young People* (pp. 29–44). Cham: Springer International Publishing. doi:10.1007/978-3-030-04960-7\_3
- Website, T. A. (n.d.). Internet safety: teenagers. *Internet safety: teenagers*. Retrieved May 16, 2022, from <https://raisingchildren.net.au/teens/entertainment-technology/cyberbullying-online-safety/internet-safety-teens>
- Weinmann, M., Schneider, C., & Brocke, J. v. (2016, December). Digital Nudging. *Business & Information Systems Engineering*, 58, 433–436. doi:10.1007/s12599-016-0453-1
- Wisniewski, P. (2017, December). BBL: Taking a Teen-Centric Approach to Adolescent Online Safety – HCIL. *BBL: Taking a Teen-Centric Approach to Adolescent Online Safety – HCIL*. Retrieved May 16, 2022, from <https://hcil.umd.edu/events/event/bbl-title-tba-5/>
- Wisniewski, P. (2018, March). The Privacy Paradox of Adolescent Online Safety: A Matter of Risk Prevention or Risk Resilience? *IEEE Security & Privacy*, 16, 86–90.  
doi:10.1109/MSP.2018.1870874
- Wisniewski, P. J., Vitak, J., & Hartikainen, H. (2022). Privacy in Adolescence. In B. P. Knijnenburg, X. Page, P. Wisniewski, H. R. Lipford, N. Proferes, & J. Romano (Eds.), *Modern Socio-Technical Perspectives on Privacy* (pp. 315–336). Cham: Springer International Publishing. doi:10.1007/978-3-030-82786-1\_14
- Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B., & Carroll, J. M. (2017, February). Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?

*Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 51–69). Portland Oregon USA: ACM.

doi:10.1145/2998181.2998352

Wisniewski, P., Xu, H., Rosson, M. B., & Carroll, J. M. (2017, February). Parents Just Don't Understand: Why Teens Don't Talk to Parents about Their Online Risk Experiences.

*Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 523–540). Portland Oregon USA: ACM.

doi:10.1145/2998181.2998236

Yevseyeva, I., Morisset, C., Turland, J., Coventry, L., Groß, T., Laing, C., & van Moorsel, A. (2014). Consumerisation of IT: Mitigating Risky User Actions and Improving

Productivity with Nudging. *Procedia Technology*, *16*, 508–517.

doi:10.1016/j.protcy.2014.10.118