To Nudge or Not to Nudge: Co-Designing and Evaluating the Effectiveness of Adolescent Online Safety Nudges

Zainab Agha
Vanderbilt University
Nashville, TN, USA
zainab.agha@vanderbilt.edu

ABSTRACT

There is growing concern regarding adolescent online risks posed by social media. Prior work calls for a paradigm shift from restrictive approaches towards strength-based solutions to online safety, that provide autonomy and control to teens. To better understand how we might design online safety interventions that help teens deal with online risks, we must include teens as partners in the design and evaluation of online safety solutions. To address this gap, my first dissertation study focused on co-designing online safety features with teens, which showed that teens often design real-time interventions that resemble "nudges". Therefore, my dissertation focuses on evaluating the effectiveness of these nudge designs in an ecologically valid social media simulation. To do this, I will conduct three studies: 1) a User Experience Bootcamp with teens to teach them design skills for co-designing online safety features, 2) a focus group study to design an ecologically valid social media simulation, 3) a between-subjects experiment within a social media simulation for evaluating the effect of nudges in educating teens and helping them make safer choices when exposed to risk. My goal for this research is to understand, design, develop, and evaluate online safety nudges that can help promote self-regulated, autonomous, and safer interactions for teens online.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

adolescent online safety, user experience, co-design, participatory design, nudges

ACM Reference Format:
Zainab Agha. 2023. To Nudge or Not to Nudge: Co-Designing and Evaluating the Effectiveness of Adolescent Online Safety Nudges. In *Interaction Design and Children (IDC '23), June 19–23, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3585088.3593923

1 INTRODUCTION

Social media has become an integral part of adolescents' lives today, where more than half of the US teens say it would be hard to give up social media [20]. Concurrently, there is growing awareness in the community about the types of risks posed by social media such as cyberbullying [4], sexual risks [8], exposure to explicit content [7], or information breaches [17]. There have been extensive efforts within the Human Computer Interaction (HCI) communities dedicated towards understanding teens online experiences and safety approaches.

A majority of research on this topic has focused on restrictive approaches such as technology restrictions [10], or parental controls [3], which are often considered privacy invasive by teens [21]. Recently, researchers have called for strength-based approaches, that can empower teens to be resilient in the face of risks [2, 9, 22]. Yet, there is still a lack of actionable teen-centric solutions that propose *how* we can help teens navigate their online risks [16]. Additionally, many researchers have recognized that it is essential to involve them as participants in the design process through co-design efforts [6, 15]. Yet, existing co-design research efforts have largely focused on either a single type of online risk (such as cyberbullying [5]) or focused on redesigning parental controls [15]. Therefore, there is a need to involve teens as equal partners in the design of online safety solutions for a variety of unsafe online experiences, that go beyond parental controls.

To address this gap, my first dissertation study focused on co-designing online safety interventions with teens that are contextualized to their diverse online risk experiences. Results from this study showed that teens designed real-time online safety interventions which often resembled "nudges". A nudge is designed to alter people's behavior through positive reinforcement without compromising their decision-making autonomy [19]. This first study highlighted the need to further evaluate the effectiveness of these nudges in helping them deal with unsafe interactions in an ecologically valid environment. Therefore, my dissertation focuses on implementing and evaluating these nudge designs with teens in an web-based social media simulation. In order to design and evaluate teen-centric online safety nudges in a realistic environment, my dissertation consists of three studies that answer the following overarching research questions:

- **RQ1:** What social media features would teens co-design when envisioning online safety tools that address their most commonly encountered risk experiences?
- **RQ2:** When designing a social media simulation to teach teens about how to navigate online risks, how can we develop ecologically valid personas, risk scenarios, and design-based nudges?
- **RQ3:** How do online safety nudges co-designed with teens influence youths' perceptions about online safety, trust, and autonomy, as well as their decisions when forming new relationships online?

IDC '23, June 19–23, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0131-3/23/06.
https://doi.org/10.1145/3585088.3593923

To answer these questions, I will conduct three research studies: 1) a User Experience Bootcamp study to co-design online safety interventions with teens, 2) a focus group study with teens to develop an ecologically valid social media simulation for evaluating nudges, and 3) a between-subjects experimental study to evaluate online safety nudges in a social media simulation. In the next section, I describe the context, methods, and contributions of these studies.

2 RESEARCH OVERVIEW

2.1 Study 1 (Completed): User Experience (UX) Bootcamps with Teens (RQ1)

My first study has been completed and published (in-press) at CSCW 2023. My previous work at CSCW 2021 investigated how we can effectively involve youth, as a vulnerable population, in sensitive online safety research, which showed that co-design with teens needs to be mutually beneficial; teens require motivation and incentives for their participation [6]. Therefore, for my first study, I designed and conducted nine User Experience (UX) Bootcamp sessions virtually via Zoom with 21 adolescents (ages 13-17) based in the United States. We taught them important UX design skills, so that they could effectively create storyboards and high-fidelity, interactive prototypes for online safety interventions. In the process, teens were asked to share and design for their unsafe experiences when using social media. The trainings and research activities were conducted over a span of two days in which teens were guided to participate in three activities: a) creating storyboards for unsafe online interactions, b) whiteboarding ideas for online safety solutions, and c) developing prototypes for design-based interventions for online safety. The novel methodology of this research was published as a case study at CHI 2022 [1].

Overall, we found that teens faced risks in private chats including information breaches from strangers, sexually inappropriate messages from strangers, and cyberbullying from people they knew. To address these risks, teens designed interventions for tackling the risks at multiple stages (e.g., before, during, after) and from the perspective of both the victim and the perpetrator. Interestingly, teens more often designed for risk prevention through features that encouraged perpetrators to rethink their actions, by blocking harmful actions from occurring, or penalizing them for inappropriate behavior (Fig. 1). Teens also designed informative alerts, personalized sensitivity filters, and guided actions for assisting the risk victim (Fig. 2). In contrast to prior approaches that protected teen victims after a risk, our paper provides novel recommendations that prevent online risks before they occur by targeting designs toward the risk perpetrator.

2.2 Study 2: Designing an Ecologically Valid Social Media Simulation (RQ2)

My dissertation Study 3 focuses on implementing co-designed nudges in Study 1 within a web-based social media chat simulation to evaluate the efficacy of nudges, where researchers will play the role of other users and introduce low-level risks scenarios triggering nudges. Yet, it is critical to design an ecologically valid social media simulation in order to effectively evaluate nudges, relevant to teens experiences without their involvement in the design process [18].

Therefore, for my Study 2, I will conduct a focus group study with N=20 youth (ages 13-21) to get feedback on and design ecologically valid user personas, risk scenarios, and nudge designs to be implemented within the social media simulation. I will develop 10 personas based on the risk scenarios shared by teens in the UX bootcamps. Moreoever, I will further refine and iterate on the nudge designs to synthesize 4 final nudges catering to 4 risky scenarios, broadly covering low-level information breaching, cyberbullying, inappropriate sexual remarks, and spam message risks. Specifically, teens will be asked questions about how they would improve the personas, risk scenarios, nudges, and to provide suggestions on how the risky interactions would evolve based on their personal experiences. Overall, the contribution of this study is that it will be the first to crowdsource experiences of youth and involve them in the design of an ecologically valid social media simulation.

2.3 Study 3: Evaluation of Online Safety Nudges in a Social Media Simulation (RQ3)

While prior work has investigated nudges in the larger context of privacy and security [13], no formal studies have attempted to evaluate the effects of nudges on coaching teens about how to identify and respond to potentially unsafe people online. Recently, Masaki et al. [14] conducted a survey-based study to understand teens preferences about privacy nudges. We build upon their work by moving beyond self-reported survey data, to a more ecologically valid approach for evaluating real-time interventions that can help teens be safe online.

I will conduct a between subjects design with N=100 teens (ages 13-17), where half of the participants will receive online safety nudges, whereas the other group will not get exposed to any nudges. This study will be conducted remotely via Zoom and an online web-based social media simulation. Teens will be asked to interacte with 10 other social media users in the simulation, with the end goal of identifying safe users on the platform. These social media users will be researchers acting as users, who will follow pre-defined user personas and scripts designed in Study 2 to interact with the participant. The researchers will follow a Wizard-of-Oz approach [11] with rule-based low-level risks introduced in the conversation which will trigger nudges. Throughout the experiment, teens will be part of a Zoom call with a researcher and asked to think aloud as they complete these tasks. Lastly, they will be asked to fill out a post-survey about how their interactions on the social media influenced their perceptions about online safety. Overall, this research will be the first to evaluate adolescent online safety nudges in a realistic simulation, with important implications about how teens build trust, respond to nudges, and form relationships online.

3 CHALLENGES TO GET FEEDBACK ON AT THE DOCTORAL CONSOTRIUM

The evaluation of online safety nudges comes with several unique challenges. While it is essential to simulate low-level risks in the simulation to effectively understand the role of nudges in online safety, I need to ensure teens well-being and safety during the

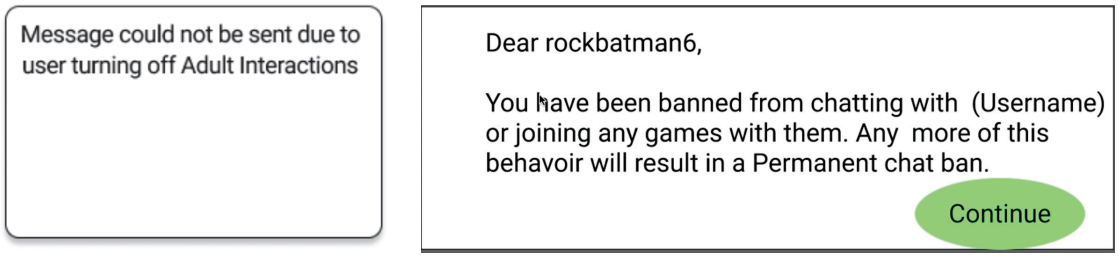

Figure 1: Left: Nudge blocking Risk Perpetrators harmful messages. Right: A nudge warning the risk perpetrator about a penalty if risk continues

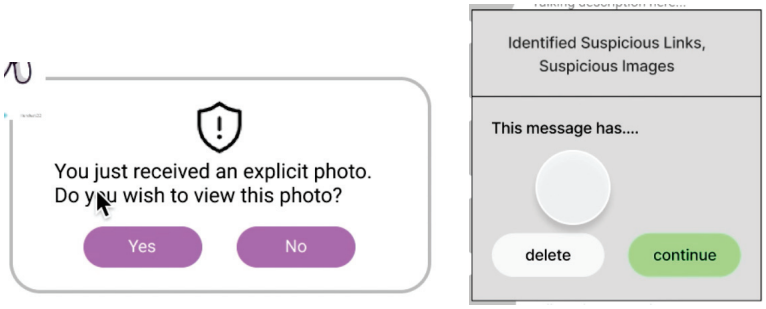

Figure 2: Left: Nudge providing informational risk warning. Right: Nudge guiding the teen towards safe actions.

study. This is also particularly important for addressing IRB ethical concerns regarding this study, as the study will require a certain level of deception to ensure that participants behave realistically in the simulation, which is challenging for adolescent research. Further, researchers who will follow defined user personas may need to deviate from the script for a natural flow of interaction. Therefore, it is challenging to figure out the right balance between a controlled environment and necessary deviations from the persona for a successful experiment. Lastly, conducting the experiment via Zoom, in which participants can think-aloud their decisions is important for getting in-depth feedback for their choices. However, being observed during the experiment may subject the participants to social desirability bias [12] and affect the ecological validity of their interactions. The IDC Doctoral Consortium provides the perfect opportunity for me to get feedback on navigating these challenges from experts in the fields of adolescent research and co-design.

4 DISSERTATION STATUS AND LONG-TERM GOALS

Currently, I am a fourth year PhD student at Vanderbilt University, advised by Dr. Pamela Wisniewski, expected to graduate in May 2024. My Study 1 has been accepted and is in-press at CSCW 2023. For my Study 2, I have obtained IRB approval and plan to complete data collection and analysis by Fall 2023. For my Study 3, I am in the process of obtaining IRB approval and developing a web-based social media simulation, which will be launched in Fall 2023 and completed by Spring 2024. Therefore, the IDC Doctoral Consortium provides the perfect opportunity for me to get guidance on the next steps of my research for finishing my dissertation.

ACKNOWLEDGMENTS

This research was supported by the William T. Grant Foundation (#187941, #190017) and National Science Foundation under grants CHS-1844881. 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.

REFERENCES

[1] 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.

[2] Ashwaq Alsoubai, Xavier Caddle, Ryan Doherty, Alexandra Koehler, Estefania Sanchez, Munmun De Choudhury, and Pamela Wisniewski. 2022. MOSafely, Is that Sus? A Youth-Centric Online Risk Assessment Dashboard. *CSCW'22 Companion* (2022).

[3] Giuseppe Amato, Paolo Bolettieri, Gabriele Costa, Francesco La Torre, and Fabio Martinelli. 2009. Detection of images with adult content for parental control on mobile devices?. In *Proceedings of the 6th International Conference on Mobile Technology, Application & Systems*. 1–5.

[4] Monica Anderson. 2018. A majority of teens have experienced some form of cyberbullying. (2018).

[5] 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.

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

[7] Yara Barrense-Dias, Lorraine Chok, Sophie Stadelmann, André Berchtold, and Joan-Carles Suris. 2022. Sending One's Own Intimate Image: Sexting Among Middle-School Teens. *Journal of school health* 92, 4 (2022), 353–360.

[8] Prema Dev, Jessica Medina, Zainab Agha, Munmun De Choudhury, Afsaneh Razi, and Pamela Wisniewski. 2022. From Ignoring Strangers' Solicitations to Mutual Sexting with Friends: Understanding Youth's Online Sexual Risks in Instagram Private Conversations. *CSCW'22 Companion* (2022).

[9] Leen d'Haenens, Sofie Vandoninck, and Verónica Donoso. 2013. How to cope and build online resilience? (2013).

[10] Douglas A Gentile, Amy I Nathanson, Eric E Rasmussen, Rachel A Reimer, and David A Walsh. 2012. Do you see what I see? Parent and child reports of parental monitoring of media. *Family Relations* 61, 3 (2012), 470–487.

[11] Thomas Grill and Manfred Tscheligi. 2013. The ConWIZ protocol: a generic protocol for wizard of Oz simulations. In *International Conference on Computer Aided Systems Theory*. Springer, 434–441.

[12] Pamela Grimm. 2010. Social desirability bias. *Wiley international encyclopedia of marketing* (2010).

[13] Athina Ioannou, Iis Tussyadiah, Graham Miller, Shujun Li, and Mario Weick. 2021. Privacy nudges for disclosure of personal information: A systematic literature review and meta-analysis. *PloS one* 16, 8 (2021), e0256822.

[14] 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.

[15] Brenna McNally, Priya Kumar, Chelsea Hordatt, Matthew Louis Mauriello, Shalmali Naik, Leyla Norooz, Alazandra Shorter, Evan Golub, and Allison Druin. 2018. Co-designing mobile online safety applications with children. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–9.

[16] Anthony T Pinter, Pamela J Wisniewski, Heng Xu, Mary Beth Rosson, and Jack M Caroll. 2017. Adolescent online safety: Moving beyond formative evaluations to designing solutions for the future. In *Proceedings of the 2017 Conference on Interaction Design and Children*. 352–357.

[17] Wonsun Shin and Hyunjin Kang. 2016. Adolescents' privacy concerns and information disclosure online: The role of parents and the Internet. *Computers in Human Behavior* 54 (2016), 114–123.

[18] Marc Steen. 2013. Co-design as a process of joint inquiry and imagination. *Design Issues* 29, 2 (2013), 16–28.

[19] Richard H Thaler and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

[20] Emily A Vogels, Risa Gelles-Watnick, and Navid Massarat. 2022. Teens, social media and technology 2022. (2022).

[21] Pamela Wisniewski. 2018. The privacy paradox of adolescent online safety: A matter of risk prevention or risk resilience? *IEEE Security & Privacy* 16, 2 (2018), 86–90.

[22] Pamela Wisniewski, Haiyan Jia, Na Wang, Saijing Zheng, Heng Xu, Mary Beth Rosson, and John M Carroll. 2015. Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 4029–4038.