Privacy-Aware Digital Mediation Tools for Improving Adolescent Mental Well-being: Application to School Bullying

Maria Gaci, Isabelle Vonèche-Cardia and Denis Gillet

School of Engineering
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland
email: maria.gaci@epfl.ch, isabelle.voneche-cardia@epfl.ch, denis.gillet@epfl.ch

***Abstract*—In human-computer interaction, self-disclosure of sensitive information regarding distressing experiences requires the establishment of a trust channel between the user and the digital tool. As privacy and security have been identified as factors that contribute to increased levels of trust, they could be utilized to design digital tools that encourage and empower adolescents to disclose school bullying. This work-in-progress paper presents an interdisciplinary research project aimed at combining appropriate levels of usability and security to design a privacy scheme for adolescents in order to provide a digital solution that will help anti-bullying intervention at schools in Switzerland and beyond. The process for designing the interaction and interface of the digital tool is presented in the context of interviews with domain experts. Furthermore, participatory design workshops with Swiss teachers and students are used to inform the key trustful features that the tool should exhibit.**

Keywords–Human-Computer Interaction; Privacy; Trust; Security; Child-Computer Interaction.

I. INTRODUCTION

Bullying, delinquency, substance abuse, depression, and social isolation are some not uncommon distressing experiences that adolescents can encounter during their development [1]. These experiences affect their mental health, development, and wellbeing. School bullying is one of the most prevalent and complex of these experiences and can result in bullycide (suicide due to bullying) [1]. According to two surveys conducted in the Swiss cantons of Valais and Geneva in 2012 and 2013, school bullying affects one to two students per class, or 5-10% of all adolescents in Switzerland [2].

By definition, school bullying is "a systematic abuse of power in interpersonal relations exerted by one or more children" [3] [4]. It is a form of violence exerted by the wrongdoers (referred to as bullies) to the target individuals (referred to as victims) through different forms, such as physical, verbal, or cyber (involving the use of electronic technology) [3]. Various approaches exist for solving bullying conflicts such as the Shared Concern Method [5], yet the identification of the conflicts still remains a challenge as it relies on self-disclosure.

One of the major issues related to the identification of school bullying is the reluctance of adolescents to report their experiences to teachers, parents, or support teams at schools. In fact, it is estimated that less than 15% of students report school bullying conflicts [6]. According to [7], students perceive several barriers to self-disclosing, such as fear that their bullies might perpetrate more frequent or severe attacks, fear of peer disapproval, negative self-thoughts (e.g., feeling weak/undermined), and preference for autonomy or "dealing with it oneself." Moreover, several studies have indicated that teacher's negligence, passive role in intervening and failure to maintain a positive classroom climate directly affect students' decisions to self-report [8]. Witnesses to bullying are also often reluctant to disclose their observations due to fear of retribution and uncertainty about intervention [9].

In spite of research demonstrating that self-disclosure reduces bullying in schools [10] and the fact that higher levels of self-disclosure are recorded when a privacy-oriented approach is adopted in digital environments [11], no digital solutions currently exist in Switzerland. As such, this project aims to design a digital tool for adolescents aged 12–16 years old that will act as a mediator for disclosing school bullying. This digital tool is being developed in collaboration with a Non-Governmental Organization (NGO), and its validation will be performed in public secondary schools in Switzerland in collaboration with their health service teams.

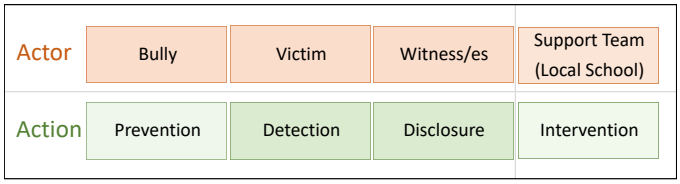

Figure 1. The actors and the actions targeted by the digital tool.

Our tool will serve as a means for the detection and disclosure of school bullying conflicts. Its goal is to encourage adolescents to self-report instances of bullying by providing an interface and means of communication with which adolescents are comfortable, i.e. a mobile application. The application should act as a catalyst to inspire students to seek face-to-face discussions with a human mediator like a teacher or member of the health services team. In order to create a closed-loop model, however, prevention and intervention will also be slightly targeted. Prevention will be targeted by increasing awareness regarding bullying in schools through testimonials. Temporary digital intervention will be provided through tailored advice available in the mobile application. The actors and the actions targeted by the digital tool are illustrated in Figure 1.

This ongoing research project aims to contribute to the field

 ISBN: 978-1-61208-761-0

of Human-Computer Interaction (HCI) and eHealth by providing insights to the following research questions: *How should digital tools be designed in order to encourage adolescents to disclose distressing experiences, such as school bullying? What are the key trustful features that such digital tools should exhibit?* The paper describes the general framework of the project, and the process followed for the initial design of the prototypes of the digital tool for self-disclosing school bullying.

The paper is organized as follows: Section II presents a brief review of literature on the topics of self-disclosure and privacy, existing anti-bullying solutions, and the current limitations. Section III presents the development methodology of the anti-bullying digital tool. Section IV presents the design of the digital tool in accordance with the feedback gathered from domain experts and the target audience of adolescents. Finally, Section V briefly summarises the current state of the project.

II. Literature Review

A. Self-disclosure and Privacy

In 2019, the world's population was composed of 1.2 billion adolescents [12]. Despite the fact that this age group comprises 16 percent of the world population, research studies have emphasized the lack of scholarly work focusing exclusively on them [13]. This insufficient knowledge results in little guidance on the unique requirements, opportunities, and challenges when designing interactive tools for this target group [14]. One of the most effective practices for understanding the needs of adolescents and designing creative interfaces for them is the organization of participatory design workshops [15] [16], an approach that aims to actively engage and include adolescents as co-designers.

Research on the topic of self-disclosure has demonstrated its importance in maintaining psychological, physical, and spiritual well-being [17]. Some of the conditions identified by [17] under which people are willing to disclose personal information are: the specific need for physical "private places", the need for privacy more generally, the identity of the person to whom one might disclose himself, and the relation between the two.

Digital interaction also favors self-disclosure [18]–[21]. Medical patients have reported a higher number of symptoms and negative behaviors when interviewed through a digital tool rather than face-to-face [19]. Furthermore, participants claim to provide more honest and candid answers when digital means were utilized [20]. Higher disclosure rates have also been recorded in studies eliciting sensitive information through digital means rather than face-to-face or using pen-and-paper [21]. The same applies to scenarios where individuals may feel particularly vulnerable to the consequences of self-disclosure [21]. Finally, rates of self-disclosure through digital interfaces have been observed to be higher for individuals who perceive their health condition to be stigmatized [11].

When comparing digital and face-to-face interactions, privacy, and anonymity were among the top factors listed affecting subjects' willingness to disclose sensitive information [18] [21]. Higher rates of self-disclosure were recorded when people communicated in a visually anonymous manner rather than non-anonymously. Furthermore, increased willingness to answer sensitive questions and decreased errors associated with sensitive topics were recorded when privacy-enhancing data collection modes were employed.

Recent studies by [22] and [23] have demonstrated a preference for embodied conversational agents (e.g., virtual agents) rather than human interviewers. Four principal reasons were determined to affect the individuals' preference: (i) lack of judgment, criticism, and reactions (verbal or nonverbal), (ii) ease of providing answers due to the digital interface (texting enables participants to formulate their answers at their own pace), (iii) personal comfort due to reduced negative feelings, such as anxiety, embarrassment, or guilt, and (iv) protection of personal information and privacy.

Individuals tend to more likely disclose their experiences if they are doing so to someone who is perceived to be trustworthy [17]. Hence, the establishment of trust is required for individuals to feel comfortable in self-reporting sensitive information. Due to the complexity of trust as a social phenomenon, major questions arise with how to establish trust and how to reliably signal that trust through different interfaces and interactions [24]. According to [25], there are four elements that ease the communication of trust through digital interfaces: (i) design quality (site organization, visual design), (ii) up-front disclosure and transparency, (iii) comprehensive and current content, and (iv) connection to the rest of the web.

As privacy is significant to encouraging self-disclosure, attempting to establish trust requires a thorough security and privacy model [26]. Trust models, like [27], which combine aspects of usability and security have been demonstrated to impact users' levels of trust. The above model is composed of six building blocks, namely: (i) security (authentication, data access control, data integrity, software change procedures, and physical security), (ii) usability (perception issues, motor accessibility, and interaction design issues), (iii) privacy (user anonymity and data confidentiality), (iv) reliability and availability (vulnerability to denial-of-service attacks, connection to the internet, quality of service), (v) audit and verification mechanisms (cryptographic methods, audit trails, use of trusted agents), and (vi) user expectations (product reputation, prior user knowledge, knowledge of technology).

B. Existing Solutions

Recently, increasing emphasis has been placed on the design and evaluation of digital tools for adolescents. These digital tools not only help adolescents understand the nature of bullying but also serve as sources of positive intervention. The existing approaches can be categorized into two major groups: (i) as approaches for the prevention and (ii) as approaches for the disclosure of the bullying cases in schools. Below, existing solutions for each approach are presented along with a brief analysis of the mechanisms utilized by each of them to give an overview of the emerging technologies in this field.

1) Prevention-oriented Approaches: FearNot! [28], StopBully [29] and #StopBully [30] are interactive mobile apps aiming to develop the behavioral competence of victims and witnesses necessary to avoid and deal with future bullying situations. All three mobile apps educate and raise awareness by presenting real-life situations that the players have to respond to. The mechanisms utilized by the apps to teach effective responses to bullying are as follows:

- *Storytelling* is utilized by FearNot! to present the user with a virtual environment that improvises real-life bullying situations. Three-dimensional agents in

 ISBN: 978-1-61208-761-0

a virtual school are designed to foster empathy and emotional involvement of the users. The story is gradually built in response to the suggestions of the users in the various episodes through artificial intelligence techniques.

- *Games* are a mechanism utilized by StopBully and #StopBully for delivering educational content to players. The games provide players with challenges that they have to solve to gain points and progress to higher levels. Both games are educational with cartoon-like characters and environments.
- *Videos*, *animated comics* and *quizzes* are utilized by #StopBully to train players on bullying. Videos and comics present the friendship of four characters, which is broken when one of them becomes a bully. The knowledge acquired through the videos and comics is put into practice through quizzes. There are different types of quizzes, such as multiple-choice, rearrange the letters, type an answer, etc.

2) Disclosure-oriented Approaches: STOPit Solutions [31] and Anonymous Alerts [32] are two similar incident reporting apps for students experiencing distress in schools. Both apps are available as mobile and web versions, and they provide digital solutions for two types of users — students and teachers — as described below.

- Students can utilize the platforms provided by the two apps mentioned above to anonymously report bullying cases by attaching videos, photos, and screenshots to the report. The apps also enable students to customize the incident type, location, and language. An anonymous messaging channel is also available to enable students to seek immediate help along with an emergency button that alerts the severity of the incident.
- Teachers can utilize the incident management platform to receive real-time updates on incident reports sent by students and parents. The platform enables them to monitor the reports and forward them to local authorities in case immediate action is required. Teachers can also run analytic and trend reports to identify patterns of bullying in their school.

These approaches, while effective in some contexts and for some problems, are not sufficient for our intended application and audience.

C. Limitations

In summary, digital solutions are available for adolescents to disclose bullying acts, yet some gaps exist: the existing solutions focus on one of the actors or specifically in one action. Moreover, there is limited research in the field of HCI on how trust can be established by design between adolescents and digital tools when self-disclosing sensitive mental health experiences. Finally, there exists limited analysis of how privacy and anonymity affect the disclosure of distressing experiences among adolescents, as they could be also utilized to promote spam or to practice additional bullying.

III. Methodology

To develop digital tools that nudge adolescents (victims, witnesses, and bullies) into disclosing bullying experiences in schools, a design thinking process [33] is being followed. The goal is to gain an empathetic understanding of the issue of bullying in schools, necessary for designing an effective solution for adolescents to disclose such experiences. The design thinking process is composed of four stages (Ideating, Prototyping, Implementing, and Validating), but as the project is still ongoing, only the initial two stages of the process are described in this paper.

Multiple iterations of this process were used to re-frame the issue of bullying in a human-centric way. Ideas regarding features and functions that the digital intervention should encompass were gathered through engagement with experts in the field as well as future users, as described below:

Initially, the process of requirements elicitation was conducted through unstructured interviews with domain experts, educators, and bullying mediators. Interviews were designed to be broad and open-ended to gain a deeper perspective on domain-specific practices, goals, and concerns related to bullying in public schools in Switzerland. The aim of the interviews was also to identify factors that might influence the adoption of new solutions in the specific context. Six one-hour meetings were held in the span of one year.

Research studies have demonstrated that focus groups are a successful data collection technique for interaction design for adolescents [14]. Thus, focus groups were organized both face-to-face and online (due to COVID-19) during the spring semester of the 2019-2020 academic year. As the project targets adolescents, the focus groups were organized with students in public secondary schools in Switzerland. The focus groups aimed to gather a wide range of opinions, viewpoints, and insights by the users, as well as raising issues that were not previously identified. Adolescents age 12–16 were encouraged to engage in interactive group discussions regarding the design of the digital tools and, participatory design workshops were organized to propose the initial paper-based prototypes. A teacher and/or mediator already familiar with the students of each focus group (hereafter referred to as teacher *T1* and *T2*) were present and moderated the sessions. Two female and five male adolescents were part of the first focus group, and four female and one male adolescent were part of the second focus group (hereafter referred to as student *S1 – S12*). Students were recruited on a voluntary basis by their teachers.

During the interactive group discussions, adolescents described the bullying situation in their respective schools mentioning the methods that they can currently use to disclose instances of bullying, i.e., talking directly to teachers, calling a mediator, and sending an email. Students mentioned an initiative for dropping anonymous letters in designated boxes in the schools' halls, yet the project was never finalized. The principal issue identified with the current approaches was the fact that "the students do not dare to talk" (*S4*), as "everyone is different and some students are at ease talking to people that they know, but some of them are at ease when anonymous" (*S10*). Hence, anonymity emerged as an important aspect when disclosing distressing experiences, as students: i) were afraid of how adults would react, ii) were afraid of the fact that adults might judge them, or iii) were feeling intimidated due to the sensitivity of the topic. As such, in addition to requesting extra-curricular programs for raising awareness about bullying in schools, participants requested an anonymous digital channel for self-disclosure.

During the participatory design workshops, adolescents were involved in brainstorming sessions regarding design

 ISBN: 978-1-61208-761-0

possibilities. Participants imagined the application would "play the role of a mediator that the students can write to, if they trust it, as the information is confidential" (*S11*). The application would further help them "feel less lonely because usually when students are bullied, they are left in the loneliness and they do not really want to have contact with people" (*S12*). Besides, the application would help students feel "reassured" (*S4*) and "send small messages every day asking how they were feeling" (*S6*). The main feature of the application was voted to be anonymity, as "an anonymous application cannot judge the students" (*S4*). Adolescents thought that the user should enter a "*pseudonym*" (*S4*) and an optional phone number, but no additional personal information or password to guarantee *anonymity by design*. When the students were asked to brainstorm on the series of interactions offered in the application, they mentioned the selection of a "role, such as victim or witness" (*S3*) and "form of bullying, such as verbal or physical" (*S5*), so that the application could pose the appropriate questions according to the needs of the student. Several students from both focus groups suggested that a chatbot might be used to customize the experience. Moreover, they suggested that the chatbot might be able to decide if the user was in danger and the appropriate measure that the user should take. The questions posed by the chatbot should be "convincing" (*S4*) so that the user is encouraged to request help from the support team, but must also be able to decide what information he/she would like to disclose. Finally, adolescents stated that they would "trust the chatbot" (*S12*) because their perception was that "Artificial Intelligence does not make mistakes" (*S3*).

IV. Tool Design

Taking into consideration the feedback from the domain experts and the ideas of the students and teachers during the participatory design workshops on the features that the digital tool for self-disclosing bullying should exhibit, a mobile application will be developed. The architecture, features that will enhance usability, and the security considerations for the application based on the trust model presented in Section II-A are presented below.

A. Architecture

The requirements elicitation process with the experts and future users enabled the selection of the appropriate type of digital tool, namely a mobile application, for two main reasons. Firstly, according to initial interviews with domain experts, adolescents in the school setting are primarily bullied by their peers, and portable devices that enable students to disclose information quickly and privately were recommended. Thus, even though the larger screen size of a computer-based tool enables better readability, better visual acuity, and higher usability, these benefits are a drawback to privacy, as bystanders may be able to read the content on the screen from different distances [34]. Secondly, according to initial interviews with the different parties, the majority of students in Switzerland have access to mobile phones, although students are not allowed to utilize them during lecturing.

A hybrid mobile application will be developed using the React Native framework to support multiple mobile operating systems. The application will be structured in two main layers of infrastructure: the mobile device and the remote school. Initially, the bullying information disclosed by the user will

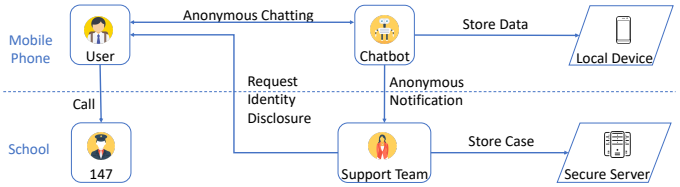

Figure 2. The architecture of the app for self-disclosing bullying in schools.

be anonymously stored in the mobile device. If the user agrees to disclose the information with the support team, their information will be shared with the support team of the specific school and stored in a secure local server. A human mediator will then contact the user anonymously and request the user to disclose their identity to provide face-to-face help. Students can also utilize the app to directly and securely contact third-parties, such as 147 (the consultation service for young people in Switzerland) or the police department. A simplified version of the architecture of the application is illustrated in Figure 2. The figure depicts: (i) the two layers of infrastructure, (ii) the interaction of users with the chatbot and the support team, and (iii) the storage of data on the local device and the remote server.

B. Usability

Chatbots were requested by students during the focus groups as a means to automate the disclosing process by "providing emergency support 24/7" (*T1*). As research has also demonstrated that they could be utilized to enable users to access information about bullying at any time and improve school cohesion [35], they will be included in the app. Students suggested that the chatbot will firstly ask generic questions to familiarise themselves with the user and then gradually perform a self-assessment task to determine the role of the user (i.e., victim, bully, witness). According to the students, this approach would also "increase the trust level" (*S12*) towards the application. The chatbot will provide immediate support in the form of advice ("Would it be possible to talk to your parents about this issue?" - *T2*) and try to nudge the user to disclose the case and seek help from a human mediator. It will never disclose information without the consent of the user unless the user explicitly mentions *suicide*, in which case due to legal requirements in Switzerland, the anonymous chat will be automatically forwarded to the support team. The chatbot will be programmed to run locally on the mobile device, so the costs and benefits of rule-based versus AI implementations are being evaluated. An illustration of a possible conversation with the chatbot is shown in Figure 3.

Avatars were suggested by the students as an option for creating immersive digital tools and have been demonstrated to be an effective mechanism to interact with users for anti-bullying education [36]. As such, users will be able to customize the appearance and characteristics (e.g., gender, age, name) of the avatar according to their preferences. The goal will be to enable users to personalize the experience in the application by creating distinct avatars that they "trust" (*S11*) and "feel comfortable having digital conversations with" (*S12*).

An interesting feature identified during the focus groups was continuously tracking the mental state of the adolescent through regular reminders. Students proposed an automated

 ISBN: 978-1-61208-761-0

notifications system that "will contact the user if he/she has not connected to the app for a specific amount of time" (*S6*) (the amount of time can be configured in the settings of the application). Through the notifications, the chatbot will ensure that the mental health of the user is not at a critical level, and it will "show interest" (*S11*) aiming to decrease the feeling of solitude and negligence. Furthermore, reminders will also be sent to teachers and the support team if the case has been neglected for a long time.

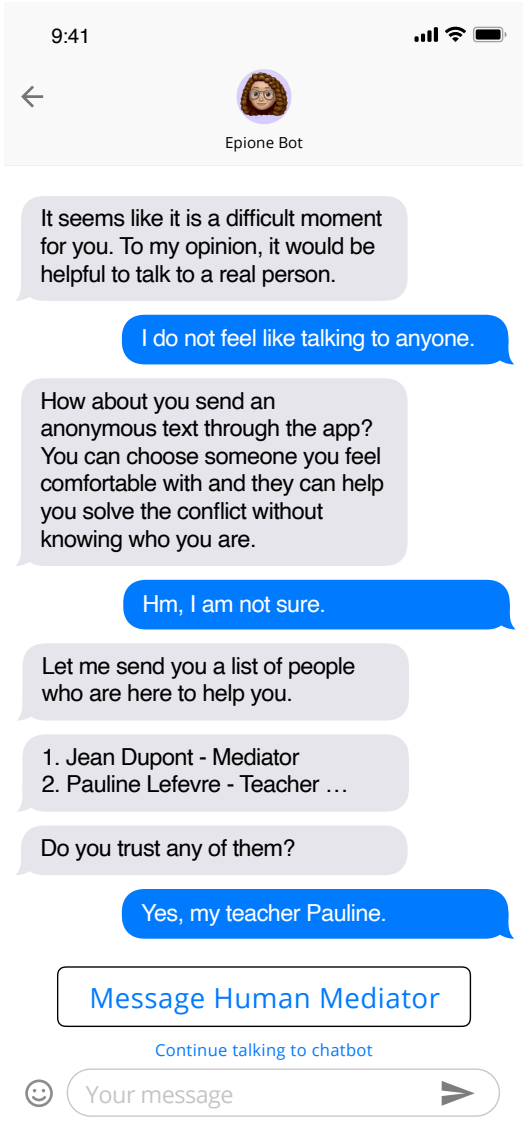

Figure 3. The User Interface of the app for self-disclosing bullying in schools. (The original prototypes were designed in French and translated to English for clarification purposes.)

The results of the interviews and focus groups gave insights into the importance of creating a community feeling. An effective way to endorse a feeling of community within the app is through sharing personal experiences, thus a *Temoignages* (testimonials) page has been created. Adolescents have the option of choosing to disclose anonymously in the school group their experience and receive support from their peers through the *heart* button. To avoid additional bullying and spamming, the shared experience needs to be validated by the team of the local school. By sharing experiences on a dedicated page, the application aims to raise awareness regarding the negative aspects of bullying, as well as to show empathy to the bullying targets.

Text and voice calls are the only modalities for sharing experiences. No videos or pictures can be shared in the application as "they might infringe the privacy of witnesses" (*T1*). Users have the option of disclosing bullying either with the support team or with a specific teacher. A list of teachers who are willing to participate in the program will be included in the app, as focus groups indicated that some students are "more comfortable and trusting with specific individuals who they already know" (*S9*). Emergency contact numbers will be included in the application if the child wishes to make phone calls and receive external help by third parties (147 or police department).

C. Privacy and Security

Initial interviews with domain experts suggested that the bullying reports should be handled by the designated team at each local school. As such, an appropriate group signature scheme [37] will be selected to ensure credential/membership authentication. The group signature scheme will ensure that the local school can verify that the bullying report was sent from an authenticated adolescent, however, it will not reveal his/her identity. A fully-anonymous system is being evaluated by distributing a group member secret key to all the identified adolescents [38] with no possibility of signature tracing through the use of special trapdoors. QR codes will be posted in each school in order to redirect users automatically to the local school support team.

The implementation of an end-to-end encrypted messaging channel will enable students and the support team of the local school to communicate in real-time, using cryptographic protocols such as Signal [39]. Thus, the support team can provide help without being aware of who the student is if the student wishes to remain anonymous during the process. Furthermore, the encryption of the chat will ensure the security of the sensitive information disclosed in the application.

Offline capabilities will be considered in order to ensure the availability of digital tools for adolescents. The offline version will offer limited functionalities, and it will enable automatic syncing once the device is connected to the internet.

An emergency delete button will be included, in case the user wishes to delete all the data saved in his local device and restart the application from the beginning. Finally, the option of deleting individual messages exchanged with the human mediator will be included to give individuals control over their personal data.

D. Trust

The authors of [27] suggested that usage of trust models facilitates the successful deployment of new technologies, hence, their trust model was referenced to design the digital tool for self-reporting bullying in schools. Usability and security were carefully incorporated, and the features of the tool are compliant with the six building blocks enumerated in Section II-A:

- Security: Integration of group signature schemes to authenticate users provides enhanced security in the app. Furthermore, users will have complete control over the personal data provided in the app.

 ISBN: 978-1-61208-761-0

- Usability: The user interface was designed in collaboration with teachers, experts, and adolescents in order to address the needs of all types of users.
- Privacy: The tool will be fully anonymous, and no identifiable data will be collected. User data will be stored locally on the device until the user agrees to contact a human mediator.
- Reliability and availability: The tool will be open-source and provided for free to students and teachers.
- Audit and verification mechanisms: End-to-end encryption of the user messages will be provided to ensure the security of sensitive data.
- User expectations: The tool is being developed by a public institution in Switzerland, and as a result, it will follow the Swiss guidelines for the protection of personal data. The simple design of the features proposed during the workshops also aims to conform with prevailing norms of mobile user interfaces.

The trust model will be validated in the future through testing with adolescents and teachers, and it will be iteratively improved to better meet the needs of the project.

E. Summary

In summary, the outcome of the focus groups and the participatory design workshops reveals that the participants, on the whole, agreed that the digital tool for bullying scenarios should be designed as a means for adolescents to receive support rather than as a means for filing bullying complaints. Therefore, participants suggested that the tool should be carefully designed to act as a companion for adolescents by providing advice through an intelligent chatbot, regularly showing interest through notifications, and ensuring privacy through anonymity.

Unlike the existing solutions presented in Section II-B2, participants suggested that the digital tool addresses all actors involved in a bullying scenario, namely the victims, the witnesses, and the bullies. The reasoning behind the suggestion was that the tool should raise awareness about bullying in schools and support should be provided to all adolescents.

The majority of participants preferred a chatbot over directly contacting a human. Several reasons were echoed for such a preference such as the lack of judgment, the general perception that Artificial intelligence cannot make mistakes, and the ability to provide immediate support. Another view presented by the participants was the personalizing of the experience in the tool tailored to one's needs through the ability to customize the behavior and the appearance of the chatbot.

The main advantage of the tool was revealed to be the fact that students can receive support while remaining anonymous. This would break the barrier of taking the first step into self-disclosing bullying and increase the level of trust between the adolescents and the digital tool. While some students indicated that they would be more comfortable talking to people they were familiar with, all participants agreed that an additional channel for disclosing bullying in their schools would be necessary, as adolescents are different from each other.

Finally, students believed that bullying prevention is best achieved through extra-curricular events in their schools, hence the main target of the tool should remain detection and disclosure. Prevention could be slightly targeted to raise awareness through testimonials. Interviews with experts revealed that intervention should be handled by a human mediator. Yet to create an incentive for the student to utilize the tool, it was decided that the chatbot would provide circumstantial advice and repeatedly encourage students to talk to a human through an anonymous channel.

V. Conclusion and Future Work

This paper aims to present the design process of an interactive digital tool for the self-disclosure of adolescents who are involved in bullying conflicts in schools. As adolescents belong to the group of the population that is not well understood, participatory design workshops were organized in public secondary schools in Switzerland to identify the key trustful features that the digital tool should exhibit. The results of this study indicate that: i) secure chatbots that assess the emotional state of the users and react accordingly ii) features that tailor the user experience such as avatars and notifications, and iii) features that foster a feeling of privacy, are aspects that should be taken in consideration when prototyping interfaces for sensitive mental health data.

The current existing solutions for disclosing bullying are not aligned with the feedback received from the students during the focus groups organized for this study. The disclosure-oriented approaches presented in Section II-B2 act primarily as a reporting system for negative activity in schools and heavily rely on the vulnerable individuals taking the initiative to file a report. Based on the interviews with the experts in the field, this approach can act as a barrier for adolescents to take the first step and disclose bullying in schools. As such, the idea is to assure users that the goal of the digital tool is to provide support to vulnerable individuals rather than solve bullying conflicts. This approach will nudge adolescents into gradually disclosing sensitive information (regardless if they are victims, witnesses, or bullies) and it was referred to be more favorable and efficient by the students who participated in the focus groups.

Based on the literature review presented in Section II demonstrating that self-disclosure reduces bullying in schools, trust is required for individuals to feel comfortable self-disclosing sensitive information, and privacy is an important aspect for establishing a trust channel, a hypothesis has been formulated in order to be tested in the future. The hypothesis is that *by designing privacy-aware medical applications, adolescents will be more inclined to disclose the distressing experiences that will contribute to improved mental health and will help develop sustainable behavior among adolescents*. Initial focus groups and participatory design with adolescents in Switzerland seem to support the hypothesis, yet additional research is required to test and validate it.

Considerably more work will need to be done in the future to evaluate the initial prototypes designed during the participatory design workshops. A natural progression of this work is to implement the digital tool based on the feedback gathered from the adolescents and the experts. As soon as a minimum viable product will be available, the features of the tool will be tested firstly with the experts and teachers who will contribute to devising several conversation scenarios for training the chatbot. Secondly, the tool will be tested with the students of the focus groups to validate the assumptions about the tool's requirements. Finally, the tools will be validated through iterations on the design of the tool from the feedback received through A/B testing in public schools in Switzerland.

 ISBN: 978-1-61208-761-0

To conclude, this research project will not only contribute to various disciplines within Computer Science such as Human-Computer Interaction, Computer Security and Cryptography, e-Health, and Social Computing, but it will also deliver digital solutions that will be available for immediate utilization in public schools in Switzerland and abroad.

ACKNOWLEDGMENT

This EPFLinnovators project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754354. Special thanks to the domain experts from the NGO and the support team for school health, teachers and the students in the two secondary schools in Switzerland for their inputs and feedback during the participatory design workshops.

REFERENCES

[1] APA, "Developing adolescents: A reference for professionals," Washington, DC: American Psychological Association, 2002.

[2] RTS. One to two students per class are victims of bullying in Switzerland. [retrived: Oct, 2020]. [Online]. Available: https://www.rts.ch/info/suisse/6718365-un-a-deux-eleves-par-classe-sont-victimes-de-harcelement-en-suisse.html (2015)

[3] K. Rigby, The Method of Shared Concern: A positive approach to bullying in schools. Aust Council for Ed Research, 2011.

[4] A. Pikas, "New developments of the shared concern method," School Psychology International, vol. 23, no. 3, 2002, pp. 307–326.

[5] J.-P. Bellon and B. Gardette, School bullying: defeating it, it's possible: The shared concern method. ESF sciences humaines, 2018.

[6] A. Castillo. How to prevent bullying at school. [retrived: Oct, 2020]. [Online]. Available: https://www.letemps.ch/economie/prevenir-harcelement-scolaire (2018)

[7] M. J. Boulton, L. Boulton, J. Down, J. Sanders, and H. Craddock, "Perceived barriers that prevent high school students seeking help from teachers for bullying and their effects on disclosure intentions," Journal of adolescence, vol. 56, 2017, pp. 40–51.

[8] K. I. Cortes and B. Kochenderfer-Ladd, "To tell or not to tell: What influences children's decisions to report bullying to their teachers?" School psychology quarterly, vol. 29, no. 3, 2014, p. 336.

[9] I. Oh and R. J. Hazler, "Contributions of personal and situational factors to bystanders' reactions to school bullying," School Psychology International, vol. 30, no. 3, 2009, pp. 291–310.

[10] W. P. Murphy, J. S. Yaruss, and R. W. Quesal, "Enhancing treatment for school-age children who stutter: Ii. reducing bullying through role-playing and self-disclosure," Journal of fluency disorders, vol. 32, no. 2, 2007, pp. 139–162.

[11] S. A. Rains, "The implications of stigma and anonymity for self-disclosure in health blogs," Health communication, vol. 29, no. 1, 2014, pp. 23–31.

[12] Adolescents overview. [retrived: Oct, 2020]. [Online]. Available: https://data.unicef.org/topic/adolescents/overview/ (2019)

[13] S. Yarosh, I. Radu, S. Hunter, and E. Rosenbaum, "Examining values: an analysis of nine years of idc research," in Proceedings of the 10th International Conference on Interaction Design and Children, 2011, pp. 136–144.

[14] E. S. Poole and T. Peyton, "Interaction design research with adolescents: methodological challenges and best practices," in Proceedings of the 12th International Conference on Interaction Design and Children, 2013, pp. 211–217.

[15] Z. Ashktorab and J. Vitak, "Designing cyberbullying mitigation and prevention solutions through participatory design with teenagers," in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016, pp. 3895–3905.

[16] G. M. McCarthy, E. R. Rodriguez Ramírez, and B. J. Robinson, "Participatory design to address stigma with adolescents with type 1 diabetes," in Proceedings of the 2017 Conference on Designing Interactive Systems, 2017, pp. 83–94.

[17] S. M. Jourard, The transparent self. Van Nostrand Reinhold Company, 1971.

[18] A. N. Joinson, "Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity," European journal of social psychology, vol. 31, no. 2, 2001, pp. 177–192.

[19] J. H. Greist, M. H. Klein, and L. J. Van Cura, "A computer interview for psychiatric patient target symptoms," Archives of General Psychiatry, vol. 29, no. 2, 1973, pp. 247–253.

[20] M. Ferriter, "Computer aided interviewing and the psychiatric social history." Social Work and Social Sciences Review, 1993.

[21] S. Weisband and S. Kiesler, "Self disclosure on computer forms: Meta-analysis and implications," in Proceedings of the SIGCHI conference on human factors in computing systems, 1996, pp. 3–10.

[22] M. D. Pickard, C. A. Roster, and Y. Chen, "Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions?" Computers in Human Behavior, vol. 65, 2016, pp. 23–30.

[23] G. M. Lucas, J. Gratch, A. King, and L.-P. Morency, "It's only a computer: Virtual humans increase willingness to disclose," Computers in Human Behavior, vol. 37, 2014, pp. 94–100.

[24] L. F. Cranor and S. Garfinkel, Security and usability: designing secure systems that people can use. O'Reilly Media, Inc., 2005.

[25] A. Harley, "Trustworthiness in web design: 4 credibility factors," Utg. av Nielsen Norman group. url: https://www. nngroup. com/articles/trustworthy-design, 2016.

[26] K. S. Jones, "Privacy: what's different now?" Interdisciplinary Science Reviews, vol. 28, no. 4, 2003, pp. 287–292.

[27] L. J. Hoffman, K. Lawson-Jenkins, and J. Blum, "Trust beyond security: an expanded trust model," Communications of the ACM, vol. 49, no. 7, 2006, pp. 94–101.

[28] M. Sapouna et al., "Virtual learning intervention to reduce bullying victimization in primary school: a controlled trial," Journal of Child Psychology and Psychiatry, vol. 51, no. 1, 2010, pp. 104–112.

[29] C. Raminhos et al., "A serious game-based solution to prevent bullying," International Journal of Pervasive Computing and Communications, vol. 12, no. 2, 2016, pp. 194–215.

[30] H.-F. Neo, C.-C. Teo, and J. L. H. Boon, "Mobile edutainment learning approach: #StopBully," in Proceedings of the 2nd International Conference on Digital Technology in Education. ACM, 2018, pp. 6–10.

[31] STOPit Solutions. [retrived: Oct, 2020]. [Online]. Available: https://stopitsolutions.com

[32] Anonymous Alerts. [retrived: Oct, 2020]. [Online]. Available: https://www.anonymousalerts.com/webcorp/

[33] H. Plattner, "An introduction to design thinking process guide," The Institute of Design at Stanford: Stanford, 2010.

[34] J. F. Jones, S. A. Hook, S. C. Park, and L. M. Scott, "Privacy, security and interoperability of mobile health applications," in International Conference on Universal Access in Human-Computer Interaction. Springer, 2011, pp. 46–55.

[35] A. Latham, K. Crockett, and Z. Bandar, "A conversational expert system supporting bullying and harassment policies." vol. 1, Jan 2010, pp. 163–168.

[36] R. Aylett et al., "Unscripted narrative for affectively driven characters," IEEE Computer Graphics and Applications, vol. 26, no. 3, 2006, pp. 42–52.

[37] D. Chaum and E. Van Heyst, "Group signatures," in Workshop on the Theory and Application of of Cryptographic Techniques. Springer, 1991, pp. 257–265.

[38] Y.-k. Lee, S.-w. Han, S.-j. Lee, B.-h. Chung, and D. G. Lee, "Anonymous authentication system using group signature," in 2009 International Conference on Complex, Intelligent and Software Intensive Systems. IEEE, 2009, pp. 1235–1239.

[39] K. Cohn-Gordon, C. Cremers, B. Dowling, L. Garratt, and D. Stebila, "A formal security analysis of the signal messaging protocol," in 2017 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, 2017, pp. 451–466.

 ISBN: 978-1-61208-761-0