Developing Design Interventions for Cyberbullying: A Narrative-Based Participatory Approach

Leanne Bowler¹, Eleanor Mattern¹ and Cory Knobel²
¹ School of Information Sciences, University of Pittsburgh
² University of California-Irvine, Donald Bren School of Information and Computer Sciences

Abstract
This paper presents a user-generated framework for designing affordances that would counter acts of cyberbullying on social media sites. To do so, we used narrative inquiry as a research methodology, which allowed our two focus groups – one composed of teens and the other of undergraduate students – to map out a cyberbullying story and overlay it with a set of design recommendations that, in their view, might alleviate mean and cruel behavior online. Four “cyberbullying stories” were constructed by the participants, each one revealing two sub-plots – the story that “is” (as perceived by these participants) and the story that “could be” (if certain design interventions were to be embedded in social media). In this paper, we describe seven emergent design themes evident in the participants’ design recommendations for social media: design for reflection, design for consequence, design for empathy, design for personal empowerment, design for fear, design for attention, and design for control and suppression.

Keywords: social computing, cyberbullying, design interventions, values in design, participatory design, narrative inquiry


Copyright: Copyright is held by the authors.

Acknowledgements: We gratefully acknowledge the participation of the young people in this study. Their contributions are invaluable. We thank Patrick Keilty, Assistant Professor at the Faculty of Information, University of Toronto, for his helpful comments and suggestions.

Contact: lbowler@sis.pitt.edu, emm100@pitt.edu, cknobel@uci.edu

1 Introduction

This article presents a user-generated conceptual framework for understanding and guiding the design of social media that counteracts or prevents mean and cruel online behavior. It does so through the use of narrative inquiry, a research method that allowed teens and young adults to map out a cyberbullying story and overlay it with a set of design recommendations that, in their view, might prevent or alleviate cyberbullying.

Two focus groups – a group of four teens in high school and a group of five young adults completing their undergraduate studies – used storytelling and sketching to elicit visual narratives that communicated their perceptions of cyberbullying and to propose design features that might shape the cyberbullying story in a more positive direction. Four “cyberbullying stories” were constructed by the participants, each one revealing two sub-plots – the story that “is” (as perceived by these participants) and the story that “could be” (if certain design interventions were to be embedded in social media).

2 Cyberbullying

Bullying as a major public health concern is a historic problem, but 21st century technologies have allowed for it to assume new characteristics and have introduced new tactics for aggressive behavior (Juvonen and Gross, 2008, p. 497). “Cyberbullying,” as a distinct form of bullying, has consequently entered the vernacular in recent years, with scholars pointing to the intentional use of technology as a means to hurt another individual. In the hands of young people who are still developing their impulse control and are particularly...
vulnerable to peer-pressure, social media can allow for 'online expressions of offline behaviors' and facilitate negative and damaging activities, one of which is cyberbullying (O’Keefe et al., 2011, p. 800).

Cyberbullying, while lacking in a consensus definition (National Science Foundation, 2011; Stewart and Fritsch, 2011), reflects the core elements of bullying as it is traditionally understood. Researchers at the National Science Foundation observe that the variety of definitions of cyberbullying “typically start with three concepts: intent to harm, imbalance of power and usually a repeated action, although some experts replace ‘repeated action’ with ‘specific targets’ ” (National Science Foundation, 2011). The use of electronic technologies to carry out this intent, display this imbalance of power, and target others repeatedly in cyberspace is, naturally, another component of the definitions surrounding cyberbullying. Brady (2010) describes cyberbullying as “the use of communication-based technologies, including cell phones, e-mail, instant messaging, text messaging, and social networking sites, to engage in deliberate harassment or intimidation of other individuals or groups of persons using online speech or expression” (p. 113). Patchin and Hinduja (2006) are briefer in their description, defining cyberbullying as the “willful and repeated harm inflicted through the medium of electronic text” (p. 152).

While the academic literature provides this definitional context, Alice Marwick and danah boyd (2011) explain that the language adults used to speak about cyberbullying may differ from the language used by the young people who are involved in or who observe this online behavior. Rather than characterizing instances of online name-calling, arguments, and discord as “cyberbullying,” Marwick and boyd explain that teens attach the label of “drama” to these incidents.

Cyberbullying has qualities that are distinct from bullying in its more traditional form of direct, face-to-face interaction between the dominant individual (the bully) and the less dominant individual (the victim). Juvonen and Gross (2008) cite the pervasiveness of Internet use, coupled with the absence of adult supervision in online environments, as creating “a fertile ground for bullying” beyond school grounds (p. 497). In this electronic environment, bullies may feel both a sense of anonymity and distance, feelings that can promote harmful behavior (Mason, 2008; Suler, 2004; Trolley et al., 2006; Williard, 2005). The online environment provides an apposite set of factors for bullying to occur. Current research in online behavior and cyberbullying suggests that people with depression (which perpetrator and the target often struggle with) tend to prefer online social interaction, which may drive more behavior into the “cyber” context of bullying (Caplan, 2003). Further, online communication can be less tempered, and more emotionally and socially charged than face to face communication. Though anonymity is frequently pointed out as a disinhibiting factor in online interactions, Suler points to five additional factors at play with relevance to cyberbullying: invisibility, asynchronicity, solipsistic introjection, dissociative imagination, and minimization of authority - each of which can take place in non-anonymous contexts such as social media (Suler, 2004).

Moreover, Rogers (2010) observes that the nature of cyberbullying allows for harassment and intimidation to gain entry into environments that were “safe” from traditional acts of bullying. She writes, “Cyberbullying can take place at any time during the 24-hour day...This can be responsible for a large part of the emotional damage inflicted on victims, who then feel they have no refuge, no one to trust and can never be safe anywhere” (p. 13). As Slonje and Smith (2008) remark, a victim of bullying in its traditional sense may be able to find solace in the knowledge that the bullying would remain on school grounds. In the case of cyberbullying, the victim may receive texts, emails, messages via social networking sites in their safe place and at all times. More recent findings by Sevcikova, Smahel, and Otavova (2012) show that adolescents (research subject group aged 15-17) perceive that online bullying is entangled with the social milieu of school and the victim is aware that a collective group of peers are bearing witness or observing the behavior – a common configuration in social media. This may render the feelings of victimization and transgression as more intense than if the bullying were enacted in physical space.
Both the scholarly literature and popular media reveal that bullying behavior may have lasting and devastating consequences. Juvonen, Graham, and Schuster (2003) characterize bullying as a “major public health concern facing youth” and describe adjustment difficulties, mental health challenges, and violent behavior as among the effects of bullying. Patchin and Hinduja (2010) speak to the relationship between cyberbullying and low self-esteem, problematic behavior in school, and family discord. Moreover, as news coverage of cyberbullying has made evident, there have been incidents in which cyberbullying victims have taken their own lives. Consequently, efforts have been made to combat cyberbullying through intervening measures. The literature suggests that these interventions can be classified into three types: 1) law and policy; 2) curriculum and campaigns; and 3) technological responses. By employing narrative inquiry to elicit user-generated cyberbullying narratives and design solutions, this study explored the third category of interventions, technological responses.

3 Methods
The study used narrative inquiry, a qualitative methodology that is most commonly employed in educational research, to uncover user-generated design interventions for preventing or countering cyberbullying behavior. Defined by Connelly and Clandinin (2006) as “the study of experience as story” (p. 477), narrative inquiry research is characterized by the narrative serving as both the object of study and the method (Connelly & Clandinin, 1990). Our implementation of narrative inquiry departed from the more common use of this approach as a means to unveil participants’ lived experiences. Aware of the pervasiveness of cyberbullying as a societal problem, our intention was not to have our participants divulge personal, sensitive, or traumatic stories to us and their fellow participants. We instead probed our participants’ perceptions of the cyberbullying experience as they imagined it would be for someone else, asking them to tell us a story about “mean and cruel behavior online.” Bowler et al. (2013) provides a further description of the procedures used in this study and narrative inquiry as a methodology.

During the Spring 2012 term, we conducted two storytelling sessions on campus with nine participants. The first session was with five undergraduate students – five females in their early 20s, all from the University of Pittsburgh. The second session, held one month later, was with four teens: one girl and three boys between the ages of 14 to 17. Convenience sampling was used to recruit participants for this study. Parental consent, as well as assent from the teens, was received just prior to the start of the storytelling session with teens. Parents did not stay in the room during the session.

The storytelling sessions took place on two Sunday afternoons in a classroom at the University of Pittsburgh’s School of Information Sciences. The first session with the undergraduate students lasted three hours, with equal portions allocated for sketching and group discussion. The five undergraduates were divided into two smaller groups – one group with three participants, the other with two. In order to better accommodate our teenage participants, we shortened the protocols with the teen group, focusing more on the sketching and less on the group discussion before and after sketching. This second session with teens ran for two hours. The four teens were also divided into two smaller groups – one with two boys and the other with one girl and one boy. In both storytelling sessions, there were three investigators in the classroom. Two investigators interacted with the participants while third investigator observed and took notes. Informed by Marwick and boyd’s (2011) findings, we aimed to allow the young people to use their own words to label the roles and events in their narratives of online conflict, rather than insert the language that is commonly used in mass media and which tends to reflect an “adult” point of view. We asked them to simply tell us a story that depicted “mean and cruel behavior online.” In total, the study resulted in four narratives telling the story of bullying in social media environments and a set of accompanying design interventions that might alleviate or even intervene in mean and cruel online behavior.
Figure 1: A cyberbullying narrative. Sticky notes indicate places where a design intervention should be inserted into the narrative.

4 Results

Four “cyberbullying stories” were constructed by the participants, two by the teen group and two by the undergraduate student group (Figure 1 depicts one of the cyberbullying narratives by an undergraduate group). While we asked the participants to frame their thoughts about cyberbullying around a narrative, their stories do not necessarily follow a traditional narrative arc, with a proper beginning, middle and end that is usually in the shape of a resolution of a problem. In fact, we found several of the stories were quite post-modern in their messy storylines and ambivalent conclusions, a reflection perhaps of the very nature of social media. The complex web of relationships and “storylines,” shifting roles, and ever-morphing outcomes in social media seems to preclude a neat and tidy ending.

After the participants had generated their visual stories illustrating ‘mean and cruel behavior online’ they were then asked to think about design interventions that would discourage or prevent such behavior. The participants wrote their ideas on sticky notes and then stuck the notes at the point in their story at which the design interventions were supposed to work. Present in the narratives and discussions of the narratives that 'could be' are the following design themes: design for reflection, design for consequence, design for empathy, design for personal empowerment, design for fear, design for attention, and design for control and suppression. Many of the design features suggested by the participants elicit a range of provocations and thus find themselves classified under more than one theme. Table 1 below describes the design themes and associated features that the teen and young adult participants highlighted.

<table>
<thead>
<tr>
<th>Design Themes</th>
<th>Design Principles</th>
<th>Design Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design for Reflection</td>
<td>Design that creates a pause, slowing users down so they can consider the ramifications of their actions.</td>
<td>Pop-up warning about cyberbullying timed to last for ten seconds so that users can stop and think. Alert boxes with reflective questions anytime one clicks “like”, asking the user “why do you like this?”</td>
</tr>
<tr>
<td>Design for Consequence</td>
<td>Design that ensures that there are consequences for bullying behavior.</td>
<td>Public shaming through a “bully button”. Facebook-imposed restrictions as a punishment for bullying behavior. Reports of inappropriate online behavior sent to perpetrator’s school.</td>
</tr>
</tbody>
</table>
Design for Empathy
Design that can make pain and sadness concrete, allowing bullies and their followers to see how victims suffer.
Design affordances such as sad music and emoticons. Design features that create a more emotive social media environment.

Design for Empowerment
Design features that redress an imbalance of power.
Adult interventions figure largely in this design feature. The system facilitates adult interaction, thereby lending the power of adults to the victim. Adults post supportive messages or warn the bullies that adults are watching.

Design for Fear
Design that harnesses the power of fear.
A “bully button” and the use of personalization, both of which send the message that “you’re being watched”.

Design for Attention
Design that catches the attention of bullies.
Anti-bullying messages that are prominent, loud, personalized, and even irritating. Bright colors should be used.

Design for Control and Suppression
Design that would trigger the suppression of content either by Facebook administrators or through an algorithm.
The system alerts Facebook staff when there are too many “likes” within a short period of time (a clue that something is going viral), resulting in the removal of offensive and cruel content. Facebook-imposed filters for offensive words.

Table 1: Participants’ Design Recommendations.

4.1 Design for Reflection
In this study, the suggested design interventions and the accompanying explanations showed that some participants were aware of the value of reflection in countering cyberbullying behavior. One group, for example, suggested pop-up messages timed to last for ten seconds (in other words, users cannot close the pop-up dialogue box until after ten seconds). The group explained that during the ten second delay, social media users would have a chance to read and process the pop-up warning about cyberbullying. While not all users would use this mandatory pause to reflect thoughtfully on their own feelings and values, this design feature would afford them the opportunity to do so, at least according to some of the participants. Interestingly, the downside of this design feature (that it is an irritating intrusion) was not explored. A similar suggestion came from the members of one undergraduate group, who proposed the inclusion of alert boxes with reflective question prompts that would appear each time a Facebook user attempts to perform an activity on the site. When asked what would trigger the alert box, one of the participants explained, “Just anytime you click 'like' or something. It’s like, 'why do you like something?' or 'why do you like this?" Presumably, during this time of prompted reflection, a potential bully could reconsider a posting a harmful message or image or a potential bystander could reconsider clicking the “like” button in response to a negative comment and joining the bullying fray.

4.2 Design for Consequence
In general, the cyberbullying narratives were absent of consequences for the bully. Three of the four stories had ambiguous endings and in none of the stories were the “bullies” shown to pay consequences for their behavior. The participants expressed a view that in order to solve the problem, it is vital to spread awareness that there would be certain consequences for individuals engaging in cyberbullying. Among the consequences identified by the participants were public shaming, getting in trouble with the school (and having to visit
the guidance counselor or principal), getting in trouble with the police and the law, and imposed Facebook restrictions.

Perhaps the most provocative suggestion was the addition of a bully button, to be activated when a damaging photo of the bullied individual was actively shared and liked. It would allow people to flag a bullying situation with comments such as, “REALLY mean comment”. The participants thought bullies would avoid the public shaming that would occur through the accumulation of “bully” points. Social media users, they suggested, would be afraid of being labeled a “bully” and avoid this consequence by refraining from engaging in acts that may be perceived (and publically dubbed) as bullying behavior. The participants believed that in avoiding this consequence, the users would refrain from engaging in cyberbullying behavior. One of the undergraduate participants compared the Bully Button to the act of “liking” something on Facebook, explaining “Like you can see how many people like something, you can see how many people bully button a comment.” She continued to say, “It’s kind of embarrassing…If I have twenty bully buttons next to my comment, it’s like ‘you’re a big jerk.’ ”

4.3 Design for Empathy

The young people in this study felt that it was important for bullies and their followers to see how victims suffer. They assumed that a design feature that could make pain and sadness concrete would lead to bullies self-regulating their bullying behavior. Thinking specifically about Facebook, one group suggested more use of sad posts, sad songs, or emoticons by the victims on Facebook, believing that bullying wouldn’t “happen if the bullies realized they’re wrong…” This realization would come if “bullies can actually see that they’re causing this pain…” While Facebook posts, sad music, and the use of emoticons are not new design ideas, it does suggest the need for a more emotive social media environment and begs the question as to whether there is a better way to bring these elements to the fore, to make them more apparent and impactful to others viewing the social media spaces of those who are bullied.

4.4 Design for Personal Empowerment

An imbalance of power seems to be inherent in cyberbullying and the participants suggested that the bullied victims need some power of their own. Bullies and their supporters have on their side the wild encouragement from within their circle, the anonymity of their acts, a lack of consequences, and the speed of networked, digital information for disseminating online bullying behavior. What power does the bullied have and how can the victim and his or her defenders’ power be embodied in design? Anonymity was a key theme here. Showing support for the bullied is predicated on the supporter having the safety of anonymity. Anonymity is therefore power. The "Bully Button," for example, would allow for observers to come to the defense of the victims by anonymously calling attention to bullying behavior.

To young people, adults often represent power, but adults took an active role in defending the bullied in only one of the narratives. Interestingly, the anonymity of adults was not seen as beneficial. Indeed, knowledge that adults can lend support to the victim was exactly the point. One group of teens described a cyberbullying story where, rather than deal head-on with bullying, adults would instead “like” Joe, the victim, in order to protect his self-esteem. They would also respond directly to the bullies with posts in Facebook like, “That’s not very nice – I don’t appreciate you saying something mean about him.” By sharing and extending their power into the social media environment, adults would empower the victim.

4.5 Design for Fear

The participants made several suggestions that seemed to harness the power of fear as a means to deter a bully and, in doing so, support the victim. Not to be confused with Design for Reflection (design that affords the cognitive practice of thoughtful introspection), Design for Fear activates mechanisms that might cause social media users to feel caution before acting. The proposed design features seem to suggest that,
at least according to the young people in this study, social media that affords fear might actually be a useful feature.

In one of the teen group's stories, the victim's defender turned into a bully and the original bully into a victim. To deal with this never-ending saga of fighting, the boys in this group suggested that pop-up messages should appear in Facebook with an ominous warning: “Stop bullying today or you could be next, Ricky” (Figure 2). This cryptic language, directed at a specific target, sounds more like a threat than a warning, and draws upon fear as a potential prompt for hesitation. It was not explained exactly how the system would know that the person is bullying, but more interesting was the use of personalization to get the user's attention. Personalization, as the teens explained, gets your attention because you know the message is designed for you. Clearly, you are being watched and knowing that someone is watching you is frightening.

4.6 Design for Attention

Participants suggested that important anti-bullying messages need to be prominent, loud, and even irritating, in order to be noticed by the “bullies.” Personalization would also catch the attention of people. The threatening message, “Stop bullying today! Or you could be next, Ricky,” would work not only because of the fear it would engender, but because it was “tailor-made.” As a male teen explained, “If it says your name in there, you’re definitely going to notice it...Like, you’re going to read it and be like ‘Oh wait, that says my name.’” There was little concern about what personalization would mean for privacy.

4.7 Design for Control and Suppression

The participants proposed designs that would trigger the suppression of content either by Facebook administrators (some participants assumed that Facebook employees, not algorithms, were making decisions about individual posts) or through an algorithm. Three groups mentioned filtering for “rude words.” One group of undergraduate females in their early 20s, was more specific about the kind of language that should be filtered, saying words like “slut” and “whore” should automatically be flagged and reviewed by Facebook. This same group also thought that images should be actively filtered by Facebook. An image that received 200 “likes” in a short period of time should raise an immediate red flag for the social media system, since (at least according to this group’s cyberbullying narrative) mean and cruel behavior happens in a rapid surge of online postings. While one group of undergraduates thought there should be an algorithm built into the system to catch certain language and trending images, the system response they suggested was entirely human - someone at Facebook should to look at the flagged messages and images and make a determination as to their appropriateness. As one participant said, “If there are more than like X number of 'likes' on a certain picture, the Facebook staff could look at it and be like ‘ok, we’re going to look at this and decide is it a good thing or a bad thing.’”

Figure 2: Design for Fear/Design for Attention: A personal message sends a threat and gets a bully’s attention: “STOP BULLYING TODAY! Or you could be next, Ricky”
5 Discussion and Conclusion

The participants in this study took their task seriously, suggesting a wide range of design themes - design for reflection, design for consequence, design for empathy, design for personal empowerment, design for fear, design for attention, and design for control and suppression. Collectively these design themes give the bullied and their supporters a range of active and passive tools that might pre-empt or push back against mean and cruel online behavior. It seems, at least to the participants, that well-being in social media environments is integrally linked to a balance of power between the bullied and the bullies.

Several of the design themes offer a clear acknowledgement of the important role of the bystander and the inherent social nature of bullying in social media (Design for Consequence, and Design for Personal Empowerment, for example). According to the narratives in this study, it is simply not possible to disentangle the bully, the bullied, and the crowd that “watches” the story unfold. These are fluid, changing roles. The message from the participants is that if we are to design for well-being, at least in terms of social media, we need to move from a view of bullying as a dichotomous relationship between the victim and bully, toward a broader, community-wide conception of the problem. This may seem obvious, given that social media is inherently social, but many approaches to the prevention of cyberbullying seem to focus overwhelmingly on the relationship between the victim and bully – who bullies, who gets bullied, what to do, how to avoid, and who to talk to – and less so on the active roles that bystanders can play.

Power is a pivotal theme that weaves its way throughout several of the design features, suggesting that well-being in social media environments is integrally linked to a balance of power between the bullied and the bullies. Empowering the victims by giving them tools to push back (some might say tools that give the bullies a taste of their own medicine) seems, according to some of the participants, important to healthy social media spaces. While not raised by the participants themselves, an interesting consideration for designers would be the consequences of social media spaces that even out the balance of power: Would this design recommendation result in a “cold war” style standoff or an escalation of bullying?

As with all design, there is the intent and the unintended consequence. The participants, laser focused on the task of ending cyberbullying, did not seem to consider the downside of their design suggestions. While this may have reflected the shape and protocols of the study, it may also reflect a gap in their understanding of the complexity of social media and the very idea that design has consequences. For example, the Bully Button, an aggressive tool for pushing back against the bullies and their henchmen, could quite easily lead to more bullying. Like the scarlet letter “A” that branded Hester Prynne an adulterer or a stockade in a village square that marked an individual a criminal or sinner, the designs for a Bully Button would allow for the punishment of the individual by the community.

Several of the design features suggested by the participants clearly targeted affective aspects of design (Design for Empathy and Design for Fear in particular) and raise some interesting questions. When designing for empathy, the participants assumed that showing the suffering of the bullied would lead to an empathetic response from the bullies and result in end of bullying. This is not always the case. It is important to note that there are two forms of empathy – cognitive and affective empathy (Karem et al., 2001). Cognitive empathy is knowing how other people feel while affective empathy is sharing other people’s feelings (or, as the saying goes, feeling their pain). Bullies can know about someone else’s pain but that may not change behavior if the bully or bullies simply don’t care. Bullies can have cognitive empathy and use it in Machiavellian ways to bully their victim in more perfect ways. So the design problem might not be to design for empathy but to design for caring.

Design for Fear raises another concern. Fear is a powerful emotion and, in an evolutionary sense, humans are designed for fear. But there is an enormous amount of discomfort associated with fear, which begs the question: how does the system (in our study, Facebook) balance the fears and anxieties of the bullied with those of the bullies? At what point does causing someone to hesitate become fear-mongering and harassment and thereby derail the point of design for well-being? At what point does this design cease
to create a healthy environment for users of social media? This question was not raised by the participants in the study.

Perhaps the antithesis of designing for fear is Design for Reflection. This includes design interventions that encourage quiet and introspective thought. Akin to Hallnäs and Reström’s “slow technology” (2001), Design for Reflection is meant to encourage mindful, self-aware online behavior. It is interesting that while both the teen and undergraduate groups said that social media users should take steps to pause and think before acting, it was only the undergraduate students (young people in their early 20s) who suggested a specific mechanism (pop-up question prompts) in their first attempt at the “design” activity. While the younger participants were aware that users should stop and think before acting, they initially did not identify a specific strategy to enable such behavior. It was only with deeper probing by the investigators that one teen group did ultimately suggest a design feature, the pop-up warnings about cyberbullying. This may point to a developmental or experiential difference vis à vis the use of social media. Future research on designing for well-being should explore the relationship between age, experience, and design recommendations.

In this study we asked participants to think about what social media might look like if it was designed to prevent or intervene in mean and cruel behavior. The design interventions suggested by the participants reflect their lived experiences, perceptions, and values related to social media environments. We are not arguing that all the design interventions suggested by the young people in this study be actualized into real designs. Rather, what we present here is a window into these young people’s fears, anxieties, values and ethics related to mean and cruel online behavior and their proposals as to how to curb it. Young people have an earnest desire for social media environments that encourage wellbeing. Designers should take from this framework lessons about the expectations that young people hold in relation to social media and build social media environments that reflect these needs.

6 References


7 Table of Figures

Figure 1: A cyberbullying narrative. Sticky notes indicate places where a design intervention should be inserted into the narrative..............................................156

Figure 2: Design for Fear/Design for Attention: A personal message sends a threat and gets a bully’s attention: “STOP BULLYING TODAY! Or you could be next, Ricky”.................................159

8 Table of Tables

Table 1: Participants’ Design Recommendations..........................................................157